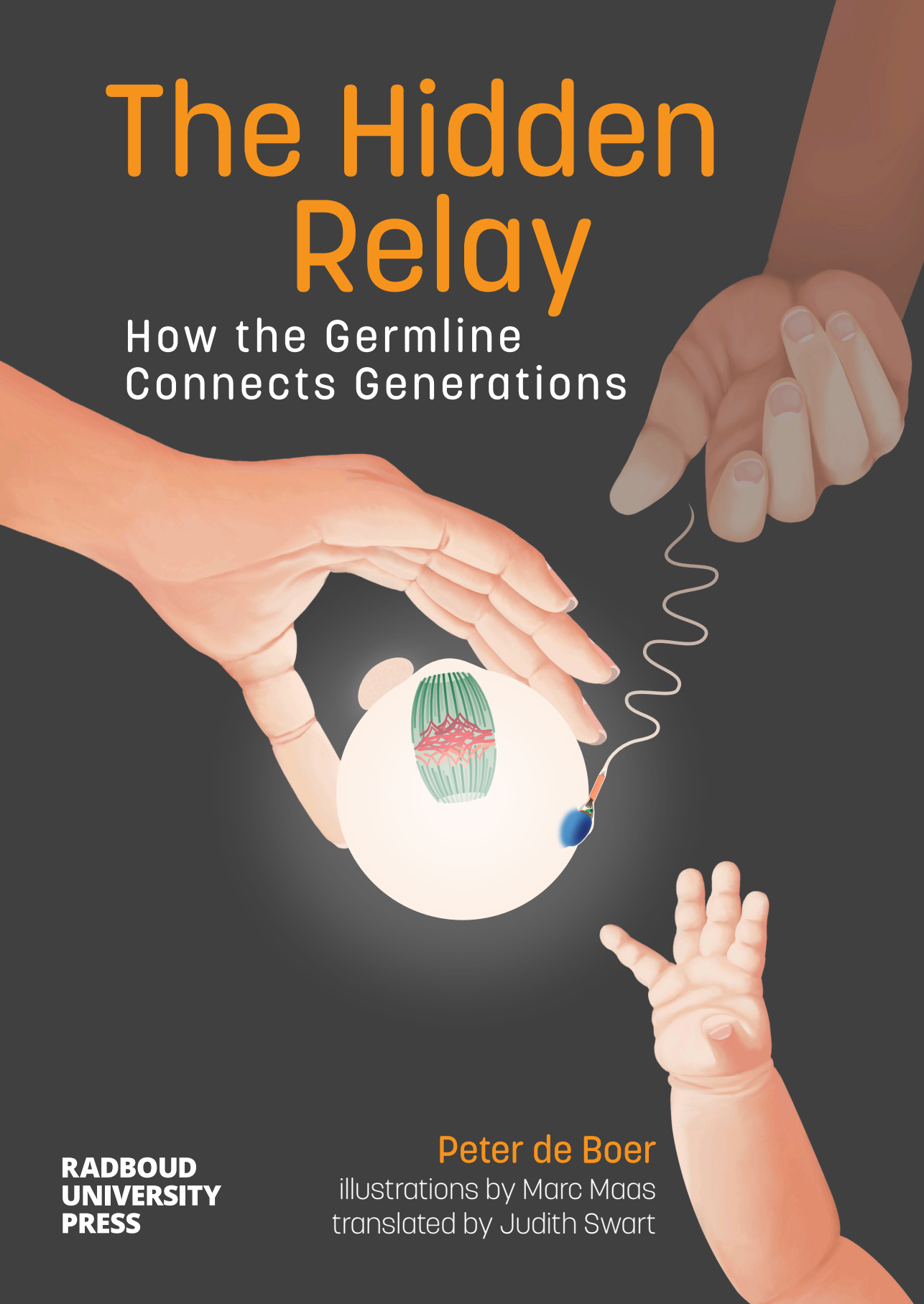
The Hidden Relay

How the Germline Connects Generations



Peter de Boer
illustrations by Marc Maas
translated by Judith Swart

RADBOUD UNIVERSITY PRESS

The Hidden Relay

The Hidden Relay

How the Germline Connects Generations

Peter de Boer
Illustrated by Marc Maas
Translated by Judith Swart

Financial support for the figures was received from the Dutch Society for Clinical Embryology (Vereniging voor Klinische Embryologie, https://www.embryologen.nl) and from the Simons Foundation Fund (https://www.nvhg.nl/simonsfonds/)

VERENIGING VOOR
KLINISCHE EMBRYOLOGIE

SIMONSFONDS

The following organisation and individuals are gratefully acknowledged for their contributions to the cost of translation: Mouse News Letter Ltd, Andrew and Jeremy Searle

The Hidden Relay. How the Germline Connects Generations
Published by RADBOUD UNIVERSITY PRESS
Postbus 9100, 6500 HA Nijmegen, the Netherlands
www.radbouduniversitypress.nl | radbouduniversitypress@ru.nl

This is a translated and updated edition first published as: *De verborgen estafette. Hoe de kiembaan de generaties verbindt*. Nijmegen: Radboud University Press, 2023

Illustrations: © Marc Maas
Translation: Judith Swart
Design: Marc Maas and Textcetera, The Hague
Print and distribution: Pumbo.nl

ISBN: 9789493296527
DOI: 10.54195/ZQDR3182

Free download at: www.radbouduniversitypress.nl

RADBOUD
UNIVERSITY
PRESS

Table of Contents

Introduction

A semi-popular science book on reproductive genetics: why, you may ask? Because reproduction is the central theme in both biology and evolutionary biology.

How can a species persist? The relay baton in the form of the fusion of the egg and sperm cell creates a new generation that gets to run another lap and pass on the baton. What lies behind this metaphor, and how does it work? What are the implications of this relay race for a species? The field of reproductive genetics seeks answers to these questions.

The themes of reproductive genetics are at the heart of the big questions of life: where do I come from, who am I, and where am I going? Issues that are not only relevant for any particular individual in the present but that also greatly involve past and future generations. The thought that the germline has never been interrupted in the time that came before us can truly make your head spin; in a way, we are as old as life itself.

In reproductive genetics, several disciplines within biology come together, including cell biology, molecular biology, genetics, cell genetics, molecular genetics, but also quantitative and population genetics. Expanding the scope slightly will also bring anatomy, endocrinology and behavioural biology into the picture. Here, we aim to map the field in such a way that high school biology knowledge (in the Netherlands: at VWO level) combined with interest, curiosity and (probably or most likely) professional involvement should be enough to read this book.

The first part (chapters 1-6), serves to refresh some basic knowledge. For this, we have drawn from textbooks such as *Essential cell biology* by Alberts et al., (2022) and *Introduction to genetic analysis* by Griffith et al., (2020).

The book then takes you on a journey through the germline, from the fertilised egg through the first embryonic cleavage divisions, to the segregation of the cells that will later give rise to

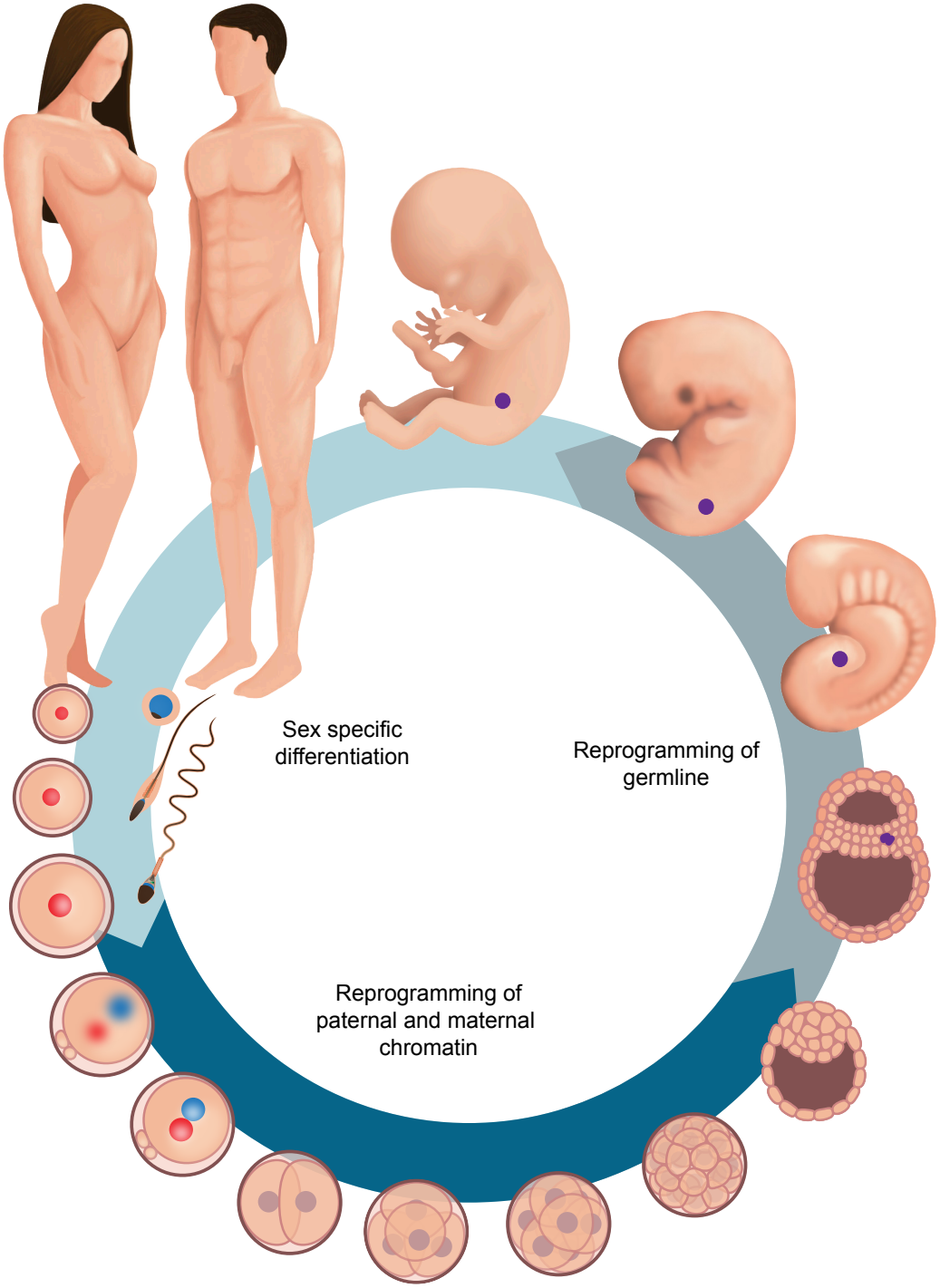

Sex specific
differentiation
Reprogramming of
germline
Reprogramming of
paternal and maternal
chromatin

FIG 1

An impression of the **germline**, illustrated in humans. **Gametes** are specialised cells and at the same time, they enable new life that can continue indefinitely. For this, it is necessary that the **chromatin** is reprogrammed (or "reset", Ch9) from the beginning of fertilisation; a new round of development begins. However, this is not enough for the germline. When the new generation of **germ cell**s, the **primordial germ cell**s, is defined just after implantation of the **embryo** (purple, fig 43), a second round of reprogramming follows, during which origin-specific chromatin markers (indicating origin from the father or the mother) are also removed. This is covered in chapter 10, "**genomic imprinting**". The chromatin to which this applies, containing a small but important group of genes, is later returned to the imprinted state. That happens during **oogenesis** and **spermatogenesis**, at which point it matches the carrier's sex. In this way, it is determined which parent's genes that are involved in this, will be "switched on" in the next generation, the other copy being inactive.

mature eggs and sperm (chapters 7-18). The acquired insights are based on a range of research techniques. Our understanding of the production of egg and sperm cells, the gametes, has greatly improved in the last ten years. In this field, analyses at the single-cell level have become increasingly important. It is another area where technological development drives research. Our understanding of what needs to happen in the germline to activate and, in particular after fertilisation, deactivate genes will continue to improve, as well as our understanding of the involvement of specific genes in these processes.

How are the characteristics of both parents combined after the fusion of the egg and sperm and the subsequent development of the embryo? How often do changes occur in the DNA and possibly in the factors controlling genetic material? What causes these changes, and is this normal? How does the clock reset with the creation of each new generation? What are the influences of artificial reproduction (IVF) on these changes in and around DNA? In order to get a grasp on all these subfields of reproductive genetics, I have extensively reviewed literature from the past ten years. Not comprehensively, which is simply not feasible, but rather guided by over 40 years of teaching and research experience in this field. The literature is an integral part of that. At times, the insights obtained from scientific production are widely shared within the community, while at other times, they may be more personal. The latter might be considered

less scientifically sound, but it is still perfectly legitimate because of the incredible complexity of cells and organisms and the interactions with their environments. There is still much to be investigated and to be learned in this regard. Besides, we can distinguish “splitters” and “lumpers”. “Splitters” follow well-established paths. They unravel every component of the cell, down to its smallest parts, and then try to determine the function of each component, e.g. a protein, and sort of prove this. “Lumpers”, on the other hand, attempt to integrate everything into a whole, pursuing a holistic view.

On a higher level than the molecular focus, there is often a situation of accumulating evidence. Observations seem to support a certain way of thinking. Ultimately, the educated guess, the lowest level of scientific certainty, can also be of value.

The aim of this book is to turn the complex events in the germline, which are repeated from generation to generation, into language. “Why” questions will be inevitable here. These are often tricky in biology. “Nothing in biology makes sense except in the light of evolution,” posited Ukrainian-American geneticist, entomologist and evolutionary biologist Theodosius Dobzhansky, who worked with fruit fly species, towards the end of his life in 1973. All the more reason to confine the area in this book to mammals such as mice, occasionally cows and pigs, but with special attention to humans. Research tools, methodologies and outcomes of scientific experiments are rapidly expanding in scope. This brings clarification but also opens up new avenues of wonder. Therefore, in my opinion, much of Dobzhansky’s statement is still true.

Genetics certainly comes with a specific language. Concepts need to be well defined, and a “genetic dictionary” already existed in 1972. Nowadays, you can find all these concept descriptions or definitions in the glossary at the end of textbooks. This is the approach I use here as well. The first time concepts are mentioned in a chapter and figure/table legends, they are highlighted in **bold**. These definitions are, as much as possible, based on the previously mentioned scientific textbooks by Griffiths and Alberts. In the bibliographies provided at the end of the book, listed by chapter and text box, those publications that are important and/or inspiring are marked with an asterisk. This also includes the publications used for the figures and tables.

The historical development of specific insights is certainly covered in the chapters. This can be informative, as applications in biotechnology often tend to precede the biological-genetic understanding of concepts. Outcomes that differ from the line of thought also occur quite frequently in the so-called translational sciences.

This book is not primarily written for direct use in education, but it may very well be used for this purpose, bearing in mind that, for example, through the increasingly wide use of DNA sequencing, the parameters for genetic variation and mutation in the population will become more and more precise. This is just one example of the technological revolution, mainly in cell biology research. It results in large databases, the spin-off of which is not always easy to fully realise and convey. The illustrations are a good starting point, and a course reader for students could be created alongside the book.

In the late 1960s, my reproduction-oriented research activities started with cows, then pigs and mice; I was an Animal Sciences student in Wageningen, the Netherlands. From 1971, the focus was mainly on mice for many years, until humans entered the picture in the late 90s. Eventually, I focused only on our own species. This was, of course, because reproductive genetics nestled itself in society mainly through the growing significance of *in vitro* fertilisation, and research questions and possibilities focused on practical applications expanded. But this was not the only reason. Take one of the quality newspapers and cut out anything that relates to the subject. Many weeks, relevant articles are found with, for instance, a peak in coverage around the public release of the first genetic changes made to the human germline via the CRISPR-Cas technique, in November 2018. However, questions also arise due to people postponing the realisation of their desire to have children, with ageing playing a crucial role in this. And what is going on with sperm production? These topics make it into the columns of our quality newspapers. Via the widely discussed declining trend in male fertility, you indirectly enter the field of reproductive toxicology, which is not only relevant for sperm development but most likely also for egg development. This book also makes small side trips towards the identification of distant relatives and into paleobiology, the reconstruction of our recent evolutionary past. On top of that are the ever-increasing ambitions in the world of artificial reproductive technologies, in which a parallel with Charles Mann's book *The Wizard and the Prophet* (2018) can easily be discovered. In that book, the biotechnologist stands as a wizard opposite the prophet, who takes spontaneous fertility as a starting point.

Reproductive genetics is a field with many faces and aspects, the majority of which have crossed my path at some point since 1968. What is the image that emerges when you try to connect these aspects as much as possible? Does that lead to a greater biological understanding? Exploring this is the real intent of this book.

Sometimes, a picture says more than 1,000 words. Marc Maas, well-acquainted with genetics and a scientific illustrator, has been eager to use his talent to fill the gap between scientific textbooks and popular scientific writing. Judith Swart has taken great care to translate the text up to the level of the meaning behind the sentences. We wish you an entertaining and inspiring journey through the germline, with an open and curious mind to all the marvellous things that take place there.

CHAPTER 1

Impression of a cell

The body consists of an astronomical number of cells, occurring in several hundreds of cell types. For a human weighing 75 kg, this number is estimated to be 37.2 x 10^{12}. One special group of cells from this enormous population takes centre stage in this book. These are the cells of the **germline**, which produce the **gametes** (eggs and sperm) that connect generations. These cells are already present during **embryo**nic and foetal development. After birth, during different stages of growth and in adulthood, they assume different forms, which is necessary to continuously connect generations through time with an almost endless number of repetitions. The cells of the male and female germline naturally have their own functions, but in certain aspects they undergo similar processes, in which the egg becomes the largest cell in the body, and the sperm the smallest. While an average **somatic** cell is about 3,000 μm^3 in size, a sperm cell is 10 times smaller, and an egg cell is 1,300 times larger (a μm is 1/1000 mm).

Despite years of research, we still do not understand everything about the processes that make cells function as the smallest independent living unit of our tissues and organs. We also still know surprisingly little about the mechanisms that cells use to remain vital and to respond flexibly to the constantly changing conditions in their environment. In his valedictory lecture, delivered on June 1st, 2018, Bé Wieringa, professor of cell biology at the Radboud University Medical Centre (Radboudumc) in Nijmegen, extensively addressed not only many well-known aspects but also the gaps in our knowledge surrounding cells. He also provided a glimpse into possible directions of future research. His presentation began with him showing us the chemical formula of the composition of the human body: this can be written as C 10^{27}; H 10^{27}; O 10^{27}; N 10^{26}; P 10^{25};

S 10^{24}; Ca 10^{25}; K 10^{24}; Cl 10^{24}; Na 10^{24}; Mg 10^{24}; Fe 10^{23}; F 10^{23}; Zn 10^{22}; Si 10^{22}; Cu 10^{21}; B 10^{21}; I 10^{20}; Sn 10^{20}; Mn 10^{20}; Se 10^{20}; Cr 10^{20}; Ni 10^{20}; Mo 10^{19}; Co 10^{19}; V10^{19}.

With a bit of math, this brings you to about 10^{15} atoms per cell, chemically linked in an unbelievably large number of different molecules, which together define life. Besides the ubiquitous carbon, hydrogen, oxygen and nitrogen, there is an abundance of phosphorus, sulphur, calcium, magnesium, sodium and potassium. The elements that are present in smaller amounts are called trace elements. The fact that they are not as abundant does not diminish their importance. For example, zinc (Zn) and selenium (Se) are often associated with the quality of sperm (Ch17.4).

Figure 2 shows a somewhat artistic and extremely simplified impression of a cell. Firstly, the prominent boundaries around the cell itself and around the cell nucleus stand out in this figure. These boundaries consist of membranes formed by a tight double layer of different types of lipids, with phospholipids, glycolipids and cholesterol being the most common ones.

Membranes play a role in many cellular activities; their main task is not only to form a barrier but also to enable communication and transport between the interior and the exterior environment. The nucleus, a storage place for our genetic material (the **DNA**) and the centre of organisation and control of the cell, even has a double membrane with special adaptations for this purpose. To facilitate transport between the nucleus and the rest of the cell, this nuclear envelope has pores (fig 2) through which even large molecular structures can be transported from the nucleus into the cell. Around and inside the cell, single membranes are mostly found, as seen in the **organelle**s, compartments with specialised functions. Every cell has several types of those organelles. Among them are the endoplasmic reticulum (fig 2), involved in the synthesis and modification of proteins; the Golgi apparatus (fig 2), working as a kind of sorting and transport station of the cell (and also involved in lipid **metabolism**); and the peroxisomes and lysosomes. The latter two are responsible for the conversion of reactive oxygen species and waste disposal in the cell, respectively. Remarkably, the mitochondria (fig 2), the powerhouses of the cell, which are also involved in other metabolic conversions and even in the regulation of cell death (**apoptosis**), are enclosed by a double membrane.

The shape and function of the cell are further supported by an internal skeleton with highly flexible properties (fig 2). This skeleton controls the dynamics and movement of the cell and plays a particularly important role in cell division. Each of these parts of the cell has already been extensively studied using a variety of techniques.

Nevertheless, our knowledge of what happens in the liquid and solid environments inside the cell – the **cytoplasm**, nucleus and organelles – is still far from complete. In order to keep a cell alive, there must actually be a continuous state of slight imbalance, which can be gauged by the concentration and behaviour of those huge numbers of molecules (**RNA**s, proteins, lipids, sugars, ions, etc.). Consequently, there is permanently a huge dynamic in the molecular composition within and between cell compartments, a dynamic that is also evident in the communication between the external and internal environment. The resulting complexity is almost incomprehensible, and "frozen" illustrations of the location and the often enormous amounts of different types of molecules, as shown in figure 2, are, of course, a radical simplification of reality.

It is crowded in a cell, incredibly crowded. This is known as molecular crowding. Just keeping a cell functional and vital requires the expression of a large number of genes. Recent estimates indicate that the RNA and protein products of ~3000 to ~7000 of the approximately 20,000 protein-coding **gene**s in humans (and other mammals) are involved in sustaining the basic functions of each cell. Each protein-coding **gene** will lead to the synthesis of several types of **messenger RNA** (box 1), and these mRNAs are translated into proteins by ribosomes, sometimes up to hundreds or thousands of times (fig 2, box 1). For this purpose alone, each cell possesses an enormous synthesis capacity, in the form of massive numbers of ribosomes in the cytoplasm and on the endoplasmic reticulum (fig 2). The total number of mRNA molecules is estimated at more than 3×10^5, resulting in an average production of roughly 10 billion protein molecules within each cell after **translation** (fig 2, box 1). About 10% of these are eventually found in the cell membrane. However, the picture presented here only provides an impression of the average process. Depending on the cell type and its specialisation, cell-specific gene products are also produced.

The numbers of RNAs and proteins mentioned here can thus vary significantly and can even be many times larger. The complexity of all of this expands not only in terms of quantity but also in terms of the types of proteins, because a very large percentage of our genes can encode multiple protein variants (including by so-called alternative mRNA splicing, box 1, fig 9). Therefore, the aforementioned 20,000 genes in our hereditary material carry the information for a multitude of different protein types. Elsewhere in the **genome**, codes are found for specific classes of small proteins (up to about 100 amino acids long). Not all of them are known yet, but they are of biological importance. Moreover, many proteins are often further modified through the either temporary or permanent attachment of various chemical groups in order to control

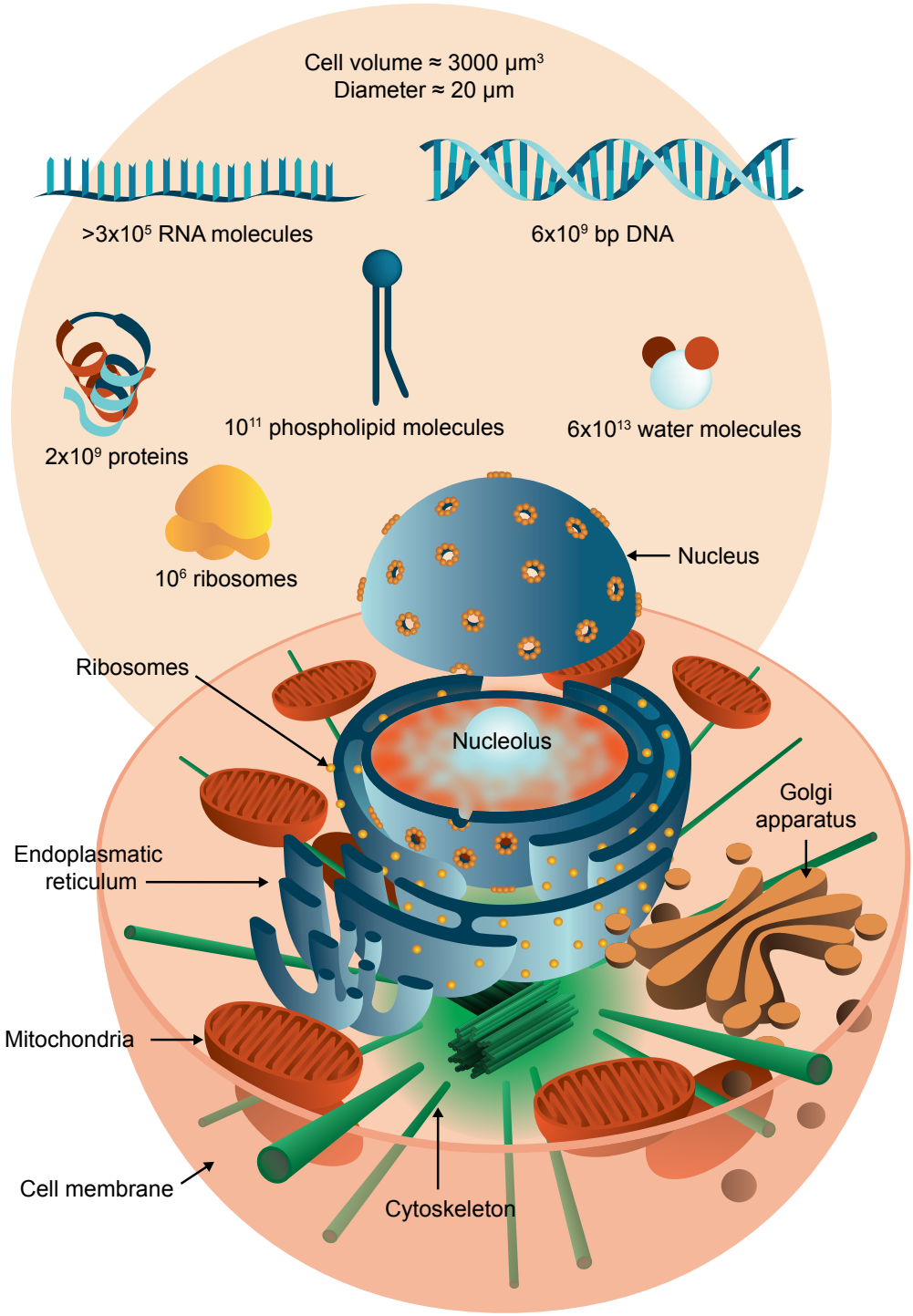

Cell volume ≈ 3000 μm3
Diameter ≈ 20 μm
>3x10^5 RNA molecules
6x10^9 bp DNA
10^11 phospholipid molecules
6x10^13 water molecules
2x10^9 proteins
10^6 ribosomes
Nucleus
Ribosomes
Nucleolus
Golgi apparatus
Endoplasmatic reticulum
Mitochondria
Cell membrane
Cytoskeleton

FIG 2

Highly schematic representation of a cell. In the nucleus, the brown **chromatin** has a higher density than the light blue (see also fig 7). The **nucleolus**, the nuclear body in which the ribosomal **RNA** accumulates after **transcription** (and where it does have a high density), is indicated. The ribosomes on the endoplasmic reticulum are yellow. The depiction of the cytoskeleton in green is highly schematic and incomplete. The large number of phospholipid molecules underlines the importance of membrane structures for the cell.

their function or life span. Well-known examples out of the more than 50 reactions known for this purpose include phosphorylation, methylation, acetylation, ubiquitination, glycosylation, hydroxylation and sumoylation. Many proteins do not function on their own and can only act when they are linked with other proteins, forming joined complexes. Their function is often only temporary. This dynamic cooperative behaviour, and thus the biological functioning of the whole complex, is constantly regulated by the addition and removal of the chemical modifications mentioned above (which are, in turn, controlled by the enzymatic activity of other proteins). Another way to highlight the mind-blowing complexity of a cell is by pointing out that approximately 10,000 different organic and inorganic small-molecule substances may be involved in all the chemical processes that collectively control the cell's metabolism.

Collectively, this unimaginable amount and variety of molecules ensures that coordination and balance are maintained in the activities of the cell, including equilibria within and between cells. This natural physiological equilibrium is also referred to as "**homeostasis**". In the regulation of this homeostasis, the margins can sometimes be quite wide. While strict rules seem to exist for the timing and quantity of production of some gene products, the generation and use of many RNAs and proteins can sometimes seem like a rather random process. Even when a cell becomes dysregulated and stressed as a result, it is not necessarily immediately fatal. In fact, cells have various ways to defend themselves against stress, allowing them to regain their balance. Excessive uniformity in cellular actions and interactions with the environment is not even good. Without any subtle disturbances of balance and gradients of molecular concentrations (between microcompartments in the cell and between the interior and exterior), nothing would flow or move, and a cell would be stone-dead. There is constant action and movement, and that makes cells alive!

Driven by technology, the insight just described has led to major changes in cell biology research. For example, the physico-chemical environment is increasingly being studied, down to the tiniest corners of the cell. This involves investigating the response to mechanical forces, the production, distribution and use of energy-carrying compounds, and the peculiar interactions that molecules can have with each other locally in the cell, where the incredible density of molecules leaves no room for water. It has recently been discovered that in such environments, “crowd control” is needed to manage the density, similar to the measures taken to manage the COVID-19 pandemic. New molecular and microscopic techniques also allow for the observation of entirely new types of associations between RNA molecules and proteins. Here, a so-called phase transition from liquid to solid occurs, and small-scale aggregates or membraneless organelles (“biomolecular condensates”) are formed. The chemical reaction conditions inside these organelles are not very well understood yet. One example of this is the **nucleolus**, known since the early days of microscopy. This structure inside the nucleus acts as a factory for the synthesis of ribosome components (fig 2), which are, in turn, essential for protein synthesis. The cell benefits from such compacted structures, as it allows for certain important reactions to take place locally and very quickly and efficiently.

Another more intensively studied new phenomenon within cell biology concerns the versatility and flexible behaviour of RNA and protein molecules, comparable to the trend in today’s society of a single job no longer being sufficient when hourly wages are lower. In some cases, this concerns molecules that are evolutionarily old and have acquired multiple biological functions over time. It can also involve protein molecules that no longer have a uniform three-dimensional shape. They can assume their active form when action is required, “as if key and lock are made to fit when you open the door,” as Wieringa put it in his valedictory lecture. For the population of RNAs in the cell, new functions continue to be described almost every month. In addition to the three classical types of RNAs, r(ibosomal) RNA, t(ransfer) RNAs and m(essenger) RNAs (box 1, fig 9), we now know many new forms of RNA, such as the s(hort) and l(ong) n(on) c(oding) RNAs and the circular RNAs, the functions of which are still far from being fully understood.

From the foregoing, it will be clear that the variation in the molecular content of cells can be enormous, even within a population of one type that looks very similar at first glance. When abnormalities become too large, causing an individual cell within the population to cross the threshold and undergo the process of naturally regulated cell death (apoptosis), it is not a disaster. In multicellular organisms, neighbouring cells in a tissue or organ can easily compensate

for this. All of this applies to gametes as well, although the concept of quality has a different and often more crucial connotation in that context. That will be addressed later. First, let's focus on the structure and function of the nucleus, the control room where our DNA is stored, embedded in the **chromatin** of the **chromosomes** (fig 2), and from where all the life functions mentioned above are directed (Ch2).

CHAPTER 2

Chromosomes

2.1 Introduction

There are four reasons why **chromosomes** are so wonderfully fascinating. Firstly, they function as the knowledge centre of the cell, and, collectively, of the individual (Ch1). Secondly, they serve as a physical stage for regulating the availability of genetic information (box 1) for the cell (and thus for the organism). Thirdly, chromosomes can copy themselves to pass on information to daughter cells and to the next generation. For this, they must be transportable, meaning they must be able to condense and have a specialised structure, the **centromere**, which can be pulled on during cell division (fig 3). Finally, the identical types of chromosomes obtained from the father and mother (the **homologues**) can find each other in the **germline** in preparation for reduction divisions on their way to **haploid gametes**, a process known as **meiosis** (Ch5, fig 13).

The number of chromosomes in a single set is indicated by the symbol N (the haploid number, fig 12). This number varies among life forms, is species-specific, and hence also varies among mammals. In humans, this set contains a solid 3 x 10 billion **DNA** bases. A **diploid** nucleus contains 2N chromosomes: 46 in humans, 38 in pigs, 40 in mice and 60 in cows. This makes relatively little difference to the total amount of DNA per nucleus, which does not vary much between these species.

2.2 Chromosomes during the cell cycle

2.2.1 The cell cycle

Over periods of time, cells are characterised by high division activity. That means that they have an active cell cycle which is subject to regulation, a separate branch in cell biology. The goal here is to introduce some concepts that we will benefit from later when we go through the germline. Figure 4 shows the cell cycle in the typical textbook-like manner.

After division, a cell may or may not decide, guided by its environment, that it has had enough. Terminally differentiated brain cells and muscle cells are examples of this. The **gene**s needed to produce daughter cells are switched off.

When newly produced daughter cells do not know when division activity will occur again (but keep that possibility open), they are in the G0 stage of the cell cycle (fig 4). Preparations to duplicate DNA are temporarily out of the picture. A G1 daughter cell has no such doubts and immediately prepares for the next cell division. In anticipation of the occurrence of **mutation**s, cells have detection systems to identify problems in the DNA in general, irregularities during DNA duplication, and the absence of a correct connection between the chromosomes, in this case, the **chromatids** (a name for the duplicated chromosomes prior to cell division) and the **spindle apparatus** (fig 3, 4). In the literature, these are referred to as checkpoints.

Detection of DNA damage (a broad concept, see fig 33) activates one of these checkpoints and leads to the “freezing” of the cell cycle: repairing the DNA damage takes priority, repair comes before progression. Later we will see how the checkpoints function in the germline; DNA repair is also active there.

When the cell considers the damage to be unreasonably large, euthanasia may be initiated, either independently or upon the order of neighbouring cells. This phenomenon is known as **apoptosis** (Ch1). A protein that plays an important role in the decision between apoptosis and repair is code-named p53 (fig 4). P53 is a **transcription factor** (box 1) and can thus contribute to the activation of genes, thereby determining the next steps. Transcription factors, proteins that bind to the **promoter** region of the **gene** directly or indirectly through other proteins (box 1), never work alone (box 1). This partly explains why p53 can drive both DNA repair and apoptosis; it depends on its partners.

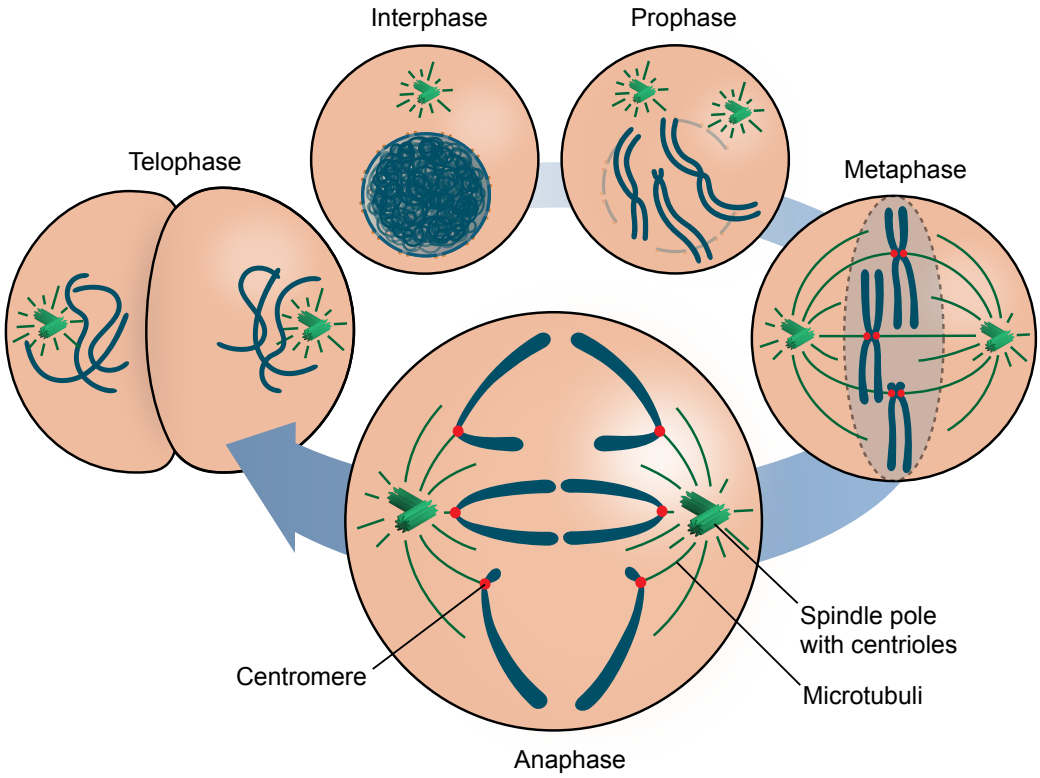


FIG 3

A visual representation of the cell cycle with an emphasis on the stages (phases) during which **chromosome**s are visible due to **chromatin** condensation; **mitosis** has begun. First to emerge from the interphase is the prophase. The **centromeres**, where the **spindle apparatus** attaches, are shown in red. The spindle apparatus is partly structured by the centrosomes with the centrioles, the “salt and pepper shaker”-like constructs that define the poles of cell division (green). Once the chromosomes have reached the plane of division through the action of the spindle apparatus, it is referred to as the metaphase of the cell cycle. The release of the **chromatids** at the centromeres marks the onset of the anaphase of the cell cycle. The stage in which chromosomes are still visible, shortly before the formation of a new nuclear membrane, is known as the telophase. In the **germline**, nuclear division is always accompanied by division of the **cytoplasm**.

When p53 chooses the path of preserving the cell, the cell cycle machinery is halted in the G1 stage until the DNA is repaired (fig 33). The signal will then be green, and DNA **replication** can start.

This step in quality control is known as the G1-S checkpoint (fig 4). If damage is noticed or caused while the DNA is being replicated, p53 will also be signalled about another life-or-death decision that must be made. In a situation where an accident strikes during the G2 phase, initiating the entire division apparatus is a risky operation (fig 4). Once again, a break is taken. The signal to prepare for **mitosis** is postponed until the chromosomal DNA is back in order. A final high-risk stage in the transmission of the correct genetic information to the daughter cells is division itself (fig 3, 4). A protein complex, the anaphase promoting complex (APC), plays a central role in preventing incorrect distribution of chromatids (fig 3, 4) among the daughter cells. The checkpoint itself is called the “spindle assembly checkpoint” (SAC).

When there is a defect in the connection of even one chromatid to the fibres of the spindle apparatus (fig 3), the cell division is frozen until the connection is repaired. Occasionally, things do go wrong, leading to a chromosome missing from a daughter cell. Often, the other daughter cell winds up with an extra copy of that chromosome. This phenomenon is known as **nondisjunction**.

2.2.2 Copying DNA in preparation for cell division

The part of the cell cycle that deals with the duplication of the genetic material is called the S phase (or synthesis phase, fig 4). The time it takes a cell to copy the DNA is at least about 5 hours and easily about 8 hours; it can vary considerably. The replication fork (fig 5) shifts at a rate of 100 bases per second. To stay within this time frame, replication must start simultaneously at many sites per chromosome; otherwise, copying the DNA would take days. This initiation occurs in clusters, meaning that the DNA polymerase machinery (fig 5) is synchronously activated on adjacent **chromatin** loops (the “looped domains” and thus TADs of fig 7).

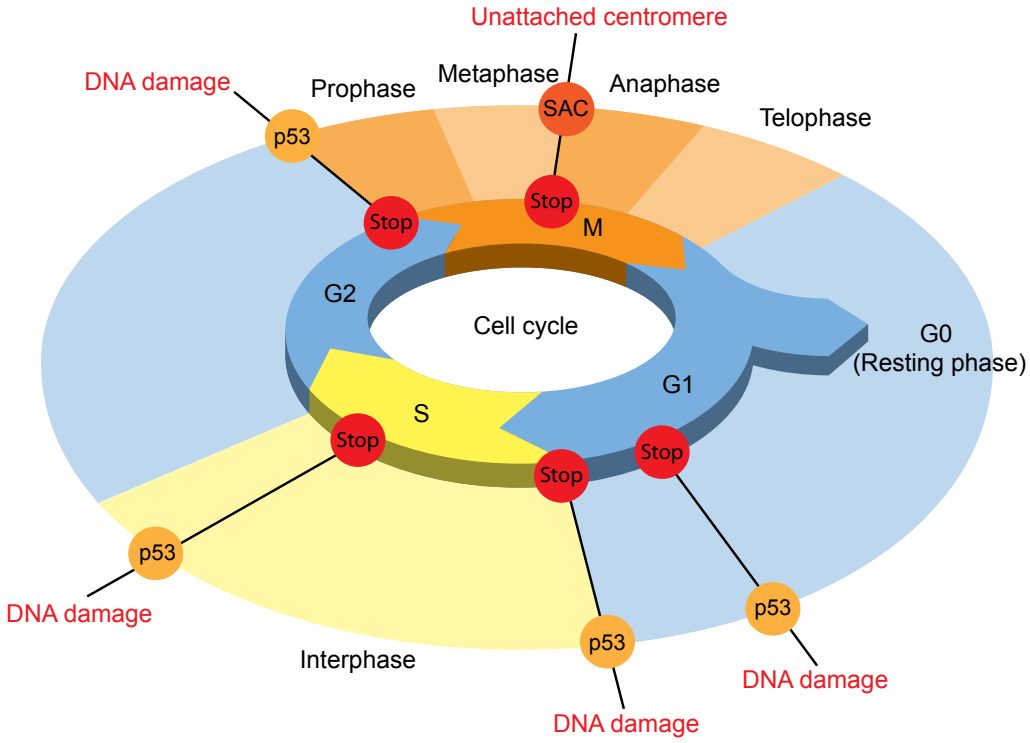


FIG 4

The cell cycle, but now with an emphasis on the interphase (G1, S and G2) in which **DNA** duplication during the S (synthesis) phase is the defining element. G stands for "gap" and M for **mitosis**, the nuclear division (fig 3), which, together with the division of the **cytoplasm**, results in two daughter cells. The checkpoints, in which the **transcription factor** p53 plays such a significant role, are indicated and collectively known as the **DNA damage response** (DDR). P53 is also described as a police officer (see also fig 33). The representation of the stages of the cell cycle does not represent the time span, which is more variable for G1 and S than for G2 (shorter) and M (even shorter). For highly actively dividing cells, a 12-hour cycle is possible, but it is usually longer. A **centromere** that is not attached to the **spindle apparatus** triggers the "spindle assembly checkpoint" (SAC) via the inactivation of the "anaphase promoting complex" (APC).

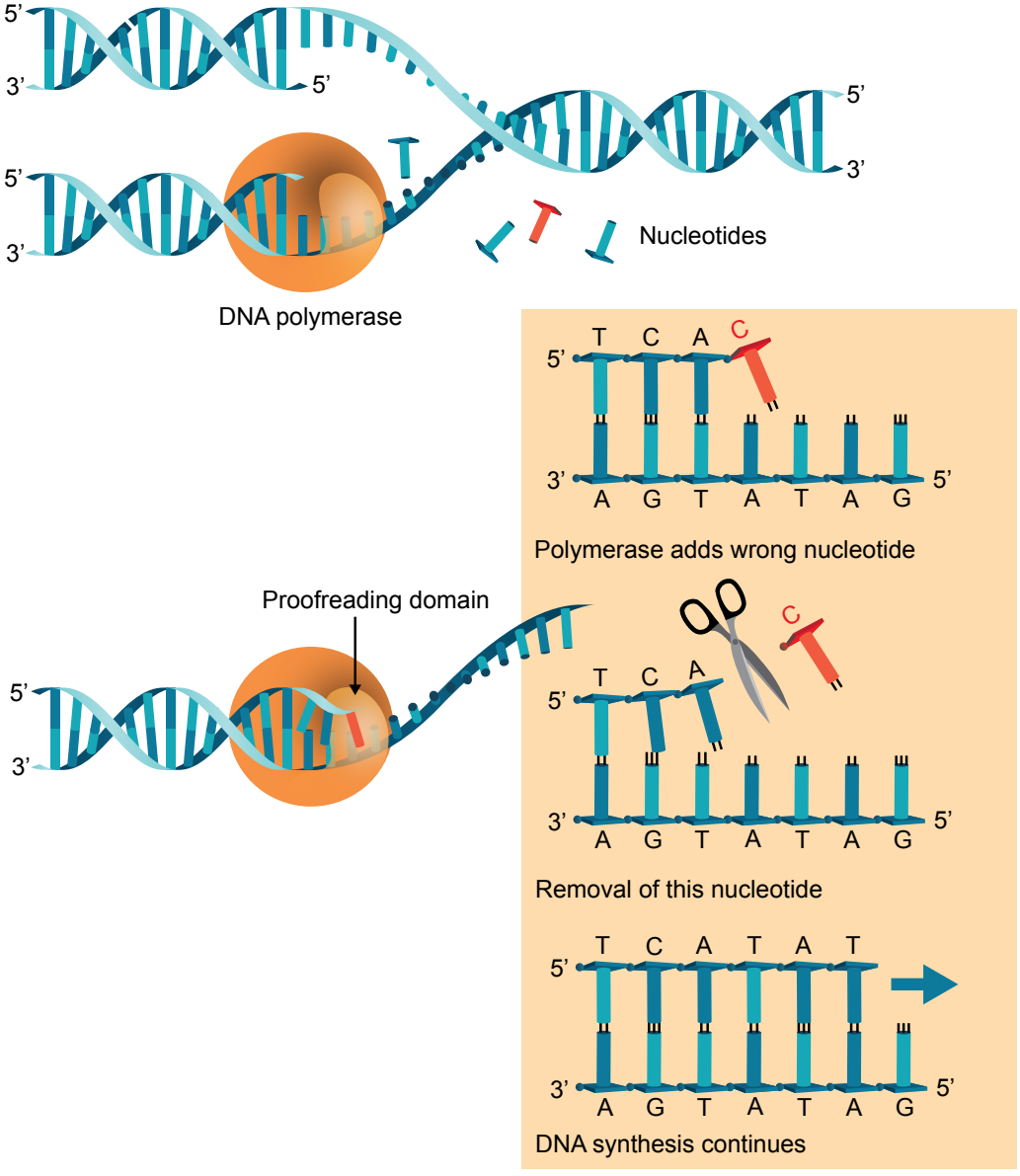

5’
3’
5’
5’
3’
5’
3’
Nucleotides
DNA polymerase
T C A C
5’
3’
A G T A T A G
5’
Polymerase adds wrong nucleotide
Proofreading domain
5’
3’
C
T C A
5’
3’
A G T A T A G
5’
Removal of this nucleotide
T C A T A T
5’
3’
A G T A T A G
5’
DNA synthesis continues

FIG 5

DNA replication. The chains of **nucleotide**s in DNA (and **RNA**) have polarity. This is indicated at the ends by a 3' and a 5', a notation derived from the molecular structure. The DNA polymerase that links the nucleotides during replication can only add them at the 3' end. Since the matching of complementary bases A-T and C-G requires that the strands of the DNA double helix run in opposite directions, the DNA polymerase also works in two directions in the replication fork. The polymerase follows the unwinding of the double helix into single-strand DNA. The ability of the DNA polymerase to proof-read, i.e. directly correct an incorrectly incorporated base (nucleotide), is too important not to mention here.

Replicating information is never completely error-free. Give the person on the left side of a fully packed front row in a theatre a short sentence and ask them to whisper it to their neighbour, and so on. Then, ask the person on the right side of this row to say out loud what they have heard. Chances are that, by then, not much of the original sentence will be left. In this regard, the DNA replication system of the cell does a much better job, whether it is the intestinal bacterium *Escherichia (E) coli* or a mammalian cell.

The DNA polymerase reads the single-strand DNA molecule to be copied and inserts a T **nucleotide** opposite an A and an A opposite a T. *Mutatis mutandis*, the polymerase picks a C when it encounters a G in the strand to be copied, and, vice versa, a G when the code contains a C (fig 5). In doing so, it makes one error per 100,000 links. This is recognised in 99 out of 100 cases by another part of the same protein, which then takes care of repair (fig 5). The remaining error frequency of one in 10 million is still not biologically acceptable. Even a simple organism like the bacterium *E coli* already has a repair system called **mismatch repair** (MMR) (fig 33). It responds to a misincorporated nucleotide, say a C versus an A instead of a T, and has survived all evolutionary upheavals: all mammals (including us) have it. Again, 99% of misincorporated nucleotides are corrected. How the system can distinguish the recently incorrectly chosen base from the template to be copied, which contains the correct information, is still not fully understood in mammals. Eventually, replication has an error rate of one in a billion bases. That is three per **genome**, so six for a diploid cell per cell division. Apparently, this is acceptable. The MMR system will appear several times in later sections of this book, due to its great importance in the germline.

Once the chromosomes have been copied, the sister chromatids remain closely connected to each other (fig 3, 6). Even before the chromosomes are copied, a protein complex is present that provides structure, like a backbone (fig 7). In the study of nuclear structure and the robustness of chromosomes within that nucleus (see below), the so-called **cohesin complex** (fig 6) plays an increasingly important role. During replication, the ring-shaped complex consisting of four proteins holds the two sister chromatids together. If this complex did not play such a significant role in the germline, it would not be introduced here.

2.3 The chromatin structure of the nucleus

2.3.1 Introduction

If you could isolate the DNA molecules of each human chromosome just after the division of a diploid cell, stretch them out and lay them end to end, you would have a total length of 2 metres of DNA per nucleus. This length is packed into a tiny ball with a diameter of 5-8 µm: comparable to 2 km of thread in a tennis ball. In an interphase nucleus, there is already an average condensation factor of 500. In order to make the chromosomes transportable during cell division, an additional factor of 20 is added to this.

To help us comprehend how all of this is possible without all that DNA getting tangled in an inextricable knot, chromosome models have been designed (fig 7). As technology advances, these models are constantly refined. By no means have all mysteries been unravelled yet, if we can even formulate them. After all, reading **gene**s (box 1), copying chromosomes (fig 5), repairing the DNA that is constantly needed (fig 33), and condensing chromosomes in preparation for cell division all involve changes in the structure of the DNA, complexed with many different kinds of proteins, collectively known as the chromatin. Its ability to be stained effectively is very helpful for recognition of the cell nucleus in tissue preparations, and of chromosomes when they are captured on a microscope slide just before the moment of division (fig 3, 12, Ch3.4).

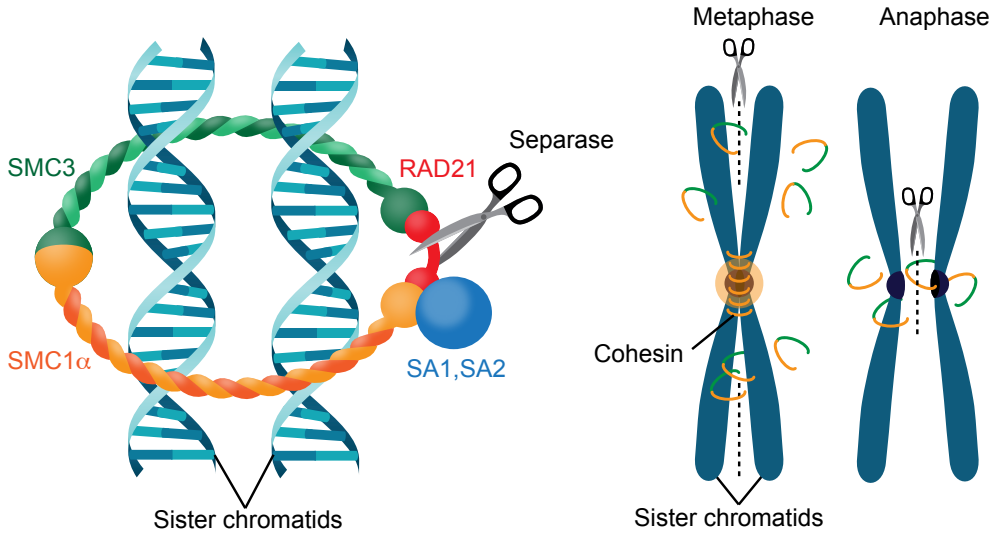


FIG 6

An illustration of the **cohesin complex** and one of its important functions in the cell cycle. The involved proteins are indicated with colours (apart from the cleaving enzyme separase). The convention is to write **gene** names in italics; all in capitals for humans and only the first letter for mice. The gene-derived proteins are completely written in capitals and are not italicised for both human and mouse. The **centromere** region is dark. Heading towards metaphase, the **chromosome** arms lose the cohesin complex first; in the centromere region, this happens when anaphase begins. The meiotic version is subtly more complicated compared to the mitotic version shown here. Whereas this complex was initially studied for its major role in the separation of **chromatids** in **mitosis** and in the second **meiotic** division, and of **homologous** chromosomes during the first meiotic division (fig 3, 13 and 30), it now receives at least equal attention for its function in organising the interphase nucleus (fig 7).

The diagram in figure 7 begins with the structure of the DNA double helix (the Watson-Crick model from 1953). For the first level of storage, cells use **nucleosome**s, which consist of 8 small **histone** proteins around which the DNA is wound (fig 7, 8). Each nucleosome contains two of the four main histone types: H2A, H2B, H3 and H4.

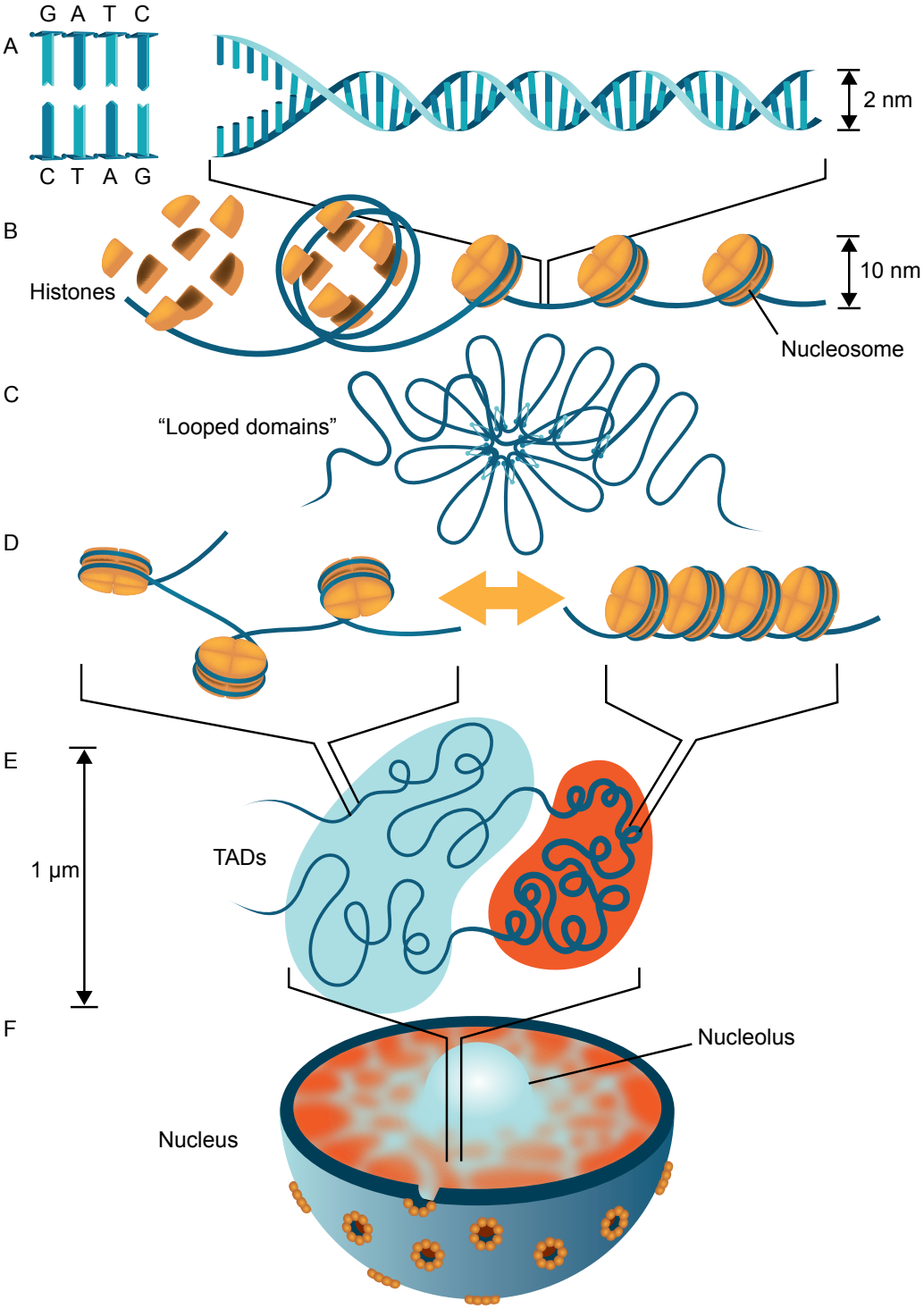

A
G A T C
C T A G
2 nm
B
Histones
10 nm
Nucleosome
C
"Looped domains"
D
E
1 µm
TADs
F
Nucleolus
Nucleus

FIG 7

Highly schematic representation of the structure of **chromatin** in the interphase nucleus. The **DNA** double helix (a) regularly winds around a complex of 8 **histone** proteins, creating the **nucleosome**s (see also fig 8) that form seemingly endless chains (b). The chains are organised into loops, often referred to as looped domains (c). The **cohesin complex** (fig 6), depicted here as very small green arrowheads at the base, plays an important role in the formation of these loops, hence representing a sort of backbone of the **chromosome**. Sections of these chains can have a looser and more open structure (d). The open structure is necessary to facilitate **transcription** (box 1). The more recently discovered key concept for understanding the structure of chromatin in the nucleus is the "topologically associating domain" (TAD, e). These can have a more open (blue, larger) or a more closed chromatin structure (reddish-brown, smaller), corresponding to the transcription activity of genes in the looped domains that populate these TADs. They are several hundred kb in size (median value 880 kb). Finally, in the nucleus (f), it is indicated that the active TADs are located more inward. The chromosomes have their own domain here, so TADs typically (but not necessarily) consist of chromatin from one specific chromosome. Besides the TADs, the nuclear membrane and **nucleolus** also play a significant role in the chromatin organisation within the nucleus.

Understandably, they have been well-conserved over the course of evolution: for example, H4 differs by only two out of 102 amino acids between peas and cows. The positions of nucleosomes on the DNA chain are not fixed; they can vary. This movement is carried out by specialised protein families called chromatin remodellers, and it naturally costs energy. It is for example needed for **transcription** (fig 9). With some clever techniques, such as allowing chromatin to unfold on a drop of water using soap, the chain of nucleosomes, wrapped around and connected by the DNA double helix, was made visible under an electron microscope in the years 1974, 1975. A nucleosome has a diameter of about 10 nm (fig 7). How the "string of beads" is subsequently folded and further condensed is a major question in current research into the puzzle that is the nucleus. It is clear that the loop formation that occurs afterwards (fig 7) is part of a mechanism that regulates whether the chromatin is transcriptionally active or inactive. The loops vary in size, with 100 kilobases (kb) often mentioned, but they can also be larger. Formerly, people in

this field only talked about looped domains, but now the discussion also includes “topologically associating domains” or TADs (fig 7) to indicate that nearby loops often have the same functional status; the genes in these loops are predominantly on or off.

It is clear that the cohesin complex, among other proteins, is an important player in determining the loop structure (fig 6). There is no complete overview of the backbone of the chromosome (fig 7), which is part of the nuclear matrix, a structure which can be thought of as the skeleton of the nucleus. However, the known functions of chromatin, such as altering its structure for the purpose of gene expression, initiating DNA replication in preparation for cell division, unravelling DNA, repairing DNA, and condensing chromatin in preparation for cell division, are likely related to it.

2.3.2 Types of chromatin

Traditionally, chromatin has been divided into **euchromatin** and **heterochromatin**. This division was strongly influenced by the staining pattern of nuclei in **histological** tissue sections observed under a light microscope. Anything that appeared darker was considered heterochromatin.

Not many researchers who have made groundbreaking and significant discoveries survive in textbooks (or on Wikipedia), but that is not the case for Murray Barr from Canada. Together with his colleague Bertram, he published in 1949 that the nuclei of nerve cells of female and male cats differ in their staining pattern. In the female nuclei, they observed a small sphere that was homogeneously intensely coloured tucked against the nuclear membrane. Today, you can still label this as a “Barr body”. In fact, it stained like heterochromatin, but what did this mean? After the famous Japanese scientist Susumu Ohno (who had relocated to Los Angeles and was known, among other things, for the idea that evolution is driven by the replication leading to duplication of genetic information) discovered in 1959 that the Barr body had to represent one of the two X chromosomes, Mary Lyon eventually hit the jackpot. In a 1961 *Nature* publication, she proposed that at some point in the development of female **embryo**s, it is decided per cell which of the two X chromosomes, the mother’s or the father’s, throws in the towel and thus stops expressing genes. This is, in a nutshell, the Lyon hypothesis. Mary based her hypothesis on work with mice. These experimental animals were abundantly available where she worked, at that time the Medical Research Council radiobiology unit in Harwell, UK. The mutation she used affects a coat colour gene on the X chromosome, and the female mice resemble tortoiseshell cats. Research in Harwell that involved **ionising radiation** and later also chemical **mutagen**s

yielded many mutants and biological genetic insights. In those days, all **mutant** lines were kept with live animals, because freezing 8-cell embryos to create an embryo bank (**cryopreservation**) was not possible yet.

The inactive X chromosome stains as heterochromatin, and always has a normally stained and functioning **homologue**. Since the X had the ability to choose early in embryonic development, we refer to the **heterochromatic** state as **facultative**, in contrast to the heterochromatin around the centromeres (fig 3, 12), which is called **constitutive**. Box 4 is entirely devoted to the inactive X chromosome.

The technological revolution that has given DNA research such momentum and scale has not bypassed the study of chromatin. The field is vast and is becoming more complicated as research techniques advance, but universally applicable rules have certainly been discovered, which can be used to describe the state of chromatin. Histones play a central role in this (fig 7, 8, table 1). During the assembly of nucleosomes, two **dimer**s are first formed out of an H3 with an H4. Together with DNA, this constitutes the basic nucleosome. Two dimers of H2A and H2B are then added to this. Particularly the histones H2A and H3 stand out for having multiple variants, each of which is, of course, linked to specific functionality. Such a variant can, for example, make the nucleosome more unstable, thereby making DNA more accessible for transcription. This is gratefully utilised in the germline. Added to this is an even more intricate system of chemical additions (Ch1), ranging from simple methyl (CH_3) and acetyl (C_2H_3O) groups (table 1) to small proteins (ubiquitination, sumoylation, Ch1) and, in the case of DNA damage, entire arrays of ribose sugar molecules like Christmas trees. In the search for ways to interpret the message of these so-called **post-translational modification**s (**PTMs**), researchers have stumbled upon some natural laws.

In 2001, a publication titled *Translating the Histone Code* by Thomas Jenuwein and David Allis appeared in *Science*. At that time, not all elements of the code were known yet, but it was already apparent that by examining molecular changes in the outward-projecting histone tails (see fig 8 and table 1), one could have an idea of the state of chromatin at that site. H3 is the histone type with the most “chromatin state” indicators. The tail of H4 also has a distinct significance in this regard. Both for ultra small-scale molecular research on chromatin (think of a gene promoter) and for microscopic-level observation of chromatin in every cell type, including the germline, it is necessary to have a set of antibodies that recognise histone PTMs. The fact that these antibodies are so incredibly specific and detect the smallest differences has always amazed

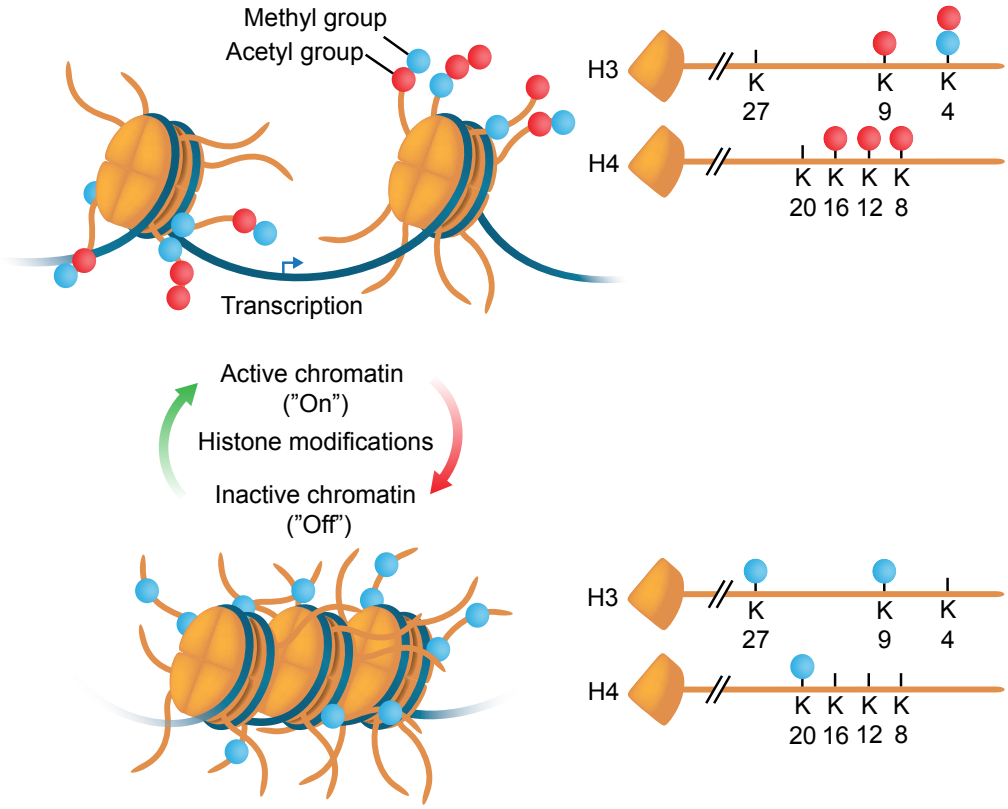

Methyl group
Acetyl group
Transcription
H3
K 27
K 9
K 4
H4
K 20
K 16
K 12
K 8
Active chromatin
("On")
Histone modifications
Inactive chromatin
("Off")
H3
K 27
K 9
K 4
H4
K 20
K 16
K 12
K 8

FIG 8

When the **nucleosome**s are assembled, tails of linked **amino acid**s protrude from them. At specific sites on these tails, the addition (or removal) of a small chemical molecule can start a cascade of reactions. This whole process also takes place at neighbouring nucleosomes. The most studied of these small chemical groups (this figure and table 1) are placed on the amino acid lysine (K), which has a suitable structure for this purpose. Depending on the type of molecule (methyl, acetyl) and which lysine is involved, the final result is either open and active **chromatin** (green arrow) or dense and inactive chromatin (red arrow). These modifications also occur in the folded parts of the **histone** proteins (here depicted for H3 and H4) of the nucleosomes. The central element in **gene** regulation is the accessibility of **DNA**. Open chromatin facilitates DNA **transcription** and dense chromatin inhibits it. Ultimately, multiple interrelated chromatin changes determine the defined locality and fine-tuning of turning a gene on or off.

me as a non-molecular biologist/immunologist. Enzymes are needed to apply the modifications mentioned above. Over the years, these have indeed been discovered and described in thousands of scientific publications. They are in turn removed by other enzymes.

Disabling or increasing the activity of these kinds of enzymes has opened up a world of experiments to discover the biological significance of this all. We will see an example of this in chapter 15. These enzymes, the regulators of the "histone regulators", are also under regulation, and so on. Such a cascade of regulation is quite dizzying (see also Ch1).

How do these histone PTMs convey their message? Figure 24, which will follow later on in this book, gives an as yet simple idea of the impressive complexity of chromatin regulation. The enzymes that place the PTMs are known as "chromatin writers". The proteins that bind to the PTMs and help determine the level of DNA activity (in transcription, transcription factors) are known as the "chromatin readers". There are "chromatin erasers", which in turn convert the local condition of chromatin into another condition. And then there are "chromatin remodelling complexes" that can shift nucleosomes along the DNA; "histone chaperones" that assist in a change within a histone class, as occurs most frequently for H2A and H3; and, on top of that, non-coding **RNA** molecules also influence the chromatin landscape. This is all necessary to fulfil the main functions of chromatin: transcription, replication, repair and **recombination**.

Table 1 A simple impression of the significance of the most common **histone PTM**s.

Histone type	**H3**	**H3**	**H3**	**H3**	**H4**	**H4**	**H4**	**H4**	**H2AX**	**H3**
Amino acid, position K = lysine, S = serine	K4	K9	K27	K3 6	K8	K1 2	K1 6	K20	S139	S10
Acetylated 1,	active	active			active	active	active			
Methylated, increasing in strength, 1,2,3	active	constitutive hetero-chromatin, inactive	facultative hetero-chromatin, inactive	active				inactive		
Phosphorylated									Mainly double-strand DNA breaks	Condensed chromatin in mitosis and meiosis

Constitutive **heterochromatin**, such as that located around the **centromeres**, is always rich in H3K9me3. The inactive X of somatic cells (box 4) is marked by H3K27me3, just like many **chromatin** domains in **autosomes** that do not undergo transcription in the respective cell type. Hence, this chromatin can be either active or inactive, with the inactive state being facultative. **Transcription**ally active chromatin may, for example, have an acetyl group on H3K9, two or three methyl groups on H3K4 and on H3K36. H2AX is an important member of the H2A family. The involvement of this histone protein in **DNA** repair processes (detectable by phosphorylation on serine 139) is most clearly visible at double-strand breaks. This is of great research interest due to the availability of suitable antibodies.

Researchers believe that in the not-too-distant future we will be able to observe what happens locally within and on chromatin in 3D, and perhaps even in 4D. Ultimately, it should now be clear that the term euchromatin is an umbrella term, within which variation is still enormous.

When a cell goes into mitosis, the chromatin landscape is automatically passed on; it is copied along with the DNA, so that the gene expression associated with the cell type is inherited. The details of this are still largely unknown. It has proven to be a great challenge to understand this “bookmarking” of chromatin for subsequent cell generations and to determine exactly which PTMs are responsible for memory maintenance. Everything included in this subject is also covered by the term **epigenetics**, which refers to the **heritability**, whether for shorter or longer duration, of the functional state of DNA.

Box 1 The central dogma of molecular biology

Proteins consist of polymer chains of different types of **amino acids**, of which we know 20. Each type of protein has its own amino acid composition and length, determined by the number and sequence of the types of amino acids that it contains. The total set of proteins of a cell is known as the **proteome**. A human cell contains around 2×10^9 protein molecules. Estimates of the number of truly different proteins within the proteome vary from 40,000 to 80,000 for humans (see also below). Most types of protein are present in numbers of about 1,000 to 10,000 molecules per cell, but some are much rarer. The **DNA** of each cell contains the complete recipe, the **genetic code** for the formation of all different types of proteins. Genes are the unit of information, with the information being encoded in the **nucleotide sequence** of the DNA of each **gene**.

When a gene is activated and expressed, the DNA is first copied into the linear single-strand nucleotide sequence of **RNA**, the primary transcript (pre-**messenger RNA**, fig 9). This process is referred to as **transcription**. For the total population of all types of RNAs formed within a single cell, the term **transcriptome** is often used. Usually, this means the RNA that is translated into protein, so the collection of mature messenger RNAs (mRNA, ~3×10^5 - 10^6 molecules per cell). This is the RNA that is formed in the nucleus after maturation of the primary transcript (fig 9). In addition, there are many non-protein-coding types of RNA. It is well known that these have significance in the construction of components (e.g. ribosomes, fig 2), in transport (e.g. of amino acids, fig 9) or in regulation of cellular processes. Their functions are currently being studied intensively. The term gene also applies to these functional RNA molecules. Ultimately, the genetic code that is transferred from DNA to RNA through transcription can be translated into the characteristic amino acid sequence of a protein (fig 9). This process is called **translation**. During translation, each consecutive set of three nucleotides (one **triplet** or codon) in the designated parts of the mRNA (the **exons**) encodes one amino acid (fig 9). With a choice of four nucleotides per position, there are 4 x 4 x 4 = 64 codon possibilities (in DNA lan-

guage the bases A, G, C and T; in RNA language the bases A, G, C and U; fig 5, 7). The entire overview of relationships between codons and the corresponding translation into amino acids is known as the genetic code. The nucleotide notation of RNA is usually used to represent this. Four triplets, namely the start codon AUG and the stop codons UGA/UAG/UAA, indicate where the translation of RNA into protein begins and ends. The start codon codes for methionine. Therefore, almost every protein starts with this amino acid; AUG is a unique code. The stop codons terminate the elongation of the chain of amino acids; they do not code for an amino acid. That means there are still 64 - 4 = 60 different codons available to incorporate the remaining 19 types of amino acids. This phenomenon is known as the “degeneracy of the genetic code”, and it ensures that a base change (a **mutation**) in the last letter of the codon in the DNA of a gene (and therefore also in its RNA copy) often does not result in a different amino acid in the protein chain. Mutations at the first or second position in a codon usually do lead to change. When the mutant codon does not result in a different amino acid, it is called a “synonymous” mutation, and when it does, it is called a “missense” or “non-synonymous” mutation. If the base change turns the codon into a stop codon, this is referred to as a “nonsense” mutation. If there is a change in amino acid but there are no noticeable consequences, this is called a silent mutation. The word synonymous is also used in this context. Not every change at the DNA level within a gene necessarily leads to a problem in the functioning of the protein, but most of the changes do. When a gene no longer codes for a functional protein, the term “loss of function” is used. This can be caused by a missense or nonsense mutation. Insertion or deletion of one or a few bases (an **indel**) can also shift the reading frame of the codons, causing them to no longer be correct. The location of these types of mutations in the gene, and thus in the replicated mRNA chain, ultimately helps determine whether a protein loses or maintains its function.

Traditionally, the concept of a gene was synonymous with the concept of a protein, and therefore it was long believed that the total number of genes in our **genome** also

determines the total number of different proteins, but that is not the case. We now know that the vast majority of protein-coding genes are made up of segments, the exons and **intron**s (fig 9). The pre-messenger RNA of the gene still contains the linear sequence of both (fig 9). As mentioned above, the introns are removed after maturation, after which the exons are found in the mRNAs as very accurately linked sections (fig 9). The alternation of introns and exons allows for variation during the maturation phase, as exons can be skipped. As such, from more than half of the pre-mRNAs that are formed, only a part of the entire initial set of exons is used. The presence or absence of exon sequences in mRNAs is determined during their maturation by a process referred to in the scientific literature as “alternative splicing”. This process can be both spontaneous and cell type-dependent and can result in the presence or absence of domains in the protein (with one exon coding for one domain) after translation. It almost goes without saying that small mutations can also interfere with the process of exon selection and can thus also be a source of pathological events in a gene.

Hence, it is now assumed that the approximately 20,000 known protein-coding genes do not code for exactly the same number of different proteins. In cells and tissues of humans and other mammals, the number of truly distinct protein types is now estimated to be between 40,000-80,000, as reported earlier in this box. The total amount of genetic information in mature mRNA molecules of a cell type, possibly a tissue or organ, is referred to as the **exome**.

As mentioned above and also in chapter 1, there is a growing focus on genes that do not code for proteins, whose primary product is not mRNA but another type of RNA. Long known are ribosomal RNA (which is incorporated into ribosomes via the **nucleolus**, Ch1, fig 2) and tRNAs that play a role in supplying amino acids during translation into proteins

(fig 9). More recent discoveries include many other non-coding (nc) types of RNA, such as “short non-coding” (snc) RNAs, “long non-coding” (lnc) RNAs and even circular RNAs with often still poorly understood roles in the regulation of gene expression or other cellular processes (see also Ch1). Research into the significance of these RNAs, of which the lncRNAs are numerically by far the largest contributors, and which collectively occur at least five times as often as the 20,000 protein-coding genes, is also a growing area of interest for the **germline**.

Regulation of transcription, a topic long studied within molecular biology, is a particularly compelling illustration of the molecular complexity involved in cell functioning. It starts with the recognition of motifs in the DNA sequence around and/or in each gene by large protein complexes formed from so-called **transcription factor**s and associated activator and/or repressor proteins. The complex molecular interactions involved ensure that **RNA polymerase**s (fig 9) start transcription at the correct location. They also regulate how often and for how long this occurs, leading to pulsatile transcription activity of a gene. The regulatory motif that determines the place where transcription starts, called the **promoter**, often lies close to this site. DNA motifs that determine the frequency of transcription can be located across a much larger area in or around the gene. How the multitude of factors ultimately determines the amount of RNA produced from a gene in a specific cell is still partly unclear. The details of this process are beyond the scope of this book.

From the above, it is clear that genetic information is transferred from DNA to RNA and then to protein. This concept is known as the central dogma of molecular biology, first formulated in 1958 by Francis Crick (co-discoverer of the double helix structure of DNA, along with James Watson). In 1970, he further supplemented details relating to this dogma in a *Nature* publication. Since then, it has played an important role in advances

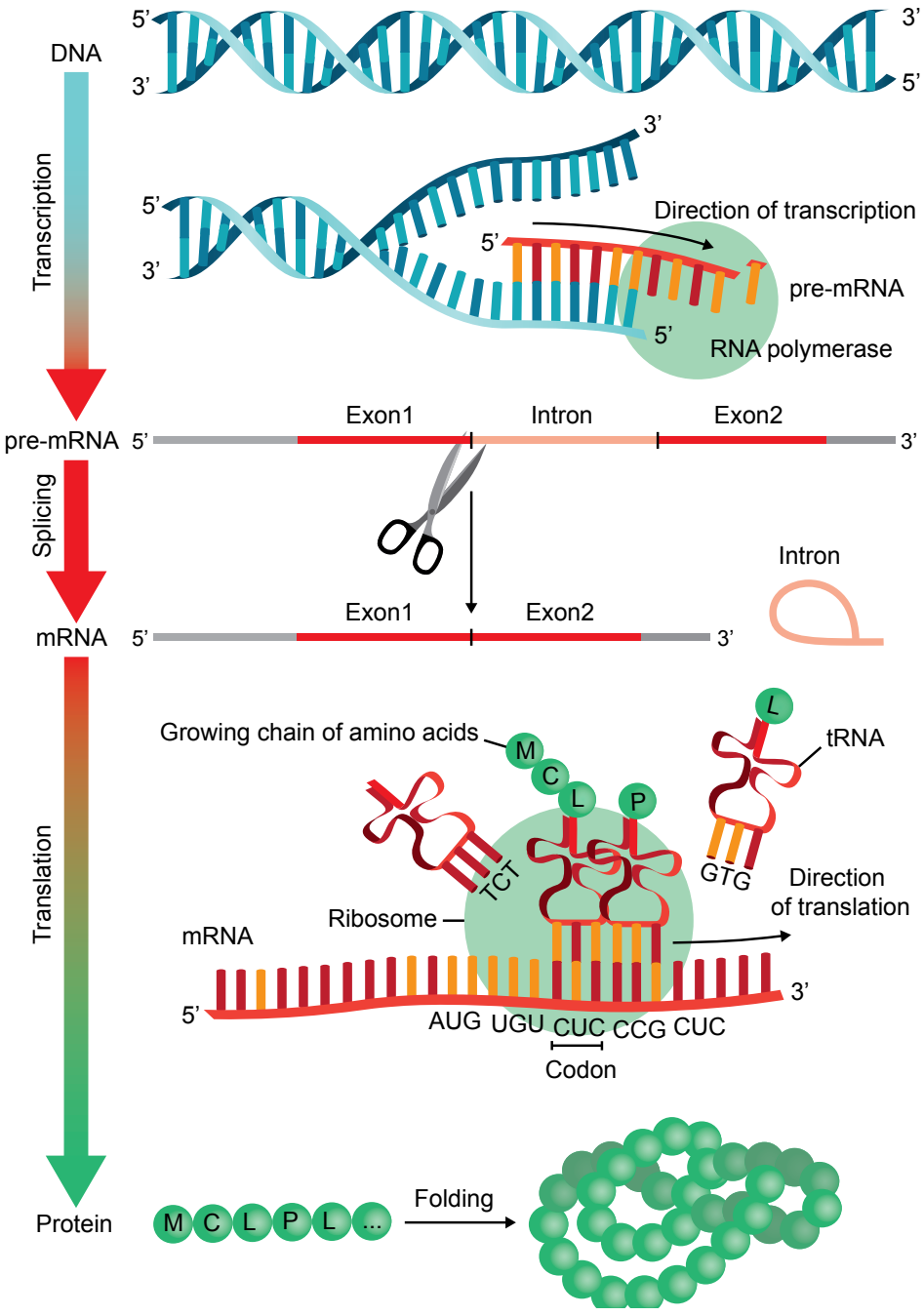

DNA
5'
3'
3'
5'
Transcription
3'
5'
3'
Direction of transcription
5'
pre-mRNA
5'
RNA polymerase
pre-mRNA
Exon1
Intron
Exon2
5'
3'
Splicing
Intron
mRNA
Exon1
Exon2
5'
3'
Translation
Growing chain of amino acids
M
C
L
P
L
tRNA
TCT
GTG
Direction
of translation
Ribosome
mRNA
5'
3'
AUG UGU CUC CCG CUC
Codon
Protein
M C L P L ...
Folding

FIG 9

The pathway from **DNA** to protein, via **RNA**. During transcription, the **RNA polymerase** extends the primary (pre-messenger) RNA at the 3' end. In the nucleus, this pre-**messenger RNA** is processed by removing the parts that do not code for **amino acid**s: the **intron**s. **Exons** remain and the mature mRNA is created. The process is known as "splicing". After this maturation process, the mRNA moves from the nucleus to the **cytoplasm** of the cell. Here, amino acids supplied by transfer (t)RNA molecules are strung together during **translation** of the **genetic code**. The determining factor in this is the "anticodon" code of the tRNA, which recognises the codon (of the **triplet** code) for this particular amino acid in the mature mRNA through base pairing. The genetic code is written in RNA language, with a U in every place where there is a T in the DNA sequence. Each amino acid has its own letter in the alphabet. A protein always starts with the amino acid methionine (M). The C in the small piece of newly formed protein stands for the sulphur-containing cysteine, the L for leucine and the P for proline.

within the molecular life sciences. Nevertheless, the all-encompassing applicability of the central dogma had to be adjusted considerably over time. After all, we will see later on that in the germline, but also for instance in processes that occur as a result of viral infections, genetic information can sometimes also flow "in reverse", from RNA to DNA.

2.3.3 Types of DNA

The number and sequence of nucleotides/bases A, T, C and G contain the entire genetic information of a species. This extensive and characteristic set of bases linked per chromosome is known as the species' genome, and the science that seeks to extract biologically relevant information from this is called **genomics**. The word genome actually refers to the haploid situation, but genomics is mainly practised in the diploid state, signifying the complete genetic information of a cell or organism. Thanks to technological breakthroughs enabling the cost-effective sequencing of bases per species and per individual ("**next generation sequencing**", **NGS**), this branch of genetics is growing enormously, which would not be possible without nearly unlimited computing capacity. In a narrower sense, genomics is mainly interested in the approximately 20,000 protein-coding genes (Ch1, box 1). The **exons** (box 1) of these make up only 1.5% of the genome.

When you focus your attention on the whole genome, the most startling fact is that about half (table 2) of all DNA consists of base sequences with motifs that occur in many copies: in tandem, like wagons of a freight train lined up one after the other, or dispersed throughout the entire genome. The length of one such basic element can vary considerably, and so can its numbers. The microsatellites (table 2), repetitions of a simple motif of 2-6 bases, which occur at multiple sites in the genome, easily mutate in a number of copies of the base motif, because DNA polymerase does not like monotony and sometimes loses count during DNA replication (fig 5).

Minisatellites are a bit larger, as shown in table 2. These motifs also easily mutate to a different number of adjacent copies, but through a different mechanism, related to meiosis (**gene conversion**, will be addressed in Ch5.1). The reason for mentioning it here is that the concept of **DNA fingerprint**ing is based on this. In 1985, Alec Jeffreys, now Sir Alec, from the Department of Genetics at the University of Leicester, discovered the large genetic variation resulting from variable minisatellites with a basic DNA motif ranging from 10 to 100 bases. A year earlier, he published the DNA sequence of such a 33-base pair (bp) sequence found in an **intron** of the human myoglobin gene. Myoglobin is an oxygen-binding protein in muscles. This "repeat" is variable in length (number of copies) and occurs at several places in the genome. DNA fingerprinting utilises this variation and, for years, it was the technique used to provide everyone with a genetic barcode. It proved to be quite a breakthrough in establishing parentage and in forensic personal identification, where it is still used nowadays. Unsurprisingly, the list of scientific awards that Alec Jeffreys has received is impressive.

At the start of this century, we visited one of his associates, Yuri Dubrova, originally from Ukraine, who had previously been drawn to Sir Alec's reputation. We saw a small, unremarkable looking person in a lab coat, deeply bent over a -80°C freezer, apparently searching for DNA samples. It remains fascinating, this contrast between day-to-day activities, the pleasure one can derive from them, and the discovery that opens the gates to so much insight with a wide range of applications in society.

Simple or larger DNA base sequence motifs that are linked in tandem in many copies play a large role in specialised functions of the chromosome. For example, the constitutive heterochromatin around the centromeres (Ch2.3.2, fig 3, 12) consists of a number of highly related motifs. The basis for this is a repeat that is 171 base pairs long in humans (table 2). In our species, this is known as alpha satellite DNA (table 2), and in mice as "major" and "minor" satellites, all historically evolved terms. The structure of the different repeats for the human centromeric regions has been resolved in the 2022 version of the human genome (Ch4.2), and the observed variation is striking. They replicate their DNA late in the S phase (fig 4). In humans, these regions are 250 to 5,000 kilobases (kb) in size. Late replication, strong staining (in mice) and compact chromatin; that does not indicate transcription. When an article published in 2002 by a renowned French group hinted that non-protein-coding (nc) RNA (Ch1, box 1) was involved in the formation of heterochromatin at the site of transcription, I really had to get used to that idea.

Now, about 20 years later, we know much more about satellite DNA. In the centre of the heterochromatic centromere region, where the connection to the spindle apparatus forms (fig 3), the structure of chromatin (fig 7) changes, allowing satellite DNA to be read. Satellite RNA is needed on site for a subsequent chromatin change, in which the nucleosomes (fig 7, 8) are locally degraded and rebuilt, this time carrying a different kind of histone 3, which is needed for transporting the replicated chromosomes to the daughter cells. This is a great example of how delicate the behaviour of chromosomes is, and how much detail is involved in it.

A very well-known motif that is repeated many times, TTAGGG, is located at both ends of each chromosome, the **telomere**s. It covers about 5 to 10 kb of DNA, equivalent to 800-1,700 repeats (table 2). In mice, which have the same basic motif, they are five to 10 times longer. You can think of this repeat as a solution to the so-called end-replication problem.

Table 2 The contribution of different types of repetitive **DNA** motifs to the **genome**. These repeats can occur predominantly in long sequences (in tandem) or dispersed. The overall influence of repetitive DNA motifs in the genome is large. Researchers emphasise that the estimate of around 50% is probably on the low side, with two-thirds often being mentioned.

Type of repeat	**Size in bases**	**Tandem**	**Dispersed**	**N copies**	**% genome**
Microsatellites	2-6	yes	yes	> 10^6	low
Minisatellites	15-100	yes locus 1-5 kb	yes		low
Telomere base sequence	6	yes	at chromosome ends	800-1700 per chromosome end	Approx. 0.01%
Centromere alpha satellites	171	yes	constitutive heterochromatin. In and around each centromere	± 10^6	5%
Transposons					
DNA transposons	Up to 3 kb		yes	500,000	3%
Retrotransposons					
SINE, SINE- Alu, of which ±1000 active	100-300		yes	1,800,000	13%
LINE of which ±100 active (6 kb)	1-5 kb		yes	950,000	20%
LTR ERV, no longer active in humans but active in mice	Up to 4 kb		yes	650,000	8%
SVA of which ±50 active	2 kb		yes	5,500	0.4%

Explanation of abbreviations. LINE “long interspersed nuclear element”. SINE “short interspersed nuclear element”. In humans, most SINEs have a recognition site for the DNA cleavage enzyme Alu. This recognition motif occurs only in primates (prosimians and apes, including humans). LTR “long terminal repeat”. **Retrotransposons** that have this motif at the ends are called ERVs “endogenous retroviruses”. SVA is an abbreviation of SINE, VNTR, Alu; it contains these motifs. Actively transcribed retrotransposons (in RNA) can be reintegrated via the active LINEs, and thus contribute to DNA variation. Microsatellites are also known by the abbreviation STR (short tandem repeats). Minisatellites also appear under the term VNTR, “variable number of tandem repeats”. This is what **DNA fingerprinting** is based on.

Normally, chromosomes become slightly shorter with each cell division, and thus with age. This is due to properties of the DNA polymerase; the last piece of DNA is not so easily copied. Without delving further into molecular details, the enzyme **telomerase**, which has properties of a **reverse transcriptase**, can extend a single-strand DNA molecule using RNA as a template. The piece of RNA used for this is complementary to the telomere repeat.

Of course, cells should not perceive the end of a chromosome as a double-strand DNA break (fig 33), as this would trigger the surveillance system as the first step in repairing the damage (Ch2.2.1, fig 4). The solution consists of placing the last piece of single-strand telomeric DNA back in a loop. Subsequently, a complex of six proteins aptly named “shelterin” makes the entire structure invisible to the double-strand break detection service.

This is another example of a situation where the effect of a single protein is minimal, while its action together with other proteins is profound. Telomeres have really captured the imagination, and the discoverers of its DNA structure, Elizabeth Blackburn, Carol Greider and Jack Szostak, were awarded the Nobel Prize for this discovery in 2009. For those who long for eternal life, attention has been drawn to this. But eternal life, aside from species extinction, can really be found in the germline.

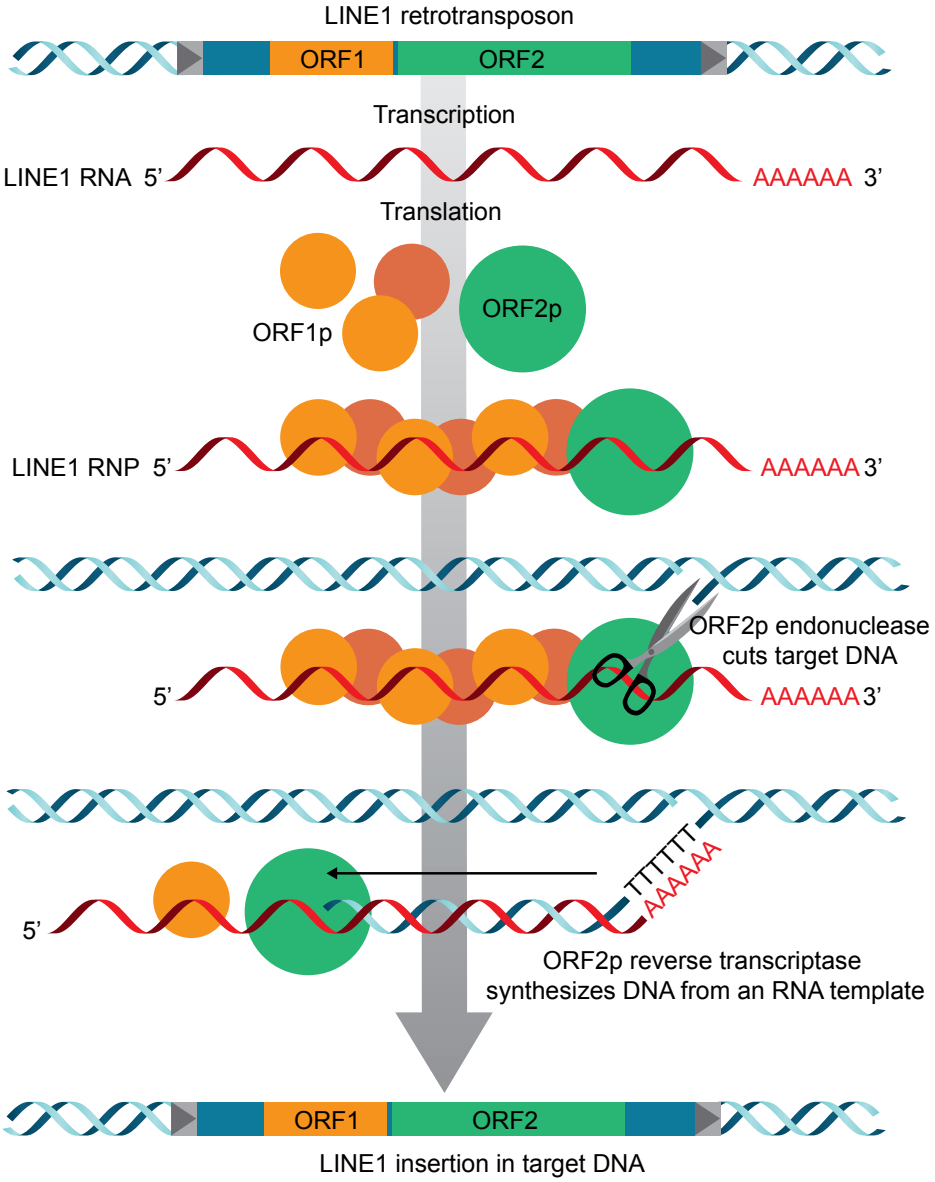

LINE1 retrotransposon
ORF1
ORF2
Transcription
LINE1 RNA 5'
AAAAAA 3'
Translation
ORF1p
ORF2p
LINE1 RNP 5'
AAAAAA 3'
ORF2p endonuclease
cuts target DNA
5'
AAAAAA 3'
TTTTTT
AAAAAA
5'
ORF2p reverse transcriptase
synthesizes DNA from an RNA template
ORF1
ORF2
LINE1 insertion in target DNA

FIG 10

Schematic representation of how active LINE-1 **retrotransposons** can copy themselves in the **genome** (table 2). Only a small minority of these retrotransposons is active, meaning they can be read (transcribed) and reinserted into an AT-rich region in the genome via the **RNA** intermediate. To mobilise a LINE and let it land as a **DNA** copy, the proteins ORF1p and ORF2p are both required.

ORF1p has various auxiliary functions, ORF2p can both cut DNA (it has **endonuclease** activity) and copy RNA into DNA (it has **reverse transcriptase** activity). This activity can also be used to serve other types of RNA, for example, that of an active SINE (table 2). The abbreviation RNP stands for "ribonucleoprotein particle", the complex of RNA with attached proteins, in this case all derived from the active LINE.

By far the largest contribution of repeat DNA structures to the genomes of many life forms, including our own, is formed by so-called **transposable elements**. These are elements that have multiplied in the genome like parasites. This can happen via DNA replication (resulting in a **transposon**), but it can also occur through the enzyme "reverse transcriptase" that turns RNA back into DNA, resulting in **retrotransposons** (fig 10). In that case, the original master copy (or copies), with an evolutionary history of approximately 100 million years, is an RNA virus. In mammals, this pathway is the most active and thus contributes most to the genome (table 2). A characteristic of these repeats is that they appear to have been dispersed throughout the genome by chance. This process is still ongoing, also in humans. Its drivers are about 80-100 active copies of the LINE family (table 2). An intact LINE1 of 6 kb contains two genes that can be transcribed, with one of them producing a protein that possesses both reverse transcriptase and **endonuclease** activity (fig 10). SINEs (table 2) are found in most mammalian genomes. The origin of the small 300 bp motif is known (it is related to a non-protein-coding RNA in the ribosomes (fig 2), which plays a role in protein synthesis). To replicate itself, it utilises the reverse transcriptase/endonuclease of the active LINEs.

As a result, one in 20 babies has an additional SINE copy and one in 100-200 a new LINE copy. Do these new pieces of DNA land somewhere in the genome completely by chance? SINEs always end up in sections that are copied earlier in the S phase, while LINEs land in the large pieces of DNA that are copied later (fig 4). Researchers view retrotransposons as a kind of chromosomal framework for the orderly expression of the genome and the functioning of the

cell. Later it will become clearer that a lot is needed to transmit the whole array of dispersed retrotransposons through the male and female germline without unwanted side effects.

Using a combination of an old technique to physically separate short pieces of DNA, and next generation sequencing to determine the DNA base sequence for each fraction, a researcher from Naples, Giorgio Bernardi, succeeded in making a clear narrative out of the organisation of chromosomes. After starting chromosome research in 1968, he never left this field and remained active for a very long time. One way of looking at pieces of DNA of around 100 kb is to determine the ratio of AT to CG base pairs. The fractions that Giorgio isolated differ increasingly in the percentage of CG base pairs. It is now apparent that parts of loops and entire loops (fig 7) with a similar percentage of CG in the DNA are often adjacent to each other. Loops with a slightly higher percentage of CG content (ranging from 46% to 59%, the highest value) contain a group of genes, about 20% to 30% of the total, that are expressed in every cell. These can be seen as the genes responsible for the “cellular infrastructure” (encompassing structures such as membranes, the cytoskeleton, the nuclear skeleton and all basic metabolic processes). This is the “old genome”, which is not involved in the numerous specific functions of cells in a multicellular organism with a lot of specialisation. These genes are also called “housekeeping genes”. These loops are replicated first in an upcoming cell division (Ch2.2.1). It accounts for about 14% of the total DNA. In CG-rich DNA, the gene density is higher. The CGs are particularly prominent in the stretch of DNA preceding the protein code, the promoter (box 1), which partly regulates whether the gene will be used (see fig 11). When this was discovered by Adrian Bird, sometime around 1985, the term “CG island” was introduced for this type of promoter. This old genome is located in the centre of the nucleus (fig 2, 7). The part of the genome that relates more to differentiation within individuals and between species, and is less rich in genes, is located more peripherally (fig 2, 7). Each chromosome has sections of the old genome and sections of the evolutionarily more active newer genome (represented by the light and dark bands in fig 12, respectively). As we have just seen, the old, early replicating genome contains the SINEs, and the newer, later replicating genome contains the LINEs, hence the idea that these elements provide structure to chromosomes and are related to gene activity. Besides all this, chromosomes have their own domain in the nucleus (fig 7), but that can never be very strictly regulated. And so, everything contributes to the image of an incomprehensible complexity, as presented in chapter 1.

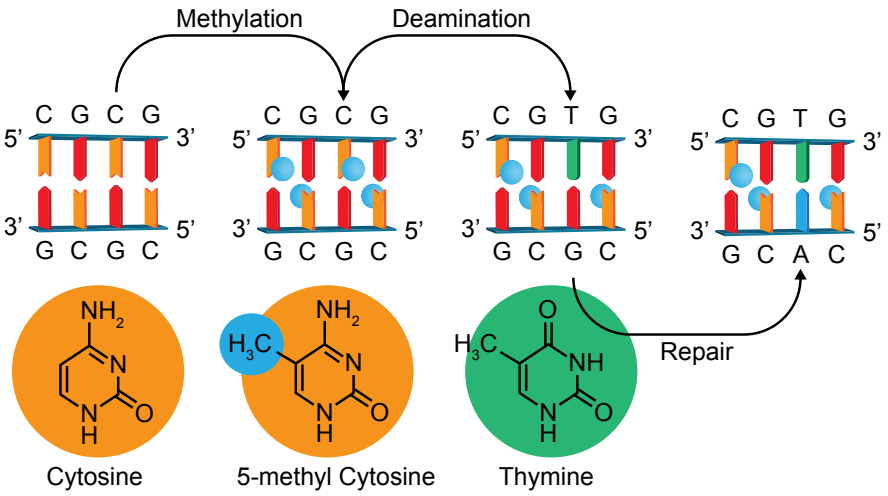


FIG 11

CpG methylation and the C > T **mutation**. The number of CGs in the human **genome** is estimated to be 29 million. In **somatic** cells, most of these are methylated, which applies to the male **germline** as well (fig 23). Their main function is to confirm the inactivation of local elements such as **retrotransposons**. This phenomenon is discussed in more detail in chapters 9, 10, 15 and 16, and in box 4. The CG islands of many (>60%) protein-coding genes, including housekeeping genes (Ch2.3.3), are unmethylated, but CG methylation does not seem to play a leading role in regulating **gene** activity via the **promoter** in those regions. Understanding the **methylome** is a huge scientific challenge. The advantages of CpG methylation for genome regulation must have outweighed the disadvantage of an increased risk of the C > T mutation. Some **eukaryotes** can function without this system (the nematode *Caenorhabditis elegans* is an example).

CHAPTER 3

Genetic variation

3.1 Introduction

Halfway through our studies of animal husbandry in the late 60s, when we wondered what genetics was really about and what the core of the field was, two words came to mind: **mutation** and **recombination**. Both contribute to the concept of genetic variation. Mutation will be discussed here, and recombination in chapter 5.

In its early days, genetics was “practised” based on a selection of observable traits. The different expressions of such a trait (known as **phenotype**s, for example coat colour) could be explained through patterns of inheritance, assuming the existence of a single **gene** responsible for that trait. A prerequisite for the many crossbreeding experiments that have been conducted is the genetic variation of that gene, i.e. the availability of at least two **alleles**, implicating the presence of mutation(s) (Ch5, fig 13). Do realise that the structural properties of **DNA** were completely unknown until the end of World War II; it is hard to imagine a greater difference from our current situation. Mutations are now simply identified by determining the DNA base **sequence**.

Starting from the second decade of this century, and following a **genome**-wide approach, the focus was mainly on the **exome** (box 1), the collection of protein-coding base sequences (**whole exome sequencing**, **WES**). Now, sequencing the whole DNA, known as **whole genome sequencing** (**WGS**), is advancing. These approaches, summarised under the term **NGS** (**next generation sequencing**) provide unparalleled insights into genetic variation in humans and animals. When it comes to non-coding DNA, there is often still little understanding of the relationship between variation and function. It is likely that a significant portion of non-coding DNA has no direct function and is “neutral”, meaning that it is not subject to any selection pressure. Our

understanding of variation in the total DNA profiles of humans and animals is not yet complete, but the outlines of the ever-expanding knowledge are beginning to emerge. There have been two dominant developments. Firstly, the number of individuals from the global population that has been sequenced with WES/WGS is increasing into the millions. Secondly, the continuous progress in technology allows for larger and larger DNA fragments to be read by the device in a single analysis step. Tables 3 and 4 are attempts to summarise genetic variation at different levels, from a single DNA base to entire **chromosomes**. The classification of types of genetic variation is strongly based on the size (in base pairs) of the altered region. The great majority of variation occurs at the level of a single base, which then changes to one of the other three bases (for example, from C to T, which is the most common change).

When their number is limited, any missing or additional base(s) are referred to as **indel**s (table 3). Changes which comprise over 50 **nucleotide**s are covered by the term structural variants (tables 3, 4). Within the structural variants, a classification scheme has been made that partly dates back to the time when chromosomes could only be studied in mitotic cells using a light microscope (the basis of table 4). Finally, restless **retrotransposons** and the phenomenon of "**copy number variation**" also contribute to genetic variation (table 3).

3.2 Variation at and around the base

As mentioned above, population-level sequencing is the main avenue for gaining insight into genetic variation in humans. In the *Nature* issue of May 28, 2020, the genome aggregation database consortium (gnomAD) provides an overview of the progress. At that date, their database contained the DNA sequence, mostly exomes (WES) but also whole genomes (WGS), of more than 140,000 people, together accounting for 270 million variants. These range from mostly base pair changes to small indels, to a selection of structural variants with a maximum size of around 10 megabase (10 million, 10 mb). A year earlier, TOPMed, an entirely American human genome initiative focused on the pathology of heart, lungs, blood and sleep, reported the first WGS analysis of the DNA of nearly 54,000 people with origins from all over the world. With a bit of calculation, you can deduce that, across the whole group, one of every eight bases shows variation. This variation has accumulated over very many generations and has continually expanded. Coincidence plays a major role in its emergence. The mechanisms of DNA repair

are not flawless (consider, for example, the errors in DNA **replication**, Ch2.2.2) and damage in DNA occurs constantly under the pressure of oxygen **metabolism** (Ch14, fig 33) and other environmental influences.

Most base pair variations are very rare. In large datasets, about 50% of all different mutations are found in only one single individual. In any genetics textbook, you will find that most mutations are detrimental, but this obviously does not apply to regions that have no specific function. Therefore, functionless regions in the genome can apparently be recognised based on the fact that the fraction of mutated bases is higher there. Conversely, you can discover in large datasets that there has been selection pressure over the centuries on mutations that would lead to a different **amino acid** (box 1); these types of mutations are becoming rarer.

The analysis of genetic variation at the base level is, of course, extremely fascinating, because it reveals where selection has acted during our history. It is a true goldmine for the various subdisciplines of genetics. Analyses on the progression of the emergence of genetic variation over the centuries will also appear, because each passage of genetic material through the **germline** leads to new mutations (table 3, 4). By applying WES/WGS in combinations of parents and children ("trio sequencing"), we now have a reasonably good understanding of the rate at which mutations arise (present in the child but absent in the parents, table 3). For base pair changes, this is around 60 per generation for parents who are about 30 years old (table 3).

Table 3 This table gives an impression of genetic variation in the human population based on **next generation sequencing** of a growing sample of many thousands of individuals, as published in 2019 and 2020 (see list of references). These estimates will become increasingly accurate as the sample size expands; there is currently a significant underrepresentation of Africa and parts of Asia. Because short read sequencing now exists in parallel to long read sequencing, the contribution of long read sequencing to our understanding of genetic variation will continue to grow. This is especially true for variation larger than 50 bases. These are referred to as structural variants/aberrations. A deletion is self-explanatory. In a duplication, a chromosomal segment has doubled locally. In an insertion, this duplication has ended up somewhere else in the **genome**. *De novo* structural chromosomal aberrations/variants generally arise much more frequently in the male **germline**.

Type of mutation	**Individual incidence**	**Total population incidence**	***De novo* mutation frequency**	**Through father**	**Through mother**
Base change	1 in 800 bases	1 in 8 bases	1.88×10^{8} per base per year: 56 (both parents 30 years old)	1.5×10^{8} per base at age 30: 45 in total	0.37×10^{8} per base at age 30: 11 in total
Indel	1 in 15,000 bases	1 in 100			
SINE insertion polymorphism	?	50,000	1 in 20 babies		
LINE SVA insertion polymorphism	?	4,000 SVA 4,000 LINE	1 in 100-200 babies		
Structural variants/ aberrations (unbalanced)*	median values				
Deletion	3505	186,000	*De novo* rate 0.3/child		
Duplication	732	6000			
Insertion	2612	121,000			
CNV	548	>12,000			

For an explanation of SINE, LINE, SVA, refer to table 2. For **CNV**, **indel** and **SNP**, see the glossary. The vast majority of mutation events shown here have no **phenotypic** consequences. The median value is the number that lies in the middle of the entire range of individual values. The publication of the nearly complete human genome by the T2T (telomere to telomere) consortium in spring 2022 provided new impetus to further develop our understanding of genetic variation in humans, for which this can act as a better reference genome compared to GRCh37 (on which table 3 is based) and CRCh38. In the meantime, a new genome consortium has emerged, called the Human Pangenome Reference Consortium. The aim is to construct a more accurate representation of genomic variation, including global genomic diversity. Of course, this raises questions about the genotype-phenotype relationship (Ch4, 6 and 18). For instance, a conclusion to date is that 48% of protein-coding genes have non-synonymous (box 1) amino acid changes between haplotypes (Ch5, fig 16).

The doubling time for the amount of paternal *de novo* SNPs within one generation is about 20 years after sperm production has started. Maternal *de novo* SNPs also increase with age but at a much slower rate. This increase is primarily due to a change from a C to a G, indicating the repair of a double-strand **DNA** break.

When comparing the rate of base changes between **somatic** cells (Ch2.2.2) and germline cells, you are bound to reach the conclusion that the germline is better protected against *de novo* SNPs than somatic cells. To properly make that comparison, you need insight into the number of cell divisions that separate the generations. As indicated for the male germline in chapter 12.3, this is not possible to determine in humans with reasonable accuracy. At the predicted rate of three base changes per **haploid** genome per division in somatic cells (Ch2.2.2), only 15 divisions are needed to arrive at the number of *de novo* SNPs from paternal origin per generation in humans (thus based on the somatic DNA replication accuracy). This points out that replication accuracy in the male germline must be higher, as 15 divisions is far too low. Recently, in 2021, this reasoning was revisited using **NGS** results, that confirm the low **mutation frequency** (27 times lower compared to the colon and rectum). However, the number of divisions between generations remains unknown. The same reasoning applies to the female germline.

* For structural variants (>50 bp), the development of analysis technologies, including long read sequencing combined with the examination of an increasing number of

parent-offspring trios, will lead to a higher estimate. In addition, more and more populations from around the world will be examined. By 2020, a number of >25,000 per individual is already mentioned. “Unbalanced” refers to the addition or loss of DNA compared to a reference genome. The numbers indicate the count of different SVs per individual and in the research population up to that point (publication 28 May 2020).

A large proportion of the variation at a single base position made its way into the population much earlier, which is extremely advantageous for the search for the **genetic background** of traits (Ch6). Geneticists speak about **SNPs**, “**single nucleotide polymorphism**s”. How do you discover such changes, preferably for many positions on the genome at the same time, without having to sequence entire genomes? This can be done with current chip technology, for example from the company Illumina. The chips used to determine SNPs resemble microscope slides. They contain about 12 fields with a size of 9 x 9 mm, each of which can easily read 500,000-1,000,000 SNPs. This is done using fluorescent light. The whole approach thus relies on a combination of microtechnology, automation and data processing. The key concept for SNP analysis is that a sequence of around 20 bases already functions as a unique address: at each position you have a choice between an A, T, C or G. Simplifying this calculation, let’s assume they all have a probability of 25% for that one position. The probability of that special base sequence across 20 adjacent positions then equals 0.25 x 0.25 and so on, equal to 0.25^{20}. This is small enough for this base sequence to be considered a unique address in the sea of 3×10^9 nucleotides per genome (Ch2.1). In reality, the frequencies of C and G are a bit lower than 0.25, and those of A and T a bit higher. In order to conduct genetic research with these data, the SNP base that is in the minority in the population must be present in at least 1 in 50 individuals. That individual is almost always **heterozygous** (and the rest are **homozygous**). Together, those 50 individuals have 100 bases for this one position, of which 1 is now deviant. In this case, the **allele frequency** is 1%, the minimum frequency that is useful for genetic analyses. About one in 1,000 base positions meet this condition. If the frequency is 1% or higher, which is usually the case, you can assume that the original mutation occurred in a very distant past, perhaps even before our departure from Africa.

Interestingly, this is not something you read much about. Another inference you can make is that, mostly, there has been no selection pressure on such SNPs, but when there is, it does not hamper the genetic analysis of the trait under study (rather to the contrary, Ch6).

SNPs are very useful for tracking changes in the population, such as the departure from Africa and the further spread of *Homo sapiens* around the world. Research on both of the aforementioned DNA banks has shown that the variation in the group that stayed in Africa is greater than in the group that left. This is because the latter was smaller in size and contained less genetic variation as a group, compared to the group that stayed. An extreme illustration of this phenomenon can be found among the Amish, who isolated themselves as early settlers in the USA about 14 generations ago and still adhere to a simple way of life. Perhaps for this reason, the *de novo* **mutation frequency** per base seems to be somewhat lower in this population.

3.3 Structural variants

Structural variants refer to all changes that affect a DNA segment (or segments) larger than 50 bases. In classical **cytogenetics** including **somatic cell genetics**, where studying chromosomes through the light microscope was the main focus (Ch3.4), one distinguishes between balanced variants (referred to as abnormalities, not involving any loss or gain of DNA) and unbalanced variants (where that is the case). The latter category includes missing segments (deletions), duplicated segments (duplications in tandem), or segments that have ended up elsewhere (and are called insertions, usually preceded by a duplication process) (table 3). In the past, due to the limited resolving power of the light microscope (chromosome segments of 5 mb DNA were roughly the smallest that could be detected), these analyses lacked sensitivity. Now, these limitations have been completely removed thanks to DNA sequencing techniques, from an SNP platform to WGS. That is where great knowledge gains lie. As it turns out, unbalanced structural variants leading to a different “dose” of DNA than the **diploid** reference state, are far from rare (see table 3). Detecting balanced chromosomal aberrations via sequencing takes a bit more.

In the most common type, two chromosomes have exchanged terminal pieces (a reciprocal translocation, fig 12 and table 4). This occurs remarkably frequently in humans and pigs, but hardly at all in cows and mice, a difference that remains unexplained to date. Reciprocal translocations can be seen as an ultimate attempt of a **germ cell** to repair one or more double-strand DNA breaks. So far, WGS has mainly used so-called “short read sequencing”, but this is quickly changing into “long read sequencing”. The DNA fragments that are physically analysed in short read sequencing are around 50-250 bases long. Therefore, “catching” the transition from the base sequence of one chromosome to that of the other is not easy. Long read sequencing will bring a change to this, but population data are not yet available for this technique. With this

method, the equipment processes long stretches of DNA, from 10 kb to as much as 2 mb, so transitions from one chromosome to the other will be better detected. The legend of table 3 discusses the implications of the addition of long read sequencing with regard to studying variation in the genome. It surely stimulates the discovery of structural variants that also play a role in evolution, despite the fact that they can be very obnoxious at the clinical level.

A special class of structural variants is brought about by the still ongoing process of jumping (inserting) LINEs and SINEs (Ch2.3.3, table 2, 3, fig 10). In humans, one out of every 100-200 newborn babies carries a new LINE insertion, and one out of every 20 carries a new SINE insertion. This process also leads to genetic differences between people, as shown in table 3. It is, therefore, not a coincidence that this type of mutation is also detected when searching for the genetic backgrounds of diseases. However, the mobility of SINEs and LINEs is not a dominant driving force behind genetic pathology.

Another source of structural variation can be found in DNA motifs ranging from 500 bases to a megabase in length, which can occur in tandem in a varying number of copies. This phenomenon is known as **CNV**, copy number variation. Table 3 demonstrates that this variation is not rare. Since 2007, CNV has attracted considerable attention due to the discovery that this process may play a role in neuropathology, as is evident in autism. Later, an involvement of CNV was also determined in schizophrenia and intellectual disability. Currently, the role of CNV in overall genetic variation and its contribution to human pathology is still not fully understood, but it is certain that it can play a role. Unbalanced structural variants, as well as changes at a single base position, are found in just one single person in about half the cases. It thus represents a unique mutational event, something that becomes more likely as the affected area grows larger.

3.4 The history of cytogenetics, imaging structural chromosomal aberrations and abnormalities in the number of chromosomes

Structural and numerical chromosomal variations were initially studied by cytogeneticists, specialists within the field of genetics who focused on studying chromosomes through the light microscope (fig 12). This approach, in particular, allowed human genetics to further develop after World War II. **Cytogenetics** may seem like a dying field, as the techniques used in it are being surpassed by DNA analysis methods, but that sells the discipline short.

From the very start, it proved challenging to routinely study mammalian chromosomes. The time it takes for a cell to divide is relatively short compared to the time that elapses between

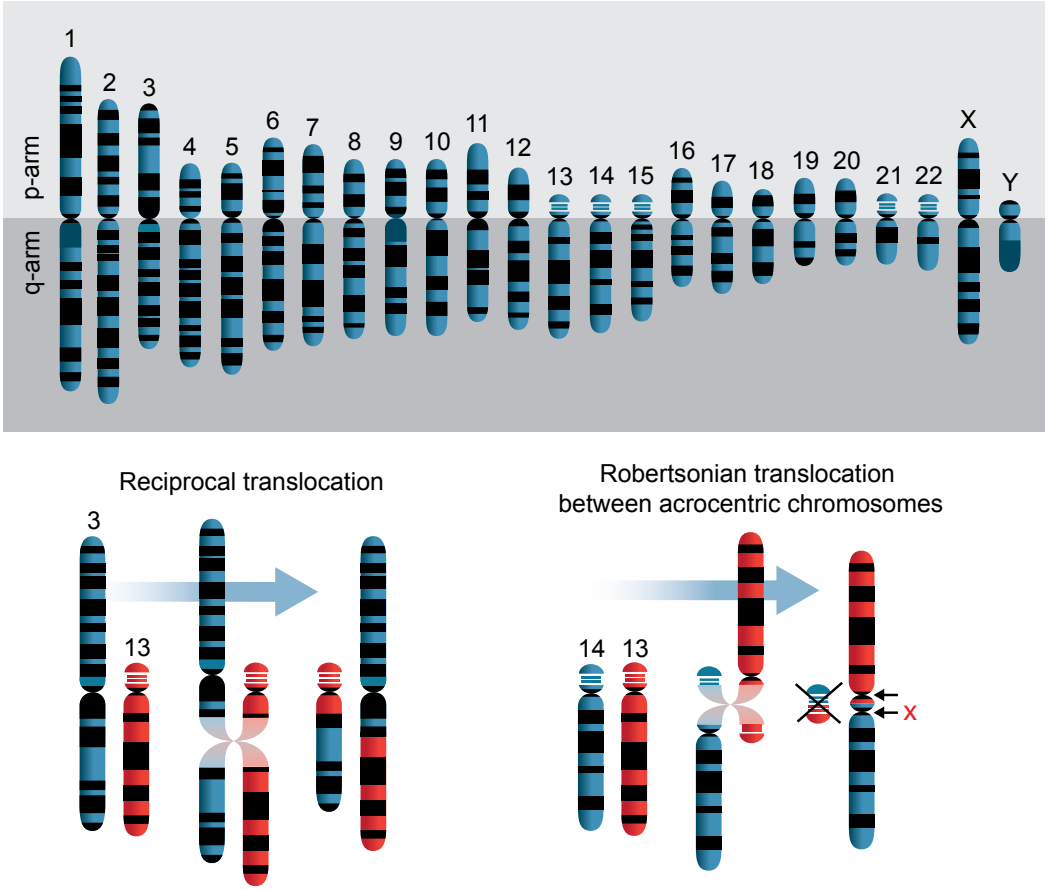


FIG 12

A human **haploid karyogram** based on G(iemsa) banding. The high resolution, reflected here by the greater number of bands, can only be achieved when the **chromosome**s are not fully condensed yet, thus on their way to metaphase (fig 3). The light bands represent gene-rich regions, the dark bands represent regions with fewer genes. Reciprocal and Robertsonian translocations are not rare in humans (table 4). The latter category includes translocations that take place in the **centromere** region with **heterochromatin** (table 1), between chromosomes whose centromere is, microscopically viewed, close to one end (acrocentric, chromosomes 13, 14, 15, 21 and 22). The most common mode of occurrence is shown here, where the two centromeres coordinate their activity or one is repressed (the red cross).

Table 4 The **cytogenetics** of **chromosomal** aberrations at birth, based on a large Danish population study (n = 34,910) from 1991, see list of references. Estimates of the likelihood of the occurrence of a new translocation stem from the same period and are from two publications (listed in the references). Note that these numbers are from before population-level prenatal diagnostics.

Structural chromosomal aberrations balanced *	**Per 1000 births**	***De novo* frequency**
Robertsonian translocations**	1.23	0.7×10^{-4}, 1.1×10^{-4}
Reciprocal translocations***	1.43	2.7×10^{-4}, 5×10^{-4}
Inversions	0.34	
Structural aberrations unbalanced *		
Deletions, duplications, markers, rings	0.29	
Numerical aberrations		
Autosomes ****	2.07	
Sex chromosomes (including cytogenetic variants) *****		
47,XXY	0.83	
47,XYY	0.60	
47,XXX	0.52	
45,X	0.26	
Total	7.57	

* An unbalanced chromosomal aberration involves the addition and/or loss of genetic material.

** Robertsonian translocations almost exclusively occur in the female germline.

*** Reciprocal translocations almost exclusively occur in the male germline.

**** Almost all of these are **trisomies** 21 (Down syndrome). In mothers under 30 the occurrence is about 1 in 1500, over 40 it rises to 1 in 100 births. About 90% of them are of maternal origin, which is generally applicable to trisomies. The exceptions to this rule are 47,XYY (all paternal) and 47,XXY (fifty-fifty paternal and maternal).

***** To estimate the frequencies by sex, the numbers can be doubled. The researchers report 51.2% boys.

two divisions (fig 4). Searching for dividing cells is a time-consuming and exhausting task. Fortunately, in 1953, Charles Ford and John Hamerton discovered the action of colchicine as an inhibitor of the **spindle apparatus** in the lead-up to metaphase (fig 3, 4). They noticed that when using this, chromosomes remained in the highly shortened, even more condensed form.

Colchicine is a so-called alkaloid. These are nitrogenous substances of plant origin, characterised by complex ring-shaped structures. Caffeine and morphine also belong to this family, as well as vincristine from the pink periwinkle plant, which also inhibits the spindle apparatus. Colchicine is extracted from autumn crocus, an autumn-flowering crocus-like plant. With the discovery and use of these types of substances, a prerequisite had been met to routinely start chromosome research. Another obstacle was related to the challenges of loosening chromosomes that are tightly packed in the nucleus before interacting with the spindle apparatus (fig 3). In the 50s, it was accidentally discovered that washing cells in a simple weak hypotonic saline solution makes them swell. This increases the distance between the chromosomes awaiting division. The trick is to retain this situation by using a fixative. A mixture of three parts methanol (or ethanol) and one part acetic acid was (and is) routinely used for this purpose. The finishing touch consists of transferring a few drops of fixative with cells to the microscope slide, and letting the specimen dry until it reaches a state referred to as “air dry”.

This is how the chromosome number of a diploid human cell was determined to be 46 in 1956, in a race that Charles Ford and John Hamerton narrowly lost. However, they were able to confirm the results obtained by Tjio and Levan from Sweden by the end of that year. After that, the first reaping of genetic diagnoses in humans began. In 1959 alone, the following discoveries were made: for the **autosomes**: Down syndrome results from the presence of one extra chromosome (47,XX,+21; 47,XY,+21); for the **sex chromosomes**: Turner syndrome is caused by the absence of one X chromosome (45,X) and Klinefelter syndrome results from the presence of an extra X chromosome (47,XXY). In the diagnosis of Turner syndrome, Charles Ford won the race. Charles was an extraordinary man, eternally driven, sharp, somewhat eccentric and incredibly kind.

In order to analyse a field of chromosomes fixed with methanol and acetic acid before the computer age, you had to photograph them, print copies, cut them out and then match the chromosomes as **homologous** pairs. This is how one creates a **karyogram**, the systematic arrangement of identical chromosomes into homologous pairs. Sometimes, the differences in shape are so pronounced that you can confidently place the copies from both parents side by side (although you obviously don’t know which chromosome comes from which parent). Only the

sex chromosomes of a male cell can often be attributed to a specific parental origin, helped by the fact that the Y chromosome is usually relatively easy to recognise.

Depending on the species, a smaller or larger part of the set of chromosomes could not be unambiguously identified. Therefore, those early karyograms did not actually amount to much.

In 1970, Torbjorn Caspersson of the Karolinska Institute in Stockholm published the results of a chromosome staining with the **fluorochrome** quinacrine mustard. Each chromosome was found to have its own fluorescence pattern when viewed under the appropriate microscope. For experienced pattern recognisers, it was no longer difficult to create a karyogram that was accurate. A year later, it was discovered that using the old-fashioned nuclear dye Giemsa (a mixture from 1904, named after the German bacteriologist and chemist Gustav Giemsa) can produce an even clearer banding pattern that can be observed under an ordinary light microscope (fig 12). Quinacrine binds better to DNA segments that are richer in AT/TA base pairs. One of the discoveries from those years was that a part of the long arm of the human Y chromosome (fig 38), the part that does not contain genes, that is **heterochromatic** and can vary quite a bit in length, fluoresces most brightly of all, and also lights up in sperm cells. This way, sperm cells can thus be sexed, but they are no longer useful afterwards.

So, the analysis method using Giemsa bands has triumphed. The longer the chromosomes, the better the results, but the more challenging it is to deposit the chromosomes side by side on the glass. An American publication appeared showing as many as 2000 bands in human chromosomes, still 1.5×10^6 (1.5 mb) per band. This way, it was possible to see how the dark bands merged as the chromosomes became shorter. Soon, it became clear that the darker bands replicate their DNA later in the S phase (fig 4) than the lighter bands. The LINEs (Ch2.3.3) are located in the darker bands and the SINEs in the lighter ones. Giorgio Benardi's CG-rich segments are in the light bands, and the AT rich ones in the dark bands (Ch2.3.3). As such, there was already plenty of speculation about the "compartmentalisation" of the genome years ago. The appearance of bands naturally fuelled the discussion, which is now being reinvigorated with the discovery of TADs and how they can aggregate into higher-order structures (Ch2.3, fig 7, 12).

In order to routinely examine the chromosomes of humans and mice, one needs a cell type that can easily be obtained. Blood contains white blood cells, but those do not divide. The substance PHA (phytohemagglutinin, an extract from the bean *Phaseolus*) stimulates a portion of these blood cells to divide, which was discovered around 1960. This paved the way for cytogenetic testing in patients: from 1978, the determination of a chromosome profile became included in the general coverage of medical procedures in the Netherlands.

3.4.1 The microscopy of structural chromosomal aberrations

Cytogenetics provided us with a first insight into structural chromosomal aberrations, especially after chromosome banding became the default for making a karyogram. Now, exchanges of chromosome segments between different chromosomes, the origin of a reciprocal translocation, were easier to detect (fig 12). When constructing the karyogram, you would end up with four chromosomes: the two original ones, and the two with exchanged segments. In humans, this actually occurs remarkably frequently (table 4). Usually, carriers of such translocations do not exhibit any visible external signs. In couples using IVF, they are more common in situations where the man is the reason for referral (Ch12.12.1, table 6). Another type of translocation, the Robertsonian translocation, is also not rare (table 4, for an explanation, see fig 12). Translocations must have played a role in evolution; otherwise, karyograms of different mammalian species would not have been so different. To observe a reciprocal translocation with chromosome banding, the chromosome segments exchanged between two non-homologous chromosomes (fig 12) must be quite large, at least in the range of several megabases.

Occasionally, such a structural chromosomal aberration establishes itself in the population. In humans, around 2% of individuals are carriers of an inversion in chromosome 9 (fig 12). The inverted chromosome segment is heterochromatic and contains the **centromere**. This does not cause any issues during **meiosis**. Cytogeneticists refer to it as a neutral variant.

In mice, reciprocal translocations hardly ever occur spontaneously, and they mainly arise after irradiation of the testis (with **ionising radiation**). Carriers of a reciprocal translocation in mice, pigs and dogs have greatly reduced litter sizes. As a result of homologous chromosome pairing during the first meiotic division, the four involved chromosomes seek each other out (Ch5.1, fig 12, 14). During subsequent meiotic reduction divisions, this leads to the production of **gametes** of which half or more than half simultaneously lack one and have a duplicate chromosome segment (fig 26). After fertilisation, this, in turn, leads to **embryo**nic mortality around, and shortly after the period of implantation. In humans, you would expect a prolonged time between successive births, which, however, goes unnoticed in most societies today. Half of the children will inherit the reciprocal translocation, while the other half will be chromosomally normal. Each reciprocal translocation is different; there are some that may remain hidden in the population for generations, with the frequency of this being unknown. There are also some that come with an increased risk of having a child with a congenital defect. This is due to the single absence and

triple presence of the exchanged chromosome segments in situations where the genome was not sufficiently out of balance to lead to prenatal mortality.

A new optical technique has recently been presented by a company called Bionano, and it bridges the gap between cytogenetics and sequence-based technologies. The technique uses fluorescence labelling of ultra-long linear DNA at a specific sequence motif and subsequent analysis of these linearised molecules in nanochannels. The resolving power for deletions and insertions is in the order of 500 bp. It is greater for reciprocal translocations, inversions and duplications, at around 40 kb.

3.4.2 The microscopy of aberrations in chromosome number

Attention to the fact of aberrations in chromosome number in humans, arising around conception, was already drawn by the well-known relationship between a mother's age and the likelihood of her having a child with Down syndrome (after the establishment of its origin in 1959, more on this in Ch11.4.1). The work of a French couple in the 70s shed new light on this matter. J and A Boué conducted systematic cytogenetic research on cell material from spontaneous abortions in humans, almost all of which take place in the first three months of pregnancy. In about 60% of this series of 1498 miscarried **foetus**es, the chromosome number was incorrect, rendering them nonviable. Most commonly, they observed **aneuploidy**, the presence of an extra chromosome (54% of abnormal **karyotype**s), and, much less frequently, a missing chromosome (always 45,X Turner syndrome, 15%). **Polyploidy** (**triploidy** and sometimes tetraploidy) was present in 26% of chromosomally abnormal foetuses. Later studies, in which all chromosomes could be distinguished, confirmed that **trisomy** 21 (Down syndrome) is not uncommon in spontaneous miscarriages (this was the case in at least 10% of trisomies).

At that time, this was a sensational discovery. The work was published in a book on ageing egg and sperm cells, which caused it to fall into oblivion, despite the fact that the Boués were the first to demonstrate on such a large scale that determining the correct chromosome number prior to and at conception (and, to a lesser extent, just after) shows that something goes wrong relatively often. The data from their work are no longer found in current overviews on this topic. As a tribute, the journal *Birth Defects Research* reprinted their publication as a classic paper in 2013.

Very recently (2023), an update was published that extends the analysis to using a genome-wide SNP profile which includes data from the parents. A series of 1,745 spontaneous abortions was analysed with state-of-the-art cytogenetics, mostly using Giemsa banding (fig 12). Again,

50% of specimens were karyologically abnormal, with 40% of them attributed to autosomal aneuploidy, 12% to sex chromosome aneuploidy (also called gonosomal aneuploidy) and 32% to polyploidy. The remaining 16% contained structural aberrations/variants and combinations of the aforementioned categories in more or less equal shares. The main question of this report was whether high-level analysis through genome-wide SNP profiling, which incorporates the genomes of the parents, could enhance the ability to detect abnormal karyotypes, including structural variants/aberrations that cause spontaneous abortion. The main conclusion of this small series of 94 specimens (that were found to be "normal" when standard cytogenetic testing was used), is that one-third of them were abnormal: more aneuploidies were detected, and in line with observations derived from ART (Ch16.3.1), it was found that aberrations can have a mitotic origin, emerging during the cleavage divisions. This finding could only be detected with SNP analysis that includes data from the parents (this way, a meiotic origin can be separated from an early embryonic mitotic origin).

From the 70s onwards, it became possible to obtain embryos in mice that either had one chromosome too many or too few for each of the 19 autosomes. The embryos with one chromosome too few all perish around the time of implantation or shortly thereafter. The survival of the embryos with one chromosome too many is much more variable, but none survive until after birth, even when it is the smallest mouse chromosome that has the extra copy. Apparently, an insufficient amount of gene product is considerably worse than an excess of it. This phenomenon is known as **gene dosage effect**s, which includes **haploinsufficiency**. Proteins invariably interact with each other in many ways, and in such cooperations, ratios matter (Ch1). The fact that, in humans, trisomies for chromosome 21, although much more susceptible to prenatal mortality than chromosomally normal embryos, can still often go through the entire pregnancy and be born, is partly attributed to the relatively low number of protein-coding genes on this chromosome.

For the sex chromosomes, it is a different story. In each cell, every X chromosome beyond the number of one is inactivated (Ch2.3.2, box 4). If this process worked perfectly, women with three X chromosomes would not stand out in any way, but that is not the case: X chromosome inactivation is not complete (box 4). Despite how interesting that is in itself, you rarely come across it in the many public journalistic investigations into the difference between men and women. This incomplete X chromosome inactivation also plays a role in Klinefelter syndrome (XXY, table 4). It also explains that Turner syndrome (45,X) is a real one with associated

phenotypic characteristics. Finally, two Y chromosomes are tolerated (table 4), with the Y being **gene**-poor (Ch12.12.2, fig 38).

In the winter of 1976-1977, I worked at the Medical Research Council Clinical and Population Cytogenetics Unit next to the Western General Hospital in Edinburgh. In Scotland, almost every baby is born in the hospital. Since the lab specialised in creating chromosome profiles, they started karyotyping more than 10,000 consecutive hospital births. Of course, there were babies with abnormal karyograms, including for the sex chromosomes (table 4). Shirley Ratcliffe worked as a paediatrician on follow-up screening of these babies. Even then, the question was what you could or would tell parents about their child's genetic makeup, when that makeup is not immediately apparent in the phenotype and a deterministic interpretation of the genetic diagnosis is undesirable. For example, they did not inform parents about a little boy with an extra X or Y chromosome, but they did want to perform follow-up tests.

In those with Klinefelter syndrome (XXY), **spermatogenesis** is impaired, which will be discussed later (Ch12.12.1). Offering reasons that were not very clear, children were lured to the hospital for check-ups. Shirley Radcliff's bifurcated skirt with Scottish tartan motif had two deep pockets. In one of them, she had a string of beads of gradually increasing size (an orchidometer). Now, the art was to palpate a testicle while, with the other hand, finding the bead of comparable size − our sense of touch is well developed! With regard to the XYY boys, she mentioned that they were more prone to having accidents due to clumsy behaviour. A publication on this came out in 1999, which was her last one related to the supposed aggression of men with an extra Y chromosome. The conclusion, based on 16 XYY boys found among the 34,380 karyotyped babies between 1967 and 1979, was tentative. Yes, there was slightly more antisocial behaviour and even a hint of criminality, but that could mainly be attributed to reduced intelligence. Furthermore, even this study was still somewhat statistically limited, she noted.

In the next chapter (4), I will try to address, as much as possible, the biological relevance of genetic variation as the basis of explaining phenotype. Afterwards the role of the germline as a driver (and/or suppressor?) of genetic variation will be described.

CHAPTER 4

From genotype to phenotype

4.1 Genetic dissection

So how does genetic variation manifest itself in what we see in another person, what we look like ourselves and how we function − in short, how does **genotype** translate into **phenotype**? This is a kind of holy grail, and it is the subject of countless studies. Most attention is focused on the involvement of genetic variation in disease and health, but also the genetic basis of intelligence, personality and BMI all pop up in the professional and popular literature. A fairly large group of **gene**s is responsible for **recessive**ly inherited disorders (“monogenic inheritance”), with most of them being extremely rare. Aberrations in genes that are **dominant**ly inherited are comparatively less common (Ch18.1). Not every **gene** leads to a distinct phenotype if only one copy is functionally present. When this is the case, it is referred to as **haploinsufficiency**, the harmful effect of a single gene dosage, which can manifest as dominantly inherited.

When the technique to knock out genes using **homologous recombination** in **embryonic stem cell**s was established in the late 80s in mice, (box 2, 3, fig 19), it marked the beginning of a prosperous time. You can now use genetic tools more purposefully to solve a cell’s biological or general physiological puzzle, thereby making a link to the phenotype. It does not matter where that gene is expressed, for example, in the reproductive organs or in the early embryo. This research approach is known as **genetic dissection**. As a PhD student, you could thus embark on a project that was based on the inactivation of a supposedly important gene, which is, for example, involved in the energy **metabolism** of the cell; it looked promising, otherwise it would not have been approved by the scientific community. When the long-awaited **knockout** mice (box 3, fig 19) were finally around, it was often the case that nothing about them seemed

different at first glance; but who knows what might be revealed when using more precise methods to determine the phenotype. In the field, this phenomenon is captured by the term "**genetic redundancy**".

"All roads lead to Rome", and the cells expressing this gene are not easily misled in their ultimate functionality. However, the opposite can also happen, finding something you were not at all looking for. For instance, the male knockout homozygotes of your favourite gene unexpectedly turn out to be sterile. The discovery that embryos that inherited defective gene copies from both the father and the mother (and that are thus **homozygous** -/-) are not viable was another surprise. In line with that theme, it has been found that knockouts for genes involved in repairing a double-strand **DNA** break (fig 33) very often produce embryos that will die around implantation. When the embryo enters the phase of rapid cell division and growth, shortly after implantation in the uterine wall, things start to go wrong. By 1994, the toolbox of gene manipulators was already extensive enough to address these issues. Ideally, the effect of the knockout should be limited to the specific process you wish to study, for example, **spermatogenesis**. Box 3 and figure 20 explain how this works. By using homologous recombination in stem cells (box 2), you can basically alter the gene of choice in any desired cell type, but the complexity and high costs of the method persist. CRISPR-Cas technology (box 3) has greatly simplified genetic manipulation, but that does not mean that it is now possible to make any tissue- or cell type-specific modifications at low cost.

We know the function of many genes, but for about half, this remains unknown. The scientific community naturally strives to obtain the most comprehensive understanding possible. A lot of information can already be deduced from a gene's DNA code. As a next step, you want to know in which cells and tissues the gene is expressed, and databases exist for that as well. However, a more definitive insight is obtained when the gene is knocked out by a **mutation** (box 3). The international mouse community now aspires to do this for every gene. The Jackson Laboratory (Bar Harbor, Maine and a very large player in mouse genetics) is a driving force in this. This global initiative is known as the International Mouse Phenotyping Consortium (IMPC). A total of 19 institutions from around the world participate in this project, with a strong presence in North America, Europe (including the MRC unit in Harwell, Ch2.3.2) and China, but there are also labs in India and South Africa. It is essential that the method of phenotyping is done according to a standardised protocol. Bob Braun, a scientist affiliated with the Jackson lab who is strongly involved with the IMPC, predicted that we would have reached half of the approximately 20,000 or so genes by the end of 2021. Up to January 2024, 8,707 genes have been examined in this

way. Of the 7,824 genes analysed up to 2021, 871 were related to reproduction, with 200 of those associated with sex determination. For 308 genes, knockout led to fertility impairment.

The step from the phenotype of the knockout to determining the precise function of the gene is still a large one.

Proteins can be incredibly versatile, and even for the product of a thoroughly studied gene like p53 (Ch2.2.1), with more than 50,000 publications in the database, new aspects are still being uncovered. A safe assumption remains that we still don't have a good idea of the function of half of the protein-coding genes, and for the remaining ones our understanding is often incomplete.

I once met Bob Braun at a Gordon Conference. Gordon Conferences, an American initiative that has been around for years, provide an informal stage for leading researchers and their young apprentices in a specialised area of science. In the mornings, there are presentations and so on, while everyone has the afternoons off. Depending on the conference venue, some people go golfing while others go swimming in a beautiful pine-lined lake with clean water. After dinner, there are more presentations and afterwards people hang out until (very) late. These types of meetings are always competitive but, at the same time, very informal, a positive side of the American science system.

Looking at the physiological role of gene functions, it is striking that the essential genes are heavily involved in development and embryology and that they are often important in protein networks. These are also genes to which genetic redundancy does not apply. In other words, there are no backup copies to compensate for damage. These backup copies fall under the concept of **paralogs**, a logical consequence of the fact that duplications of genetic material played a role during evolution (Susumu Ohno, Ch2.3.2). Our **genome** is simultaneously very robust − flexibility in DNA is tolerated − and very fragile, as a single incorrect base can already have dramatic consequences.

4.2 Genetic dissection in humans

In 2000, the first version of the base composition of the human genome was published. A famous photo shows a proud President Clinton standing next to Craig Venter of Celera Genomics, the company he had founded to accelerate the pace of the **Human Genome Project** of the 90s. On Clinton's other side stands Francis Collins as a representative of the public funds of the NIH.

Since 2000, several editions of the DNA base composition, our genetic identity, have appeared. The spring 2022 version was produced by the T2T consortium, where the T stands for

telomere. The combination of long read sequencing with the more traditional approach of DNA base sequencing has filled in the gaps where the repeat structure (see, for example, table 2) of the genome was a hindrance, such as at the **chromosome** ends, in and around the **centromeres**, and in the short arms of the acrocentric chromosomes 13, 14, 15, 21, and 22 (fig 12). The human genome is now almost completely mapped, and the published 2022 (female) version can serve as a reference for the notation of variants (table 3). The sequence of the Y chromosome, which is even more resistant to the earlier sequence technology, was published in 2023, and more attention to the **genomics** behind this will be given in Ch12.12.2.

Assisted by the ever-expanding human presence on Earth, the advancement of the combination of "medical science − genetics" and the international nature of scientific research, we ourselves are conducting genetic research on a large scale. And all of this is due to the never-ending process of spontaneous mutation in the **germline** (Ch3, table 3). For the detection of single gene-based (monofactorial, monogenic) genetic disorders, the hospital (the combination of specialists and clinical geneticists, fuelled by technological advances in DNA analysis) works as a filter. And since the number of births in countries with good medical care is high around the world (ranging from 6-20 per 1,000 inhabitants per year), biologically extremely interesting, but personally touching phenotypes are left on this filter. The technological developments in genotyping that have happened between the 70s and now, culminating in the application of **WES** and increasingly **WGS**, are reflected in the number of genetic diagnoses in the Netherlands, from less than 1,000 before 1978 to 125,000 in 2018 (a more recent update could not be obtained). So this development is brought about by the interaction between genome technology and the knowledge of the genes responsible for "single gene phenotypes" after mutation, see below. However, the elucidation of the **genetic architecture** of **polygenic** traits (Ch6), if not too genetically complicated, will increasingly add to this trend. One may wonder if there is an end to this development, knowing that it is difficult to influence the environment as a variable in the expression of polygenic traits (Ch6); examples are nutrition and low-dose complex chemical pollution (Ch17).

In the 70s, Victor McKusick of Johns Hopkins University in Baltimore began cataloguing human genetic disorders caused by a defect in a single gene. He also listed the mode of inheritance, recessive, dominant or intermediate. We were still in the era of card index boxes, and he worked on this in the early hours of the morning, aside from his regular job as a cardiologist. The list was known as "McKusick's list" and I remember that at the time, around 1980, it contained approximately 400 genetic defects. Every summer, the professor would travel to Bar Harbor, Mount Desert Island in Maine, for high-level science in the breathtaking setting of Aca-

dia National Park. On this island – actually a peninsula, connected to the mainland by a bridge, which was a playground for wealthy families such as the Rockefellers in the 30s – the Jackson Laboratory was founded in the year 1929. This laboratory houses the world's largest collection of mice with genetic abnormalities, alive and/or frozen as embryos in **cryogenic** vessels with liquid nitrogen, in addition to numerous inbred mouse strains and combinations of these inbred strains. The data storage of genetic knowledge of the mouse is also primarily in the hands of the Jackson lab. Now, when researchers need a mutant for a specific gene (most commonly a **knockout**, box 3), they can browse through the collection of the Jackson laboratory. However, it will often be quicker to create the mutant themselves using CRISPR-Cas technology in the desired background genotype (box 3, and especially when inactivation of multiple genes is desired).

"McKusick's list" is now called OMIM, Online Mendelian Inheritance in Man, a catalogue of human genes and genetic disorders. In February 2024, this database already lists 4,526 monogenic disorders (see also Ch18.1). As the analysis of the base composition of the human genome is now more or less complete, it is expected that this number will increase even further.

An example is "entry 176670", Hutchinson-Guilford progeria syndrome (HGPS), a dramatic ageing disease that is very rare and occurs spontaneously in the majority of cases. It is said that the disease is "dominantly inherited" and that the mutation originates in the germline of either the father or the mother. It occurs in one in 4-8 million births, and is hence extremely rare. But, coincidentally, a pupil at the preschool of one of our children happened to have this. Children with this form of progeria usually would not live past the age of 12, but that is slowly changing now, thanks to medical treatment of the symptoms: a lifespan of 20 years is no longer an exception. These children are psychologically normal and exceptional, but the deterioration actually begins in their first year of life. People with this HGPS are spread all over the world. Initially, the gatherings where families and children could meet would only be held in the USA. However, since 1997, there has also been a European Progeria Family Circle that aims to bring European children and their families together every year. Despite all the poor prospects and daily problems, you witness a fairytale world at such gatherings, especially when it takes place in the Efteling (a fairytale-themed amusement park in the Netherlands). Children with progeria are endearing, and the discovery of the affected gene in 2003 has raised hopes for a cure. This gene that is responsible for premature ageing has attracted publicity and is intuitively linked to a longer life. Two groups released their findings around the same time. There was an indication that the gene for this form of progeria lies on chromosome 1 (fig 12). Through the arduous procedures of those days, in which polymorphisms for microsatellites (Ch2.3.2, table 2) played

a major role, an American group came to the conclusion that this was a gene from the lamin family, specifically the gene encoding lamin A. The results were published in *Nature*. The European path to solving this genetic puzzle was quite different. In Marseille, a clever doctor noticed that the jawline of the children with progeria resembled something he knew from the so-called laminopathies. There are three types of lamin genes, and a gene defect can be less severe, so that it does result in a phenotype, but not in the striking pattern of early ageing. At that point, the researchers immediately started looking for mutations in the lamin genes in the children with progeria, and they arrived at the same result as the Americans. The publication followed in the same year in the journal *Science*, in the scientific world as highly regarded as *Nature*. This is truly a fine example of serendipity. The protein lamin A plays a role in the contact between **chromatin** and the nuclear membrane (fig 2, 7) and that contact is extremely important for DNA stability. The defective protein progerin, resulting from the most common mutation, dysregulates the cell in many ways, with DNA repair being an important component of this. There are now mouse models in which gene therapy for the most affected tissues can be tested, and a breakthrough that happened in early 2021 could be the first step towards therapy.

Currently, the findings of large-scale genome studies that give us an insight into DNA variation (Ch3) are being merged with information from all other sources, two of which have just been mentioned, concerning mice and humans, respectively.

At this point, I would like to revisit the concepts of haploinsufficiency and genetic redundancy. A measure of a gene's indispensability is whether loss of function of a single **allele** results in haploinsufficiency, manifested in a phenotype. In 2018, Craig Venter co-authored a widely cited paper that raised the question of how many of our approximately 20,000 genes we really cannot spare one of the two copies. That question is not easy to answer, not even in mice as the genetic model for mammals, as we have just seen. However, with the advent of large genome projects such as gnomAD, and with the help of statistical methods/bioinformatics, it is possible. The estimated number of genes with this sensitivity is 3,000. The presence of a defective gene copy is referred to as "loss of function". This is, of course, what is searched for in human genetics, as well as in farm animals. The number 100 appears to be an estimate of the number of "loss of function" variants/gene copies/alleles per individual. Most of these likely go phenotypically unnoticed. The number of variant alleles that are not transmitted (due to mortality and infertility) in the homozygous recessive state is estimated at 0.6 per person. A measure of dispensability (genetic redundancy) is the presence of one or two defective gene copies without a phenotype.

Another indication of lower importance is the absence of selection, reflected by the presence of more alleles with mutations for that particular gene and genetic variation around it in the population. When this is higher, the selection pressure is lower. A clear picture of an actual number of genes that can be missed in the homozygous state seems difficult to obtain, although a number of 3,000 is mentioned by Venter. Involved factors are the non-severity of the phenotypic consequence(s) of missing the gene (for instance, for olfactory genes), and the fact that evolution by gene duplication can yield copies that have not yet evolved into separate functions, so they can still compensate each other to some degree.

The relationships between the genome and phenotype can be viewed as solving an infinitely complex puzzle, of which the true number of pieces and their interactions is unknown. However, as we progress, strongly stimulated by the technology to determine DNA base sequences, computer intelligence and linked, to that, the methods to better "visualise" structural variants, our understanding will steadily increase.

NB: an approach that has been used since the emergence of DNA technology in the 80s to determine the relationship between genotype and phenotype and which has become increasingly refined, is briefly discussed in chapter 6 under the abbreviation **GWAS ("genome-wide association study")**.

CHAPTER 5

The meiotic divisions

The change from a **diploid** to a **haploid** stage, and then back to the diploid stage, is intricately linked to the concept of the alternation of generations. It turns out that the duration of the diploid versus the haploid stage can vary extremely between the many organisms that reproduce sexually. When you look up the concept of alternation of generations on Wikipedia, for instance, you enter a complex world of classifications. Mosses are an example of a life form that primarily exists in the haploid state.

For plants, invertebrates, and vertebrates, the diploid phase in life is by far the longest (a bit of an open door). To transition from the diploid to the haploid stage, cells in the **germline** go through meiotic or reduction divisions (**meiosis** fig 13). Compared to **mitotic** division, the whole process takes a long time, and it can be “frozen” at a certain stage in order to give priority to other processes, such as during the meiotic prophase of **oogenesis**. The same is true for the meiotic prophase of **spermatogenesis**, albeit to a lesser extent. In this chapter, however, we will focus on the basic principles that are the same for both sexes.

5.1 The pairing of homologous chromosomes

When **homologous chromosome**s have to part ways in an orderly manner, some preparation is in order. All of this happens in the prophase (fig 14) of the first meiotic division, which is quite lengthy for that reason. Cytologists and cytogeneticists have divided the prophase of meiosis into stages based on the behaviour of the chromosomes. During genetics lectures, this was one of those learning sequences that was quickly forgotten (fig 14). The process begins with the shortening of the chromosomes which, since chromosome duplication has taken place, consist of two

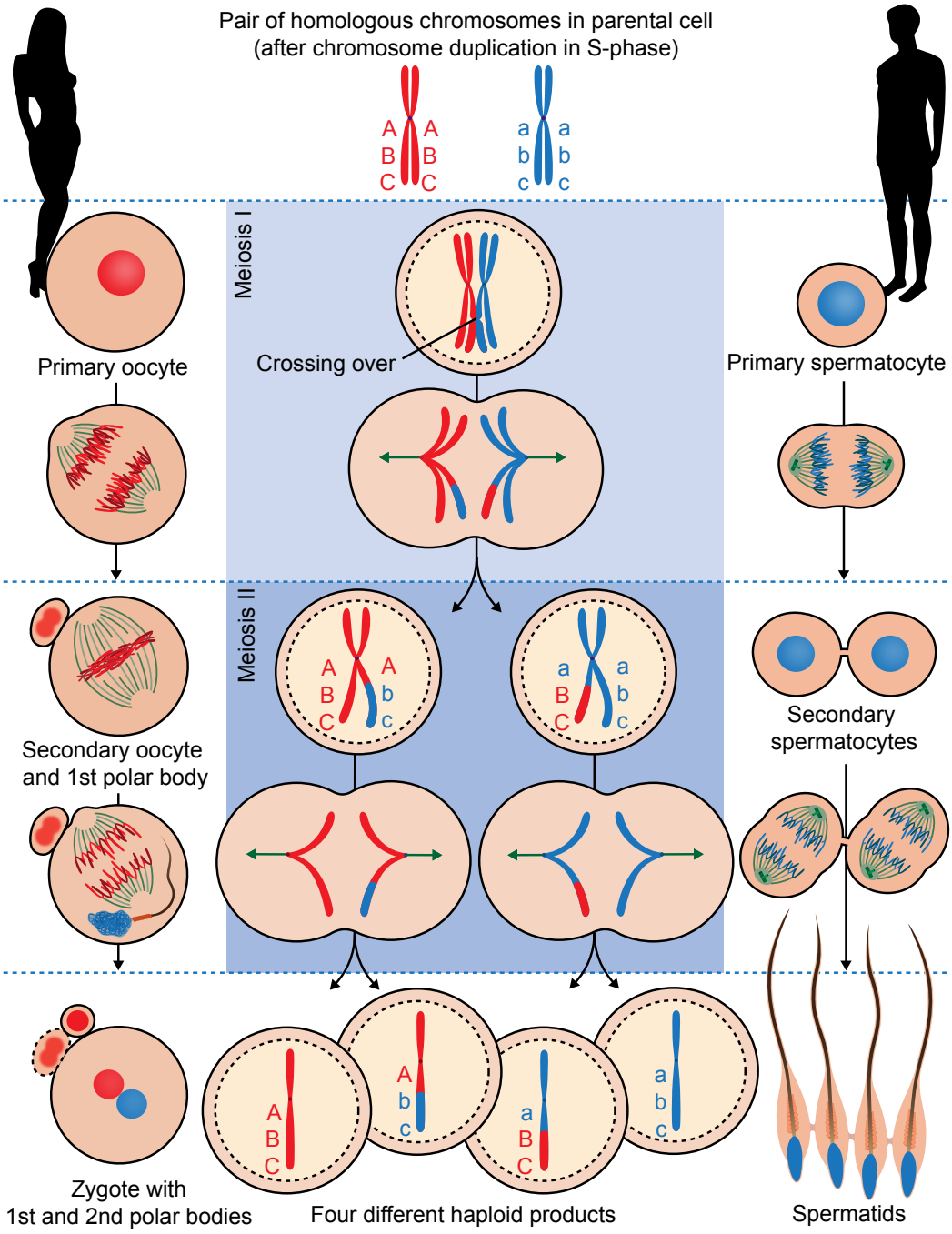

Pair of homologous chromosomes in parental cell
(after chromosome duplication in S-phase)
A A
B B
C C
a a
b b
c c
Meiosis I
Meiosis II
Primary oocyte
Crossing over
Primary spermatocyte
Secondary oocyte
and 1st polar body
Secondary
spermatocytes
Zygote with
1st and 2nd polar bodies
Four different haploid products
Spermatids

FIG 13

Summary of **meiotic chromosome** pairing, which allows **crossing over**, and the two meiotic divisions for a single **bivalent**. The meiotic stages of **oogenesis** and **spermatogenesis** are called oocyte and spermatocyte, with the prefix “primary” for the first meiotic division and “secondary” for the second meiotic division. For the sake of simplicity in the figure, only a single crossover between **loci** A/a and B/b is shown (with **allele**s A,a; B,b). This leads to **recombination** within the chromosome, represented by a colour difference between the **homologues** (red for maternal, blue for paternal). The first and second meiotic divisions happen from the beginning of the anaphase (see fig 3). The figure also serves as an illustration of Mendel’s first law. Because **bivalents** in the first meiotic metaphase and chromosomes in the second meiotic metaphase randomly orient themselves in the **spindle apparatus**, the four meiotic products for each bivalent have an equal chance of being transmitted from the oocyte into the **zygote** (the second meiotic division is completed after the secondary oocyte (the egg cell) is activated by the sperm cell, and a zygote is formed, see Ch13.5.1 and fig 42). The male **germline** is shown up to the stage of elongating spermatids (fig 31), each of which originates from a single spermatocyte with equal probability. This equality of chance remains intact until fertilisation (but see Ch12.3).

chromatids, as in **mitosis**. The process of finding the right partner is referred to as “homology search” in the literature. This is one of the most intriguing and completely miraculous processes that occurs in the germline. At that point, the nuclei are remarkably small. You might expect that a bit of space would make the search easier, but apparently, it does not. Somewhere at the beginning of this journey, the ends of the chromosomes, the **telomere**s (Ch2.3.3), locate each other. To further add to the crowding, they all tend to accumulate in one sector of the nuclear membrane, making it just like a tangle of strings of different lengths of which you are gathering the ends. These are discoveries from the early 90s, when antibodies against proteins involved in the regulation of telomeres became available and could be used to microscopically visualise the beginning of homologous chromosome pairing. The **DNA** code for telomeres is the same for each chromosome end (Ch2.3.3), so that will not help to locate the homologous chromosome. However, that is not intelligent speculation. At the domain where the telomers cluster, they attach to the nuclear envelope and communicate with the adjacent **cytoplasm**. Jeopardising either this

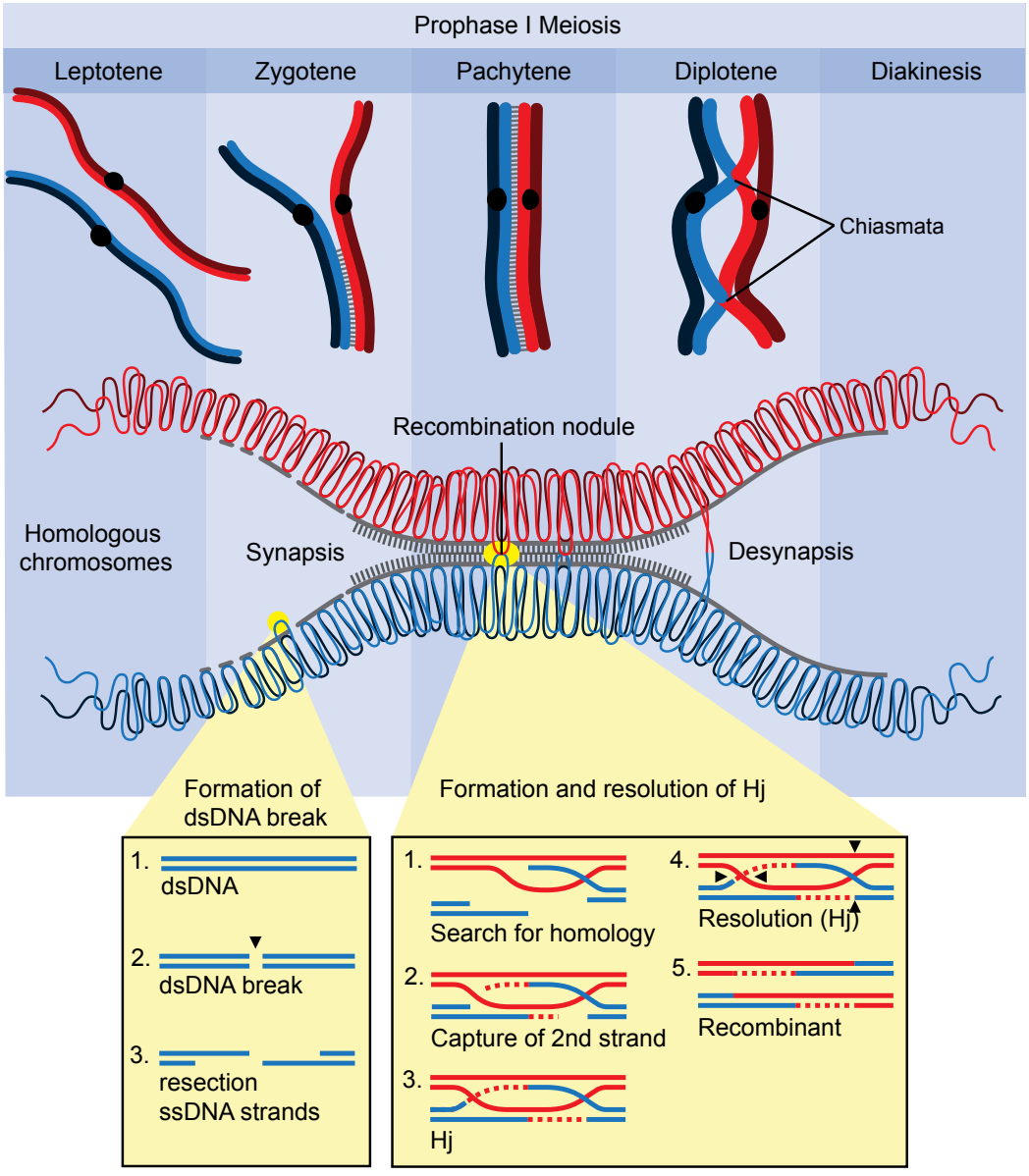

Prophase I Meiosis
Leptotene
Zygotene
Pachytene
Diplotene
Diakinesis
Chiasmata
Recombination nodule
Homologous chromosomes
Synapsis
Desynapsis
Formation of dsDNA break
1. dsDNA
2. dsDNA break
3. resection ssDNA strands
Formation and resolution of Hj
1. Search for homology
2. Capture of 2nd strand
3. Hj
4. Resolution (Hj)
5. Recombinant

FIG 14

Compressed view of the classical prophase stages of the first **meiotic** division. The two **chromatids** of the parental **chromosome**s are depicted with a subtle colour difference. The chromosome pairing (synapsis) is shaped by the **synaptonemal complex**. The **recombination** nodules can be stained with an antibody against the **mismatch repair** protein MLH1, which is a way to determine the distribution of recombination events across the different **bivalents** (see fig 15). The convention is that in humans, **gene** and protein are written in capital letters, with italics used when referring to the gene (and this also applies to the mouse; in humans all capitals, in mice only the first letter). The yellow boxes provide a simple representation of the accepted molecular **crossing over** model, where the so-called "Holiday junctions" (Hj) are the core concept. Essential is the search for **homology** of the single-strand (ss) **DNA** filament in the right panel. Arrowheads indicate where the DNA needs to be cut and where the chain needs to be repaired to rejoin the two double-strand (ds) DNA molecules. DNA synthesis is an integral part of crossing over. Due to condensation of the bivalents, crossing over becomes visible as a **chiasma** in the diplotene. The diakinesis can still be distinguished on the way to the final stage of condensation, the metaphase (fig 3, 30).

attachment (using a mouse knockout) or "the answer" of the cytoplasm near the centrosome (Ch2, fig.3) (knockouts in zebrafish) disrupts the onset of homologous pairing. **Homologues** may already come into contact with each other during the telomere movement, and a meiotic version of the **cohesin complex** (fig 6, Ch2.2.2, Ch2.3.1) plays a role in this process.

The next stage of this mysterious process involves the creation of a large number of double-strand DNA breaks, which are distributed throughout the **genome**, albeit not randomly. For a better understanding of how these breaks are used to further tighten interactions with the homologous chromosome, figure 14 is a great resource. With some assistance and at appropriately prepared sites, the enzyme SPO11, in collaboration with another protein, creates a cut in the double-strand helix. You would imagine that these breaks are then repaired using the homologous DNA base sequence of the intact sister chromatid, which is nearby, but that pathway is actually blocked.

Before the break can be introduced, the **chromatin** must be locally folded into a suitable open state. This involves the action of enzymes that methylate Histone 3 at Lysine (K), among

others at position 4 (table 1, fig 8). When Japanese researchers published an article in 2005 about a sterile mouse with a **gene** knockout for the meiosis-specific protein with this enzymatic activity, it was a difficult finding to explain in lectures. After all, a link between a DNA-cleaving enzyme and an enzyme that modifies a **histone** (table 1, fig 8) had never been determined before.

It all became a bit clearer when it was revealed that this protein not only possesses the desired H3K3 methyltransferase activity but also a domain, referred to as a "zinc finger", which can bind to stretches of DNA with limited variation in the base sequence. In other words, zinc fingers bind to a specific type of DNA motifs. This meiosis-active histone methyltransferase with the additional search function for DNA motifs is now known as PRDM9 and has become part of the theory of speciation (it is a protein that evolves rapidly, like other proteins involved in reproduction). You would think that one cannot do without an enzyme that plays such a central role in **gamete** formation. It is evolutionarily conserved and seems to be just as important for *Homo sapiens* as it is for *Mus musculus*, the rodent we are so familiar with. That is indeed true, but sometimes you still encounter mysteries.

In 2016, a genetic study of 3,222 adults born out of cousin marriages was published in *Science*. The researchers had focused on Pakistani families with a Muslim background living in the United Kingdom. The reason, of course, is that within this community, marriages within the family are common and, as a result, children born from these marriages share a part of their genome. If there happens to be a gene **mutation** on those segments which impedes the functioning of the protein, that defect can come from both parents (+/- x +/-, resulting in a frequency of a quarter according to Mendel's first law). This way, it can be determined whether the protein can be missed or not. To the researchers' astonishment, they identified a woman who was **homozygous** for a *PRDM9* knockout mutation, and yet, she was fertile.

Dogs and birds actually also lack this enzyme, so they must have a different solution for determining DNA breakpoints. Apparently, nature remains unpredictable, even for something as essential as a core process in meiosis.

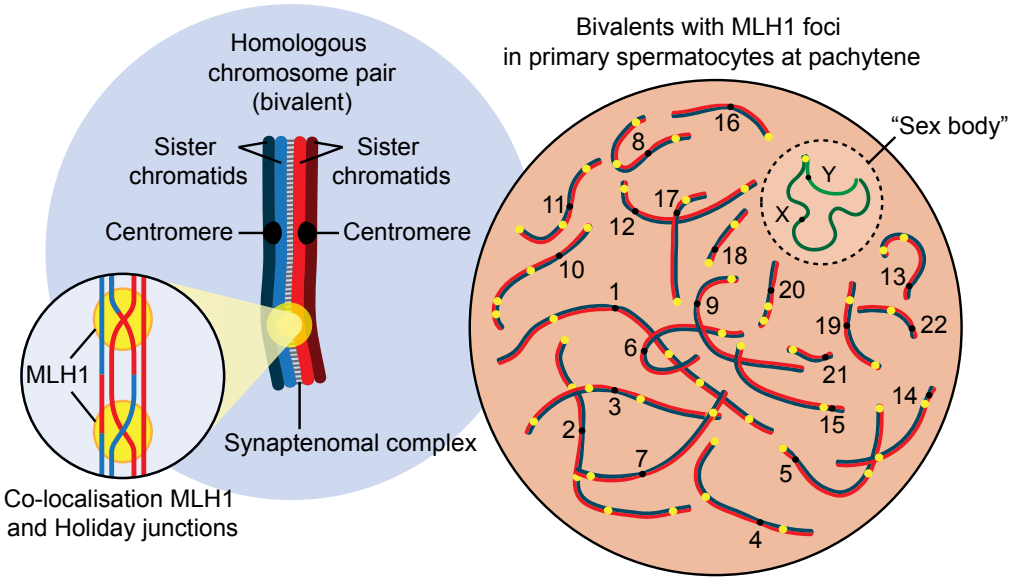


FIG 15

Impression of determining numbers and positions of **crossovers** per **bivalent** in the pachytene stage (fig 14) of the first **meiotic** division in primary spermatocytes (fig 13) in humans. The numbers of crossovers predict the number of **chiasmata** that become visible when the **bivalents** condense (fig 14). The positions of crossovers are visualised using **immunofluorescence** with indicator protein MLH1, represented here in yellow. Recognition of the bivalents is based on **FISH**. The number of MLH1 signals per meiotic cell is around 50-53 for male **meiosis**. In female meiosis, the number is higher and varies more (for most primary oocytes (fig 13), the number ranges between 56 and 86).

Double-strand DNA breaks are quickly detected by the chromatin surveillance team (Ch2.2.1, fig 4), and subsequently, the chromatin around the break is prepared for repair. An important part of the preparatory work is reflected by a member of the Histon2A family (table 1) called H2AX. H2AX is important in the germline and is evolutionarily older than the H2A found in textbooks (Ch2.3.1). We will return to H2AX in a moment, but to appreciate its significance, there is something to address first. A class of proteins that is very important for regulating various processes in the cell is collectively referred to as **protein kinase**s. These are proteins that, as

a **post-translational modification** (PTM), attach phosphate groups to a set of target proteins, which then alter their function and, for example, become active (regulation by phosphorylation). Two of these kinases are crucial for detecting DNA damage. The first is named Ataxia-Telangiectasia Mutated, abbreviated as *ATM*. This gene is responsible for a recessively inherited disorder called A-T, Ataxia-Telangiectasia (OMIM 607585, Ch4.2). Patients with a mutation in the *ATM* gene develop a whole range of defects, and a characteristic feature is their high sensitivity to **ionising radiation** (which is reflected in cell cultures of heterozygotes). The gene appears to be able to go out of control in many ways, and many types of mutations have been found in it. A particular mutation is prevalent among the Jewish population, one that must have originated long ago and can be traced back to Morocco and Tunisia. Mice that are homozygous for the "knocked out" *Atm* gene are sterile. In 1996, this report from two competing research groups appeared in the top literature. Another protein kinase gene that is indispensable in keeping DNA healthy is abbreviated as *ATR* (from "Ataxia-Telangiectasia and Rad3-related"). The first knockout mice were published in 2000: the homozygous -/- **embryo**s for *Atr* died around the time of implantation in the uterine wall, so around the age of 5 days in mice. This gene can also be found in the OMIM database, under number 601215. ATM is primarily involved in dealing with double-strand DNA breaks, while ATR is involved in single-strand DNA breaks (fig 33). As a kind of superconductor, ATM and ATR phosphorylate all kinds of proteins, including H2AX. ATM does this so diligently that a chromatin region of approximately 2 mb of DNA is modified in this way. That can be stained using a good antibody against gammaH2AX (H2AX with a phosphate group at that particular site) and observed through the **fluorescence microscope**. This way, you can already see double-strand DNA breaks developing during the further preparations for homologous chromosome pairing, but you cannot determine their number. Figure 14 gives an impression of how, starting from a double-strand break, a search can be conducted for an identical piece of DNA on the homologous chromosome, which serves as a template for repairing the break.

In this process, the protein Rad51 plays a key role, and it simultaneously marks the pieces of single-strand DNA used in "searching for homology at the base level". It is no wonder that counting the breaks involved in this process is successful with the use of an antibody against Rad51. At the beginning of the pairing, this number is around 250 in the male mouse. Also, the cutting sites themselves have surfaced. In a magnificent publication from 2016, Scott Keeney of the Memorial Sloan Kettering Cancer Center in New York explains how they achieved this. When SPO11 makes a cut and before it leaves the DNA (fig 14), it takes along a piece of DNA from the crime scene, typically fragments that are approximately 20 to nearly 40 bases long. Determin-

ing those fragments using sequencing was enough to identify the cutting site in the genome in almost 70% of cases. In cases where this was not successful, it was caused by the large number of “repeats” (Ch2.3.3, table 2) present in the genome. The now-located breaks were situated in or near a so-called “hotspot” in 60% of cases, the locations where the likelihood of a crossover (fig 14) is greatly increased. Sixty to 70% of all crossovers occur in such a hotspot of about 2 kb. And this aligns with the DNA motif of the 36 bp “zinc finger” binding site of PRDM9. Another demonstration of what is currently possible with molecular biology technology: it has been found that SPO11 usually cuts in pieces of “linker DNA” (fig 7) (i.e. between **nucleosome**s, which are likely temporarily shifted by a chromatin remodelling complex, Ch2.3.2). On theoretical grounds, this already seemed more convenient, but now, this has been confirmed to be the case.

The process of exchange of genetic information between homologous chromosomes (i.e. **crossing over**) is guided in parallel by the formation of protein axes along the chromosomes. They can use some support, and that comes in the form of a protein structure with the unpronounceable name **synaptonemal complex**. Under the electron microscope, the synaptonemal complex resembles a Swiss rack railway, where the rails represent the protein structures that provide stability to the homologous chromosomes, and the toothed rack rail in the middle represents how those two protein axes give each other a hand when they meet (fig 14). Research on meiosis really accelerated when genes encoding proteins of the synaptonemal complex were isolated. In the 90s, this was done by, for example, Christa Heyting at the Department of Genetics of what is now WUR (Wageningen University & Research). The associated antibodies were in high demand internationally and thereby contributed to the laboratory’s reputation. It is a rule in research that when tools like these antibodies are published, they are accessible to colleagues worldwide. Soon enough, many of these antibodies directed at investigating meiosis were also commercially available.

Once the homologous chromosomes are finally completely zipped together, we refer to them as **bivalents**. Naturally, it makes sense that the number of bivalents in a species is equal to the haploid chromosome number: 20 for mice, 19 for domestic pigs, 30 for cows (and zebu cattle and yaks), 23 for humans and 24 for chimpanzees, gorillas and orangutans.

The key question, of course, is what the outcome is of repairing the double-strand DNA breaks during the pachytene stage of meiosis (fig 14). It is clear that certainly not all of them are used for DNA exchange between homologous chromosomes. In this process known as crossing over, a segment with **allele** A from one homologue (from the mother) is linked to a segment with allele b from the other homologue (from the father), and vice versa, a segment with allele B

(from the mother) is linked to a segment with allele a (from the father) (fig 13, 14). Crossing over, a molecular process, ultimately leads to genetic **recombination** within the chromosome, the exchange of linear DNA information (**homologous recombination**).

During meiosis in a father mouse, only between 20 and, at most, around 30 double-strand DNA breaks are used for crossing over. In a mother mouse, this number is about 10% higher but still much lower than the estimate of 250 breaks that complete chromosome pairing. In humans, the difference between the sexes is greater (Ch11.1). It remains important that at least one crossover per bivalent is needed to halve the chromosome number during the first meiotic division (fig 13).

Let's go back to the "homology search" when crossing over comes into play. The single-strand DNA filament covered with proteins, including Rad51 (fig 14), thus searches for a complementary DNA sequence on the homologous chromosome. This initially occurs at the chromosome ends. That is probably the logical consequence of the onset of chromosome pairing, the process in which all chromosome ends end up in a subregion of the nuclear membrane. In a population with genetic variation, the homologous DNA fragments that meet do not necessarily have the exact same DNA base sequence. When there is a difference at a base position with the searching single-strand DNA segment (think of a length of 900 bases), for example, a C meets a T, this is called a mismatch. The repair process that follows is known as **mismatch repair** (Ch2.2.2, fig 33). This form of DNA repair does not slow down the progression of meiosis but is a normal part of it. How was the importance of mismatch repair for meiosis discovered?

In the first half of the 90s, the technology to knock out a gene via **embryonic stem cell**s (box 2, 3) was operable in all leading mouse labs. In the USA, labs with extensive knowledge of **gametogenesis**, with a special focus on meiosis, were relatively scarce. In other labs, molecular research on DNA repair mechanisms, including mismatch repair, flourished. Whenever a functional gene was isolated, producing a knockout mouse strain was part of the process.

At the same time, in Wageningen, we were well on our way to developing a research protocol to let nuclei of cells spread out into a thin layer, like pancake batter, on glass or ultra-thin "plastic" (for the electron microscope). Such protocols had existed for a long time, but they were somewhat unreliable, and the results were too variable. For both spermatogenesis and oogenesis, the improved technique worked perfectly, and has now become the standard. During the same period, Terry Ashley from the Department of Human Genetics at Yale University in New Haven, Connecticut, USA, paid us a visit. I knew Terry as a meiosis researcher, mainly in mice, from her years in Tennessee.

For as long as I know, it has been common for students in Wageningen to go abroad for an internship. This is how graduate student Annemieke Plug ended up at Yale, just at the time when knockout mice for genes involved in mismatch repair became available. She, of course, brought with her the improved dispersion technique, of which PhD candidate Antoine Peters was the architect. The first knockout indeed showed that mismatch repair was involved in the "homology search". That immediately led to a publication in the top journal *Cell*, with an image on the cover that was made using our technique. In that image, you can see that meiotic chromosome pairing does not progress well when this particular mismatch repair protein is absent.

Even stronger evidence that mismatch repair plays a central role in homologous chromosome pairing came from the NKI (Netherlands Cancer Institute). Because abnormalities in the mismatch repair system play a role in several forms of cancer, Hein te Riele's group worked on this topic. The protein family that detects the mismatch has two members that are exclusively found in the germline (for enthusiasts, MSH4 and MSH5). After that publication from Yale, it was a fairly obvious idea to create knockout mice in which the genes for these proteins are no longer functional. Hein told me, somewhere abroad, that they now have a homozygous (-/-) knockout for one of these two meiosis-specific mismatch repair genes (*Msh5*) in Amsterdam and that these mice, both males and females, are sterile.

This led to a collaboration that produced even better results than those reported in the 1995 *Cell* article from Yale. Completely in line with the idea that this protein helps address the mismatch that emerges when a piece of around 1 kb of single-strand DNA searches for the site with an almost identical base sequence (fig 14), the study produced results that were even clearer. All meiotic cells, both in the foetal ovaries (fig 28) and in the testes of sexually mature males, died after a few days because they failed to complete the zipping up as part of the formation of the synaptonemal complexes (fig 14). Some of the chromosomes even paired with two partners, thus also with another non-homologous chromosome (in the literature, this phenomenon is called "partner exchange"). It may be true that MSH4 and MSH5 determine mismatches at the molecular level in the search for the homologous chromosome, but these proteins also stabilise the fragile structure of the dangling single-strand DNA of figure 14; they are still objects of research when it comes to refining our understanding of crossing over.

The double-strand DNA breaks that are not further processed in the absence of (homologous) chromosome pairing activate a checkpoint (Ch12.5.3, fig 4), and then the primary oocytes and spermatocytes (fig 13) are eliminated.

It has long been known that the choice of a DNA break to repair itself using the DNA of the homologous chromosome (via a crossover, fig 14) cuts off that route for other nearby breaks.

Sometimes, those other breaks do explore the homologous chromosome, even incorporating a piece of the homologous base sequence into their own DNA, but a true crossover does not occur. This then results in three of the four chromatids at that site having the same **genotype**, and you will find a 3:1 ratio in the **gametes**. This process is known as **gene conversion**, which is common and also serves the purpose of meiotic pairing. Moreover, it leads to more genetic variation. To further complicate matters, gene conversion also occurs at the site of crossing over, and the entire process turns out to be surprisingly **mutagenic**; in humans, 4.4% is associated with a form of *de novo* mutation, and this is particularly noticeable in the gametes of ageing women (Ch11.3). Using the fact that the DNA sequence of the hotspots for meiotic double-strand DNA breaks (opening the possibility for crossing over) are now known in humans and that relatively many of us have their DNA sequenced by **WGS**, it is now possible to search for the mutagenic aspect of the repair of these breaks by matching these potential breakage sites with their genetic variability in the population. The results of this approach, feeding those about 28k DNA sequences in a database of gnomAD (the hotspots), confirmed the statistic described above (work published near the end of 2023). Apart from base pair changes, structural variants were not rare at all. The most interesting aspect was the mutational footprint hinting at the activity of relatively sloppy repair pathways, lacking accuracy. Furthermore, genes were implicated to have lost their function this way.

5.2 The meiotic divisions themselves

Figure 13 provides a schematic overview of this. Returning to the structure of the bivalent: first of all, the zip of the synaptonemal complex opens (fig 14), resulting in the separation of the homologous chromosomes, except at the sites where the crossovers become visible as **chiasmata** (a historical name, singular chiasma, fig 14). You can count them under a microscope, for a long time a specialised and, in humans, laborious task.

One of the mismatch repair proteins has the name MLH1. In 1996, a publication appeared in *Nature Genetics* in which it was shown that a fluorescent antibody staining on dispersed pachytene spermatocytes (fig 14, 15) produces a dot pattern on the synaptonemal complexes, thereby revealing the crossovers which correspond in number to the chiasmata. These MLH1 dots are easy to count, and thus, a new field of research opened up: determining the numbers of crossovers in men and women and their distribution within and between bivalents. The second author of this publication was Annemieke Plug, and the second-to-last was Terry Ashley. Crossovers

occur spread across the chromosomes, but there are clearly more of them towards the ends. They also tend to be located away from each other (fig 15). When studying the distribution of crossovers across the bivalents, the patterns are subtly different between men and women. As mentioned earlier, women have more of them, and they can lie somewhat closer together. They are also less concentrated at the ends of the chromosomes.

To enable the first meiotic division, the sister chromatids must relinquish their close bond (in the form of the cohesin complex, fig 6) so that the homologues (which will have undergone at least one crossover) can be separated. It is important to wait with this until the bivalents have oriented themselves in the centre of the **spindle apparatus** (fig 3, 13). During metaphase, the chiasmata lie on the equator, held in balance by the pulling force of spindle fibres attached to the **centromere** regions (fig 3). However, the sister chromatids remain connected at the centromeres. There, the dissolution of the cohesin complex must be delayed until the second meiotic division. Crossing over is thus essential for halving the chromosome number. As mentioned just before, it is essential that at least one crossover is present per bivalent. The great importance of the cohesin complex in the meiotic divisions of the oocyte will be discussed in chapter 11.

5.3 The genetic consequences of meiotic divisions

In the 60s, the basic genetics course at the former Agricultural College of Wageningen, the Netherlands, was marked by a high level of text retention. It was of great benefit if you could literally reproduce the key sentences during the oral exam. These key sentences served as the framework on which the lectures were built. The standard formulations certainly applied to Mendel's laws. The first law postulates that an allele pair for gene A with genotype Aa is separated during meiotic divisions: 50% of the gametes contain one allele A and 50% contain the other allele a (small a, as emphasised in the lecture). Intuitively, you might think that this occurs during the first meiotic division when the bivalents are separated, with the result that the chromosome number is halved. But when you look at figure 13, you will find that that is only half the truth, or even less. It applies to genes that are located between the centromere, where the fibres of the spindle apparatus act, and the first crossover. For genes that are located beyond the point of crossing over, the sister chromatids are no longer true sisters because recombination has occurred within the chromosome. This chromosome, now entering the second meiotic division, is **heterozygous** from the crossover point for all genes for which the individual is also heterozygous (fig 13). The reciprocal situation for this chromosome can be found in the other daughter cell (fig 13).

Altogether, you can see four genetically distinct chromatids in both daughter cells of the first meiotic division for this example chromosome. These chromatids will be the chromosomes of the future gametes. In order to get the chromatids into the gametes, the second meiotic division must be completed. Only then do the alleles of the genes that lie beyond the crossover site, counted from the centromere, segregate. The phase in which the physical segregation takes place does not affect the final result. The famous one-to-one segregation of Mendel's first law remains valid. Moreover, it forms the basis for the proposition that, as an individual, you inherit half of your genetic material from one parent.

Next, we come to Mendel's second law, known by the beautiful expression "independent assortment". This second law relies on the fact that the bivalents, as they lie on the metaphase plate (fig 3) of the first meiotic division, are not actually aware of each other. What I mean by this is that when, for example, the centromere of bivalent 1 inherited from the mother moves to the left, the centromeres of the neighbouring bivalents are oblivious to this and just proceed to do their own thing. The orientation of the bivalents is completely random when it comes to the origin (maternal or paternal) of the chromosomes. This is the foundation of Mendel's second law. This law underpins the unprecedented number of **gamete** genotypes produced by meiosis. An often unrealised consequence of this law is that, as a grandparent, you cannot assume that 25% of your genetic material will be found in your grandchild. Together with your partner, the percentage comes to 50%, but for each grandparent, most variation lies between 18% and 32%.

We also learned that in these kinds of crosses, where one parent was heterozygous for two genes (AaBb x aabb), the fact that the gamete types AB, Ab, aB and ab were found in equal frequencies in the offspring did not automatically mean that these genes were located on different chromosomes. When a crossover always occurs between two genes on the same chromosome, the famous AB, Ab, aB and ab combinations will appear in the gametes in equal frequencies, exactly the outcome of "independent assortment".

However, when two genes lie closer together and a crossover between them certainly does not occur in every meiotic prophase cell, the two allele combinations created by recombination each drop below the 25% proportion. In 1911, a student of the renowned fruit fly geneticist Thomas Hunt Morgan discovered how to use the information from such crosses to create a **gene map** for each chromosome, in which the measure of distance was the probability of crossing over (the greater the distance, the higher the probability). If, for instance, out of 1000 offspring, 10 come from gametes carrying a recombined chromosome for the alleles of these genes, the recombination rate is 1%: this is called 1 centiMorgan "genetic distance". The student who came

up with all of this was called Alfred Sturtevant, and he fortunately became a famous *Drosophila* geneticist later on.

Before the era of "whole genome" DNA projects, the gene map based on recombination by crossing over was the major reference point. You knew in which order and on which chromosome genes were located (many for morphological traits, but also protein variants that you could observe in their "allele form" in the electrophoresis gel with simple molecular biology techniques). The third edition of the mouse bible *Genetic variants and strains of the laboratory mouse*, dating back to 1996, contains a 258-page table that shows per gene what the recombination rates are with the other genes on that chromosome. In total, this table provides around 12,250 estimated distances. Considering that one first had to create the mouse strains that were homozygous **recessive** and **dominant** for the genes of interest, it is easy to realise the astronomical numbers of mice that were used to do this. Most of this work was done in the second half of the previous century, and in the first part of this period, social accountability for animal experimentation did not play a significant role. The gene sequence on the DNA map has not changed, but the relative distances between genes have (distance is now an absolute value). As mentioned earlier, crossovers are not randomly distributed throughout the DNA. In 60-70% of cases, crossovers occur at so-called "hotspots".

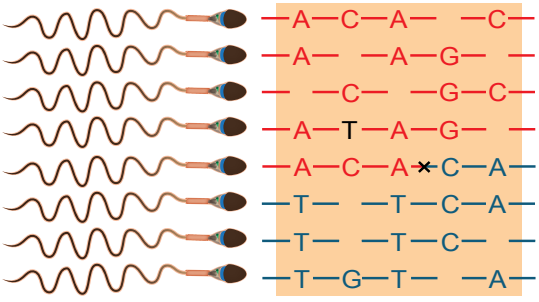
Outside those "hotspots" (i.e. "hotspots" that are less "hot"), crossovers are not equally common in all regions of a chromosome. However, viewed over many generations, they occur pretty much everywhere. Crossing over greatly increases the number of combinations of alleles of the roughly 20,000 genes (adding to the effect of "random assortment"). Recombination between and within chromosomes is always seen as a beneficial aspect of sexual reproduction, because it generates variation and therefore increases the possibilities of selection, that is, the ability to adapt to new conditions.

Now, if right next to an allele for gene A, which provides only advantages to the carrier, a new mutation emerges in a nearby gene B, which negatively affects the individual, then there will clearly be less benefit from that favourable allele A. After all, most mutations are not advantageous. Through recombination (fig 13), you can get rid of this situation again. This results in the creation of "favourable" segments of chromosomes that natural selection can work with. It can mean that individuals with such favourable contiguous pieces of genetic material (for such a contiguous piece, the word **haplotype** is used, fig 16) have more offspring and thus a higher "**fitness**". Fitness is the driving force behind evolution. Sexual reproduction also has advantages when the **mutation frequency** is a bit higher, and the environment is a bit more unstable. In that

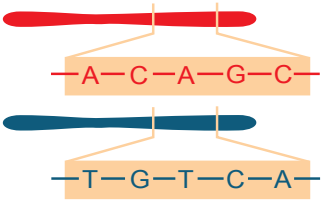
1. “Unphased” SNPs

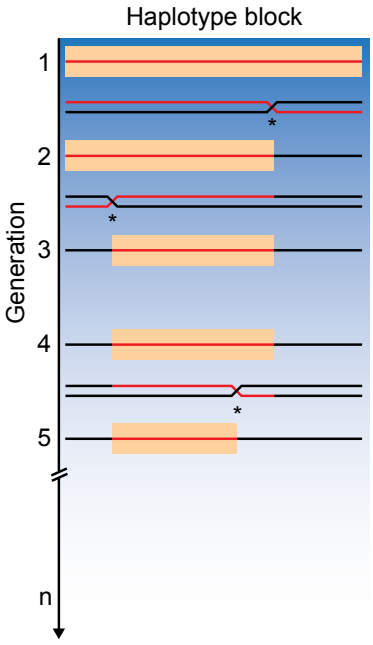
Chromosome pair

2. DNA analysis of individual sperm



3. Deduction of “phased” haplotype per chromosome





Progressive shortening of haplotypes through recombination*

FIG 16

When performing **DNA** analysis using, for example, **SNP**s, you can determine **heterozygosity** or **homozygosity** for each position, but you cannot see which **allele**s are located together on the same **chromosome**, forming a **haplotype**. In a DNA analysis of individual sperm cells, you can in fact observe this. Comparing the patterns that the SNPs make together across multiple sperm cells allows you to also see the consequences of **crossing over**. With the method described here, it is not possible to determine which (non-**recombined**) haplotype originates from the sperm donor's father and which from the mother.

DNA screening of the donor's parents is needed to determine that. The process of assigning chromosome segments to the maternal and paternal copies is known as "phasing". In the example presented here, knowledge of the SNP patterns of the father and the mother was already available (compare with fig 13). The right side of the figure is an illustration of the fact that each time a haplotype passes through **meiosis** during the succession of generations, it once again has the chance to become shorter due to crossing over. Statisticians with knowledge of SNP **allele frequencies** and an understanding of population genetics have studied how short a haplotype can be for it to still function as the marker of that particular ancestor. This means that the SNP code across that chromosome segment must be unique. Since this is not easy for a small haplotype in the case of distant relatedness, this kind of analysis is performed across the entire **genome**, allowing for the elimination of small uncertainties per haplotype. The decreasing cost of **NGS** has reinvigorated this field, and the improvement of statistical models has the same effect. NB the black 'T' of sperm cell four is a technical error.

situation, adaptation is the way to go. Developing resistance to parasites is one example. Sex is thus well-suited to the Darwinian school of thought (Charles Darwin, *On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life*, 1859). If you can combine the good alleles on a chromosome segment through recombination, you can also combine the less favourable alleles in the reciprocal crossover product (fig 13), which might lead to prenatal mortality or infertility and, thereby, a total loss of "fitness". Jaap van der Veen, professor of Genetics in Wageningen from the late 60s to 1988, referred to reproductive failure as "genetic killing": the terminology in population genetics can be a bit harsh.

5.4 The applications of meiotic outcomes

5.4.1 Determining kinship

The reason for going into such detail here is that society is currently inundated with massive amounts of genetic data. Many people send their DNA to companies like 23andMe, Ancestry and MyHeritage, and a readout in the order of 700,000 **SNPs** (Ch3.2) is done at these companies for a small amount of money. These SNPs meet the condition that two alleles (read: bases, e.g. C/T) occur at this position in the genome and that the least common one is still found, on average, in one out of 50 people, meaning the **allele frequency** is at least 0.01. The analytical possibilities at the population level are impressive. Let's start with a case of more distant kinship. When using a 700k chip, there are about 30,000 SNPs on the average chromosome. Due to genetic variation, you will get a unique pattern for that chromosome for each individual over a series of adjacent SNPs, but you cannot predict in advance which allele lies on which of the two homologous chromosomes. In meiosis, chromosomes and chromosome segments segregate. Now imagine following a piece of chromosome through a few generations. Figure 16 illustrates the idea behind this: per meiosis and thus per generation, a crossover-free segment is likely to become shorter, as new crossovers occur.

Then, the haplotype also becomes smaller. As long as it is large enough and contains an adequate number of SNPs, even much shorter chromosome segments will still exhibit a unique pattern of SNP combinations. Think of this as a barcode representing the haplotype. In principle, the entire genome can also be identified per chromosome segment by the allele pattern of the SNPs. This "tracking" is somewhat like advanced mathematics because, as noted above, you initially do not know which alleles are located together on the same chromosome in the common ancestor. However, due to the segregation in meiosis, you can identify this in a family tree using some calculations. What lies on the same chromosome segment has stayed together. In each meiosis (so in each generation), the homologous segments separate again. In this way, the number of segments with which similarity is still found decreases as you delve deeper into the family tree to find a relative through (a) common ancestor(s). The number of passages through meiosis is determined by tracing back to the common ancestor in the family tree and then going forward to the distant relative. Using the recognition of allele patterns across the entire genome, it is possible to identify a third cousin with a reasonable degree of certainty (60%). If you follow a straight line far up the family tree, then "genetic identity" is preserved a bit longer. The book

Who we are and how we got there by Harvard researcher David Reich, published in 2018, contains a wonderful figure illustrating this principle. According to his assumptions, after 7 generations, the probability that the genetic influence of one of 128 ancestors is lost increases. Of the 1,024 ancestors from 10 generations back in time, only about half still contribute to your DNA.

Determining your **genetic background** in terms of regions, countries and continents is based on the probability of an allele per SNP per region or continent. This provides information about ancestry based on how people have spread across the world. Information about all SNPs is added to this. As more people around the world submit their DNA, these estimates become more accurate, and you receive an update. This, too, is data processing, another combination of biological and statistical insights.

Appendix I further discusses the value of an SNP profile for estimating genetic predisposition for a trait based on multiple genes, which includes the most common diseases.

5.4.2 Identifying a delinquent

An evaluation of the possibilities offered by a database like the one from 23andMe appeared in *Science* in late 2018. This publication connects 1,28 million human genotypes with each other. The DNA is primarily from the descendants of European emigrants to the USA. The researchers predict that when 2% of the population has donated DNA and has made it accessible for research, almost everyone can find a distant relative. This has other implications when DNA is left at a crime scene (assuming that this DNA is related to the crime). The then traceable distant relative, in combination with genealogical data, gender, age (surveillance camera images) and location of the crime (the article reports that most crimes are committed within a 160 km/100-mile radius of the place of residence), can greatly narrow down the number of suspects. The societal implications are obviously large, and all of this was prominently reported in the news in 2018 when, based on population genetic research in the USA, enabled by a public database, the serial killer “Golden State Killer” was apprehended after 40 years. Another example is the investigation into the perpetrator of the murder of Marianne Vaatstra, who was killed in the Netherlands in 1999 at the age of 16. A population genetic DNA study among approximately 6,000 men from the vicinity of the crime in North East Friesland, which focused on the Y chromosome, led to the resolution of this crime in 2012. Based on the slowly mutating **SNP**s (table 3), which represent the entire ancient history of the carrier, it could be ruled out that the perpetrator came from a nearby asylum seekers centre.

Using the much faster mutating microsatellites (table 3, from 10^{-4} to 10^{-2} per gamete per generation) on the Y chromosome, the analysis could be shifted to the level of kinship, and the perpetrator was identified. Socially, we are primarily individuals, but genetically, we are simply part of the population.

5.4.3 The remains of *Homo neanderthalensis*

Where do these SNPs actually come from, and how old are they? This has already been addressed in Ch3.2 and table 3. At a reproductive age of 30 years, around 60 new SNPs are introduced per generation (table 3). Fortunately, the base composition of the genome does not change very quickly. Since the SNP platforms only contain SNPs with a minimum allele frequency of 1%, you can conclude that these genetic variants are old. How old exactly? That goes beyond the scope of this section, but taking a look at our genetic history may give you an idea. Thanks to the significant improvements in techniques used to isolate and purify DNA from palaeontological remains, a development primarily introduced by Svante Pääbo in Leipzig, the genome of *Homo neanderthalensis* could be published in 2010. The last common ancestors of *Homo sapiens* and *Homo neanderthalensis* are estimated to have lived around 770,000-550,000 years ago (according to the aforementioned popular science book by David Reich, you can position the split between *Homo neanderthalensis* and *Homo sapiens* at around 700,000 years ago). The genetic difference with *Homo neanderthalensis* is large enough to detect very small segments of *Homo neanderthalensis* in our genomes based on these SNPs (it is known which base, or allele, comes from *Homo neanderthalensis* and which one from *Homo sapiens*). There is not a large amount of Neanderthal DNA present in the genome of each individual, but everyone carries some. In “Caucasians”, the term typically used in the literature to refer to the combination of Europeans and people from the Near East, around 2% is still found per individual. Overall, about 30% of the Neanderthal genome is present; we do not all have the same genetic Neanderthal contribution.

Since the period of interbreeding between *Homo sapiens* and *Homo neanderthalensis*, roughly from 65,000 years ago until their disappearance around 40,000 years ago, the influence of *Homo neanderthalensis* in the genome has diminished through negative selection. In genes related to keratin proteins (in nails, hair and skin), there is still more than a coincidental “Neanderthal” influence which indicates positive selection, and there will be more examples like this. What this period of interbreeding has done to us remains intriguing. Based on the differences in

the skull − the sloping forehead of *neanderthalensis*, and the more rounded back of the head of *sapiens* − you can, for instance, explore whether you can correlate the variation in skull shapes within *Homo sapiens* with a possible residual influence of *neanderthalensis* genes still present in the population. These genes can thus be recognised based on Neanderthal alleles for SNPs. Would this allow us to identify genes related to skull shape? In January 2019, a paper was published in the journal *Current Biology*, in which they did everything possible to explore this. Of the 27 authors, four had a connection with Radboud University, Nijmegen, the Netherlands.

In 4,468 individuals, the researchers found two small pieces of DNA, characterised by five, six SNPs, a haplotype with a very high likelihood of Neanderthal origin. One fragment is about 200 kb long and lies on chromosome 1; the other is about 500 kb and is located on chromosome 18. All of this is a result of crossing over and segregation over the many generations that have passed since the period of sexual interactions in prehistoric times. For those who are interested, with allele frequencies of 0.044 for chromosome 1 and 0.055 for chromosome 18, respectively, it is not surprising that Neanderthal homozygotes for these fragments are very rare. The genes that have been identified in this way are associated with neurogenesis and the formation of the myelin sheath around nerve fibres (comparable to the plastic insulation of an electrical cord). These findings suggest that these genes contribute to skull shape during development.

The discussion regarding the influence of interbreeding with *Homo neanderthalensis* on fertility is addressed in chapter 12.5.1.

CHAPTER 6

The transition to quantitative genetics

The simple insight provided by Mendel's laws combined with **chromosome** behaviour during **meiotic** divisions eventually leads to a mind-blowingly complex picture. With a single **crossover**, without which the transition from **diploid** to **haploid** would not occur properly, you will obtain four products for a particular chromosome, each of which will have a different "**allele** bouquet" for the **gene**s located on it (fig 13). However, the same will happen for the other chromosome pairs. In this case, the number of combinations of allele bouquets across all chromosomes is four to the power of the number of **bivalents**, i.e. for human **autosomes** 4^{22}. Longer chromosomes undergo more **crossovers**. Crossovers, in general, occur at variable positions (Ch5), and the number of possible allele bouquets (read: the **genotype** of the **gametes**) that an individual with a reasonable degree of **heterozygosity** across their genes can produce is, in fact, many times larger. Therefore, the complexity lies in the almost infinite number of possible combinations of alleles across the **gene** set.

Thanks to **SNP** technology and **whole genome sequencing (WGS)**, genetic variation is now directly analysable at the population level. This makes it possible to make predictions about an individual's genetic predisposition for each trait. While the possibilities of analysis at population level are already impressive, advancements in technology and knowledge enable the move towards the individual level (but see later in this chapter and the appendix). This also applies to the commercial aspects of individual genetic analysis.

6.1 The concept of heritability, the h^2

In the Department of Animal Breeding and Genomics at Wageningen University & Research, there was a small yet significant kind of excitement in the autumn of 2018: the new edition of the quantitative genetics textbook with the broad title *Evolution and selection of quantitative traits*, a volume with a spine width of 6.5 cm (2.5 inches), consisting of 1,459 numbered pages, was about to be released. This book is the life's work of Bruce Walsh and Michael Lynch. The latter will reappear in chapter 18. While going through the book, you encounter formulas based on statistical analysis all over the place, and the text is also not always very accessible. Quantitative genetics turns out to be a super specialised field.

The traits of which the **heritability** is studied in quantitative genetics are influenced by many coding genes with multiple alleles for each **locus**, in addition to an even larger number of locations in the **genome** that can be non-coding. They are referred to as **polygenic** traits. While obvious now, these were initially assumptions. Knowledge of the genes underlying a trait and the extent of their involvement in determining the **phenotype** is not necessarily required, but this knowledge develops from the genome projects. The great achievement of people like Robert Fisher and Sewall Wright is that they combined Mendel's insight into the inheritance of one or a few genes with multiple alleles with statistical methods such as correlation calculations, analysis of variance and regression analysis. I have heard from colleagues that the original work of Ronald Fisher (UK, 1880-1962) and Sewall Wright (USA, 1889-1988) is hardly readable. However, it has been picked up and translated for teaching and research purposes, for instance, by the American Jay L. Lush (1896-1982), known as the father of modern livestock improvement.

Heritability (h^2) is a central concept in quantitative genetics, but what exactly is it? The theory behind it is based on the fact that parents transmit half of their genetic material to their children and that in this process, for **loci**/genes with two alleles, one allele ends up in one **gamete** and the other allele in another, as per Mendel's first law.

The central equation in quantitative genetics is P = G + E. Here, P stands for phenotype, G for genotype and E for environment (fig 17). In quantitative genetics, you work with average values. At the population level, P, G and E in this formula represent the entire distribution (P) and its components (G, E). Simply put, the formula breaks down the distribution into a contribution of genes and a contribution of the environment (fig 17). The contribution of genes is represented in this statistical model as a sum of the transmitted alleles for many loci. Take growth in height as an example. Some alleles for certain genes involved in this slow down growth, while other

alleles, either for the same or for different genes, accelerate growth. With more accelerators and fewer decelerators, you will grow taller than average (fig 17) and, conversely, with more decelerators than accelerators, you will become shorter than average. The same reasoning applies to the environment. By definition, the h^2 is that part of the phenotypic variance from the mean that is caused by the sum of the genetic variants. It is, therefore, expressed as a fraction and refers to the population (see fig 17 for the formula). In order to estimate the h^2, you need to use individuals that are related in a known way. For example, you might use parent-offspring pairs that are compared in a range of environmental conditions: this is the domain of statistical model builders. Of course, Mendel's first law, which states that an individual gets half of the genetic material from each parent, is incorporated into this. Identical twins who grew up in different environments can also be used. The more individuals you can include in the family relationships, the more reliable the estimate of the h^2 will be. Family relationships encompass everything that a family tree can offer; the models are able to deal with that. The formula also implies that the estimate is only valid for the population in which it is performed and for the environmental conditions under which the trait was determined. Proper "phenotyping", the determination of traits, is therefore very important. In terms of this field, an h^2 below 0.2 is considered low, between 0.2 and 0.4 is average and above 0.4 is high. Traits such as the amount of black in a black-and-white cow have a high heritability. This also applies to the likelihood of a cleft lip in humans and to morphological traits related to the skeleton, such as height (human) or stature (cow). A trait like BMI (body mass index) has an average h^2, fertility parameters have a low one. Milk production in kilograms used to have an h^2 of about 0.25 and is now calculated to be a bit higher. Experts have not been able to conclusively indicate the reason for this. Fewer measurement errors, a different statistical model, or are we currently keeping dairy cattle under more uniform conditions?

Another method to estimate an h^2 makes use of a selection experiment. For this, you need to know to what extent the, say, top 15% "best" (or "worst") animals selected for reproduction deviate from the average, and how this manifests phenotypically in the next generation. To what extent do the offspring deviate on average for the trait that is selected for? The fraction of these two average deviations is known as the realised h^2. Now you can see even more clearly what happens in real life. Walsh and Lynch's standard work provides a number of examples in which the two estimates are compared. They are of the same order of magnitude, but they do not match perfectly, and each estimate also has a confidence interval. The h^2 can be described as the most powerful tool of the livestock improver. The higher the h^2, the easier it is to select for a trait. In practice, this happens for many traits at the same time. This includes traits with a low h^2, such as

fertility. Selection can still be successful in such cases, but with the condition that the phenotypic distribution is large, which is the case for fertility.

6.2 The polygenic risk score (PRS) and genome-wide association studies (GWAS)

When a saliva sample is sent to, for instance, 23andMe, a commercial organisation for **DNA** research, it is theoretically possible to receive information not only about one's ancestry, but also about genetic predisposition to diseases. Diseases based on a single gene (monogenic, Ch4.2) are not really meant by this. In the context of this chapter, this refers to polygenic diseases. As the genetic basis becomes more complicated, predictions naturally become more difficult. Many genes and their variants (alleles) and other non-coding loci will be involved in such cases, and their identification is an ongoing but comparatively slow process.

Furthermore, these diseases are influenced by various environmental factors (e.g. lifestyle). This also makes prediction a lot more challenging (see appendix 1). The measure used for this is known by the abbreviation **PRS, "polygenic risk score"**. It entails predicting the risks for the individual based on the SNP polymorphisms. The estimate is obtained from an unselected population of many individuals who have self-documented their medical history and who have had their SNP DNA profile determined. What all this means is further explained in the appendix. This appendix also explains why these estimates quickly become problematic, although progress is made in this context as well. So, it is not fundamentally impossible to make estimates, especially when the genetic basis for a trait becomes more evident, as is, for instance, the case for both types of diabetes.

In the livestock world, the counterpart to an SNP-based estimate of genetic predisposition for a trait of choice is known as the genomic breeding value. This is very practically measured as, for example, the additional litres of milk produced in the offspring of a particular bull compared to a reference population. Cattle breeders and farmers love this, as it allows them to rank prospective artificial insemination (AI) bulls for each trait. The PRS and genomic breeding values are two different forms of presentation based on the same basic information. In principle, this information can also be used to calculate an h^2. That h^2 has the suffix m, for "markers" (h^2_M). The Dutchman Peter Visscher, who now works in Australia, has played a major role in developing the mathematical techniques for all this, especially in transferring these techniques to human populations.

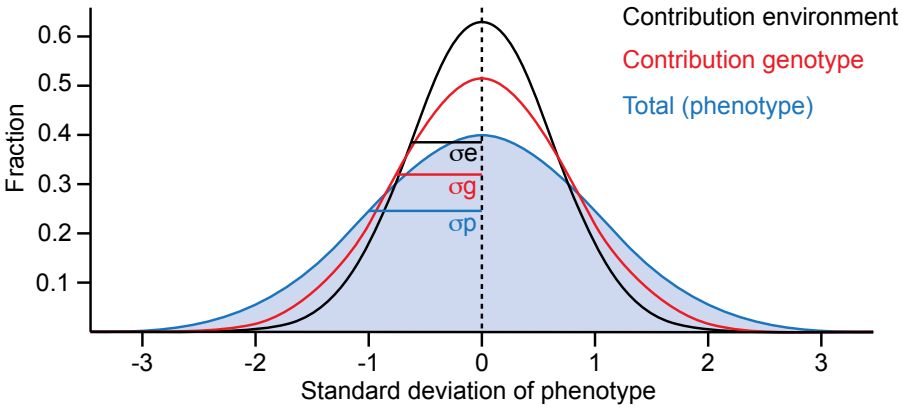


FIG 17

Gaussian curves depicting the distribution of a trait and its division into components. The distribution of the **phenotype** (blue, P for "phenotype") is divided into the contribution of the **genome** (the **genotype**, red, G for "genotype") and the contribution of the environment (black, E for "environment"). The units on the X axis indicate the standard deviation of phenotype. The formula for **heritability**, the h^2 is σ^2g/σ^2p, where σ denotes the standard deviation, and its square indicates the variance. Both are a measure of distribution. The situation shown here corresponds to a relatively high h^2 of 0.6. To represent the phenotypic distribution per genotype in this model, you would need to have access to a cloned population, which, as a thought experiment, you would subject to a range of environmental conditions; this would leave you with the σe Gaussian curve. Here, you can see this for an imaginary clone with the average genotype and thus phenotype.

One particular difference between the human population and cattle populations is their entirely different structure. The number of individuals from which the entire population of approximately 25 million Holstein Friesians worldwide descends is actually quite small. Some bulls have produced hundreds of thousands of offspring through frozen sperm and AI. To make calculations based on this, we use the concept of **effective population size** (**Ne**), developed in the 1930s.

For Europeans, a figure of 10,000 has been mentioned; for Holstein Friesians that is 50-100. The number Ne is directly related to the likelihood that at conception, so during the formation of the **zygote**, genome segments that have the same ancestry will meet in the new individual. The smaller this number, the greater the chance, and the higher the level of inbreeding.

There is another way to recognise that we are more related to each other than one might suspect with some knowledge of the family tree. For every generation that we go back, the number of ancestors in the family tree doubles. If you go back 20 generations, that number is more than a million, and for 25 generations, it is 33+ million. Let's assume that the parents of all children were, on average, 25 years old at the time of their birth (the **generation interval**), then going back 25 generations will take you back just over 600 years. It is then around 1400 AD, in the period when the European population may have halved from a previous peak of 75 million. Around 650 AD, the European population is estimated at 18 million. There is no doubt that many of us have common ancestors, based on the family tree, far back in time.

The difference in population structure matters when it comes to making phenotype predictions based on the genome. Returning to figure 16, throughout human history, crossovers near a gene that contributes to the studied trait have occurred much more frequently than in cattle. As a result, many more SNPs are needed per genome to ensure that there is one that is both unquestionably linked to a gene that may contribute to the phenotype and also **heterozygous**. Determining the contribution (as a variation from the mean) of basically every chromosomal segment to that phenotype is known as the **genome-wide association study (GWAS)**. The larger the study population, such as the population available in the UK biobank, the more accurate this estimate will be. In humans, you will easily need half a million to a million SNPs, whereas in cows, 10% of that amount is sufficient. In the case of a polygenic trait, the number of sites in the genome that can be associated with that trait is astonishingly large. To give you an idea of the numbers, think of around 50,000 in humans and 5,000 positions in cows. That is 5-10% of the number of SNPs included on a chip. If you set a threshold for significance, you can estimate from the UK biobank material that there are 850 locations on the genome that account for variation in height and 160 that relate to variation in BMI. Developments in this area are progressing at a remarkable pace. Height is regarded as the workhorse of GWAS analysis. This is covered in a study that now (as of 2022) encompasses over 5.4 million individuals, which results in an SNP-based heritability (the h^2_M) of 0.4 that is explained by 12,000 SNPs. This accounts for approximately half of the h^2 based on family relations (around 0.8), with the difference being known as the "missing heritability". This is the difference between "inheritance in practice", determining

the heritability (h^2) in the classic way using relatives and statistical methods, and determining heritability through the statistics behind an SNP analysis at the population level. Peter Visscher, the main author in this area, emphasises in a review published in 2023 (with collaborators) that a comprehensive understanding of the significance of GWAS is hampered by the fact that there is still a lack of data on a population-wide scale (so sampling from populations of non-European ancestry): “findings from GWAS conducted in one group are not always transferrable to another group”.

Now that determining the base sequence of the whole genome has become affordable, you may, of course, wonder whether it is better to use whole genome sequencing (WGS, Ch3). The idea behind it is that this way, sooner or later, you will be able to identify the (parts of) genes, and in fact, all non-coding sequence, that really contribute to the variation in a trait and hopefully also the DNA changes that make a locus a locus and an allele an allele. The reason behind this formulation is that, as determined by GWAS, the total risk burden (for common diseases) originating from coding sequence is smaller than from non-coding sequence (box 1). The literature refers to this as determining the “**genetic architecture**” of a trait. This means that the part of the genome that is involved in the emergence of the trait during development, the maintenance of the trait, and the degree of variation per locus are known. This idea has been tested on 25,465 unrelated individuals of European ancestry from the TOPMed program (Ch3). Now, rare **mutation**s in protein-coding genes with a relatively large effect could be included in the heritability, leading to an estimated h^2_M of 0.6-0.7 for height. This means that a portion of the heritability as determined by family relations is still missing, according to the authors; this is the “missing heritability” mentioned before.

The latest development here is that WGS data from an even larger population (i.e. 76,156) allow for a search for the significance of non-coding sequence. This is done by estimating the selection pressure (observed as a “depletion of variation”) on these sequences as a measure of their importance. Through this metric, attention is drawn to sequences involved in the regulation of coding genes. For instance, mi**RNA** loci (Ch12.5.2, Ch15.2.3) now come into view.

Does the above, a tendency to expand genetic analysis across the entire genome with the involvement of more and more individuals, help in predicting the individual risk of disease? “He who fears he shall suffer, already suffers what he fears” is a well-known quote from Michel de Montaigne. Let us return for a moment to polygenic risk scores (PRS) based on an SNP profile. For a trait with a relatively high h^2 (0.6-0.8) like schizophrenia, an h^2_M of 0.24 was obtained based on 53,386 patients and 77,248 controls of European descent. This study from 2022 led

to the detection of 106 protein-coding genes as candidates to be involved in schizophrenia. So, in order to detect informative DNA mutations, a very large number of people should be included in the DNA analysis, since the majority of these mutations are rare. There is another important reason why these polygenic risk scores, which are individual scores, do not mean much yet, particularly for non-extreme values. Every estimate of a contribution to the total variance also needs an indication of the degree of accuracy of that estimate. For an illustration of this, see the appendix. For most traits, the vast majority with an average h^2, a reliable estimation of the phenotype based on a commercial SNP analysis is still a rather distant prospect. However, the situation may change when the number of genes that very significantly contribute to the trait is not too large, in other words, when the genetic architecture of the trait is better known (see the appendix).

As previously stated, quantitative genetics is a field for specialists. An aspect of this field that helps assess the value of genetic estimates of an expected phenotype can also be illustrated in another way: by looking at the large genetic variation that is also present within families. Statistically, this is half of the total genetic variation. Large differences in the offspring are, therefore entirely, normal. This is also why a disease with a strong genetic component, such as cardiovascular disease, can suddenly emerge without a family history over three generations. And this is due to the segregating genetic variation during **meiosis** on the way to the haploid level, as well as the emergence of variation at the diploid level at fertilisation, occurring over all those generations. The large “within family” genetic variation is also still true for the Holstein Friesian dairy cattle population. Even at a low effective population size, this remains high. Otherwise, the response to selection shown in figure 18 would not have been possible. On to the **germline**.

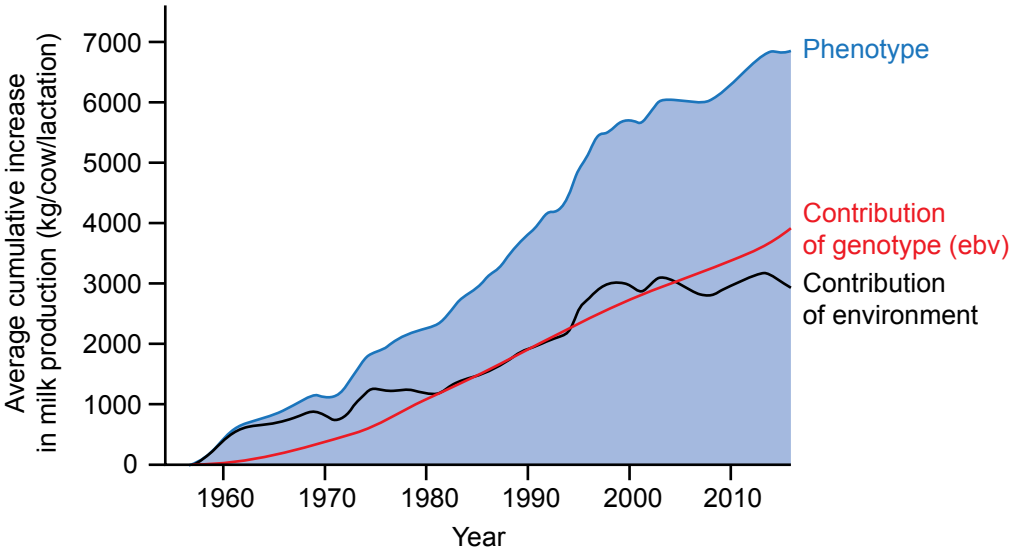


FIG 18

The selection response in milk production of Holstein Friesians in the USA between 1957 and 2015. In animal breeding, this is expressed as the “estimated breeding value” (ebv). It accounts for about half of the total increase in milk production. The other half is a result of improved care, including nutrition. How long can this increase physiologically continue, and what are the implications for the environment, both locally and elsewhere? After all, milking means feeding.

CHAPTER 7

The origin of primordial germ cells in the early embryo

The cells that make it possible to produce offspring are called **germ cells**, from their initial development in the early **embryo** to the formation of egg and sperm cells that are ready for fertilisation. In the form in which they initially appear, they are called **primordial germ cell**s.

Embryologists have, of course, long wondered when germ cells first become distinguishable in mammalian embryonic development. Most of the research in this area, initially descriptive and later experimental, has been conducted on mice. For the mouse, we do not have a complete atlas of embryonic development starting from fertilisation, as exists for the nematode *Caenorhabditis (C) elegans* (a simple model organism). So, there is no atlas that describes the fate of each cell and its progeny. Nevertheless, thanks to many publications, an increasingly complete picture is emerging for the mouse, revealing what drives the embryonic cells in the creation of the body plan that marks the end of the embryonic period. At that point, all organs are present in their early stages of development (human), or organogenesis is largely finished (mouse). From that moment onwards, we talk about **foetus**es and about the foetal period, which will last until birth. Just after implantation, the so-called **germ layers** are formed, and the longitudinal axis of the embryo becomes visible. The earliest stages of development around and just after implantation in the uterine wall are the most difficult to approach experimentally. Yet, these are precisely the most interesting in terms of the appearance of primordial germ cells.

Even in the 70s, there was a strong indication that during the period of formation of the germ cells, there are very few of them. In that decade, techniques were developed to create one embryo out of two embryos, resulting in a **chimera**. The oldest technique involved letting two genetically different 8-cell embryos grow adjacent to each other. To do this, you remove the layer of glycoproteins, the **zona pellucida** (fig 39, 42), and then, under a dissecting microscope, you

nudge the clumps of cells towards each other until they sort of fit together and stay “stuck”. The next morning, you will see an embryo shaped like a short dog bone, which can be transplanted into the uterus of a pseudopregnant mother mouse. This technique dates back to 1961. The second method involves injecting cells from the “**inner cell mass**” (**ICM**) of one embryo into the cavity of another **blastocyst** (fig 43) with a different **genotype**. With the injection technique, you can also use **embryonic stem cell**s in which a particular **gene** has been switched off or altered through genetic manipulation (boxes 2 and 3). The word “chimaera” finds its origin in Greek mythology, in which these creatures appear as a fusion of a lion’s head, a goat’s body and a snake’s tail. They had a bad reputation and were believed to be monsters. But in reality, the cells of the two embryos quickly intermingle: cell mobility is high during embryonic development. As a result, all organs contain cells of both genotypes. It was soon noticed that this was not always the case for offspring, even when the analysis was limited to fusions where the genotypes that were combined were of the same sex (XX/XX and XY/XY).

The most reasonable explanation for this is that the number of cells from which the **germline** develops is not large and originates from a small region. This increases the likelihood that one of the two cell lines will be absent in the offspring. In humans, chimeras resulting from embryo fusion are rare, if there is a reason at all to have the chance to discover them. For the germline, this has never been reported.

Initially, when certainty arose about the presence of primordial germ cells in mice, these were not surrounded by cells predestined to contribute solely to the developing embryo. Instead, they were found to be located outward, near the embryonic membranes. Using a classical enzyme staining technique from 1954, they could be detected as a group of around 40 cells next to the developing 7.5-day-old embryo. We had to wait until 2005 before a little glimpse into the origin of primordial germ cells was revealed from a molecular cell biology perspective. In that year, the group of Azim Surani and Mitinori Saitou from the Gurdon Institute in Cambridge published the involvement of the protein BLIMP-1 in the formation of the mouse germline, shortly after implantation of the embryos and just before the start of the execution of the general blueprint of the embryo with the three germ layers. In 6.5-day-old embryos, they counted between eight and 24 cells in the **epiblast** that stained for this protein. BLIMP-1 is short for “B-lymphocyte-induced maturation protein-1”. It was an eye-opener for me that a protein which later, among other things, plays a role in the differentiation of lymphocytes in the immune system, is used at an earlier stage to initiate the succession of generations.

But what does BLIMP-1 do? In the meantime, it has been renamed PRDM1, a relative of the protein PRDM9 that we encountered in **meiosis** (Ch5.1). We already know that this family of proteins has a piece (a domain), with the so-called "zinc finger" motif, which can selectively browse the vast sea of **DNA** to bind to appropriate base sequences. These proteins thus aid in activating a group of target genes, which, in this case, are required for primordial germ cell differentiation; they can be viewed as **transcription factor**s. Surani and collaborators "**knocked out**" *Prdm1* to obtain the evidence for this (box 3). In the meantime, entirely as expected, a few more proteins/genes have been discovered that are equally indispensable in the formation of the germline. Not too long ago (2015, 2017), the Cambridge-based group published that PRDM1 is also essential for the development of germ cells in humans.

Human primordial germ cells emerge from approximately day 12, which is also after the start of blastocyst implantation on day 7-8 (fig 43). Measured against the stage of embryonic development, the specification of these cells occurs slightly later than in the mouse. This does not make research any easier, due to ethical, moral and legal restrictions in Europe. The experimental study of human embryonic development around implantation *in vivo* is also impossible for biological reasons. In the UK, the boundaries of early human embryology research have been extended, and embryos generated in the IVF lab (Ch16) can be studied up to day 14 after a form of *in vitro* artificial implantation.

In recent years, interest in the early embryology of our own species, including the origin of primordial germ cells, has been on the rise. This is strongly related to the isolation of embryonic **pluripotent stem cell**s from human blastocysts (fig 43, box 2) and the possibilities for researchers to differentiate these cells *in vitro* into all cell types that make up the embryo around and after implantation. Subsequently, they can let them interact with each other to "mimic" embryonic development and thereby gain a better understanding of it. This is the route that Azim Surani and collaborators used to study the differentiation of primordial germ cells in humans. The epiblast, "the embryo in the narrower sense" (fig 43), is also, in our species, mentioned as the site of origin of germ cell precursors. The membrane of the amnion is an even stronger candidate (fig 43). Peculiarities of early embryology in primates can explain this.

When the first differentiation steps to the primordial germ cell are taken, and the cells stain for PRDM1, you will observe that, in the mouse, they are located near the periphery of the embryo, before they start their journey towards the developing gonad while dividing. In mice, this happens from day 7.5 when there are only between 27 and 52 of these cells. At this point, the gonad does not yet have a sex (Ch8), and neither does it when the first primordial germ cells

arrive. In mice, that differentiation step takes place in embryos from 11 days old, and in humans it occurs from the 8th week of development (see Ch8, when the SRY gene starts to be expressed and fig 22). In both mice and humans, the first primordial germ cells have already reached the “gonadal ridge” at that stage (fig 22). They can move on their own but are also propelled by cell movements, partly induced by the growth of the embryo.

Box 2 Pluripotent stem cells

Totipotent refers to the ability of a cell to give rise to all cell types that develop from a **zygote** after fertilisation, including the cells that will form the **embryo**nic membranes and the embryonic part of the placenta. Depending on the species, this ability is lost during or shortly after cleavage divisions; more on this will be covered in chapter 13.6.1. Once the preimplantation embryo in the **blastocyst** stage has formed a cavity and a group of cells referred to as "**inner cell mass**" has accumulated on one side (fig 43), **totipotency** has already come to an end for these cells. The same is true for the other surrounding cells of the trophoblast; they can now only contribute to membranes and placenta (fig 43). Just before implantation, in the late blastocysts, a group of cells can be distinguished within the inner cell mass, forming the **epiblast**. In mice, these make up about 10% of the ±90 cells that the embryo is then composed of (fig 43). At this point of development, we have reached the stage of **pluripotency**. Within the embryo's blueprint, these pluripotent cells can still develop into anything and thus contribute to any type of tissue or organ. They are unbiased in their choice and also contribute to the **germline** (Ch7). The epiblast cells function as the **embryonic stem cell**s (ESC) *in vivo*. The fact that differentiation into different cell types is also possible *in vitro* is essential for embryological, genetic, and medical biological research.

The history of this discovery dates back about 40 years and occurred around the same time in San Francisco, USA (Gail R. Martin) and in Cambridge, UK. Martin Evans and Matthew Kaufman from Cambridge made it into the top journal *Nature* with this, and the techniques they used are also worth mentioning here. Matthew Kaufman is an expert in the field of mouse embryology, and the atlas he created is a standard reference book for developmental biologists. When I was in contact with him in the second half of the 80s, that was his main preoccupation. Evans and Kaufman knew that the inner cell mass increases sharply in size around implantation, while you do not yet see signs of cell differentiation

in tissue sections under the microscope. Because implantation is influenced by ovarian steroid hormones, they decided to remove the ovaries of pregnant female mice 2.5 days after the onset of fertilisation. This delayed implantation, resulting in the development of more mature blastocysts that could be flushed out. They have a larger inner cell mass and more epiblast cells. For these experiments, the researchers used an inbred strain, 129 SvE. The embryonic cells were placed on a bed of skin cells (a so-called "feeder cell layer"), which could no longer divide due to pre-treatment with a strong **mutagen**. In the outgrowths of the embryos, they noticed the formation of spheres, the epiblasts, which they could harvest and further culture to obtain large cell numbers on the layer of feeder cells. They used a test for pluripotency that involved subcutaneously injecting about a million epiblast cells into a mouse of the same inbred strain (the 129 SvE mice from the Jackson laboratory in Maine, see also box 3). The injected cells developed into **teratocarcinoma**s, tumours in which various embryonic tissues can be recognised in a chaotic arrangement.

In 1968, Richard Gardner published that when cells of the inner cell mass are injected into the cavity of a blastocyst, they participate in embryonic development (Ch7). Like attracts like; the cells attach and intermingle while dividing as cells of the epiblast. Using genetic markers that distinguish them from the recipient blastocyst, it can be determined that the injected cells contribute to all differentiated cell types. The result is a chimeric embryo. The two **genotype**s usually (but not always) also contribute to the formation of the germline. This is the test, the gold standard for pluripotency. Back when genetic manipulation (box 3) did not yet exist, those markers for recognising the genetic origin of cells in chimeric mice posed quite a challenge. Initially, in 1975, a mouse with a reciprocal translocation, characterised by a **chromosome** with an abnormal length, was used for this purpose (fig 12). The research was thus limited to dividing cells, in which the marker chromosome could be recognised using a cytogenetic method (Ch3.4). The first analysis of this kind was carried out by Richard in collaboration with Charles Ford and Ted Evans

in Oxford. The technique of creating chimerism through injection later paved the way for precise genetic manipulation. This involves taking advantage of the ability of pluripotent epiblast cells to be further cultured and genetically altered *in vitro* while remaining in an undifferentiated state. We will not go into details of the culturing methods here. The fundamentals of genetic modification are discussed in box 3.

In addition, unravelling the **gene** expression patterns associated with pluripotency has ultimately led to protocols through which virtually any differentiated cell in an adult body can be reprogrammed into a stem cell. For this purpose, the expression of a cocktail of four genes involved in regulating the **transcription** of a specific set of target genes (*Oct 4, Sox 2, Klf4 and c-Myc*, see also Ch9) is upregulated. The stem cell generated through this pathway is known as an "induced **pluripotent stem cell**" (iPSC). These iPSCs are currently making their way into research (for instance, search the web for organoids) and regenerative medicine. The question whether *in vitro* differentiation of these stem cells into **germ cell**s opens the door to guaranteed fertility for everyone is addressed in chapter 16.7.

Lastly, a little more about the aforementioned teratocarcinomas. These were previously accidentally discovered in 1953 at the Jackson laboratory by Leroy Stevens while he was working on the biological effects of cigarette smoke. He used the same inbred mouse strain 129 for this purpose. This strain appeared to be susceptible to developing tumours in the testis, which could cause the organ to swell terribly. In addition to the product of a chaotic embryonic development, these tumours also contained a type of embryonic stem cells that could regrow into a tumour after transplantation.

His discoveries, which improved the definition of the term embryonic stem cell, proved fruitful for Leroy Stevens when it came to acquiring research grants. Consequently, we were able to attend a short course in his lab with a group of foreign visitors in the summer of 1980, focused on "creating chimeric mouse embryos by fusion" (Ch7). This was

possible because there were enough dissecting microscopes to allow four foreigners to practise, and, of course, because his analyst was capable of doing this.

But where did these teratocarcinomas come from? We now know that the gonocytes in the foetal mouse testis (Ch8) are still close to the pluripotent state: the door to pluripotency is, so to speak, not yet sealed with a double lock, and the cells still have some of its molecular features. These molecular features are the same as those found in preimplantation embryos and in the epiblast, with the characteristic expression of the genes for the **transcription factor**s OCT4, NANOG and SOX2 (Ch9). Around day 14 of embryonic development, the expression of these genes decreases in the gonocytes, and the germ cells enter a temporary state of rest. Teratocarcinomas thus arise from an error in the programme that must be followed in the male germline to transform a gonocyte into a prespermatogonium (Ch8). When they fail to do this correctly, there is a risk that the precursors of testicular germ cell cancer will develop. This is further discussed in chapter 12.10.

Box 3 Genetic manipulation

Few tools used in genetic research capture the imagination as much as genetic manipulation, genetic modification or **gene** "editing". The latter term emerged with the introduction of CRISPR-Cas, a technique that has made targeted **genome** alteration much more practical and widely applicable.

In 1981, two papers appeared in *Nature* and *Science*, respectively, describing the ease with which foreign **DNA** is incorporated into the genome of the larger male pronucleus of the mouse **zygote** (Ch13.5, fig 42) after injection with a micropipette. Whether the DNA is of the same species or not does not matter for the integration process. Before being incorporated at a relatively random position somewhere in the mouse genome, "trains" of several to many copies of the injected DNA segment are usually formed in the nucleus. The integration is stable, permanently appears in all **somatic** cells and the **germline** of the developing **embryo**, and is eventually also transmitted through the **gametes** to subsequent generations. Usually, the injected gene's DNA is expressed, but its expression may gradually be lost in successive generations.

We are now talking about the generation of **transgenic** mice, a technique that has since also been used with other animal species. At the time, this was a small technical (but certainly also an ethical-moral) revolution. Personally, I felt a form of disappointment because the male genome was so "promiscuous" towards "foreign" DNA (the response of the female genome will not be very different, but see Ch13.5.2). Undoubtedly, this is a consequence of anthropomorphic thinking. It is reasonable to assume that the presence of double-strand DNA breaks greatly enhances the integration process. Such breaks occur throughout the cell cycle of the zygote, with a sharp increase during S phase. In the male pronucleus, the number of double-strand DNA breaks is higher than in the female pronucleus (Ch13.5.2, fig 42). The formation of these breaks is always attributed to a

replication fork (fig 5) that momentarily halts as it encounters an obstacle for the DNA polymerase, which makes it fragile and easy to break. Through, for example, **homologous recombination** (HRR) and the less precise “non-homologous end joining” (NHEJ) (Ch13.5.2, fig 33) repair mechanisms, the zygote has the ability to repair these breaks, either correctly or incorrectly, and in this process, “foreign” DNA can be incorporated into the genome.

Many believed that all avenues, from gene analysis to rapid genetic improvement, were now open. The species barrier had disappeared, and a pharmacological application, as was pursued in the genetically manipulated and famous Dutch bull Herman, generated venture capital. In this case, it involved the production of the anti-inflammatory protein lactoferrin in the milk of the daughters, something that never materialised. Herman can still be seen in preserved condition in the museum Naturalis in Leiden, the Netherlands.

The real work, however, was yet to come at that time. As mentioned in box 2, the discovery of *in vitro* maintenance and propagation of **pluripotent stem cell**s from the **epiblast** of early embryos took place in the same period. Especially the early embryonic epiblast cells of the mouse inbred strain 129 proved to be very suitable for this (box 2). In 1987, a publication from the University of Utah (Salt Lake City, USA) with Mario Capecchi as the last author described experiments showing that these pluripotent **embryonic stem cell**s (referred to as ESCs) can “exchange” an offered DNA fragment with their own homologous base sequence (fig 19) at a sufficiently high frequency via homologous recombination (fig 14). This had previously been demonstrated for other cell types by Oliver Smithies. Since it was already known at that time that *in vitro* cultured pluripotent embryonic stem cells can fully contribute to cell types of the mouse after being introduced into the cavity of a blastocyst, the path to permanent and controlled genome alteration in the

germline was open. Mario Capecchi, along with Martin Evans and Oliver Smithies, was awarded the Nobel Prize in Physiology or Medicine in 2007 for this discovery.

The exchange of the offered fragment with the chromosomal fragment is similar to **crossing over**, with the difference that two crossovers can now actually occur close to each other (Ch5.1, fig 15). The length of the base sequence-identical segments (fig 19) in which this occurs is important: it should not be too short, so at least a few kb long, and “mismatches” (base differences, Ch5.1) should be avoided. Given their low level of genetic variation, mouse inbred strains form a suitable starting point for this purpose.

Describing the surrounding techniques and tricks to isolate and characterise the cells with successful integration goes a bit beyond the scope here. Homologous recombination generally occurs in a single gene or **locus**, thus making the stem cell **heterozygous** for the genetic change. Technically, any desired modification at gene level can be achieved, from altering a single base to introducing deletions or insertions (table 3). In jargon, there are two main classes: **knockout** means that the gene product is no longer functional; **knock-in** indicates a completely altered or entirely new (usually related, e.g. human) gene function at the site of choice. As indicated earlier in chapter 4, most applications of this are aimed at genetic analyses such as **genetic dissection** of a trait. The ambitious project to create a knockout for every gene in the mouse will also benefit from it (Ch4). Of course, proposals for such experiments are subject to review by an animal experiments committee and a genetic modification committee (the Netherlands). The trade-off between biological insight and ethics is not necessarily an easy one. However, legislation on this and hence the approach to this trade-off varies from country to country.

A refinement of the technique, which allows the genetic modification, here typically a knockout, to be specifically introduced into a particular cell type, can address this: the genetic defect is not present in every cell but only in the cell type to be studied. This is a

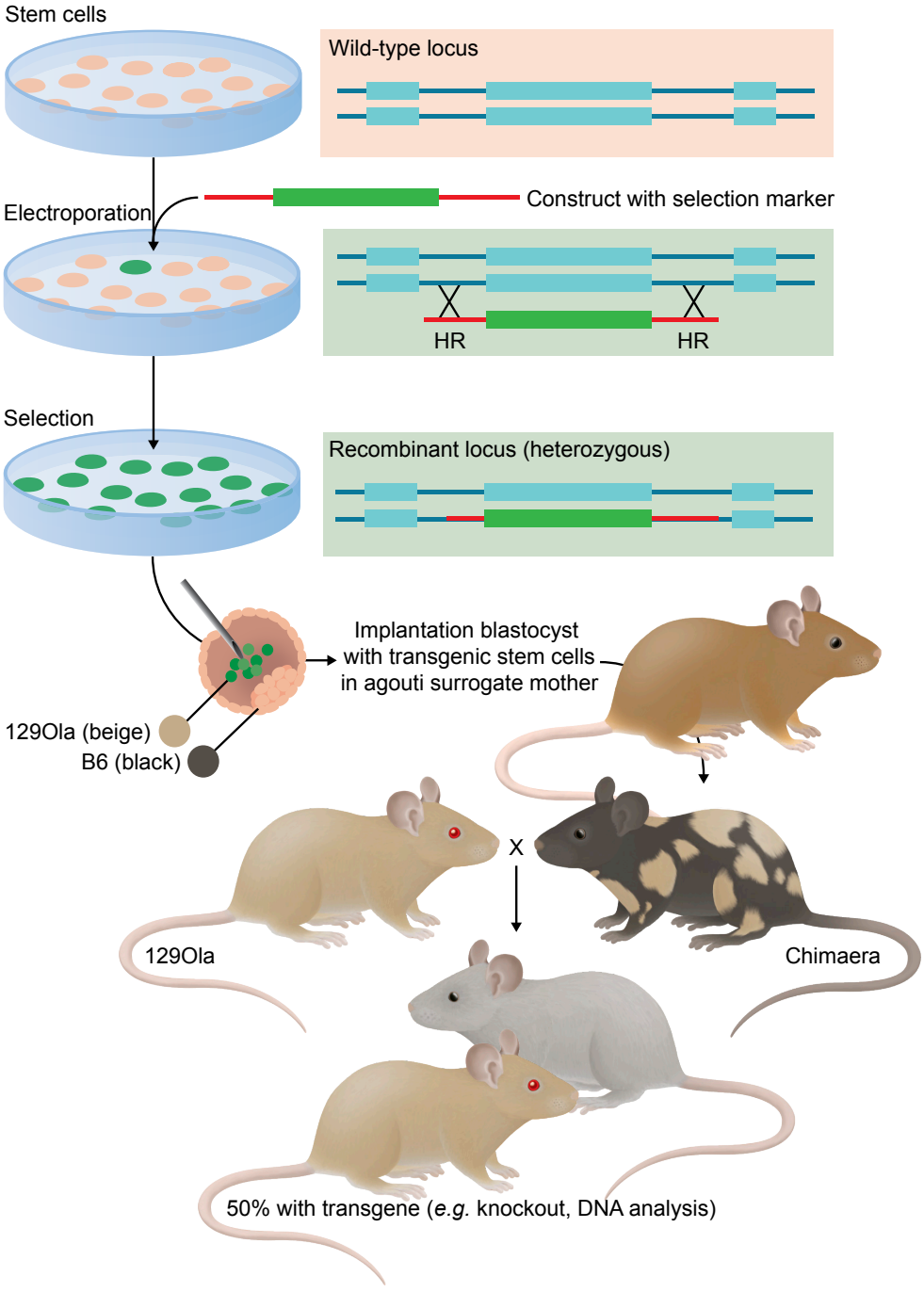

Stem cells
Wild-type locus
Electroporation
Construct with selection marker
HR
HR
Selection
Recombinant locus (heterozygous)
Implantation blastocyst
with transgenic stem cells
in agouti surrogate mother
129Ola (beige)
B6 (black)
X
129Ola
Chimaera
50% with transgene (e.g. knockout, DNA analysis)

FIG 19

Simplified step-by-step representation of genetic manipulation in the mouse **germline**. When the intention is to disable the function of a **gene**, it is referred to as a **knockout**; when an alteration is made with the aim of further studying the **gene** and its product, it is referred to as a **knock-in**. In the manipulated stem cell colonies of the 129 inbred strain that can be selected, one **allele** will have been knocked out (+/-). These cells are injected into the cavity of a C57BL6 ("black six", B6) **blastocyst** (fig 43). In principle, any female mouse can be a surrogate mother; here, an animal with wildtype coat colour (agouti) was chosen. When the **chimeric** mouse is backcrossed with 129 mice, its 129 part will result in the birth of 129 offspring of which 50% are **heterozygous** for the knockout or knock-in (the 129 genetically manipulated stem cell is heterozygous (+/-). This is typically followed by an outcrossing programme, for example with the inbred line B6.

valuable addition to research on the germline. It is also the way to go when the knockout is lethal in the **homozygous** state during embryonic development, which is relatively common for genes involved in DNA repair, for example.

This technology, called Cre/lox, is based on tools used by a bacteriophage (a bacterial virus called P1) to exist in a circular form. A gene of P1 encodes for the **recombinase** Cre, which is specific to and can bind to two identical 34-base motifs in the phage's DNA, known as loxP sites (fig 20). In 1994, the way was paved for the application of Cre/lox in genetic manipulation in mice. The first step involves using stem cell technology (as above) to place the *Cre* gene under the control of a cell type-specific **promoter**, so that expression of the recombinase is restricted to that particular cell type. In the second step, two *loxP* sequences are inserted into/around the gene that is to be disabled using the

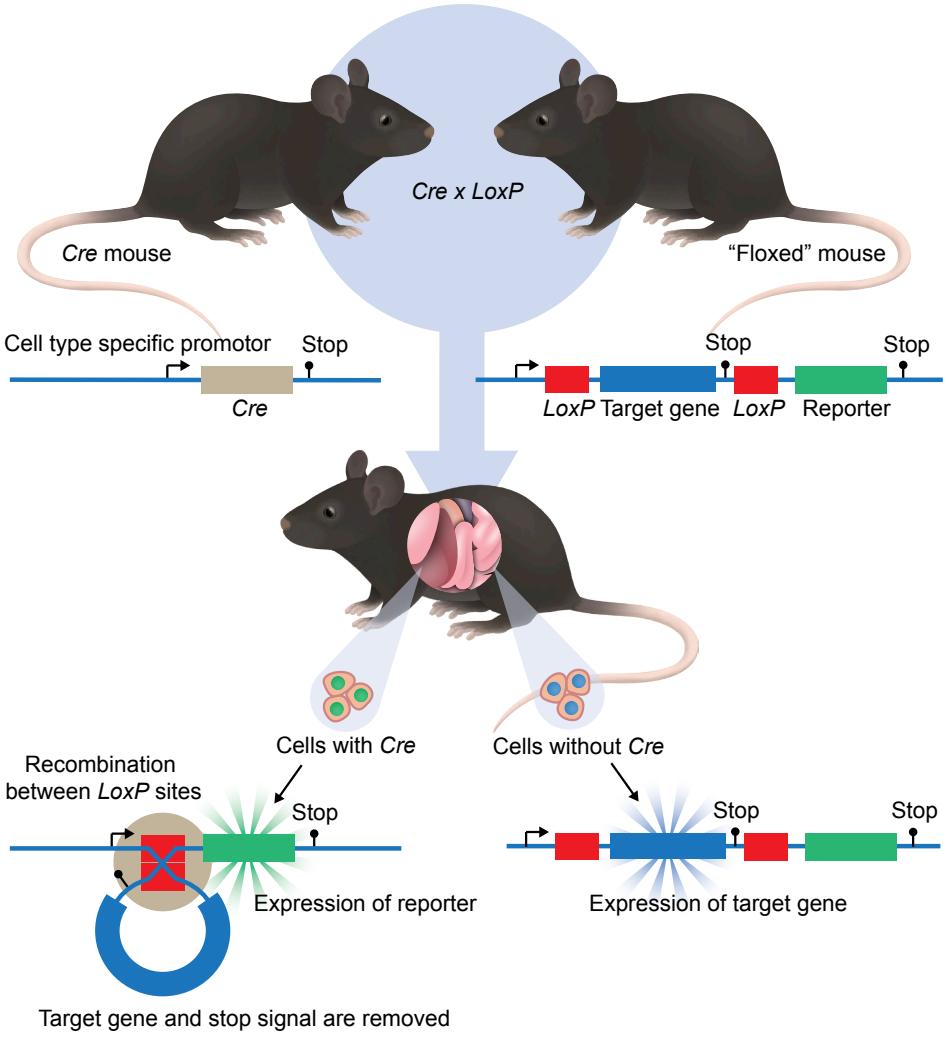

Cre x LoxP
Cre mouse
"Floxed" mouse
Cell type specific promotor
Stop
Cre
Stop
Stop
LoxP Target gene
LoxP
Reporter
Cells with Cre
Cells without Cre
Recombination
between LoxP sites
Stop
Expression of reporter
Stop
Stop
Expression of target gene
Target gene and stop signal are removed

FIG 20

The best technique to inactivate a **gene** per cell type. The **promoter** of a gene that is expressed only in that cell type functions as a gatekeeper. When applied in the **germline**, one can target a gene that is, for example, only expressed in spermatogonia, only in **meiotic** spermatocytes, or a gene that is only expressed in spermatids after **meiosis**. Mice that are **transgenic** for the *Cre* gene fused to that cell-specific promoter ("*Cre*/+") are crossed with mice with the two *loxP* sequences at strategic positions around/in the gene, in order to delete a part of it that is usually disabling. The *loxP* mouse, also obtained through genetic manipulation, must be **homozygous** for this.

same homologous recombination technology. Through crossbreeding, the two genetically manipulated genomes are combined in one cell, where the two *loxP* motifs must be present in a homozygous state (fig 20). The knockout of the gene results from Cre-driven recombination between the two "*loxP* sites" (fig 20). Mouse strains with a proper cell type-specific *Cre* insertion are highly valuable research tools and are thus in demand. This elegant multi-step technology understandably never made headlines as massively as CRISPR-Cas9, the technique for which pioneers Emmanuelle Charpentier and Jennifer Doudna were awarded the Nobel Prize in Chemistry in 2020. The origin story of this latest branch on the genetic modification family tree is extensively described in the popularised report by Jennifer Doudna and Samuel Sternberg: *A Crack in Creation: The New Power to Control Evolution* (2018). John van der Oost, professor of microbiology at Wageningen University, the Netherlands, who played a pioneering role in the discoveries that made the CRISPR-Cas technology (fig 21) applicable, is also featured in Doudna and Sternberg's book. Once again, intriguing discoveries from microbiology are combined with the desire to genetically intervene faster than is possible with traditional breeding methods.

CRISPR-Cas appears to have originated in bacteria as a defence mechanism against viral infections, complete with a genetically encoded memory library (the so-called “clustered regularly interspaced short palindromic repeats” sequences, which was shortened to the acronym CRISPR in Utrecht and Alicante). This way, the DNA of invading bacteriophages can be quickly recognised and selectively cut with the Cas enzyme (an **RNA**-guided DNA **endonuclease**). The great strength of the CRISPR-Cas machinery is the high degree of site specificity with which a cut in DNA can be made. But it remains astonishing. This specificity depends on the so-called “guide RNA”, a piece of at least 20 bases that, through RNA-DNA base pairing, provides a sufficiently accurate indication of position determination in the genome (fig 21). The Cas enzyme cuts the double-strand DNA helix with unprecedented accuracy, at a location specifically determined relative to the paired guide RNA.

Soon, in 2014, it became clear that the mouse zygote could be genetically altered very specifically in this way, eliminating the need to use detours via stem cells. Compared to stem cell techniques, several advantages are speed and cost, animal welfare, and the broad spectrum of possibilities. However, as highlighted in a 2017 review from the NKI (Amsterdam), one should evaluate which technique “leads to Rome” for each desired outcome. The main reason for this is that, with the current state of the CRISPR-Cas technique, a double-strand DNA break is the starting point for the genetic change. The zygote is not a very accessible cell system for research on DNA repair, but it is clear that this cell has many intricacies that have yet to be fathomed (Ch13.5.2): the molecular details of the repair of the break are not fully understood. They also change during the cell cycle of the zygote. In order to obtain more control and with the specific outcome of the repair in mind, a single-strand DNA “template” is added, which has homology to the DNA region in which the break was introduced and contains the desired base sequence at the correct position.

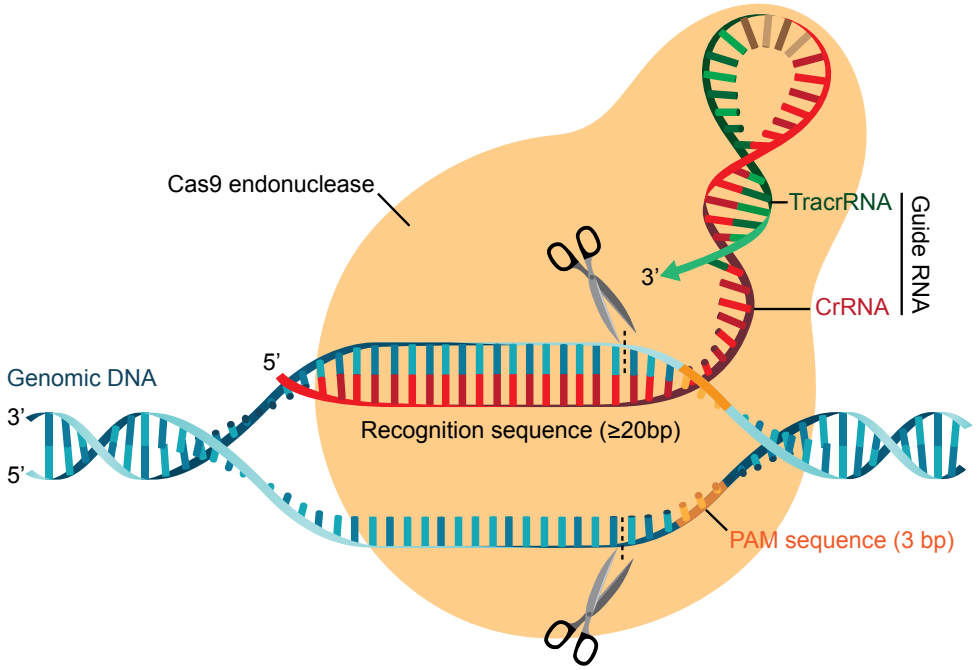


FIG 21

Illustration of the basic principle of the 20 base pair specific recognition (sufficient to approximate essentially any **genomic** location) and the way in which the Cas**RNA** complex can apply a double-strand break with accuracy down to the base. The Cr and TracrRNA come from the bacterial system, the recognition sequence follows the base sequence of the **genome** of the "host cell" that is to be manipulated. The PAM site, NGG, is genomic (N can represent any base). These occur frequently enough, also near the target site, to not hinder a researcher's plans. The system has further evolved, the double-strand **DNA** break can be exchanged for a single-strand one ("nicking") by a Cas variant. Base editing enzymes can be added to exchange one particular base ("base editing"). Alternatively, to make small insertions possible, a modified guide RNA is added in conjuncture with a **reverse transcriptase**. This module is named prime editing.

However, much more fancy modifications that are based on a single-strand break have been developed, and more are likely to come (see the caption of fig 21).

From around 2015, publications have begun to appear in which the CRISPR-Cas technology is applied to human zygotes to repair defective genes. The majority of these are from East Asia, especially China. As a rule, the mysteries of DNA repair in the zygote are treated as a black box, which in turn leads to reactions from DNA repair specialists from, for example, the USA, who point out that the discovered outcome is not actually biologically possible, thereby further increasing the mystery.

Stimulated by dr Jiankui He's experiments that led to the birth of genetically edited babies, for which he was sentenced to 3 years in prison in late 2019, a futuristic vision of germline modification in humans enabled by CRISPR-Cas has made its way from the Dutch ART field into quality newspapers. This seems premature now. It is clear that a double-strand break in the zygote is not a risk-free start of a **mutation** correction. More research on this is currently being conducted. Researchers are particularly interested in the so-called "off-target" effects − how infallible is that 20-base postal code − and are curious about variations in the process around the DNA repair work resulting from the double-strand break introduced by Cas9. Both aspects are being discovered. Benjamin Davies from London, an active mouse researcher in this field, analysed the human situation for the European ART community for ESHRE 2019, the annual major international congress dedicated to artificial reproductive technologies in Europe. He presented an update of his summary for the gathered Dutch IVF biologists in a fairly dazzling lecture in early 2020 in Zeist, the Netherlands. What became evident earlier, and also followed from his experiments in mice, is that when the Cas enzyme is still active in at least the 2-cell stage, it sometimes leads to mice being **mosaic** for the introduced mutation. Another concern is that CRISPR-Cas often cannot properly distinguish between the mutant **allele** in need of repair and the wildtype, healthy allele. For this, the difference between the

maternal and paternal alleles in the DNA base sequence used to design the “guide RNA” (fig 21) needs to be sufficient to selectively target one of the two. When this is not possible, the essentially unnecessary double-strand break that was introduced in the healthy allele may result in a defective protein when repaired incorrectly (Ch13.5.2). Benjamin Davies’ “take home message” was: it is not yet safe (see also Ch18.2), but the likelihood of it becoming safe is increasing. For example, a double-strand DNA break might be avoided by using the tool of active **retrotransposons** (Ch2.3.3, fig 10). When it comes to introducing a single base change, “base editors” exist, for C to T and for A to G. Another recent development is that of the so-called “epigenome editors”. These are enzymes that allow you to (locally) modify the methylation status of DNA (fig 11) or the **PTM**s of **histone**s (Ch2.3.2, table 1), thereby influencing the **transcription** of a gene. This is not (yet) the place to elaborate on this. Another bottleneck remains: mature eggs and high-quality zygotes are costly, sometimes very costly. In every case, the outcome of the repair will have to be monitored later in the preimplantation stage, with the possibility of mosaicism continuing to be an issue. Whether curating is better than selecting against the mutation is further discussed in chapter 18.

CHAPTER 8

The development of the embryonic gonad

Already in the early days of **cytogenetics** in humans, from the late 50s onwards (Ch3.4), everything indicated that the Y **chromosome** is involved in the development of the testis. The presence of a Y chromosome is decisive, and the number of X chromosomes does not matter. Also, in mice, XXY individuals are male and XO individuals are female.

From the very beginning, that is in the 70s and 80s, there has been some sort of postulate that in mammals, one single **gene** is responsible for the development of the testis. In the language of that time, development into an ovary was considered to be a more or less autonomous process (the “default pathway”), the basic pattern from which the **embryo** could only be diverted by intervention from the Y chromosome. Without a **foetal** gonad, the development is female, as was already shown in 1947. The hormonal aspects of the development of the excretory ducts of the gonad and the external sex characteristics are beyond the scope of this book.

The hunt for this “masculinity gene” on the Y chromosome was “hot” for years. There were some candidates, but in order to provide evidence, insertion of this gene into a pronucleus of the **zygote** and subsequent incorporation into a chromosome must produce a male, even if the sperm carries an X chromosome (box 3, fig 42). **Mutation**s in this gene can lead to (sterile) XY females. Also, men with two X chromosomes, who are rare and, like Klinefelter XXY men (Ch3.4), sterile, are interesting research subjects in this case. Moreover, you would expect the gene to be functionally conserved in mammalian evolution; the main domain of the protein that ultimately determines gonadal sex differs in only one **amino acid** between humans and chimpanzees. When **DNA** analysis gained momentum in the 80s, this paved the way (by using these sex differentiation mutants) for the isolation of a candidate gene in mice and in humans, with the possibility of testing this gene in the mouse model. The gene, called *Sry* (written as *SRY* in

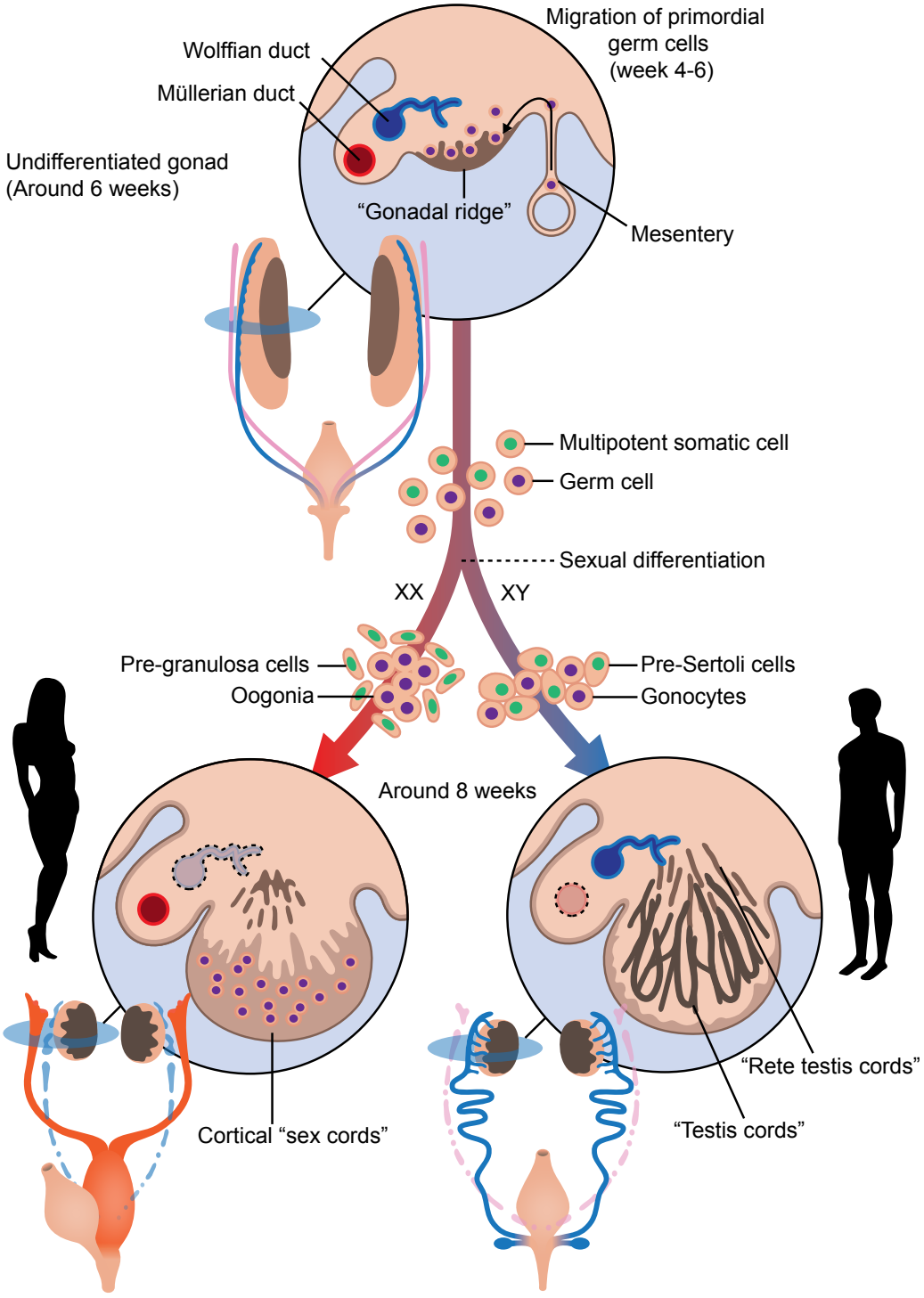

Migration of primordial germ cells (week 4-6)
Wolffian duct
Müllerian duct
Undifferentiated gonad (Around 6 weeks)
"Gonadal ridge"
Mesentery
Multipotent somatic cell
Germ cell
Sexual differentiation
XX
XY
Pre-granulosa cells
Oogonia
Pre-Sertoli cells
Gonocytes
Around 8 weeks
Cortical "sex cords"
"Rete testis cords"
"Testis cords"

FIG 22

Schematic representations of the development of the **embryo**nic gonad in humans with minimal ages post conception. Sexual differentiation of the gonad begins after 6-7 weeks due to the expression of *SRY*, which is transient. By this time, the first **primordial germ cell**s have already reached the gonadal ridge. Duct formation for both sexes is present. After differentiation of the gonad, one of the two duct systems must regress; for the female sex, this is the Wolffian duct, and for the male sex, the Müllerian duct. The rapid transition of **germ cell**s to **meiosis** in the ovary occurs much more synchronously in mice than in humans. This tighter control is also evident in the embryonic/**foetal** testis; in mice, the germ cells have a temporary **mitotic** arrest, which is defective in humans. The rete testis tubules are part of the developing drainage system of the testis.

humans), was inaugurated in 1990 in *Nature*, coming from London, for both humans and mice. The genetically engineered XX mouse that turned male also appeared in *Nature* a year later. The insertion of the *Sry* gene into the DNA of the zygote was done in a very primitive way in those days. This particular DNA fragment was injected into the male pronucleus (fig 42) of a mouse zygote in a large number of copies using a micromanipulator. When a double-strand DNA break occurs somewhere in a chromosome, such a piece of DNA has the chance to be incorporated into the genomic DNA of the cell, free-riding on the DNA repair systems of the zygote (Ch13.5.2, box 3).

Sry appears to be a rather unusual gene; it has a distinctive DNA binding motif, which is functionally conserved within mammals. In mice, it is only expressed over a relatively short period of time in the developing embryonic gonad (fig 22). The expression begins at 10.5 days, peaks between 11 and 11.5 days and then vanishes again. In humans, it appears 42 days after fertilisation and decreases from day 53. SRY is a **transcription factor** of a specific class. The actual work for the development of the testis is carried out by another member of the same family of SOX proteins (SOX9) that is regulated by SRY. *Sry* is the founding father of this family, which has 20 members. In turn, SOX9 elevates the expression of the genes required for this process: *SOX9* must remain “on” for sex differentiation at the gonadal level to proceed correctly.

An embryonic gonad that can still develop in two directions is called bipotential. In addition to the arriving **germ cell**s, this gonad consists of 3 types of cell lines: (a) the cells that will directly

support the germ cells and are in intimate contact with them, (b) the cells capable of producing sex hormones, and (c) the cells that provide the general structure. The cells that will directly support the germ cells determine the sex of the gonad through the expression of *Sry* and *Sox9*. When these cells differentiate in the direction of a testis, they first form pre-Sertoli cells, which give rise to Sertoli cells (Ch12.2, Ch12.4, fig 22). In the female differentiation towards the ovary, these cells develop into pre- or foetal granulosa cells and later into the granulosa cells found in the primordial follicles (**oogenesis**, Ch11.1, fig 28). The activation of the *Sry/SRY* gene thus leads to the differentiation of Sertoli cells. The fact that this occurs over a relatively short period, of half a day or slightly more in mice, remains remarkable and makes this a delicate process. At the onset of sex differentiation in mice, in 11-day-old embryos, there are around 20,000 germ cells according to the literature. Such estimates have a large confidence interval. The capability to eventually become a pre-granulosa cell persists for two more days. The germ cells that follow the female trajectory diligently divide and form the pool of oogonia that is at the base of the potential supply of eggs for the entire lifespan (Ch11.1, fig 28). The germ cells directed towards **spermatogenesis** also divide, and they are named gonocytes.

The first insight into the dominant position of the pre-Sertoli cell for the development of the testis came from studying the gonadal development of chimeric mice (Ch7): work done by Paul Burgoyne in the laboratory of Ann McLaren, the “grand old lady” of early embryology in England, in the 70s-90s. When you create **chimeric** mice, the **sex chromosomes** of the embryos to be fused are unknown. Statistically, you will obtain 25% XX/XX, 25% XY/XY and 50% XX/XY, XY/XX. When visually sexing these mice, it is noticeable that more males than females are born and that very few mice have intersex characteristics. It seems that the XY cell line largely determines the fate of the embryonic gonad, due to the early activation of the *Sry* gene which results in the development of pre-Sertoli cells. In the chimeric XX/XY males, the Sertoli cells predominantly had XY chromosomes. At the same time, the molecular machinery necessary for the development of an ovary is inhibited. The precursors of the testosterone-producing Leydig cells can have either XX or XY chromosomes, and the same is true for the germ cells. If, due to its **genetic background**, the *Sry* gene of the Y chromosome is switched on late during embryonic development, thereby delaying or even indefinitely deferring testis development, an ovary with XY oocytes may develop. This occurs in an outcrossing programme of, for example, *Mus poschiavinus* wild mice (from Italy, the males) and females of the well-known C57BL6 (B6) inbred line. By backcrossing onto B6 through males, starting with the **F1** male, the POS Y chromosome makes its way into this genetic background. It also happens because the SRY transcription fac-

tor of *Mus poschiavinus* does not interact well with the *Sox9* gene of B6; this interplay is quite variable. With this, it became possible to systematically study XY oocytes. I thought about this when I recently came across some researcher's musings raising the possibility that same-sex partners could conceive their own biological child (Ch16.7). If the partners are male, could an XY egg obtained through stem cell technology (box 2) be an option?

Recently (2017, 2021), a Canadian group conducted extensive research on those XY oocytes in mice, using a combination of molecular and microscopy techniques. **Meiosis**, especially the second meiotic division, does not proceed smoothly, and the early embryonic development thereafter is catastrophic. They also observed the absence of a crossover between the X and Y chromosomes, causing them to be unpaired during the first meiotic division. In April 1980, I had a look at an XY oocyte from the *Mus poschiavinus* B6 crossbreeding programme at the Jackson lab (Maine). The oocyte was in the metaphase of the first meiotic division, and the X and Y chromosomes were unpaired. This was not new and had already been described in 1977, when the Italian mice were not yet in the picture. At that time, the observation came from the MRC Radiobiology Unit Harwell, describing a rare XX/XY chimeric female. However, scientists are not the ones to be baffled by the laws of nature. With the inefficient *in vitro* oogenesis protocol at hand, the Japanese research group led by Katsuhiko Hayashi reported in 2023 that utilising mouse **pluripotent stem cell**s, an XY **karyotype** can be converted into an XX karyotype. It has long been known that, at least in the mouse, the Y chromosome is less stable in mitosis; a spontaneous loss is not rare. The chance for the remaining X to be doubled after that is lower. Here, it is helpful to invalidate the SAC (see fig 4). *In vitro* oogenesis after sex reversal will result in offspring that originate from two XY germlines, proof of principle delivered. The prospect of selfing is not brought up here, however interesting this may be for the more theoretical geneticists. More intricacies of *in vitro* oogenesis are given in chapter 16.7.

How does gonadal development proceed from the moment of sex determination (for mice starting at 11 days, for humans from week 8, see fig 22)? In the female pathway, the amount of cell proliferation of the supporting pre-granulosa cells is limited, in contrast to the behaviour of the male pre-Sertoli cells. The oogonia organise themselves at the edge of the gonad and form connections with each other, which are called syncytia or cysts (fig 22). Later, during the formation of the primordial follicles, these connections break down (Ch11.1). From 12.5 days of age onwards, meiosis can be initiated, guided by vitamin A. In mice, this is a fairly synchronous process, unlike the situation in humans. There, meiosis begins after 12 weeks, but this process is far from accurately synchronised (see also Ch11.1), and a group of oogonia seems to escape it.

When pre-Sertoli cells start to dominate in 11-day-old mouse embryos, testicular tubules containing the germ cells begin to emerge a day after (fig 22). These gonocytes multiply exponentially up to 13.5 days of age but enter a state of dormancy a few days later. They are located in the centre of the testicular tubule and move towards the wall shortly after birth. This change in location marks the transition to cells that can initiate spermatogenesis: the stem cells.

In the development of the human testis, fully developed testicular tubules with gonocytes (fig 22) are present after 11 weeks, gradually crawling towards the wall from that time onwards. They are then referred to as prespermatogonia, and they already tend to stick to each other after a division using a **cytoplasm**ic bridge: cell division is not completed entirely, which results in the formation of syncytia here as well. This is a very important characteristic of spermatogenesis, which is not omitted in chapter 12. As in mice, the germ cells that do not settle in the tubules by moving towards the wall are lost and undergo **apoptosis**. In mouse **foetus**es, gonocytes cease to divide altogether after 15.5 days, and this process only resumes shortly after birth. In humans, this development takes place over a much longer timespan, and there is always a population of prespermatogonia that is dividing. During the first months after birth, the lumen of the tubules still contain gonocytes, which can then still transition into prespermatogonia. Throughout this process, the germ cells retain some of the characteristics of a pluripotent stem cell (box 2): the loss of **pluripotency** occurs heterogeneously and progresses very gradually (see also Ch12.10).

CHAPTER 9

"Resetting" chromatin in the germline

The succession of generations is often depicted as a chain, a commonly used metaphor for a great miracle. The **embryo**nic origin of a new generation of **germ cell**s can be thought of as the narrow neck of an hourglass. Very few cells receive the signal from their environment to develop into **primordial germ cells** (Ch7).

During their formation, primordial germ cells stand out from the cells that surround them because differentiation into **somatic** embryonic cell types is blocked; they thus exit the mainstream of somatic cellular development in a timely manner. In mammals, this "escape" occurs relatively late; in humans, it happens slightly later than in mice (Ch7, fig 43).

After being destined to become a primordial germ cell, the **chromatin** undergoes a "reset". This process is completed when the primordial germ cells have reached the embryonic gonad. Prior to this overall reset, there is an initial reset that occurs in the embryo that has not implanted yet. This step is necessary to make the transition from the chromatin structures that are inherent to the formation of oocytes and sperm to the chromatin that enables the development of the preimplantation embryo. The early embryonic reset leads to the achievement of **totipotency** and **pluripotency** (box 2) and is completed in the **blastocyst** stage (fig 43).

Within a cell, there is never a greater difference in the structure of chromatin than between the two nuclei, the pronuclei of the **zygote** (fig 42). This is discussed in more detail in relation to fertilisation (Ch13). Making the chromatin of the paternal and maternal chromosomes more "equal" is another very important aspect of the initial chromatin reset. It takes some time, up until the **epiblast** stage (fig 43), before the maternal and paternal chromatin have become virtually "equivalent".

The chromatin parameter CpG methylation (fig 11) was first used to visualise the “resetting” during the development of the preimplantation embryo. This became possible when an antibody against the methylated form of C(ytosine) became available in the 1980s, and simultaneously, cell staining using **immunofluorescence** methods became more applicable. When the initial findings from the Medical Research Council “mammalian development unit” in London appeared in the literature in the second half of the 80s, it was a big eye-opener. More modern **NGS**-based techniques that read CpG methylation at the base pair level have only confirmed this picture later on (fig 23). The first thought upon seeing the graphs from Marilyn Monk and her fellow researchers was that, up to the blastocyst stage, the main focus is on preparing the possibilities for embryonic development around and after implantation. Observations from research and practice focused on IVF (Ch16.3.1 and Ch16.3.2) support this notion.

During the initial reset, the **DNA** of both the sperm and the oocyte loses the methyl group on the C (fig 11, 23). For the paternal DNA, this is an active process that begins immediately after **gamete** fusion, one of the many tasks of the egg. This process is still not entirely resolved, but it is clear that DNA repair systems play a role in it (BER, fig 33). For the female **genome**, it seems to be a matter of thinning the level of methylation. During semi-conservative DNA **replication** (Ch2.2.2, fig 5), this involves “forgetting” to add a methyl group to the newly incorporated Cs when a CpG and a GpC separate. In the blastocyst, the methylation levels in both genomes are equal, apart from a few exceptions. The most well-known of these exceptions are the “imprint control regions” (ICRs) involved in **genomic imprinting** (Ch10, fig 27).

These DNA regions are “reset-resistant” and remain methylated on the CpGs. Genome-wide demethylation does not affect ICRs in the preimplantation embryo; the egg has the tools for this (fig 24). A very nice example of this is the availability of a stock of **mRNA** of the maternal effect **gene** *Trim28*. The protein encoded by this gene acts as a “matchmaker” in a complex assembly of proteins with various tasks involving chromatin, including the protection of ICRs from demethylation (fig 24, see also Ch13.5.3). It consists of a cocktail of **transcription factor**s, “chromatin writers” and “chromatin readers” (Ch2.3.2, fig 24).

Also in humans, paternal DNA is actively demethylated in the zygote. However, thanks to the much greater resolving power of current techniques, it has recently been discovered that methyl groups are subsequently reattached to Cs, and that this occurs in conjunction with the inactivation of **retrotransposons** (LINEs and SINEs) (Ch2,3,3, table 2). CpG methylation (fig 11) serves as a confirmation signal for the locally inactivated state of chromatin. At this stage of human embryonic development, this also applies to repetitive elements such as LINEs and SINEs.

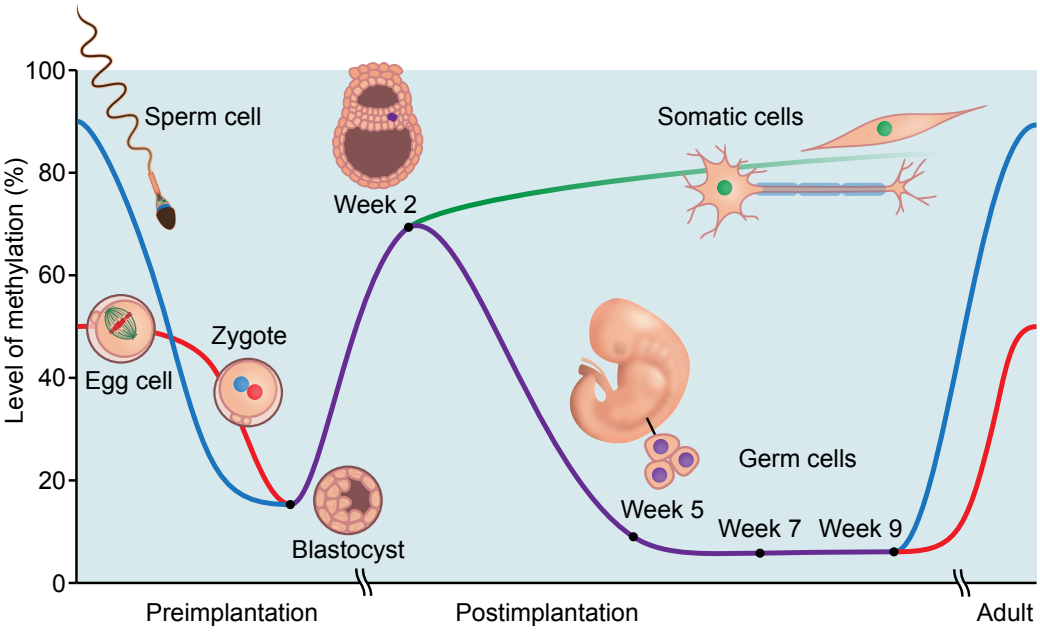


FIG 23

In the literature, the methylation degree of cytosine (usually in a CpG **dimer**, fig 11) is used to illustrate the resetting of **chromatin** in the **germline**. The female and male **gametes** are strikingly different in this regard. In the **blastocyst**, the overall methylation level has greatly decreased and is uniform. This is part of, and symbolises, the first reset. After that, at the onset of **embryo**nic development, rapid remethylation occurs, which is a dynamic process. Following the specification of **primordial germ cell**s, this process is reversed to an even lower methylation level, referred to as the second reset in the germline. The female **genome** is shown in red, the male genome in blue. This figure is inspired by the situation in humans.

CpG methylation remains a good parameter to describe the resetting in the **germline** (fig 23). But in reality, there is much more that needs to happen in the transition to the totipotent and pluripotent state. Nowadays, the complexity of chromatin remodelling in the preimplantation embryo can be investigated with great resolving power, and the data obtained this way are hardly comprehensible. The techniques are becoming increasingly reliable, but currently, it still takes 300 eggs (mouse) to obtain an impression for each individual gene **promoter** over the whole genome. The whole process is extremely fascinating, but its details go beyond the scope of this book. The resetting of chromatin is meant to lead to the activation of genes that are essential for totipotency and pluripotency, so they must be "on" in the **pluripotent stem cell**s (box 2). The transcription factors OCT4, SOX2 and NANOG are crucial in this process. The genes encoding these proteins are switched off during **oogenesis** and **spermatogenesis**, but they are not additionally sealed; there is no CpG methylation in the promoter region. This may be a harbinger of their rapid activation during the early stages of embryonic development. Throughout this entire operation, the retrotransposons (Ch2.3.3, table 2) must not get the opportunity to become active again. Their expression is, therefore, partly suppressed by the TRIM28 complex (fig 24).

As mentioned at the beginning of this chapter, the second reset takes place during the transport and proliferation of primordial germ cells on their way to the gonad (Ch7 and Ch8). Primordial germ cells tackle the tasks vigorously and remove the imprint codes of both sexes; the ICRs (fig 27) are demethylated in this process. This is necessary so that the "imprint" corresponding to the sex can be reestablished later during **gametogenesis**. The overall DNA demethylation is carried out more thoroughly than during the first reset, which primarily serves to initiate embryonic development (fig 23). The chromatin is also reorganised again. During this whole process, evolutionarily relatively young retrotransposons must be prevented from becoming active. Therefore, they retain their CpG methylation.

Each generation of germ cells thus starts with a clean slate. How clean that slate actually is will be discussed in chapter 15. After all, it remains questionable whether the whole process is really so straightforward. For now, it is important to note that exceptions have been found regarding CpG methylation at the level of individual genes, a discovery made in 2015 by Azim Surani from Cambridge. This may be an aspect of **epigenetic** inheritance, which is covered in chapter 15.

A final aspect of the second reset occurs only in XX primordial germ cells of female embryos because they were formed just after the period of random X chromosome inactivation (box 4). The two X chromosomes locate each other at the beginning of **meiosis** to prepare for the

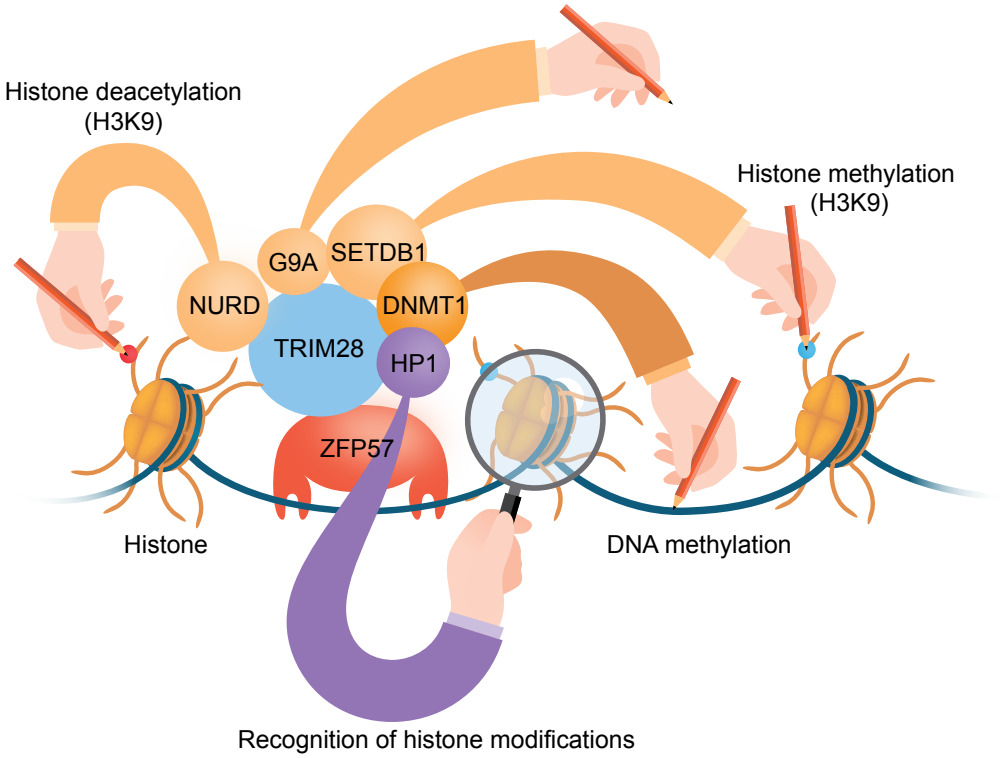


FIG 24

An artistic impression of an example of a multi-protein complex that regulates **DNA** activity at the **chromatin** level. In this case, the integrator of the functions of this complex, which becomes functional shortly after fertilisation, is the protein TRIM28. The whole complex is also named after this protein. Among the functions that need to be coordinated, those of the chromatin "writers" (orange) and "readers" (purple, concepts introduced in Ch2.3.2) are the most important. The writers perform their function through **histone post-translational modifications** (the **PTM**s in table 1) and through DNA methylation (fig 11). As a vital link in the regulation of **gene** activity, the protein ZFP57 (shown here in reddish-brown, a suppressor of **transcription**) is of great interest. It is a "zinc finger" protein (as PRDM9 from Ch5.1) and it can thus arrange positioning in the **genome**. The suppression of **retrotransposons** and protection of the "imprint control regions" (the ICRs from Ch10 and fig 27) against genome-wide demethylation of CpG **dimer**s (fig 11) from fertilisation to the **blastocyst** stage (fig 23) are very important functions of the TRIM28 complex.

process of **crossing over** (Ch5.1, Ch11.1). Any existing difference in chromatin organisation between the **homologous** X chromosomes at that time would hinder this process. Therefore, it seems logical that the recently inactivated X chromosome must be reactivated.

CHAPTER 10

Genomic imprinting

Secondary oocytes (fig 13, i.e. egg cells, Ch13) do not necessarily need to be fertilised to start **embryo**nic development. This can also be achieved with a trick, such as treatment with a bit of alcohol. Embryos created in this way are called **parthenogenote**s, and they can be **diploid** when the products of the second meiotic division have remained in the egg (in that case, the second polar body has not formed, fig 42). The maximum development of these "immaculately conceived" embryos is already reached before the organs are formed, so in the embryonic stage.

After fusion of the egg and sperm, two nuclei (called pronuclei, Ch13) arise in the now-formed **zygote**: one male, from the nucleus of the sperm, and one female, from the second meiotic division of the egg (Ch13.5 fig 42). Under the right microscope illumination, such mouse pronuclei radiate like tiny micro suns. In the early 80s, it became technically possible to exchange pronuclei between zygotes (fig 25).

By using a fungus-derived substance to "relax" the cytoskeleton (fig 2) of the zygote, you can extract the pronuclei using suction. Essentially, the removed nucleus is a small cell with a cell membrane. Using a short electric pulse, this mini cell can be fused with a zygote from which a pronucleus has first been removed. This is the protocol for constructing embryos with either only genetic material from the mother or only from the father (fig 25). The results of this procedure were reported by two research groups at about the same time, in 1984: the group in which Azim Surani worked at the Animal Research Council unit of Reproductive Physiology in Cambridge, and Davor Solter's group worked at the Wistar Institute in Philadelphia. The embryos with two female nuclei are called **gynogenote**s, and those with only male nuclei are called **androgenotes**. The difference between the aforementioned parthenogenotes and these gynogenotes is not that large. After transplantation into a pseudopregnant mouse, the gynogenotes develop

in the same manner; development of the best ones will halt during organogenesis. For the androgenotes, the outcome is even sadder; some may still implant, but embryonic development ceases a few days after implantation, which is when their maximum is reached.

The development of entirely “male” embryos resembles a form of spontaneous abortion in humans, with the abnormally shaped embryos referred to as “hydatidiform moles”. These consist predominantly of extraembryonic structures, resulting in termination of the pregnancy in the 4th-5th month.

In the years before the pronucleus exchange experiments, evidence that paternal and maternal **chromosome**s are not equal was obtained via an alternative route that, in principle, leads to the regions in the **genome** where differences in expression of the maternal and paternal **gene** copies can be found. Figure 26 explains the idea of these tests, for which reciprocal translocations are most informative (fig 12). When a chromosomal segment originating from the mother is equivalent to the same segment from the father, the origin does not matter, as long as it is present in double quantity. This is possible when the other parent does not supply that particular segment. But how do you recognise the offspring arising from such a complementary fertilisation (2 + 0 instead of 1 + 1)? That became the experimental condition that researchers like Tony Searle, Colin Beechey and Bruce Cattanach from the MRC radiobiology unit in Harwell (UK) had to fulfil in the 70s and 80s. **DNA** markers did not yet exist at that time; those could only be used from the early 90s onwards. The researchers approached this cleverly by building upon the many mutants with a **recessive** inheritance pattern, whose **homozygous** recessive **genotype** could be identified by appearance, preferably already showing a **phenotype** at birth. And, of course, one had to know on which chromosome and where on that chromosome the gene in question was located. The creation of strains with the reciprocal translocations (but also Robertsonian translocations, fig 12) and the **phenotypic** marker genes on the chromosomes involved in the translocation was the most costly in terms of the number of mice and the time spent (fig 26). The first somewhat larger study in which **heterozygotes** were crossed, conducted by Tony Searle and Colin Beechey, was published in 1978, albeit without immediate interpretation that the remarkable failure of **complementation** was an example of **genomic imprinting**. Due to the persistent work on this project in Harwell, a reasonably clear idea about the chromosomes (and segments thereof) for which male and female descendants are not equivalent had emerged by 1985. This led to the “mouse imprinting map”, in which Colin Beechey was involved until 2011. At first, regions of chromosomes were identified that had to come from both parents to prevent the development of an aberrant phenotype. Over the years, the genes located in these regions

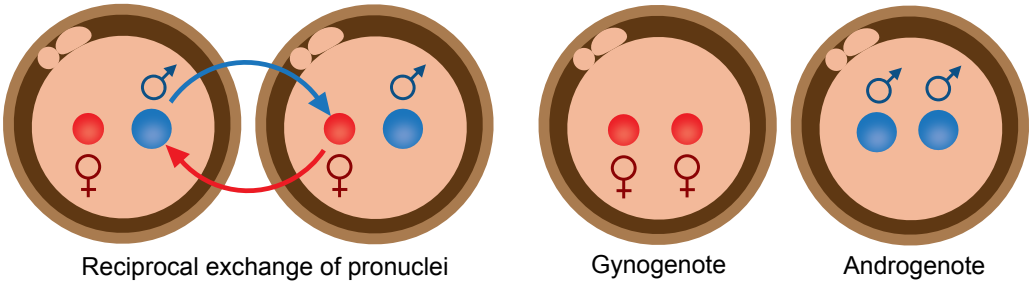
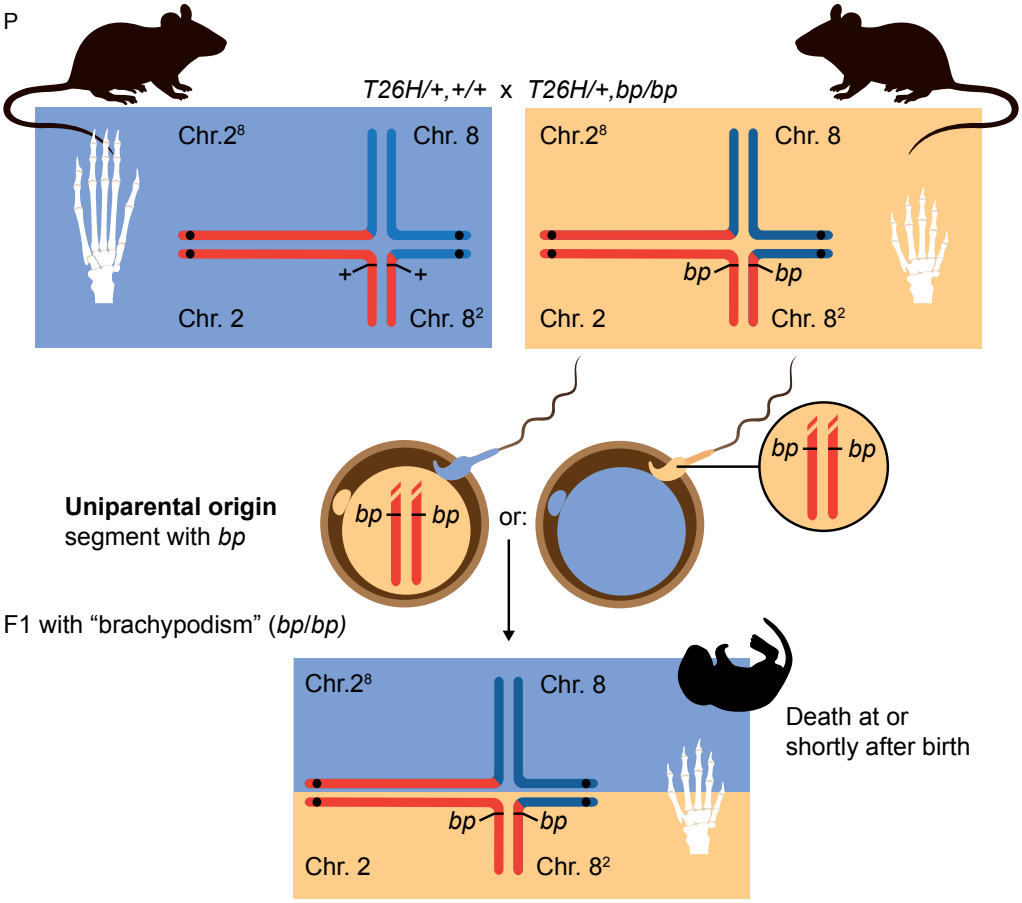


FIG 25

Illustration of the mouse experiment with the exchanged pronuclei, which gave the phenomenon of **genomic imprinting** a major research boost.

for which complementation is not possible were slowly isolated as well. These genes were found to be expressed in a single dose, either originating from the father or the mother (fig 27). Gradually, the function of these genes was elucidated. There was great interest in the **epigenetic** mechanism hidden behind monoallelic expression. Meanwhile, pathologies that appeared to be related to imprinting had also been identified in humans, such as Prader-Willi syndrome, Angelman syndrome and Beckwith-Wiedemann syndrome. Due to these epigenetic mechanisms, men and women cannot do without each other when it comes to reproduction, and what could be more beautiful than that!

What is the insight so far? Jo Peters, also from Harwell, has made great contributions to the further unravelling of the epigenetic mechanisms behind genomic imprinting during her active research career. In this, she collaborated intensively with Gavin Kelsey from the Babraham Institute in Cambridge. The following is based on their reviews. Delving into genomic imprinting directly leads to the resetting of epigenetic information in the **germline** (Ch9). The sex-specific father-mother information must be preserved during the time span between fertilisation and the onset of embryonic development and must then, in principle, be retained in **somatic** cells (the ICRs of Ch9, fig 27). However, during the second reset in the germline, when the pool of **primordial germ cell**s is formed, this information must be erased; these cells have not yet been "converted" to a sex at that time (Ch7, Ch8, Ch9, fig 23). Once that has occurred and the choice between **oogenesis** and **spermatogenesis** has been made, the correct instructions for the ICRs of the imprinted gene clusters are reinstated, corresponding to the sex (fig 27).



Example of quadrivalent derived by complementation in female descendant

FIG 26

One of the first **complementation** trials performed at the MRC radiobiology unit in Harwell (UK) in the 70s. The chromosomal segment running from the breakpoint of the *T26* translocation to the end of **chromosome** 2 is marked by the **recessive mutation** *bp* which, in the **homozygous** state, leads to severe shortening of the bones in the limbs (especially the phalanges of the feet). Otherwise, these mice are fully viable and fertile. In reciprocal translocations (fig 12), the search for **homology** during the first **meiotic** prophase in heterozygotes leads to the formation of a quadrivalent, as depicted here. Explaining the distribution of chromosomes among the secondary oocytes and spermatocytes (fig 13) goes beyond the scope of this section. Importantly, the appearance of two copies in a **gamete** is far from rare, especially for the segments between the translocation breakpoint and the chromosome end (away from the **centromere**, marked here by *bp*). This can be compensated for at fertilisation by a complementary gamete that lacks this particular segment. Beige represents the gamete with the duplication (contains chr 2 and chr 8^2) and lilac represents the gamete with the deficiency (contains chr 2^8 and chr 8) for the segment of chromosome 2 marked by *bp*. The figure shows that it does not matter which parent is homozygous for *bp*. All homozygotes (*bp/bp*) die at birth or shortly thereafter. In this way, the entire mouse **genome** has been mapped for **genomic imprinting**. The fact that one parent is responsible for one set of **gene** products and the other parent for another set hinders complementation, but the effects of double gene doses (where it is typically single due to genomic imprinting) persist through the complementation trials. P is “parental”. **F1**, see the glossary of terms.

The origin of imprinting must have something to do with the development of a uterus in mammals and with viviparity, the development of the placenta. But it also occurs in numerous other life forms. Shortly after the idea of a contribution to the next generation confined to either sex had reached theoretical biologists, they started developing concepts that accommodate this phenomenon. The most well-known of these is the “parental conflict hypothesis” (but there are more). This hypothesis assumes that the father and mother are in competition, because the mother is responsible for nourishing the embryo, **foetus** and infant, as well as herself. For the father, it might be advantageous to strive for heavier offspring, even if it comes at the expense

of the mother. After all, there is competition for nutrients between mother and foetus. A single, uniformly conclusive theory that explains imprinting will not exist and would also be difficult to devise, given the multitude of bodily functions that are affected by it.

The phenotypes of the androgenotes and the gynogenotes seem to be each other's opposite when it comes to the balance between the embryo and the embryonic membranes. Later, at birth, a reflection of this can be observed for multiple small chromosome segments that either originate exclusively from the father or exclusively from the mother (fig 26). The Harwell group referred to them as "opposite phenotypes". An example: when the duplication comes from the mother and the deficiency from the father, the mice are very thin at birth, will not suckle, and die within 24 hours. In the opposite situation, they are fat and hyperactive but also die somewhat later. In 2014, Jo Peters summarised the (patho)physiological processes in which genomic imprinting is involved in a brilliant paper published in *Nature Reviews/Genetics*. The mentioned processes and related conditions were: (a) prenatal and postnatal survival and growth; (b) **metabolism**, including obesity; (c) neurological and behavioural effects, including sleep behaviour and social behaviour, for example, caused by psychiatric abnormalities. And finally, cancer. For metabolism, aspects such as glucose tolerance and insulin sensitivity are mentioned. Especially the neurological and behavioural aspects of genomic imprinting are currently drawing attention. A more recent review indicates that this side has also attracted more research in mice. Apart from behaviour (e.g. risk-taking and impulsivity), cognitive processes may also be influenced by imprinting. In humans, Prader-Willi and Angelman syndrome (also used to illustrate the human manifestation of the opposite phenotypes), are a good illustration of the involvement of imprinting in brain function, given the intellectual disability that occurs in these diseases.

In an animal experimental approach to determine the direction of paternal and maternal effects, **chimeras** are created of normal preimplantation embryos and gynogenotes or androgenotes, after which aspects such as brain development are examined. When the balance shifted towards the androgenotes, the chimeras became larger, but their brains were smaller. In the reverse situation, with a greater contribution from the mother, the exact opposite was observed. Now, the offspring were smaller, but their brains were larger.

In mice, around 260 genes are now known to be regulated sex-specifically (by origin). For comparable chromosomal regions in humans (the **orthologes**), a rough number of 228 is given. They are found on almost all chromosomes. Completely in line with the theory and experimental findings surrounding the phenomenon of genomic imprinting, it has been shown that the placental **transcriptome** is the richest in imprinted genes.

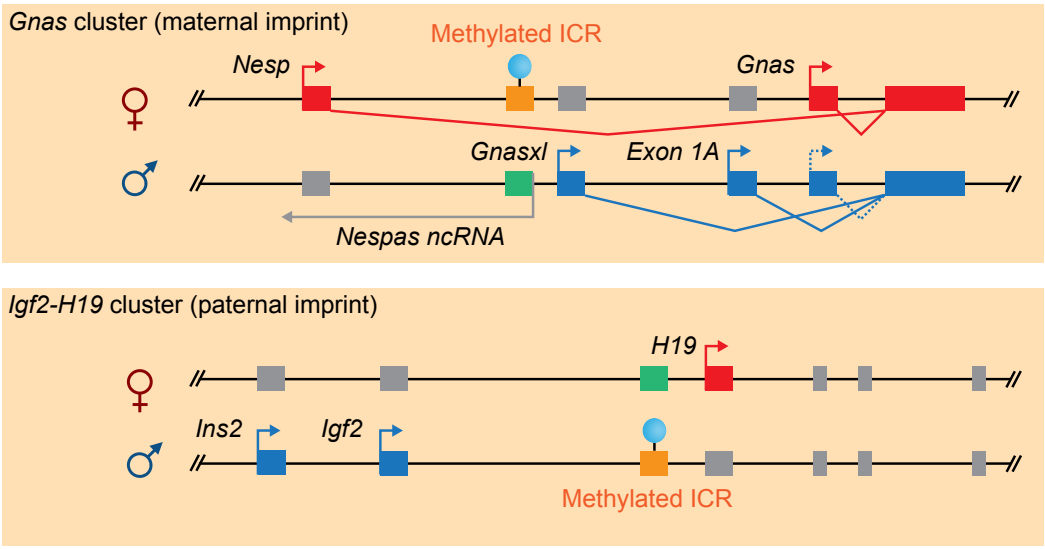


FIG 27

A simplified representation of **gene** regulation around an "imprint control region" (ICR, shown in orange here in the methylated inactive form and in green in the active form). The addition "paternal"/"maternal" indicates the sex of the **germline** in which the ICR is methylated (the imprint). The *Gnas* cluster is located in the same segment of **chromosome** 2 as the gene *bp* from fig 26. The **transcription** direction is indicated by an arrow. Genes with maternal expression are shown in red, genes with paternal expression in blue. How **exons** are utilised in the *Gnas* cluster is indicated by connecting lines (box 1). Among other things, the *Gnas* cluster codes for *Nespas*, as well as the product of *Exon 1A*, a long non-coding (lnc) **RNA**. The product of *H19* in the paternally imprinted *Igf2 H19* cluster is also a non-coding RNA. The grey blocks depict genes with suppressed expression.

In the investigation of the mechanism(s) behind gene expression based on the sex of origin, researchers came across the phenomenon called "differentially methylated region" (DMR). These are regions in which the methylation level of C(ytosine) differs between **homologous** sequences in normal body cells (these Cs are almost always followed by a G, see fig 11). There are quite a few of such regions, and the significance of many is still unknown. However, when they are located in a chromosomal segment for which genomic imprinting effects have been

demonstrated and when they are involved in the regulation of this, they acquire significance and are given a different name as a result. They are then called “imprint control regions” or ICRs (Ch9, fig 27). Figuring out how the regulation of gene activity is redacted from ICRs is an example of modern research for which genetics and molecular biology currently have the technical capabilities. In mice, three regions with a paternally methylated ICR and 20 regions with a maternally methylated ICR have now been identified. The transcriptional activity of a relatively small number of nearby genes on the chromosome, and also on the homologous chromosome segment, is regulated by these ICRs. For further clarification, see the caption of figure 27. Whether there could be communication between the homologous segments under ICR control is an aspect that has been touched upon and that may receive attention again in the future. The methylation status of the ICR does not automatically predict which of the genes that it controls in its vicinity are on and which are off. For the majority of imprinted genes in humans, the paternal **allele** is expressed (76%). In mice, the ratio between the expression of maternal and paternal alleles is slightly more balanced (45% vs. 55%, respectively).

It is not the intention of this book to zoom in on individual genes more than is necessary, but two genes encoding functionally linked proteins are worth mentioning. *IGF2*/*Igf2* is a gene of which the paternal allele is expressed (fig 27). It is one of the main growth factors for prenatal development. The expression of the *Igf2* receptor (*Igf2r*) on the cell membrane originates from the maternal gene in mice; the paternal *Igf2r* gene is switched off. IGF2R transmits the IGF2 growth signal from the external to the internal side of the cell via interaction with the growth factor. In an analysis of the **genetic architecture** of stature (back height) in cattle, regulators of IGF2 were identified as relatively important. Is this an example of a “balancing act” between the sexes? In humans, the imprint status of the IGF2 receptor is evolutionarily fading, but there are plenty of imprinted genes in the placenta of which the maternal copy is expressed, thereby helping to maintain the balance between mother and foetus. Effects on other physiological processes and neurological functions are also currently receiving ample attention in research on the role of imprinting in humans. For example, not long ago, in 2018, the role of epigenetic changes on gene expression was examined more closely in Iceland, with a genome-wide focus and with emphasis on the Prader-Willi region. The CpG methylation landscape (fig 11) could be separately reconstructed for the paternal and maternal genomes, and the **transcript**s were also distinguished by origin. This provided a much more detailed picture. Various ratios of paternal to

maternal CpG methylation were detected, extending beyond the regulation of classically known genes subjected to genomic imprinting. Hence, the situation is not always as black-and-white as had been assumed until then; as is often the case, the reality is more complicated.

Box 4 More on X chromosome inactivation

While describing **heterochromatin** in chapter 2.3.2, X **chromosome** inactivation was already briefly brought up. In genetic jargon, this solution to the difference in the number of X chromosomes between females and males is referred to as **dosage compensation**. It has already been mentioned in chapter 3.4.2 that **embryo**s with either one chromosome too few or too many are generally not born due to the numerous **gene dosage effect**s. The general purpose of X chromosome inactivation is to prevent a difference in gene dosage between males and females for the **genes** located on the X chromosome. Since the X and Y chromosomes, which originated from a homologous pair, have still maintained some "homology" from their common past, inactivation is not necessary when both gene copies are present. This more profoundly applies to the whole **DNA** base sequence of the relatively small homologous chromosome segments that are known as the **pseudoautosomal region**s (or PARs). They escape dosage compensation because it is not needed, and they are essential for enabling **crossing over** (Ch12.5.1, fig 37). The simple view that it is sufficient to just turn off one set of X-linked genes by inactivating an entire chromosome has recently been challenged by the study of the expression of X-linked genes in subjects with **sex chromosome aneuploid**ies. The take home message of this paper is that besides X chromosome inactivation (inactivation in cis), there is also X-linked gene regulation in trans, occurring between the X chromosomes. This is not rare, with more than 121 genes involved, of which 10 are most likely to have an impact on the **phenotype** of sex chromosome aneuploidies (table 4). Again, the genetic molecular approach shows that nature acts in a much more detailed and interactive way than textbook pages can convey.

In mice, the process of inactivation in female embryos commences very soon after fertilisation, at some point during the 4-cell stage (fig 43). The X chromosome from the father is inactivated first; this is called imprinted inactivation (for **genomic imprinting**,

see Ch10). This changes around implantation, and a selection process takes place in each cell of the **epiblast** (fig 43): either the maternal X or the paternal X will become inactive. To enable this, the imprinted activation of the paternal X has to be undone. Eventually, across the entire population of **somatic** cells, the maternal X chromosome is active in approximately half of the cells, and the paternal X chromosome in the other half. This leads to several distinct phenotypes, of which the tortoiseshell/calico cat is by far the most well-known. The gene for orange (O/o) is located on the X chromosome. In the presence of the active allele (O), “orange” dominates the dark colour encoded by the **autosomal** genes, while this is not the case for the inactive allele (o). Therefore, when only one X chromosome is present, you will not find a tortoiseshell/calico pattern: these are more often male cats.

In humans, the imprinted inactivation of the paternal X chromosome is skipped, but the final result is the same. Around implantation, a one-to-one distribution has also been established between both the activity and, *mutatis mutandis*, inactivity of the maternal a paternal X chromosome.

The inactivation in both mice and humans is controlled by the **orthologous** *Xist/XIST* genes. When this gene is active on one X chromosome, it must be switched off on the other. The convention around the denotation of genes and their products is that in mice, gene names starts with a capital letter and are written in italics; genes in *Homo sapiens* are also italicised and written using all capital letters (in both mice and humans, names that are not italicised and written in all capitals refer to the protein encoded by the corresponding gene). The *Xist/XIST* gene contains the base sequence for a long non-coding **RNA** (lncRNA) with around 15 kb of **nucleotide**s (a XIST protein does not exist). Ever since the evolutionary divergence of the lines that eventually gave rise to mice and humans, the basic design of the gene has been preserved. However, as expected, changes have occurred, especially in a part of the gene where this could be tolerated. It is now fairly well understood how this XIST RNA works, partly as a result of the work

of Joost Gribnau from the Department of Developmental Biology at Erasmus Medical Centre, Rotterdam. From the *XIST* gene on the X chromosome, the XIST RNA covers its environment and, in cooperation with proteins, it spreads over the entire chromosome. As a result, the X chromosome then adopts an inactive **chromatin** status. Local gene activity is thus necessary to render an entire chromosome inactive. It is believed that the RNA subsequently serves as a kind of platform to provide access to various players in the chromatin field. This way of inactivation is initially very similar to the regulation of genomic imprinting (Ch10, fig 27). As expected, the inactive X chromosome is rich in H3K27me3 (table 1), and the **promoter**s of the genes that are to be inactivated are methylated on the CpGs, which is, of course, not the case for the active *XIST* gene (fig 11). Chromosome **replication** occurs late in the cell cycle. The *XIST* promoter of the active X chromosome is methylated, and the gene is turned off. However, despite these suprachromosomal measures, there are still genes that have escaped these **epigenetic** controls. There are many more of those genes in humans than in mice. Out of the approximately 850 X chromosome-localised genes in mice, 17 always escape inactivation, and for 20 it is not entirely straightforward. Of the 639 human genes with a known inactivation status, 80 are steadfast refusers, and another 93 exhibit variable behaviour. One conclusion that could be drawn from this may be that males and females differ more in humans than in mice when it comes to the genetic activity of the X chromosome. This could be explored by comparing the development of female embryos with only one X chromosome in humans and mice. In humans, this is known as Turner syndrome (45,X), a relatively common abnormality of the number of sex chromosomes (1 in 2,000-4,000, see also table 4). The spontaneous occurrence of XO mice is rare, but good models exist for this. From these models, we know that increased prenatal mortality only occurs when the X chromosome comes from the father. When the X chromosome originates from the mother, this increase is not observed. In humans, the overall picture is much more dramatic, with 99% prenatal mortality. In mice, XO females grow more slowly and remain somewhat smaller, but they are fertile. I still remember a talk that Mary Lyon gave at the laboratory in Harwell in the

autumn of 1971, in which she reported that the fertile lifespan of XO female mice was slightly reduced, but not very much; it was a matter of a few months at most. Depletion of the stock of oocytes could be a plausible reason. The situation is very different in girls with Turner syndrome. Apart from detection of the syndrome based on stunted growth, girls with Turner syndrome are also identified because the menstrual cycle does not commence during puberty: the ovaries do not harbour any **germ cell**s or follicles (fig 28, 29). Women with a Turner syndrome diagnosis who become pregnant after all almost always have tissues with a combination of 45,X and 46,XX cells, a **mosaic** that is presumably of very early embryonic origin. This means that an X chromosome was lost during the division of a **blastomere** (fig 43), and the cell line formed from this was not subsequently removed through selection.

The relatively large difference in the 45,X, XO phenotype between humans and mice is consistent with the smaller number of genes that escape X chromosome inactivation in mice: physiologically, female mice are more similar to male mice in terms of gene functions originating from the X chromosome. These “escape genes” are, of course, extremely interesting in the search for genetic differences between men and women, especially when they are orthologes in mice and humans and when they are, for example, involved in **histone** demethylation (table 1, fig 8).

Having seen that women and men have different levels of expression in a subset of these genes on the X chromosome, an obvious question is whether this could also have an effect on the functioning of the brain, which is actually a side path in this book. One indication is that women with three X chromosomes (XXX, where two X chromosomes are inactivated, leading to a higher activity of the escape genes) have a phenotype in which the psyche is explicitly affected. It therefore seems that some of these escape genes have a function in the brain.

A very recent illustration of the power of these escape genes is the gene *TLR7*, which plays a role in the difference in the response to a COVID-19 infection between men and women. The higher dose of the protein encoded by this gene contributes to women’s

better chances of fighting the virus. A reliable radio station in Germany (WDR5) managed to explain this perfectly.

X chromosome inactivation is something that is visible around us, as you can tell from the story about the tortoiseshell/calico cat at the beginning of this box. My mentor in science in Harwell, Tony Searle, scored his first widely cited paper with this in 1949. After his time as a Japanese prisoner of war (in the period 1941-1944; this included a brief stay near the Siam Burma "death railway"), Tony became interested in genetics in the post-war years and pursued his studies in London. There, he came into contact with JSB Haldane, one of the "founding fathers" of population genetics in the years before World War II. Haldane is still cited today. He was the first to recognise that **mutation**s in genes occur much more easily in the male **germline** than in the female germline (table 3). Haldane's second scientific legacy, which is also still cited, concerns the sterility that occurs in the offspring of partners that have evolutionarily drifted apart. Haldane's rule, formulated in 1922, states: "When, in the **F1** offspring of two different animal races, one sex is absent, rare, or sterile, that sex is the **heterozygous** sex (heterogametic sex)." Mules and hinnies are the best-known examples of that phenomenon, and chapter 12.5.1 will provide more information on this.

Haldane, partly of Indian descent, had the characteristics of an eccentric Englishman. His office at the University College in London was overcrowded, and when Tony Searle asked the secretary during his first visit to the eminent scholar where he could sit while waiting, she said, "on a chair". The only free chair was Haldane's own, but Tony did not argue with the secretary. That first meeting did lead to Tony starting a population genetic study on coat colours and patterns in the cats of London, alongside his PhD research in another department. These cats were spotted in the night, and deceased animals were sometimes kept in the freezer at Searle's house. The most famous part of this project: the cats with two or three colours with as many "orange" as "non-orange" sons. And there

were also female offspring with “orange” and “non-orange” coats. With a bit of math using the Hardy-Weinberg rule, known all too well in population genetics, you can easily determine whether the cats choose each other based on appearance, or if that does not matter to them (which is the case). If you spot a tortoiseshell/calico tomcat, then this is an XXY “Klinefelter” male, meaning he is sterile (Ch12.12.1).

Could chromosome inactivation be the reason why women suffer much less from monogenic X chromosome-linked disorders than men? An explanation for this would be helped by the assumption that cells in which the “wrong” X is active have a selective disadvantage compared to the “right” one in competition for replication. Another possibility is that cells with the correct X chromosome provide “neighbourly assistance” in tissues or organs, thereby reducing the expression of the phenotype (i.e. the physiological defect).

Well-known X-linked disorders where this comes into play are colour blindness, Duchenne muscular dystrophy, **androgen** insensitivity and X-linked intellectual disability. These diseases are relatively common, in part because women can unknowingly be carriers of the mutant allele. In boys, Duchenne muscular dystrophy occurs in about 1 in 3,500 births. Only 2.5-8% of female carriers of the mutation show symptoms of the disease. It appears that an unbalanced distribution of the active and inactive X chromosome is associated with this, leading to a reduced number of muscle cells producing the correct dystrophin protein.

While muscle problems are less frequently encountered in women carrying a Duchenne allele, carriers of the X-linked intellectual disability allele are more likely to exhibit a phenotype. Estimates of the population frequency of this condition vary somewhat, but around one in 1,250-4,000 men have the phenotype compared to one in 2,500-8,000 women. This suggests that symptoms of the disorder manifest in about half the carriers, who are thus heterozygous. However, women are less strongly affected and are also considerably less likely to be diagnosed with ADHD. Here too, it seems that the severity of the syndrome is related to the extent to which the mutated X is also the active chromosome.

CHAPTER 11

Oogenesis

11.1 The overview

When the **primordial germ cell**s arrive in a gonad that "opts" to go into the female direction (due to the absence of the **gene** product of *Sry* (human *SRY*), Ch8), they continue to divide a few more times and are then called oogonia. These divisions are incomplete so that **cytoplasm**ic bridges persist between the germ cells; cysts are formed. In mice, the oogonia begin **meiosis** fairly synchronously, triggered by a vitamin A signal. As a result, about 25,000 primary oocytes are in the zygotene or even already in the pachytene stage by day 15 post coitum (p.c.) (fig 14). In humans, about 600,000 oogonia can be found from 8 weeks after the onset of pregnancy, and they continue to divide after that (up to the 6th month).

After meiosis has started in the primary oocytes around week 12 (again guided by vitamin A), the first pachytene stages can be found from 14 weeks onwards. Just like in mice, the number of **crossovers** can be counted at this stage using MLH1 **immunofluorescence** (Ch5.2, fig 15). As mentioned above, oogonia can continue to divide for weeks, which is why there are still oocytes in the pachytene stage at 25 weeks. The degree of asynchrony in meiotic development is so different in humans compared to mice that it continues to astonish researchers. That astonishment extends to the extreme variation that is seen among **foetus**es when it comes to the number of primordial follicles formed after the pachytene (something that is not reflected in fig 28). At about 22 weeks, the number of germ cells peaks at approximately 7 million, after which mortality rapidly dominates (fig 28). This phenomenon was a mystery for a long time and it actually still is. For oocytes, the normal course of events is to enter a special form of the diplotene after the pachytene (fig 14, 28). Thereby, the structure of the **bivalents** locks in place, so to speak. This

can be observed from week 18 onwards, and it is the key stage during which, a little later, the formation of primordial follicles occurs, and in which oocyte development takes place until right before ovulation. This bivalent form, with its associated **chromatin** organisation, is the starting point for all processes required to produce a mature egg. From the diplotene onwards, the cytoplasmic bridges between oocytes are abolished, and each oocyte must figure out how to attract supporting cells, the pre-granulosa cells discussed in chapter 8. At this point, the formation of primordial follicles with granulosa cells has begun (fig 28). During this process, the newly formed follicles move further away from the peripheral edge of the ovary. This phase begins with a large wave of oocyte mortality (fig 28). It appears that the maintenance of control over active **retrotransposons** (table 2) is an important cause of this. Indeed, in 2020, in a research project that was demanding in every respect, Alex Bortvin and collaborators from Baltimore discovered in mice that unwanted activity of LINE retrotransposons plays a role in the high oocyte mortality during the phase in which follicles are formed and when the pachytene comes to an end. In mice, this occurs around the time of birth.

Baker's famous graph from 1963 (fig 28) illustrates the large prenatal decline in the number of germ cells in humans. This phenomenon has prompted a group of seasoned meiosis researchers, of the type who continue working on this topic all their lives and for whom retirement does not exist, to thoroughly compare male and female meiosis in humans once again, hoping to gain more insight this way. In 2017, they reported on this in the top journal *Cell*. For their study, they used immunofluorescence with antibodies against several proteins involved in **crossing over** (including MLH1, Ch5.2 and fig 14, 15) to map this process for our species. They paid particular attention to **chromosome**s 16 and 21. The reason for chromosome 16 is that this is where errors occur most frequently during the separation of the **homologues**. Chromosome 21 was monitored due to its causal involvement in Down syndrome (Ch3.4.2). It had been noticed earlier that the number of crossovers in humans is much larger (x1.5) in female meiosis than in male meiosis, as well as the variation. The final conclusion of this study is that about 25% of crossovers are not completed. This is another reason for **apoptosis** of the oocyte. The beginning of female meiosis in humans is thus characterised by a large variation in fate among oocytes.

Let's go back to the primordial follicles in figure 28. In humans, there are about 2 million of these around the time of birth, with great variation among individuals. This is the stock for a woman's entire reproductive life. Right away, primary follicles (fig 28) develop from these primordial follicles. The granulosa cells change shape and come into close contact with the oocytes. They start dividing, resulting in primary follicles developing into secondary follicles. Finally, a

small cavity forms: the tertiary follicle is born (fig 28). In all animal species, including humans, the follicle is then between 200 and 300 µm in size, approximately a quarter of a mm. From that point on, the granulosa cells are also surrounded by theca cells, the steroid hormone-producing cells of the follicle.

After a primary follicle is formed, the oocyte starts to grow, and a gelatinous substance, the well-known **zona pellucida**, is deposited on the outside (fig 28, 39, 42). Around the time that a cavity appears between the granulosa cells, during the formation of the tertiary follicle, the oocyte is fully grown (fig 28). Throughout the follicular phase, granulosa cells maintain intensive contact with each other and with the oocyte. Cytoplasmic extensions emerge, making contact with the oocyte membrane through the zona pellucida. That allows them to communicate back and forth, and it also enables the granulosa cell to take over tasks from the oocyte. This concept is known as “bidirectional communication”, which is quite normal among cells that are part of a tissue, but is implemented in a more detailed manner here. Naturally, it also means that there is interdependence, for example, with regard to carbohydrate and lipid **metabolism**, which involves the exchange of intermediate compounds. For energy metabolism, the emphasis in the granulosa and cumulus cell (fig 28, 29) is on the breakdown of glucose into pyruvate and lactate, substances that are supplied to the oocyte. Therefore, the oocyte mainly uses the citric acid cycle in the mitochondria for energy production. However, the supporting role of the granulosa and later cumulus cells is not limited to supplying energy; the collaboration is endlessly complex.

In this mutual communication, the oocyte is ultimately in charge. For instance, the oocyte protects the granulosa cells and stimulates cell division, and follicle growth is influenced by signals from the oocyte. Another demonstration of this principle is the role the oocyte plays in the appearance of the steroid hormone-producing theca cells at the time of differentiation into the tertiary follicle (fig 28). Eventually, the granulosa cells are, in turn, involved in maintaining the meiotic arrest in the distinctive diplotene stage (fig 14, 28) of the oocyte, although this happens at the command of the oocyte. Oocytes can even signal granulosa cells to go into apoptosis. In doing so, they also eventually self-destruct. The entire follicle becomes **atretic**, taking the oocyte along in its deterioration. Since only about 400 eggs at most are destined for ovulation in a woman’s lifetime, this fate befalls almost all oocytes.

Initially, the tertiary follicles are not yet large enough to become responsive to the hormones that the pituitary gland secretes in a cyclic pattern, the **gonadotropic** hormones (fig 29). Further growth is usually necessary for this, and the size of the follicle at which this happens varies between species: in mice and rats, that size is about 0.25 mm, in pigs 1 mm, in sheep 2 mm, in

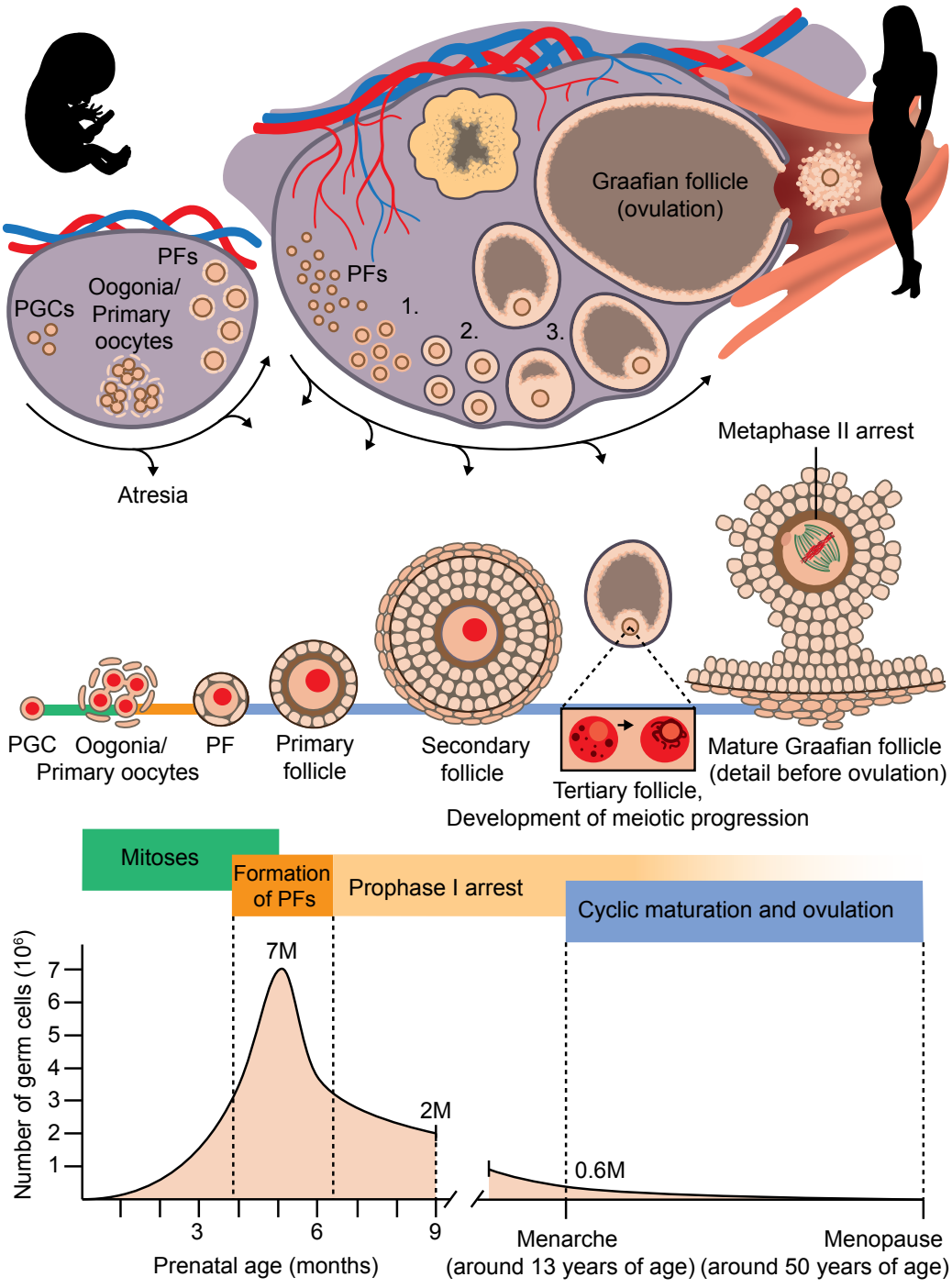

Graafian follicle
(ovulation)
PFs
Oogonia/
Primary
oocytes
PGCs
PFs
1.
2.
3.
Atresia
Metaphase II arrest
PGC
Oogonia/
Primary oocytes
PF
Primary
follicle
Secondary
follicle
Tertiary follicle,
Development of meiotic progression
Mature Graafian follicle
(detail before ovulation)
Mitoses
Formation
of PFs
Prophase I arrest
Cyclic maturation and ovulation
Number of germ cells (10^6)
7
6
5
4
3
2
1
7M
2M
0.6M
3
6
9
Prenatal age (months)
Menarche
(around 13 years of age)
Menopause
(around 50 years of age)

FIG 28

An overview of **oogenesis** from the **embryo**nic stage to adulthood. The **foetal** stage is characterised by the prophase of the first meiotic division in the primary oocytes (fig 13) and by the formation of the primordial follicles (PF). PGC is "**primordial germ cell**". When the primordial follicles are formed, a variant of the diplotene stage (fig 14) occurs, called the "dictyate stage" in the literature. This stage is specific to oogenesis. In the overview of the adult ovary, the numbers 1, 2 and 3 represent primary, secondary and tertiary follicles, respectively. In the tertiary follicles, cumulus cells arise from the granulosa cells around the oocyte, causing it to detach more from the edge of the follicle. After the secondary follicle, the diameters in the drawing are no longer in proportion. The formation of the **zona pellucida**, which first appears in the primary follicle, is indicated by a dark line. The loss of primary oocytes is a continuous process with a strong peak during the formation of the primordial follicles. The loss continues after that and manifests as follicular **atresia**. One of the few clear morphological aspects in the nucleus of the oocyte that has been associated with this is the arrangement of **heterochromatin** (Ch2.3.2) around the **nucleolus** (fig 2), illustrated here in the magnified image by the tertiary follicle. Oocytes that do not exhibit this reorganisation of the nucleus are largely eliminated. At the time of ovulation of the Graafian follicle (>18 mm), the oocyte is in the metaphase stage of the second meiotic division (see also fig 39). Around and after ovulation, the then empty follicle transforms into the corpus luteum, Latin for "yellow body" (top). Birth is indicated by an interruption in the timeline that shows the change in the number of oocytes (data all human).

humans and cattle 3-5 mm and in horses 10 mm. Since follicles are only rescued in this way from puberty onwards, those that develop at a younger age (which indeed happens) are doomed. This explains the approximately 80% reduction of the follicle pool that occurs between birth and puberty in humans (fig 28).

At the beginning of puberty, the reserve of follicles consists of a pool of primordial follicles beneath the surface of the ovary. This pool is called the baseline reserve. The other reserve, called the dynamic reserve, consists of young tertiary follicles (with a small cavity, fig 28, 29). The oocytes to be ovulated are selected from this reserve. Altogether, at this stage of life, there are around 300,000 oocytes per ovary. Researchers are under the impression that the size of

the baseline reserve can vary considerably from individual to individual. Moreover, this population appears to consist of oocytes that started meiosis later in the foetal period and that formed a follicle at a later stage.

To get an idea of the size of the reservoir over the course of life, ultrasound imaging can provide an impression of the number of follicles that are 1 mm or larger; this is referred to as the "antral follicle count" (fig 28, young tertiary follicles are about 0.4 mm in diameter). There seems to be a fixed relationship between the number of follicles in the baseline reserve and the dynamic reserve. At a young age, many more primordial follicles are mobilised for each ovulation than at an older age. In order to still reach the species-specific number per cycle, the weapon of follicular **atresia** is used more heavily. At an older age, when the baseline reserve has significantly diminished, the species can no longer afford such a heavy selection procedure; there would be nothing left for ovulation (fig 28, 29). The eggs will also be of lower quality due to ageing.

What is involved in this whole selection process is only partly understood. Follicles will compete for follicle stimulating hormone (FSH) from the pituitary gland (fig 29). If more of that is given, more eggs ovulate. Superovulation is a standard part of IVF treatment (Ch16.3). Another technique, "ovum pick up", involves harvesting oocytes surrounded by cumulus cells (fig 29) from the larger tertiary follicles during a normal cycle. This technique is used in the artificial reproduction of cattle and horses. Also, in that case, only one egg would be released at ovulation under normal conditions. With both superovulation and ovum pick up, follicles that may have already been nearing their end are saved. What we know about the variation between these eggs in superovulation and ovum pick up will be discussed in chapter 16.3.2.

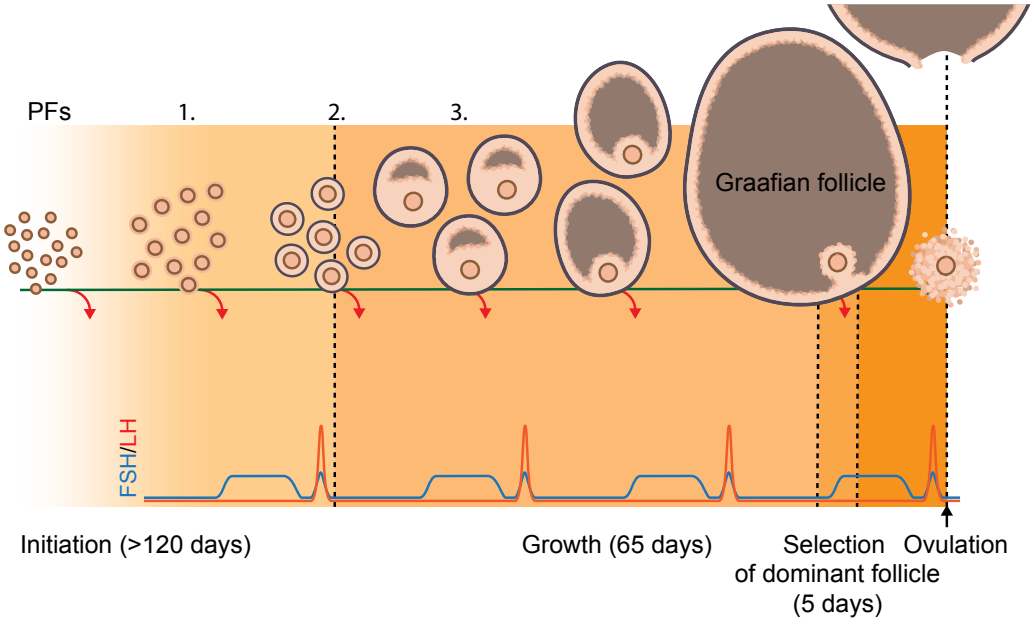


FIG 29

The course of follicle growth and selection (manifesting as follicular **atresia**) in relation to the menstrual cycle. PF primordial follicle. As in figure 28, primary, secondary and tertiary follicles are indicated as 1, 2 and 3, respectively. The oocyte is depicted in a simplistic way. The menstrual cycle is schematically represented by the concentration fluctuations of the **gonadotropic** hormones "follicle stimulating hormone" (FSH) and "luteinising hormone" (LH), the latter acting as "Zeitgeber" for ovulation. The cumulus cells of the mature tertiary follicle which surround the oocyte expand like a cloud before ovulation, which explains the term cumulus cell expansion (see also fig 39).

11.2 Development stages in the oocyte

Since it is difficult to purify oocytes by developmental stage, researchers using mice as experimental animals (which is where most of our insight comes from) have focused on the first surge of follicle growth after birth to study the different stages of development. An oocyte does not need to have completed its development to move from the recently described "frozen" diplotene into the metaphase of the first meiotic division (legend fig 14). As noted above, the granulosa

cells halt the progression of meiosis at the command of the oocyte. When you release oocytes from a young tertiary follicle (fig 28), rescuing them from being blocked in the diplotene, they can rapidly progress to the metaphase of the first meiotic division. When this trick is repeated with a slightly larger tertiary follicle, the first meiotic division of the primary oocyte can be completed, and chromosomes quickly organise into the next division shape, that of the second meiotic metaphase of the secondary oocyte (fig 13, 28, 39, 42). These eggs can be fertilised, but **embryo** development stops at the 2-cell stage (fig 43). In mice, at this stage, "the baton is passed on". This means that the maternal and paternal chromosomes start expressing genes: the first **mRNA** of embryonic origin appears. In the literature, this phase is known as the "**maternal to zygotic**" transition (Ch13.5.3). In humans, this occurs in the 4-8-cell stage. To overcome the blockade, oocytes from larger and more mature follicles are needed.

It should be clear that a lot of protein is needed for the tremendous growth of the oocyte; numerous genes are expressed for this purpose. When the follicle has become tertiary and the oocyte is fully grown, this **transcription** stops and somewhat later, the structure of the nucleus changes in a similar way in mice, pigs, monkeys, cattle and humans. The **constitutive heterochromatin** of the **centromeric** regions (Ch2.3.2, fig 3) starts to surround the **nucleolus** (fig 28). Oocytes that fail to achieve this are eliminated. From then on, the opportunities for embryonic development supported by the egg start to unfold. Since no more genes are transcribed at this point, and even in mice, a few days may pass before the egg will ovulate, mRNA molecules that encode proteins needed to complete the meiotic divisions are kept in a kind of dormant state. This also applies to the mRNA of genes that are essential for initiating embryonic development after the entry of the sperm cell, here collectively called the "maternal effect" or "maternal to zygotic" genes (Ch13.5.3). The regulation necessary to stabilise and temporarily disable this mRNA collection involves RNA binding proteins that enclose RNAs in large complexes called "mitochondria-associated ribonucleoprotein domains" (MARDO). Recently, in 2022, these have been described in more detail. It is beyond the scope of this book to go into the details of this astonishingly complex regulation process. The oocyte stores proteins that are needed at the onset of embryonic development in so-called cytoplasmic lattices. These are structured like periodic filaments, characterised by their big surface. They are especially important for the management of chromatin, harbouring the tools (i.e. the enzymes) that are needed for this. But before that phase, the chromatin of the oocyte also has to be prepared in such a way that it facilitates embryonic development after activation of the secondary oocytes through contact with the sperm (fig 28). In another recent mouse study, a more unifying principle around the maternal

preparation of chromatin for early embryonic development is offered for the first time. The starting point of this research is the regulation of the methylation status of H3 at K36 (table 1). It is essential that this site is demethylated for gene activity to remain silenced. However, methylation of the **promoter** (the CG islands of Ch2.3.3, fig 11), a nearly final shutdown, would be excessive, as gene activity is required after fertilisation. In line with this concept, the overall methylation remains relatively low during **oogenesis** (fig 23). The eye-opening insight is that the enzymes that prevent methylation of H3K36 also prevent the *de novo* CG methylation enzyme from gaining access to the **DNA** (not the DNMT1 of fig 24, but DNMT3A). On top of that, these enzymes also connect with another major regulator of gene activity (for insiders: the Polycomb Repressive Complex 1, PRC1). This level of regulation is installed when the oocyte is fully grown.

11.3 DNA repair in the oocyte

How is the oocyte able to protect itself and its chromosomal DNA from things like environmental damage? How does the oocyte handle oxidative damage induced by the high metabolic-mitochondrial activity required during the enormous growth at the beginning of follicle development (fig 28)? An initial insight comes from research on the effects of **ionising radiation** on the emergence and "survival" of **mutation**s. The aftermath of the nuclear bombing of Japan at the end of World War II, the introduction of nuclear power, and the potential side effects of using imaging with ionising radiation in hospitals are all reasons why a lot of research was devoted to the genetic risks of radiation for offspring, especially in the second half of the last century. For this purpose, large numbers of mice were commonly used as experimental animals. As **cytogenetics** advanced, it became easier to look at individual cells in the **germline**, which reduced the number of experimental animals that were needed to some extent.

Sometime in the autumn of 1971, when I was following an intensive training period at the MRC Radiobiology Unit in Harwell, the renowned researcher Mary Lyon mentioned a stage of oogenesis in mice that appeared to be unreasonably sensitive to radiation. As is common in laboratories where science is a full-time occupation, she brought up this observation over morning coffee. Mary Lyon had seen that radiation damage at this stage caused a severe decline in litter size that did not recover. Later, the same phenomenon was found in other rodents, such as the Chinese hamster, albeit not to the same extreme extent as in mice. The moment of greatest sensitivity apparently coincides with the organisation of primordial follicles around and just after birth. Now, 48 years later, one can speculate that this phenomenon may have something to

do with the defence against retrotransposon integrations, as recently discovered in Baltimore (Ch11.1). However, the mysteries surrounding this temporal hypersensitivity in oogenesis are still far from being clarified.

In the stage of primordial follicle organisation, the whole system of DNA damage detection is on high alert, and any damage is barely tolerated (Ch2.2.1, fig 4). This means that when damage occurs, the oocyte will mostly opt for apoptosis, and the primary oocytes in primordial follicles will continue to carry this ability throughout life. When it comes to the detection of DNA damage and the response to this in the oocyte, an important role in surveillance and regulation is fulfilled by the protein p53, along with the oocyte-exclusive isoform of the protein p63, another member of the p53 family (Ch2.2.1, fig 4, p53 is depicted as the police officer in fig 33). An important step in the elucidation of the **DNA damage response** (DDR) of the primary oocytes in the pool of primordial follicles has recently (in 2023) been made by Ewelina Bolcun-Filas' group at the Jackson laboratory. Further up the reaction chain that leads to the DDR, it was found in Maine that the control of p53 and p63 by "upstream" CHEK2, an important protein specialised in dealing with double-strand DNA breaks, is paramount for the choice of apoptosis (over DNA repair, thereby safeguarding against mutation induction). However, not every form of DNA damage leads to a choice for apoptosis. In the baseline reserve (Ch11.1) of primary oocytes, break repair does occasionally take place. In the field of gynaecology, it has been noticed that women who are carriers of a *BRCA* mutation have a shortened reproductive lifespan. Since *BRCA1* (and *BRCA2*) are involved in the repair of double-strand DNA breaks, a mutation in a single copy of *BRCA1* (in addition to *BRCA2*) is a genetic predisposition for a greatly increased risk of breast cancer. In *BRCA1/Brca1* **heterozygotes** in humans and mice, more double-strand DNA breaks are detected in initially surviving oocytes as age increases.

A certain level of tolerance where apoptosis is avoided is thus present in these oocytes. At the same time, the surveillance by p53 (fig 4) and the oocyte-specific form of p63 is responsible for the disastrous effects of chemotherapy on female fertility.

At the beginning of this century, it was found that the age of menopause (and thus the size of the oocyte population) has a high degree of **heritability** (Ch6.1). With the use of contemporary **GWAS** research (Ch6.2), there is now a high chance of detecting specific gene loci involved in this phenomenon. A very large study utilising this approach was published in *Nature* in the summer of 2021. Among the 290 genomic positions where genetic variation could be related to menopausal age, as many as 35 genes were found to be involved in the DDR (damage detec-

tion, repair and checkpoint control execution, fig 4). Still, we actually know surprisingly little about the DNA repair of primary oocytes in growing follicles.

In growing oocytes, at least in mice, the apoptotic pathway is less strictly executed after the induction of DNA damage. But with all that naturally occurring follicular atresia (Ch11.1, fig 28), wouldn't there be selection against follicles containing an oocyte with DNA damage? In 1974, Mary Lyon published an experiment with one of her colleagues that, although now largely forgotten, contributes to our understanding. Mature ovaries received a substantial amount of ionising radiation (4 **Gy**), and the effects of this were studied in the first week and third week afterwards. You can get an impression of the damage this does to the bivalents by freeing the oocytes from the more mature tertiary follicles and, after meiosis has resumed, observing the condensing bivalents of the first meiotic division under a light microscope. Some oocytes did not survive irradiation, while those that did contained relatively high numbers of chromosomal aberrations both in the first week and after 2-3 weeks. An explanation for this is faulty or failed repair of double-strand DNA breaks. Among other things, you can see, for example, bivalents that are connected to each other due to incorrect DNA repair. The oocyte and follicle do not seem to have a very active policy to counteract these chromosomal aberrations. The oocytes from the irradiated younger follicles had even more abnormalities than those released from the more mature follicles within a week after irradiation. About 10 years earlier, a first hint was already obtained about the oocyte being rather relaxed about the regulation of DNA damage-related checkpoints in the cell cycle shortly before ovulation (Ch2.2.1 and fig 4).

In 1963, a paper was published by Robert Edwards, one of the pioneers (and Nobel Prize winner for this work) of the test tube baby, and my great inspirer Tony Searle from the MRC Radiobiology Unit in Harwell (Ch10). Female mice were injected with a hormone extracted from the blood of pregnant mares (PMSG). This hormone mimics the action of the gonadotropic follicle stimulating hormone FSH (fig 29). After 48 hours, the follicles have grown sufficiently and the mice can be injected with human chorionic **gonadotropin** (HCG) as a next step. HCG is the substitute for the body's own LH (fig 29). Like LH, HCG stimulates the follicles to ovulate after the oocytes have completed the first meiotic division and are arrested in the metaphase of the second meiotic division (fig 28, 39, 42). The entire process takes about 12 hours in mice. This protocol of using gonadotropic hormones to obtain more oocytes in a "timed" manner had already been published by Ruth Fowler and her husband (Robert (Bob) Edwards) in 1957. Edwards' and Searle's research revealed that from the time that meiosis is restarted by HCG, oocytes become more and more sensitive to radiation with increasing chromatin condensation. This

was observed as a peak in embryonic mortality from implantation onwards. At a radiation dose of 2 Gy, only 25% of normally fertilised eggs survived until the end of the embryonic period. In 1991, we repeated Edwards' and Searle's trial in Wageningen and examined the maternal chromosomes at the time of the first cleavage division of the **zygote** (fig 42). We observed that this stage was reached at the normal pace despite the introduction of double-strand DNA breaks by radiation shortly before ovulation. "Repair" of the breaks occurred, but this very often resulted in a structural chromosomal abnormality (Ch3.4.1). As mentioned before, these almost always hinder subsequent embryonic development.

It had previously been discovered in Harwell that the "repair" of these often complicated breaks can already take place in the hours after irradiation, while the bivalents shorten and prepare for the first meiotic division. Although it certainly exists, we now know that the control of the cell cycle (Ch2.2.1, fig 4) during the meiotic divisions of the oocyte is regulated in a very special way. This is something that will be revisited later in the context of **nondisjunction**.

All in all, we understand relatively little about the ways in which the oocyte deals with threats to the integrity of its DNA over the course of its often long lifespan, especially after follicle development has started and checkpoints are less strictly executed. Nowadays, when research on mutations is conducted, it is done using **NGS (WES, WGS)**, the determination of the base sequence of the **genome** (Ch3.2 table 3). Jacob Goldmann, from the Department of Genetics at the Radboud University medical faculty in Nijmegen, was awarded his PhD in 2019 for his thesis entitled "Characterisation of *de novo* mutations in the human germline". The experimental design that was systematically used is known as "trio sequencing". For this, you need DNA from the father, the mother, and one or more of their children. By comparing the children's DNA with the parents' DNA base by base, you can determine the positions where the child has a different base than the parents. Using this approach, you can essentially see the emergence of **SNP**s (Ch3.2 table 3). Similar research is now being conducted in different parts of the world, and the results are all consistent. The findings seem reassuring because the changes at the single base level in the germline are not very large per generation (table 3). The uniqueness of Jacob Goldmann's work was that he had a sufficiently large dataset of trios (1291), which allowed him to estimate the influence of the mother's age (the influence of the father's age on the number of *de novo* base pair changes is covered in Ch12.3 and Ch12.13, see also table 3). In children with slightly older mothers, he noticed something strange. Firstly, more mutations were present than in children with younger mothers. Moreover, his analyses (which are far too complex to present here) indicated that these mutations seemed to have arisen as a byproduct of the repair of a meiotic

double-strand DNA break. This finding is also more widely shared now. A puzzle that remains to be solved concerns the phenomenon that with age, the number of crossovers in the egg that will be ovulated increases, and the fraction of those that is “complicated” also continues to rise. One of the statements accompanying his dissertation addresses our limited understanding of the DNA maintenance of the oocyte and the response to that in the follicle. This is because it remains complicated to conduct DNA-level research on individual oocytes in relation to the physiological status of the follicle.

11.4 Nondisjunction

Nondisjunction refers to the failure of chromosomes to separate properly during the first meiotic division and the failure of **chromatids** during the second meiotic division or during a mitotic division. This leads to abnormalities in the number of chromosomes in the daughter cells, known as **aneuploidy**.

A sort of general rule in gynaecology is that when all recognised pregnancies are considered together, around 15% (10-20%, increasing with the mother’s age) of them end in spontaneous abortion. The higher the number of children, the more likely it is for a family to experience this. In the mid-70s, it was first shown that embryos with an abnormal number of chromosomes strongly contribute to this phenomenon (Ch3.4.2). Currently, it is estimated that around 30% of the total number of spontaneous abortions are caused by meiotic and early embryonic mitotic (Ch16.3.1) errors resulting in **monosomies** (45,X) and trisomies (especially for chromosomes 15, 16, 21, and 22).

11.4.1 Age effects

If anything has stimulated research into the meiotic behaviour of the oocyte, it is the image published by the English researcher Lionel Penrose in the journal *The Lancet* in the mid-50s. This graph illustrates the relationship between a mother’s age and the chance of having a child with Down syndrome (**trisomy** 21), showing a marked increase in that likelihood from the age of 35. Between the ages of around 25 and over 40, that chance increases 15 times. In the 70s and 80s, counting the chromosomes in eggs that were ready to be fertilised (fig 42) clearly revealed that the age effect on nondisjunction (fig 30) also exists in mice, although it is much smaller. This diminished the value of mice as an experimental model for research on nondisjunction in

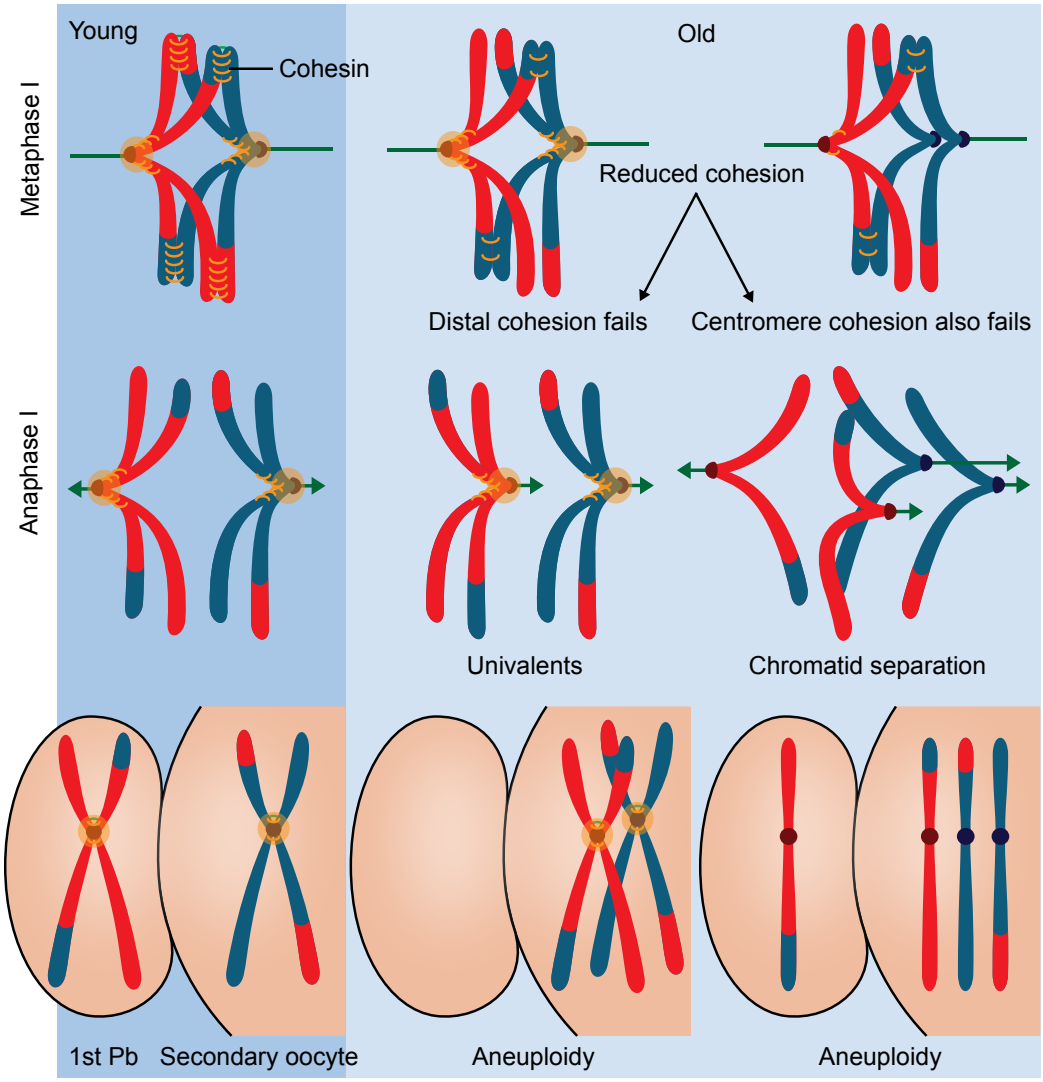

Metaphase I
Young
Cohesin
Old
Reduced cohesion
Distal cohesion fails
Centromere cohesion also fails
Anaphase I
Univalents
Chromatid separation
1st Pb
Secondary oocyte
Aneuploidy
Aneuploidy

FIG 30

An as yet simplistic representation of the mechanics behind the increased risk of **non-disjunction** with ageing of the oocyte in the ovary. Central to this is the deterioration of the functionality of the **cohesin complex** (see fig 6), depicted here as a thin yellow ring. Normally, this complex remains present at the **centromere** during the first meiotic division. Due to reduced maintenance of the cohesin complex, partly as a consequence of ageing, the **bivalents** are at greater risk of either partial or complete premature disintegration (right), and the centromeric regions become susceptible to premature splitting into **chromatids**. Which numerically abnormal **chromosome** complement ends up in the polar body and which in the egg is assumed to be based on chance.

humans. For very understandable reasons, it is difficult to obtain an impression of the age effect on nondisjunction in human embryos around implantation when all chromosomes are included in the analysis. However, this is now possible thanks to artificial reproduction techniques. In 2013, a publication out of the USA described that a trophoblast biopsy was taken from 15,169 consecutive **blastocyst**s (fig 43) in the clinic, which was subsequently analysed for the presence of aneuploidy using SNP technology (Ch3.2). The age of the mothers involved in this analysis ranged from 22 to 45 years. The percentage of aneuploid blastocysts was between 20 and 90%. There is a trough between the ages of 25 and 30, followed by a gradual increase that is almost linear, reaching 90% around the age of 45. It was already known that the increase in nondisjunction can follow a different course over time for different chromosomes; compared to chromosome 21, chromosome 16 becomes a risk factor much earlier.

In 1968, Robert Edwards and his colleague Henderson postulated that the mouse oocytes that are the first to prenatally commence meiosis in the newly formed ovary are also the first to engage in follicle development and ovulation. This became known as the “production line hypothesis”. The scientific significance of this postulate was closely related to the outlined relationship between the mother’s age and the chance of having a child with Down syndrome. The researchers looked at the morphology of the bivalents of the oocyte around the time of separation of the homologous chromosomes, a few hours before ovulation. The technique used for this was already mentioned in chapter 3.4, but it becomes more challenging when dealing with a single cell (or a few cells).

What was immediately visible under the microscope after staining was the big difference between the bivalents of egg cells from young and old females. Those from young females are much longer and more detailed, the **chiasma**ta (fig 14) are easier to count. In contrast, those of older females appear much fuzzier, and the chiasmata at the end of a bivalent are often no longer visible. Sometimes, the homologous chromosomes are situated separately but still opposite each other (fig 30), not resistant to the forces exerted when the oocyte is dispersed on the microscope slide. Henderson and Edwards were the first to observe this, and they hypothesised that this was due to a lower number of crossovers in oocytes that come later in the "production line" of the foetal ovary. Now that we can simply count the number of crossovers under the microscope using MLH1 immunofluorescence (Ch5.2, Ch11.1, fig 15), we know that this does not hold true in humans. Even more intriguing is the recent finding reported earlier that, in our species, as age increases, more crossovers are actually observed in oocytes that are about to ovulate.

11.4.2 Mechanisms

Some clarity, but no more than a hint, emerged in 1991 with a publication by one single author. Ruth Angel worked at the MRC Clinical and Population Cytogenetics Unit next to the Western General Hospital in Edinburgh (Ch3.4.2). She had access to an IVF clinic and had permission to look at the chromosomes of unfertilised eggs. In 38 chromosome portraits of the metaphase of the second meiotic division (fig 13; the chromatids are only attached to each other at the **centromere**), she found four with an abnormal appearance with single chromatids. Could those have originated from the separation of a homologue into chromatids during the first meiotic division? In one of them, that was indeed the case (fig 30).

In 2005, it became clearer what was going on here. In Pat Hunt's group, at that time based in Cleveland, Ohio, genetic engineering (box 3) was used to create a mouse strain with a defect in a gene encoding one of the proteins of the **cohesin complex** (Ch2.2.2, fig 6), the protein complex that holds the two sister chromatids together. They observed a disruption of the bivalents that was strongly dependent on the age of the mutant mice. In the eggs of relatively young 4-month-old females, these bivalents were split into four separate chromatids. But technology has advanced: now, oocytes can be monitored *in vitro* under the microscope around the time of the first meiotic division. This was reported out of Sweden in 2015. The researchers used oocytes from young and old mice and women. In the oocytes from old mice and women over 35 years old, a chiasma is more easily released (fig 30). The **univalent**s that are formed this way

then become susceptible to splitting into chromatids: in fact, this happens one division too early. It is evident that the ageing of the oocyte plays an important role in this.

We used to think that to have, for example, a child with Down syndrome, it was necessary to manoeuvre the entire bivalent of chromosome 21 into the daughter cell (the secondary oocyte), but this is not always the case. The location on the chromosome where crossing over takes place also contributes to the risk of hazardous meiotic behaviour. Sometimes crossing over is absent (chromosome 18), but even a single crossover located close to the **telomere**s, along with ageing, predisposes the chromosomes to nondisjunction (fig 15). This is common for chromosome 16, the chromosome undergoing the most frequent separation failures in female meiosis. Chromosome 21 has all the conditions that increase the likelihood of an incorrect distribution between the oocyte and polar bodies: crossing over close to the centromere is also a factor here.

In chapter 2.2.1, the cell cycle checkpoints G2 > M and the spindle assembly checkpoint (SAC) were discussed (fig 4). Surely, one could assume that these control mechanisms come into play when things go wrong during female meiosis? “State-of-the-art” molecular and cell biological work in recent years, primarily performed in mice, has shown that something peculiar happens in the oocyte. The oocyte is, in fact, able to check whether the centromeres are connected to the **spindle apparatus** but is relatively lax in doing so. It does not respond to a few bivalents that are not orientated correctly in the spindle apparatus of the first meiotic division, even though, judging by the behaviour of proteins involved in SAC, the checkpoint does exist. Moreover, the effectiveness of the SAC is highly age-dependent, as it is much less effective in old mice. Even reasonably high levels of double-strand DNA breaks inflicted on the contracting bivalents do not disrupt the system. In the meantime, while quite a few of the breaks are being repaired, either correctly or incorrectly, the meiotic divisions continue.

From all this research, it appears that the first meiotic division is the main source of origin of abnormalities in the number of chromosomes in the egg. Apparently, the cohesin complex (fig 6, 30) is poorly maintained with age. However, physiological factors, such as hormone levels combined with age, also matter. Something else that also plays a role in human oocytes is the relatively long duration of the formation of the spindle apparatus (16 hours), as well as its instability, even in connection with the centromeres. The communication of the oocyte with the cumulus cells (fig 28) also does not improve as age increases. In 2022, another wonderful overview of this topic was published by Melina Schuh’s group at the Max Planck Institute in Göttingen. It illustrates the various aspects of meiotic nondisjunction in oocytes in a particularly beautiful way.

What stands out is the attention now being given to **oxidative stress** as a factor in irregularities in meiotic divisions. Can a primary oocyte do anything about this?

Another important publication in which this is addressed was also released in 2022. In the primary oocytes of primordial follicles (fig 13, 28), an enzyme complex of the mitochondrial respiratory chain (complex I) is altered in such a way that it remains inactive, thereby reducing the production of oxygen radicals (Ch14).

Ageing is a relative concept, and the rate at which it occurs varies from individual to individual. The stock of primordial follicles at birth is variable, and so is the rate of its depletion until puberty. This results in a fairly large spread in the age of the last ovulation, from 35 to 45 years. It is not clear whether and how an earlier depletion of the oocyte stock affects the likelihood of errors in the first meiotic division. Apparently, there has never been much selection pressure in our evolutionary history to have children at a slightly older age.

Let's briefly return to Robert Edwards and the "production line hypothesis". According to the wonderful review published by Danielle Monniaux of INRA Nouzilly in France in 2014, this phenomenon does exist, but, as mentioned earlier, there is no reduction in the number of crossovers in foetal oocytes with a later onset of meiosis. Hence, what Edwards and Henderson primarily observed was the declining functionality of the cohesin complex in the bivalents, causing the homologues and the sister chromatids to separate too early as the oocytes age (fig 30).

11.4.3 External influences

With a system as fragile as the meiotic divisions of the oocyte, one wonders how easily external substances can affect it. This book is not a toxicological-genetic treatise, but I want to provide two examples as they are so closely related to everyday experiences. One example of a widely used chemical compound is bisphenol A, BPA for short, which is used to harden plastic. Pat Hunt's group discovered the meiotic effect of BPA by accident. In 2003, they described problems with the arrangement of bivalents in the metaphase in mouse oocytes in their laboratory. Usually, this did not involve total chaos, but there were irregularities in which nondisjunction occurred. The mice were kept in old containers made of Makrolon, a hard polycarbonate plastic produced using BPA. During use and cleaning, the containers become slightly damaged, releasing BPA into the environment. This explained the observed meiotic effect, as the irregularities disappeared completely after new containers were purchased. In follow-up experiments, the researchers administered BPA in different doses and looked at the effect of both duration

and degree of exposure. The concentrations at which there was already a noticeable effect corresponded to concentrations that are not uncommon in human settings. The results were published in 2003 in *Current Biology*, a renowned journal. This work attracted special attention because BPA has many applications and because low concentrations of it can be detected in most people. We now know that BPA recognises the receptor protein of **oestrogens** on the cell membrane of granulosa cells (fig 28). This makes it an "**endocrine disruptor**".

Next, the second example. In 1982, Matthew Kaufman (box 2) published that an ovulated egg (fig 42) can be tricked with alcohol. The egg becomes activated and embarks on the path of embryonic development as **parthenogenote**, without the baggage of the sperm cell (Ch10).

Various outcomes are possible; the embryo could become **haploid**, **diploid**, or something in between. Ultimately, this led to another new discovery, that of an effect of alcohol (ethanol) on the correct distribution of chromosomes in the second meiotic division of the egg. Was this due to the way of activating the oocyte (Ch13.5.1), or is it an effect of alcohol on the spindle apparatus? The latter seems more likely. In 1987, the same Pat Hunt published the result of a study on ethanol and nondisjunction during the first meiotic division in male mice. Although this process is not as sensitive as the female meiosis, there was an observable effect, namely occasional abnormal outcomes of meiosis. At the time, this was a sensational discovery. But don't be alarmed: social drinking does not pose this kind of risk. The alcohol levels that these researchers used in their experiments were far higher.

11.5 Oocyte quality

Prenatal mortality, in most cases embryonic mortality, is a normal phenomenon in mammalian reproduction. Livestock breeders viewed it as an undesirable loss while working towards achieving faster gestation in cows and larger litters in pigs. In mammals, embryonic mortality is determined by counting the number of corpora lutea (fig 28) on the ovaries and comparing this with, for example, the number of implantations. Corpora lutea arise from the emptied follicles that remain after the release of eggs and thus provide an indication of the number of eggs released (fig 28). In my MSc thesis from 1971, an estimate of 25% is given for prenatal mortality in gilts (sows having their first litter). The vast majority of eggs are fertilised, and most cases of mortality occur in early gestation. The still unimplanted 8-11 day old embryos, on which I attempted to conduct chromosomal research, had been obtained from sows, aged one to a few years old

and normally no longer used for breeding, that were mated again. The mortality was extremely variable, ranging from under 10% to almost all embryos.

Using the assumption that preimplantation mortality is an extension of oocyte quality, we wrote an internal discussion paper on this topic in Wageningen (at the Agricultural University) in the second half of the 80s. We worked with mice, rats and pigs. Determining the "quality" of each individual oocyte, while they all look more or less the same under the microscope, was a great challenge at the time, and it actually still is. We provided the following indications for the level of prenatal mortality: the rodents mice, rats, hamsters and rabbits have the lowest mortality rate, at 15% or below. In cattle and pigs, we found a mortality rate of 20-40%, and in humans this was 25-50%. Most of this occurs before weeks 8-9 of gestation and goes unnoticed. A concept like the "maternal to zygotic" transition (Ch13.5.3) was hardly known, and molecular techniques to address this were not available in the 80s for material as scarce as an oocyte. In species that typically give birth to a single offspring per cycle, fertilisation does not necessarily lead to offspring. The idea that variation in the maternal to zygotic transition (Ch11.2) is behind this is further discussed in Ch13.5.3 and Ch16.3.2. However, this concept can easily be broadened to the general state of protein and RNA reserves and to the preparatory state of maternal chromatin that enables correct gene expression after fertilisation.

11.6 Oogenesis *in vitro*

In a commentary from 2018, John Eppig from the Jackson Laboratory Bar Harbor, Maine, USA, a well-known and highly successful researcher on many aspects of oogenesis, mentions two reasons why *in vitro* follicle development would be so valuable. (a) It can help women preserve fertility and protect against infertility after chemotherapy, and (b) it could be used to save animal species from the devastation caused by humans on Earth. The *in vitro* application mentioned under (a) will have become less important by now due to new possibilities to freeze the tissue layer beneath the **epithelium** of the ovary, where the primordial follicles are located, and transplant it back after a hopefully successful treatment, if having children is desired. Even in the Netherlands, one can unsuspectingly watch the national news on TV and hear a gynaecologist in a white coat mention that *in vitro* folliculogenesis is imminent for humans. This is another path John Eppig has taken as a trailblazer. The research field in reproductive sciences has always pursued *in vitro* models to enhance the possibilities of studying life processes, and also to reduce the use of experimental animals, now that stem cell techniques and cell differentia-

tion methods are available (box 2). So, what explains the strong drive of a class of researchers towards expanding the *in vitro* trajectory in human reproductive biology? It is probably our medical focus on the individual.

The second application, mentioned under (b), concerns attempts to prevent the extinction of animal species, of which the northern white rhinoceros is a good example. If primordial follicles from the two remaining females of this species could be cultured, and it was then just a matter of waiting before mature oocytes could be harvested, there might have been a high chance of obtaining embryos of both sexes. It remains to be seen whether this could be the case for the two frozen blastocysts that were obtained from the northern white rhinoceros and are now preserved, as has been featured in the media. Also, as many speakers with common sense will point out, trying to preserve the white rhino would only make sense if you can guarantee a safe habitat. In the meantime, publicity that surrounds this ultimate biological experiment to revive the species from an extremely small gene pool is generated as a sign of hope − or is it too little, too late? In chapter 13.7 and chapter 16.7, *in vitro* fertilisation and embryo techniques used for both humans and animals will be revisited, and in chapter 16.7, *in vitro* **gametogenesis** is discussed in more detail.

CHAPTER 12

Spermatogenesis

12.1 Introduction

The study of mammalian **spermatogenesis** has a dedicated community of practitioners. For some of them, this work is even a lifelong addiction. Since the arrival of the **knockout** mouse models in the 1990s (Ch4.1, box 3), the majority of research on the many phenomena associated with spermatogenesis has been conducted in this species. Earlier, endocrinology (the study of hormones) was an important point of focus, while the response of spermatogenesis to **ionising radiation** also received considerable attention, starting in the 1960s. One might also presume that the field of contraceptive research has actively focused on spermatogenesis, but this is not so much the case. Articles with relevance to this application do mention temporary blocking of spermatogenesis, but it generally remains silent after that. Recently, the use of the plant compound triptonide has been mentioned in work from the USA/China in this field (more in this in Ch12.6).

Spermatogenesis can easily be divided into three sub-processes: 1) the **mitotic** multiplication of the **diploid** spermatogonia originating from the spermatogonial stem cells, 2) the transition to the **meiotic** cell phase via a final mitotic division; the spermatogonia have now become spermatocytes, and finally 3) the **haploid** phase after the meiotic divisions, in which the spermatids transform into spermatozoa through a process called **spermiogenesis** (fig 31). When the final stage of that process has been reached, the cells detach from the wall of the seminiferous tubules (in fact, the cell membrane of the Sertoli cells) and enter the lumen, a process referred to as **spermiation**. Then, the transport of the sperm cells to and through the epididymis begins. Once that long journey has been completed, they reach the vas deferens (fig 31). During

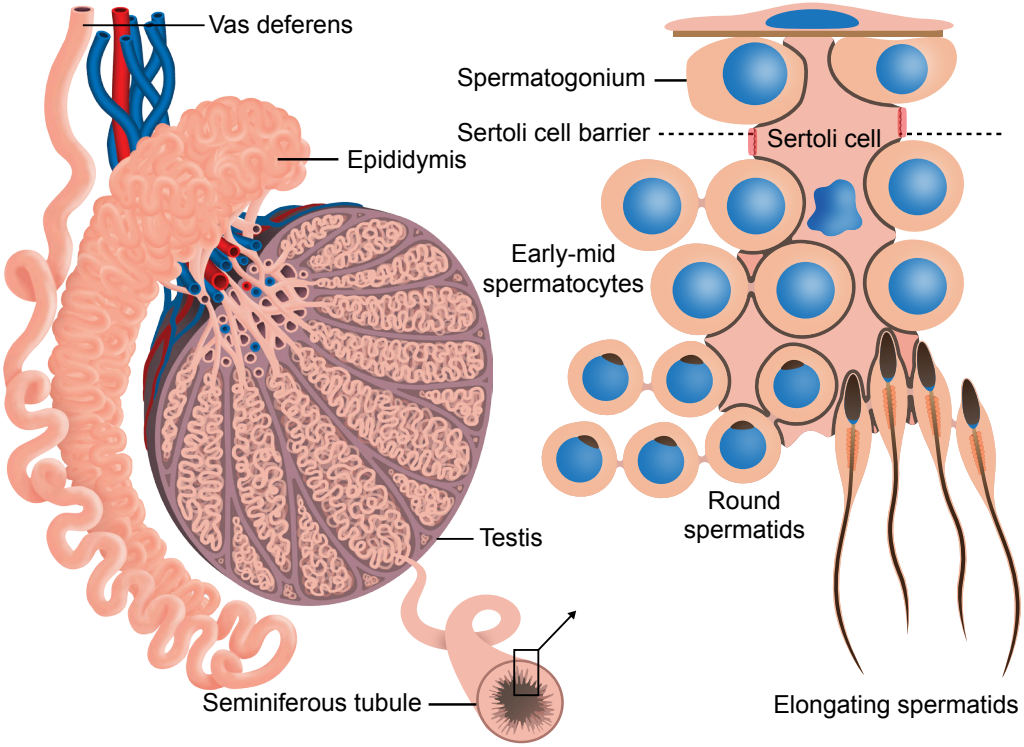

Vas deferens
Epididymis
Testis
Seminiferous tubule
Spermatogonium
Sertoli cell barrier
Sertoli cell
Early-mid spermatocytes
Round spermatids
Elongating spermatids

FIG 31

Schematic representation of the testis with an emphasis on the seminiferous tubules, the drainage to the epididymis and the transition from the epididymis to the vas deferens. The blood circulation is also depicted (inflow red, outflow blue).

The process of **spermatogenesis** takes place from the wall of the seminiferous tubule, where the stem cells of spermatogenesis are located. The comprehensive role of the Sertoli cell is highlighted. Spermatids whose nucleus becomes oval are known in the literature as elongating spermatids. After a division, the spermatogenic cells hold on to each other using a **cytoplasm**ic bridge. The Sertoli cells that do not divide do this with the help of a dense connection between the cell membranes (in red). Mammalian spermatogenesis is divided into stages based on the development of the **acrosome**. Certain morphologies of the acrosome (from a vesicle to a cap on a round nucleus to largely covering an elongating nucleus, shown here in black) are associated with specific spermatocyte developmental stages and specific spermatogonial differentiation stages.

These fixed combinations are known as the stages of spermatogenesis and all have a certain duration. Naming these stages using microscopic analysis is a specialised task in any species, but especially in humans. The experts will recognise stage IV of humans in this figure. The term **spermiation** is used for the release of the spermatids from the **epithelium**, which can now be called spermatozoa or sperm cells. The efferent ducts from the testis to the epididymis are known as the rete testis (see also fig 22).

ejaculation, male **gametes** are mainly recruited from the vasa deferentia. This involves the innervation of smooth muscle cells in the wall of the vas deferens. From the moment that stem cells form differentiating spermatogonia, which begin their journey of multiplication and development, the entire process is strictly time-regulated. The time course of the different phases varies between species. When conducting research on spermatogenesis, you need some sort of timetable for this. In mice, the trajectory from stem cell to sperm cell that is released into the testis is estimated to take 34.5 days, and in humans, this is 74 days.

12.2 Evaluating spermatogenesis

The stem cells of spermatogenesis are, of course, the source around which everything revolves. When they are not sufficiently protected, and/or when their daughter cells too frequently choose the path of differentiation, the capacity for sperm production declines faster over the course of life.

In humans, the quest to identify the one and only spermatogonial stem cells has not yet been completed with full collegial agreement at the time of writing this, but the picture is becoming clearer. Before continuing, let me first introduce testis **histology**, the widely used approach to gain a better understanding of spermatogenesis through the analysis of tissue sections (fig 31). When you look at a seminiferous tubule, you can see that spermatogenesis generally takes place in the direction from the periphery towards the centre. Spermatogonial stem cells are situated along the wall, in relatively small numbers in mice but in large numbers in humans. These cells, like cells at any stage of spermatogenesis, need support from the Sertoli cells; they are fully dependent on them (fig 31). Once the Sertoli cells (which we encountered earlier in the decision of sex differentiation towards male during **embryo**nic development, Ch8) are formed during development, they no longer divide. They cannot function without the signal of the steroid hormone testosterone, of which the supplier is fortunately nearby. Testosterone is produced by the so-called Leydig cells, located in the tissue between the seminiferous tubules. However it may be, spermatogenesis is entrusted to the good care of the Sertoli cells throughout the entire process, from the stem cell stage to spermiation. How directing is a Sertoli cell really? First of all, Sertoli cells are known as the suppliers of vitamin A, which is necessary for the differentiation of spermatogonia and also as a signalling molecule for the initiation of **meiosis**. So these are, so to speak, phased assignments, which change per single location as the stages of spermatogenesis progress over time (fig 31). If you were to be seated on the edge of a tubule, you would have to wait 8.5 days in mice (the cycle time) before you would see the same composition of cell types of spermatogenesis again as “a snapshot”. In mice, we have divided this ongoing, continuously repeated display into 12 stages based on microscopic screening of histological sections (fig 31). The shape of the developing **acrosome** is leading in this classification system. In this way, a diagram of spermatogenesis in mice was created in 1956, which is forever associated with the name of Eugene Oakberg of the Oak Ridge National Laboratory in Tennessee. Evaluating spermatogenesis based on this diagram does require practice. However, an experienced researcher can then provide a quantitative estimate of any problems, if there are any.

To determine the entire cycle duration (i.e. the 8.5 days mentioned just before), the researchers injected the **DNA** building block thymidine labelled with radioactive hydrogen into the abdominal cavity. At desired times thereafter, mice were sacrificed and histological sections were made, which were then covered with photographic film. The progeny of the cells, which had incorporated the radioactive thymidine (represented by the T in fig 5) shortly after injection during DNA duplication for cell division, were covered with silver grains after development of the film, which makes them clearly distinguishable. As time passes, you can observe the appearance of this signal in increasingly further-developed spermatogenetic cells. Based on this, it is possible to estimate the entire cycle duration. For the human version of this analytical method, testis biopsies from American prisoners were studied in the early 60s. This way, the cycle length of spermatogenesis in humans has been determined to be 16 days. While histological research on the course of spermatogenesis generally started smoothly, it proved to be a much more laborious undertaking in humans. Instead of the 12 stages observed in mice, only six were apparent in humans for many years. The reason for the difficulty of studying human spermatogenesis lies in the fundamental difference between mice and humans in the strategy to achieve a very large production of sperm cells. In mice, there are relatively few spermatogonial stem cells and many spermatogonial divisions before meiosis begins. In humans, the opposite occurs: the number of stem cells is large, and there are relatively few mitotic divisions. In mice, there are seven to 10 divisions, so each stem cell that advances along the path of differentiation into a sperm cell produces a large family of up to 1,000 spermatocytes if **apoptosis** does not occur. This number is small in humans, as mentioned in the most detailed overview published in 2017: a family of spermatocytes counts a mere 8-16 members. As a result, you can distinguish up to three different stages in a histological section of a single cross-sectioned seminiferous tubule in humans. The analysis is then reserved for those biologists who are truly addicted to human spermatogenesis, and there are very few of them. Guided by Dutch researcher Dirk de Rooij, human spermatogenesis has also been categorised into 12 stages as of 2013, using the development of the acrosome of spermatids as a reference point (fig 31). As indicated before, this analytical methodology is too specialised for clinical research on impaired spermatogenesis. The most widely used and relatively simple evaluation system is the “Johnson score”, named after a Danish pathologist who simply assessed the presence of spermatogonia, spermatocytes, round spermatids, and elongated spermatids (fig 31) in 40 histological sections per testis biopsy. When spermatogenesis synchronously halts at a certain stage, known as a “maturation arrest” (Ch12.12), it can be identified in this way.

12.3 Spermatogonial multiplication

As noted earlier, there is a big difference between humans and mice in the way spermatogonial multiplication occurs. In rodents, the stem cells, of which mice have about 35,000 per testis, divide about two to three times during each cycle of the **epithelium**. Humans have a much larger population of stem cells, which must have been established starting from the formation of the testis until puberty. This extensive supply leads to a high occupancy of stem cells along the wall of the tubules. They are not optimally used for differentiation. For each stem cell, this happens less than once per cycle duration of 16 days. The intellectual father of this model from 2017 is the aforementioned emeritus professor Dirk de Rooij, who conducted this research in collaboration with a great laboratory team in Rome. There, they discovered that only 15% of undifferentiated spermatogonia, which can now be classified as stem cell candidates, are in an active cell cycle. Even with this knowledge, it is still impossible to predict how many sperm cells that will eventually yield. Initial ideas about the continuation of spermatogonial multiplication, based on work in mice, were, on reflection, quite simple. It was believed that when an isolated living stem cell divides and the two daughter cells remain connected by a **cytoplasm**ic bridge, a clone of differentiating spermatogonia is formed. When this does not happen, two new stem cells are created. However, a slightly less systematic representation of the process is now preferred. An important question in this regard is to what extent a differentiating stem cell can still backtrack. In any case, the size of the stem cell population must be monitored. When the products of spermatogonial divisions remain connected by cytoplasmic bridges, the cells move along the path of further differentiation towards the formation of spermatocytes. Cytoplasmic bridges are also formed during the meiotic divisions of spermatocytes (fig 31).

Spermatids, which are haploid in the nucleus, share their cytoplasm with many other spermatids, in principle allowing them to still have access to all products of the initial **genotype**. Physiologically, they still mostly behave like diploid cells (fig 13). This mechanism may play a role in any potentially preventable consequences of the separation of the X and Y **chromosome**s. Moreover, the entire system facilitates the synchronous development of the **germ cell**s and prevents a phenotypic effect at the haploid level: it is an axiom (that is not frequently expressed) that the genotype of a spermatid has no effect on the functionality of the sperm cell that develops from it. Fortunately, technology has now advanced to the point where **transcript**s can be recognised for each individual spermatid. When the genetic difference between the **homologous** chromosomes is large enough, the transcripts can be classified by origin per **homologue**, and

you would expect to find them in equal amounts in each spermatid. Research published in 2021 largely confirmed this expectation. However, a different picture was found for about 13% of the genes and their products. This group was enriched in genes with a role during spermiogenesis and can thus lead to an effect of the genotype of the individual sperm on the likelihood of fertilisation. The assumption that follows from Mendel's first law may then no longer apply. In the scientific literature on the subject, this phenomenon (a deviation from a 1:1 split) is referred to as "transmission rate distortion". This subtle way of still being able to distinguish between **allele**s expressed during spermiogenesis seems to offer an evolutionary advantage for male fertility.

Understanding spermatogonial multiplication is not only relevant for our knowledge of sperm production. It is also important for other aspects. For example, we know that chemotherapy for cancer treatment is often harmful to human spermatogenesis and that recovery from this damage is not a given. Often, sperm production does not properly resume afterwards (Ch12.7). It is also known that the number of divisions, in this case, DNA duplications, that occurs between generations plays a role in determining the amount of new **mutation**s for a new generation. DNA **replication** is inherently not error-free (Ch2.2.2, table 3, fig 5), and this is also true during spermatogonial multiplication (but see the caption of table 3). This is one reason why children of older fathers are found to have considerably more mutations, in this case, single-base mutations (*de novo* **SNP**s, Ch3.2), than children of younger fathers. The number roughly doubles every 20 years as a man ages.

In addition to inaccuracies in DNA replication, other factors come into play: a change from a C to a T (fig 11) and a C that is replaced by a G (together accounting for a quarter of *de novo* SNPs) are events that are not initially linked to DNA replication, but attributed to the environmental influences (see also Ch12.13 and the caption of table 3).

12.4 The Sertoli cell barrier

The comprehensive role of Sertoli cells in spermatogenesis has already been discussed above (Ch12.2). Sertoli cells are connected to their neighbouring Sertoli cells via a special circuit between the cell membranes. Together, they form a contiguous layer in the seminiferous tubule, the "blood-testis or Sertoli cell barrier" (fig 31). One possible function of this barrier is to keep out harmful substances that could damage the DNA of the spermatogenic cells. In addition, the barrier also helps prevent autoimmune reactions. The interfaces between Sertoli cells are very extensive. When a spermatogenetic cell, which has just started meiosis, has to pass through

the Sertoli cell barrier, the process resembles a passage through a narrow lock with very thick gates. Sometime in the early 70s, I saw those “lock gates” in a large electron microscopy photo on a poster displayed at a Dutch natural science meeting in Utrecht. The poster was made by Mebius Kramer, Dirk de Rooij’s teacher. Kramer did research on testicular material from bulls. I felt struck by a sensation of “this is actually true”, not yet realising at the time that all those spermatogenetic cells are attached to each other in shorter or longer chains and that they will really not let go of each other when they are pushed into the lock and come out again.

12.5 Male aspects of meiosis

Male meiosis differs quantitatively and also qualitatively from female meiosis. The processing of double-strand DNA breaks, which are involved in the search for the homologous chromosome and **crossing over** (Ch5.1), occurs in a more uniform manner than in female meiosis (Ch11.1). The number of crossovers is also lower (Ch11.1), and their distribution along the chromosome is subtly different. In males, crossovers are found closer to the chromosome ends. Consequently, the loci between which **recombination** has actually occurred are also further apart.

Nondisjunction leading to an extra chromosome is rare in the first meiotic division in mice, with an occurrence of around 1%. This number is barely higher for the second meiotic division. In humans, the **cytogenetics** of spermatogenesis has always been a laborious occupation, with few researchers venturing into this. Counting chromosomes during the second meiotic division was difficult. The first estimates made in populations of sperm cells were based on **fluorescence in situ hybridisation** with chromosome-specific pieces of search DNA (“probes”). This led to an estimate of around 0.1% **disomic** spermatozoa for each chromosome, approximately 2.5% in total. A magnificent paper published in the summer of 2020 presented a method that actually leads to an accurate estimate. At the Department of Genetics at Harvard Medical School, a DNA technique was developed to determine the sites of crossovers per individual sperm cell based on 1% of the **genome** and also to read out which chromosomes are present in duplicate and which are missing. In this study, more than 31,000 sperm cells from 20 young donors between the ages of 18 and 38 were examined in this way. It was found that only 0.7% of all sperm were disomic and that slightly more (1.8%) lacked a chromosome. The second meiotic division was the most error-prone. Abnormalities in the distribution of **sex chromosomes** are the most interesting (table 4), and those appear to occur somewhat more frequently during the first meiotic division. In previous work with chromosome-specific fluorescent probes, an age effect for nondisjunction

in men was never found. However, it is true that when spermatogenesis is worse, increased frequencies (about two to four times higher) of nondisjunction are observed.

12.5.1 "Sex bodies"

The major difference between female and male meiosis is the presence of two different sex chromosomes in the latter. We know almost precisely down to the DNA base which part (or parts) of the X and Y chromosomes are perfectly homologous. These are the DNA regions that are necessary to enable crossing over (fig 15, 38). Due to the **chiasma** that results from this, the X and the Y chromosomes are connected to each other until the end of the metaphase of the first meiotic division (legends fig 13, 14 and 30). They separate in the anaphase, thereby ensuring a sex ratio of one in the offspring. In humans, two homologous "**pseudoautosomal region**s" (PAR1 and PAR2) are involved in the positioning of the crossovers. The longer one (PAR1) is about 2.6 mb in length and is located at the end of the short arms of the X and Y chromosomes (fig 37, 38). In this region, there is always one crossover. The other region (PAR2) is much shorter (fig 38), is located at the end of the long arms of the X and Y chromosomes, and is much less frequently involved in a crossover (only in about 25% of cases). In mice, the PAR is somewhat longer. There is no place where the crossover incidence per mb of DNA is higher than in the male PAR. These relatively small pieces of homologous X and Y DNA search for each other along the nuclear membrane, which can be considered a master test for homologous meiotic pairing (Ch5.1). It occurs slightly later than the pairing of the **autosomes**. Once the PARs have found each other, the sex chromosomes immediately form a **chromatin** domain that then starts to behave very differently than the chromatin of the other **bivalents**. **Gene** expression ceases, and the region that is commonly referred to as the "**sex body**" (fig 15, 37) becomes filled with proteins that protect the DNA, for example by being involved in its repair. In chapter 5.1, we discussed the large amount of double-strand DNA breaks that are inserted into DNA at the beginning of meiosis to contribute to the "homology search" and to enable crossing over (fig 14). This also occurs in the sex chromosomes, and these breaks need to be repaired as well. For the majority, this can be done using the DNA of the sister **chromatid**.

Possibly due to the fact that the PARs are not long, small differences in their base order that have appeared over the course of evolution seem to quickly affect the meiotic pairing of the sex chromosomes. In the late 70s, we received some Japanese house mice (*Mus musculus molossinus*) from Professor Alfred Gropp in Lübeck. When we crossed them with our albino Swiss

laboratory mice, we encountered fertility issues which, among other things, were related to the fact that the Japanese Y chromosome would not always cross over with the X chromosome of the Swiss mice, causing them to be univalents. This resulted in reduced sperm counts in the epididymis (fig 31).

This is an example of the onset of **hybrid** sterility. The X and Y chromosomes act as some sort of gatekeepers to keep the species "pure". The fact that hybrid sterility first affects the sex with two different sex chromosomes was formulated as early as 1922 by the famous English population geneticist JBS Haldane (box 4). This is one of those old rules in genetics that has stood the test of time. Just search the internet for the Liger, the hybrid offspring of a male lion and a female tiger. My barber from Kurdistan brought this up once, and he went on to say: "Have you still not looked it up yet?". So Haldane's rule works, among other things, via the meiotic behaviour of the X and Y chromosomes due to the decreased homology between the PAR regions. Could that have played a role in the outcome of the encounters between *Homo sapiens* and *Homo neanderthalensis*? There are indications that intercourse between modern and older hominids would not have been beneficial for the fertility of the offspring. Generally, there are no or only very few traces of the Neanderthal genome left in regions richer in **gene**s involved in male (and also female) meiosis. The contribution to the X chromosome is also lower than that to the autosomes. Of genes related to meiosis, alleles of Neanderthal origin appear to have disappeared through natural selection. Remarkably, the Y chromosome is never mentioned in reflections on the genetic signature that these sexual encounters have left in us. It may well be that the legacy of Neanderthal DNA (Ch5.4.3) was initially established through, or mainly through, the female **F1**s.

12.5.2 Chromatin changes and gene expression during meiosis

Slightly more than 2,000 genes are involved in spermatogenesis, as they have a greatly increased **RNA** expression compared to other organs. It is estimated that at least 500 of these genes are exclusively involved in this process. During the pachytene stage (fig 14), the chromatin changes, testis-specific **histone** types emerge (which also need to be incorporated in the **nucleosome**s, Ch2.3.1), and the nuclei become larger and looser. **Transcription** (box 1) increases sharply and stays high after the meiotic divisions. This was already discovered in the 60s by incorporating radioactive uridine into RNA, followed by visualisation using **autoradiography** and light microscopy. Far from all **mRNA** molecules are translated into protein. Now, with the application of **next**

generation sequencing (NGS) to the **transcriptome**, the amazement about the level and variation of RNA production that is presented here only increases (consider, for example, the use of **exons**, box 1). Retired reproductive biologist Anton Grootegoed from Rotterdam sometimes compared it to a “test run”. The engines for RNA transcription and RNA processing are cranked up once more before the path to differentiation into sperm cells is initiated after meiosis. The biological significance of this exuberant transcription is not fully understood yet.

Of all the chromatin remodelling that occurs during the male pachytene, that of the sex chromatin in the sex body (fig 15, 37) is the most profound. Depending on the species, the nucleosomes are more or less rigorously broken down and reformed again, this time with a different H3 variant (fig 8). This process is part of the transcriptional inactivation of the sex chromosomes within the “sex bodies”. After meiosis, when they have ended up in separate spermatids, the repression of gonosomal gene activity is somewhat less strictly regulated, but the sex chromosomes remain largely inactive.

While it doesn’t seamlessly fit into this section, we must also pay attention to the following in this chapter: the expression of small RNA molecules that do not encode proteins, the so-called small non-coding (snc)RNAs. Three subclasses of these are miRNA (19-23 **nucleotide**s), siRNA (idem) and piRNA (26-32 nucleotides). Interest in the biological function of these RNAs is growing, and this has yielded entirely new insights for cell biologists and geneticists, among others. This is certainly true for their role in **gametogenesis**; it is now assumed that the importance of sncRNAs in spermatogenesis is much greater than in **oogenesis**. In fact, piRNAs are expressed almost exclusively during spermatogenesis. During the formation of the testis and thereafter, they particularly suppress the activity of **retrotransposons**. This is expressed in the form of CpG methylation (fig 11). During meiosis, the piRNA-mediated silencing of retrotransposons also plays a role. However, sncRNAs do more than that. Among other things, they have a function in the regulation of (hetero)chromatin and in clearing excess RNA later in spermiogenesis. The miRNAs and siRNAs are known to be specifically involved in controlling the expression level and **translation** of protein-coding mRNAs during spermatogenesis. More and more reviews are appearing about their great biological importance and about the clinical expression(s) that abnormalities in populations of sncRNAs may cause. As such, knowledge of sncRNA also significantly contributes to the growing understanding of the incredible complexity of spermatogenesis.

12.5.3 Checkpoints in meiosis

In mice, two distinct checkpoints (fig 4) during the first meiotic division have been identified: the pachytene checkpoint and the metaphase I checkpoint. The pachytene checkpoint was first observed in 1983 in a mouse from a radiation experiment conducted at the MRC Radiobiology Unit (Harwell). In this mouse, half of chromosome 7 had ended up in chromosome 1, and large unpaired segments of chromosome 7 were observed in every pachytene spermatocyte. The female meiosis could still handle this in some way, but the male meiosis apparently could not. Halfway through the pachytene stage, just before the time of chromatin change and increasing transcription (Ch12.5.2), the spermatocytes collectively underwent apoptosis and were "eaten" by the Sertoli cells. This turned out to be the precursor of a **phenotype** that was observed more frequently later, when knockout mice (with defects in single genes, box 3) were introduced. When the invalidated gene has an essential function in homologous chromosome pairing, in this case, the repair of double-strand DNA breaks, the pachytene arrest is activated. The arrest was, in fact, already addressed in chapter 5.1, in the description of the role of **mismatch repair** genes in the formation of **synaptonemal complex**es and, hence, bivalents. It is also characteristic that the gene activity of the X and Y chromosomes in the "sex bodies" (Ch12.5.1) cannot really be reduced. This alone would be sufficient to trigger the pachytene arrest in mice, which is also effective in humans. I have always been surprised that the collective clearance of the entire production of meiotic cells at the pachytene checkpoint does not lead to a chaotic histological image. If you could find a way to interfere with chromosome pairing, you would essentially have an ideal contraceptive for men. Ideas about this were dropped at the former Organon company (the Netherlands) in the late 90s. To further develop plans in this direction, an *in vitro* test for the functioning of a candidate protein involved in chromosome pairing had to be available. Obviously, that protein should only be present in meiotic cells. The search for an inhibitor for such a protein could then be done by testing compounds from a whole library of chemicals. However, the right *in vitro* test did not exist back then. In hindsight, the time was not yet ripe for an approach like this.

A second checkpoint during spermatogenesis occurs in the metaphase of the first meiotic division. This checkpoint was already hinted at before (Ch12.5.1), in the story about the disrupted pairing behaviour of the X and Y chromosomes and the reduced sperm counts in hybrids between the Japanese house mouse and Swiss laboratory mice. Separately located X and Y chromosomes will activate the SAC, the spindle assembly checkpoint (Ch2.2.1 Ch11.4.2, fig 4).

As in primary oocytes, the presence of univalents predisposes to an incorrect distribution of chromosomes among secondary spermatocytes (compare with fig 30). The SAC cannot completely prevent the formation of spermatozoa with numerical abnormalities (Ch12.5), including aberrations regarding the sex chromosomes. The source of the XXY **karyotype** (Klinefelter men, table 4) can be found in both female (two X chromosomes) and male meiosis (X plus Y chromosomes).

Something resembling these checkpoints, but with a different origin, is related to the process of shielding the sex chromosomes in the sex body. In both mice and humans, disrupting this mechanism can seriously affect fertility, with an example of this provided in figure 37. In the presence of a structural chromosomal aberration (Ch3.4.1, fig 12, 37), the search of the involved homologous chromosome segments for each other often proceeds less smoothly. Segments that have difficulties with this alignment appear to be “attracted” to the X and Y chromosomes, which slightly lag behind in their search for homologous pairing. The unpaired chromosomal segments simply join the X and Y, leading to the loss of the strict separation between sex chromatin and **autosomal** chromatin. As a result, unpaired autosomal chromatin is also directed towards inactivation, and conversely, sex chromatin is activated. All of this has been most thoroughly studied in mice. The consequences, which become more severe as the number of sex bodies “infected” with autosomal chromatin increases, can, in extreme cases, be evident as **azoospermia**. Often, the homogeneity of the otherwise immaculate-looking mouse sperm is also lost. A wide variety of shapes appears, especially in the head of the sperm. The tail is also affected (fig 37), and motility decreases. In short, the mouse sperm starts to closely resemble human sperm (Ch12.9.1, table 5, fig 34, 37).

12.6 Spermiogenesis

Once the second meiotic division has ended, shortly after the first meiotic division, spermiogenesis begins: from the morphogenesis of the round spermatids until the release of the sperm cells into the lumen (fig 31) of the seminiferous tubules. The formation of the acrosome and the tail starts shortly after meiosis has been completed (fig 31, 32). The cells are active in transcription and protein synthesis. Many things need to happen and this requires extensive regulation. Part of the gene activity required for this process has already been initiated during meiosis. The mRNA molecules can be stored in an unused state for a long time before actual translation into protein takes place (box 1, fig 9). This principle is also useful later on; a large amount of protein

is needed when the DNA is packed into the head of the future sperm cell, but the transcription of mRNAs for this purpose has already ceased during the elongation of the nucleus.

Spermiogenesis can be easily followed by looking at the formation of the **acrosome** (Ch12.2, fig 31, 32). It starts as a vesicle that gradually sinks over the nucleus and makes contact with it, comparable to a knitted egg warmer covering a boiled egg. Upon the first contact of the vesicle with the nucleus, the chromatin within the nucleus initiates its transformation into a much more compact form (fig 32). This is something we accidentally discovered around 2010, when we were studying human spermatogenesis in Nijmegen, the Netherlands. The chromatin remodelling proceeds from the front of the nucleus towards the tail. We imagine that the chromatin is stored like rope on rolls, finding support from a sort of skeleton at regular intervals (fig 32). To transition from the **nucleosomal** chromatin structure (fig 7, 8) to a much denser composition, a substantial number of double-strand DNA breaks has to be made. This likely serves the purpose of preventing various pesky knots in that endlessly long DNA. There are not many researchers who study these intriguing DNA breaks. One might ask, for instance, whether these breaks are randomly introduced or if there is a site-specific mechanism behind this (which seems to be the case). An additional question is how these breaks are repaired, because there are not many left once the chromatin is unpacked in the egg (Ch13.5.2). The small protein protamine, rich in the **amino acid** cysteine and therefore containing many SH groups, plays the main role in condensing and preparing the chromatin for transport (fig 32). However, not all DNA is involved in this transformation. In mice, around 1% of the DNA remains in the more open nucleosomal structure, and in humans, that is around 10%. A subsequent process that more or less seals the condensation of chromatin begins in the testis and is completed in the epididymis (fig 31). By forming disulfide bridges (-S-S-) between the SH groups of the cysteines, adjacent protamine molecules become interlinked (fig 32), tightly stabilising the bound chromatin DNA. There is unanimous agreement that this provides the best protection for the chromatin until it can be received by an egg. Moreover, a compactly structured nucleus with a minimal amount of cytoplasm is beneficial for the streamlined shape and swimming speed of the sperm cell.

12.7 DNA repair, damaging agents

A first impression of the radiation sensitivity of young spermatogonial stem cells was documented in 1973 by Paul Selby of the National Laboratory in Oak Ridge, Tennessee, at that time a giant mouse laboratory. Selby needed no fewer than 55,456 mice for his research, animals

born after their fathers had been irradiated earlier, around the time of their birth. The young spermatogonia are very sensitive to DNA damage and usually choose to undergo apoptosis. This occurs despite high levels of proteins involved in DNA repair. By mating these wildtype (+/+) males that had been irradiated as baby mice with females that are **homozygous recessive** for seven genes with an observable phenotype (-/-), you could easily see the radiation-induced damaging mutations (often deletions). This manifests as the appearance of the recessive phenotype in offspring which would normally be wildtype (+/-) like the father. However, mutational damage in these seven genes occurs sporadically after exposure to radiation, which explains the need for the large numbers of mice in this study. Sometimes, multiple offspring with the same recessive phenotype for the same gene were born: a mutated spermatogonial stem cell had permanently settled in the epithelium of the seminiferous tubule and could then generate multiple offspring. This so-called "specific **locus** test" was perfectly accepted at that time. The few laboratories in the world conducting such experiments used a high number of mice, a practice that is difficult to maintain nowadays due to improved attention to the usefulness and necessity of animal experiments. The spermatogenesis precursor cells are susceptible to cell death, which, together with the effective action of DNA repair proteins, may indicate an evolutionarily evolved strategy to control the load of new mutations as generations pass.

Even when spermatogenesis is ongoing, DNA-damaging agents and ionising radiation have a negative effect on stem cells. In mice, research in this area, piggybacking on the large-scale radiation studies of the 1950s-1980s, was mainly carried out by Tony Searle and Bruce Cattanach at the MRC Radiobiology Unit in Harwell UK. Bruce Cattanach wanted to determine whether fertility could recover when increasing doses – of up to 10 **Gy**, which cause males to become completely azoospermic – were used. As it turned out, a surprisingly simple mathematical principle underlay this: the time that elapsed until fertility returned was directly proportional to the dose of radiation that was received. It is, of course, remarkable that a dose of 10 Gy, which causes about 320 double-strand DNA breaks per G0, G1 stage (fig 4) **somatic** cell nucleus *in vitro*, is survived at all by a small fraction of the stem cells. The conclusion drawn from this observation was that after damage to the spermatogenetic epithelium in mice, repair of that epithelium occurs first, and spermatogenesis resumes afterwards. In primates, including humans, the regulation seems to be less strict. The development towards sperm cells actually starts too quickly, and this comes at the expense of the stem cell reserve and the eventual sperm production (Ch12.3).

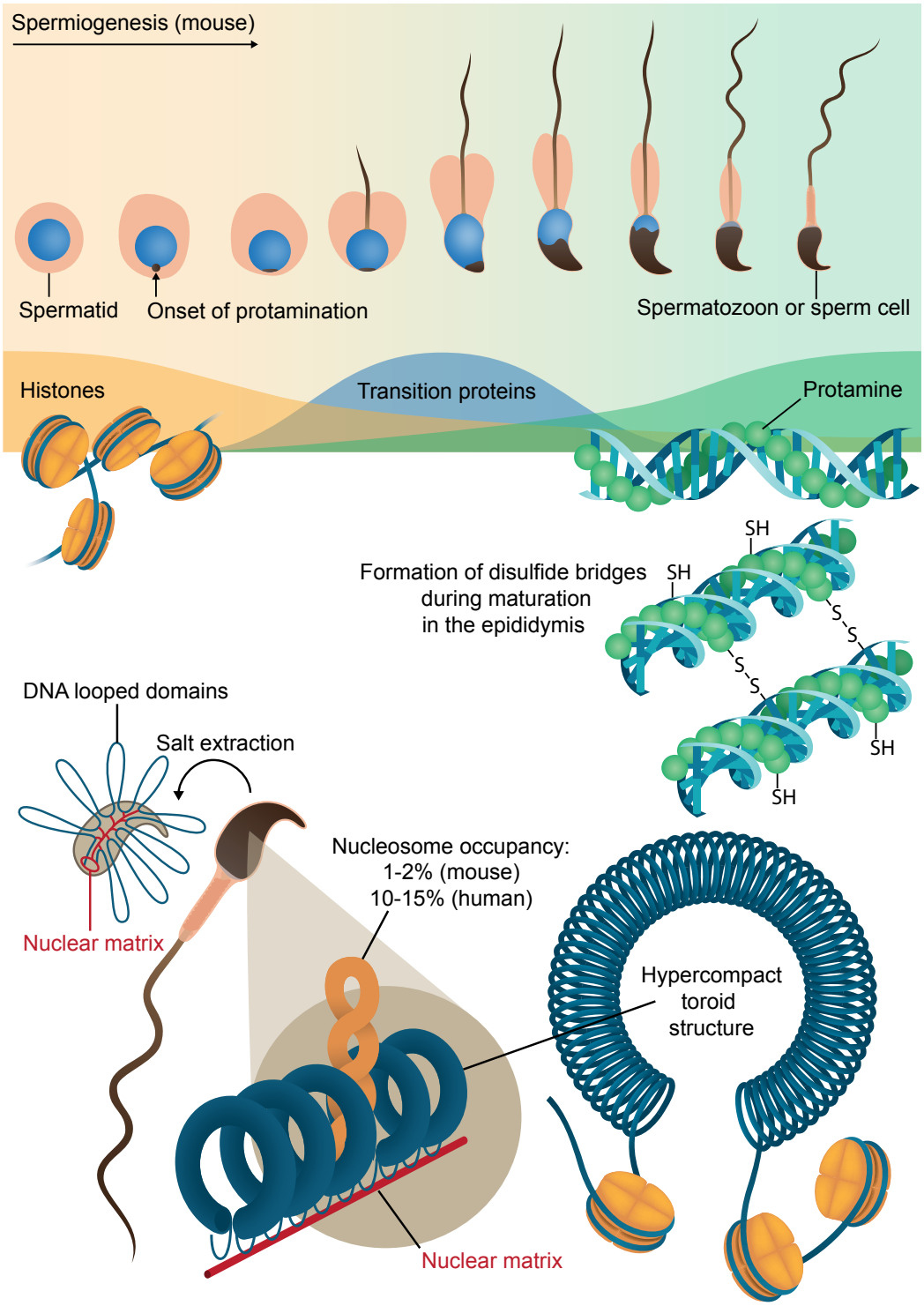

Spermiogenesis (mouse)
Spermatid
Onset of protamination
Spermatozoon or sperm cell
Histones
Transition proteins
Protamine
Formation of disulfide bridges during maturation in the epididymis
SH
S
DNA looped domains
Salt extraction
Nuclear matrix
Nucleosome occupancy:
1-2% (mouse)
10-15% (human)
Hypercompact toroid structure
Nuclear matrix

FIG 32

The widely accepted model of the **chromatin** changes that occur during **spermiogenesis**. These changes take place against the background of a shape change of the nucleus, from round to elongated (fig 31). The transition from the **nucleosomal** structure to the compact protamine structure proceeds via the intermediate phase of transition proteins, about which not much is known yet. The nucleus establishes an eccentric position in the **cytoplasm**, which eventually descends over the tail and largely disappears. The formation of the tail already starts at the beginning of spermiogenesis, just like the development of the **acrosome**, shown in black here. The looped domains (Ch2.3.1, fig 7) form a remarkable element in the structure of the nucleus. They attach to the nuclear matrix (see also Ch2.3.1) and give the nucleus a high degree of stiffness. Investigating the chromatin organisation of the sperm cell is extremely difficult. While the egg can "unpack" the sperm without damage (naturally, after fertilisation), it is almost impossible for researchers to do so.

The very early radiation experiments of the 60s and 70s, conducted in mice, already suggested that stem cell spermatogonia produce few structural chromosomal aberrations per double-strand DNA break (Ch3.4.1). Reciprocal translocations are by far the most common of this type of aberrations. They arise during an extreme attempt by the cell to repair those life-threatening breaks, which involves using another chromosome as a last resort (fig 12).

The "state-of-the-art" way to observe the occurrence of reciprocal translocations was to wait until the progeny of the stem cells reached the metaphase of the first meiotic division (legends fig 13, 14). There, you would see four chromosomes attached to each other instead of the normal situation where two chromosomes have formed a bivalent (fig 26 depicts this situation in the pachytene (fig 15)). In the early days of research on the mutational sensitivity of spermatogonial stem cells to radiation, the entire field was heavily dominated by the search for standards for radiation exposure, for workers in the nuclear industry, for example, but also for men and women in the general population (in which case stricter standards apply). Extrapolating the results of mouse experiments to the human situation was a tricky issue in this process. However, recruiting inmates from American prisons as test subjects provided a solution, with little attention to ethical considerations surrounding this type of research at the time. The findings, coming out of the aforementioned Oak Ridge National Laboratory, were published in 1975 in the prestigious

journal *Nature*. Also in men, irradiation and repair of DNA damage in stem cells led to reciprocal translocations. At a radiation dose of 2 Gy, they observed some more of these translocations than in mice, while there were relatively fewer at a dose of 6 Gy. Apparently, human stem cells are more sensitive to radiation and those with extensive DNA damage tend to die earlier than in mice.

Despite the fact that years have passed and significant advancements in DNA techniques have taken place, we still have relatively little specific knowledge about DNA repair in the spermatogonial stem cells in mice (but see the caption of table 3). In any case, the process occurs differently than in somatic cells. When there is damage, apoptosis is a normal way out, but this is not the uniformly chosen solution. That is because there is also a group of radiation-resistant stem cells. These may be cells that are in a favourable stage of the cell cycle. This explains why fertility in male mice still recovers after some time following a higher dose of radiation. By then, they will certainly no longer be free of chromosomal abnormalities. Studies in mice from the 70s and 80s clearly showed that the irradiated males could transmit reciprocal translocations to their offspring up to an old age, although with decreasing efficiency. Since it is unclear how many new mutations stem cells can tolerate, this has clear clinical implications for humans. It will always be better, and also ethically preferable, to try to "secure" future fertility for men undergoing cancer therapy with DNA-damaging agents by freezing sperm in advance. Do not rely on the chance that spermatogenesis will slowly recover. The additional risk that sperm cells will carry new mutations as a result of chemotherapy and/or radiation is simply too high, and the evidence for this has been provided very recently. This came from a large study that searched for higher spontaneous **mutation frequencies** for base pair changes in 21,879 trios ("trio sequencing", Ch3). This phenomenon, rare on its own, occurred after paternal chemotherapy before conception (questioning the "spontaneity" of the mutations). For further discussions on the effects of ionising radiation on offspring in humans, see also chapter 17.3.

The more rapidly dividing differentiating spermatogonia are even more sensitive to radiation than the stem cells; they massively undergo apoptosis. Tony Searle and his assistant Colin Beechey found a profound dip in mouse sperm production 5 weeks after an acute dose of 2 Gy (a typical daily dose of radiation in tumour therapy). This dip (down to 10% of the normal level) did not prevent the fertilisation of virtually all eggs after mating with females.

Remarkably, apoptosis of the differentiating spermatogonia is also employed as a mechanism to adjust their number to the supporting capacity of Sertoli cells. It is unknown whether

spermatogonia with mutations are particularly sensitive to this and if they are the first to be selectively eliminated. When additional double-strand DNA breaks are introduced in primary spermatocytes during the long lead-up to the first meiotic division, these are “repaired”. The repair capacity is, however, more limited. When damage is inflicted that affects the double helix more subtly, the type of damage caused by, for example, ultraviolet radiation but also by some chemicals, the outlook is even more disheartening. Hardly anything is done. One possible explanation is that the NER process involved in this (fig 33) cannot be active if other mechanisms involved in normal double-strand DNA break repair are active at the same time (during meiosis for crossing over events). In 2015, the leading expert in this complex field, Francesco Marchetti from Ottawa, Ontario, Canada, published that in mice, DNA damage in the male **germline** caused by the chemotherapeutic agent Melphalan (after exposure during the long meiotic prophase) is only repaired in the **zygote**, resulting in structural chromosomal aberrations. Melphalan is a so-called bifunctional alkylating agent.

The situation becomes even more complex after meiosis, during spermiogenesis. If something is wrong with a chromosome at that time, which is then present in a single form and consists of just one chromatid, there is no copy available to retrieve the original information. Although some repair of double-strand DNA breaks still occurs in round spermatids, as spermiogenesis progresses, it appears to become increasingly attractive to leave things as they are and await potential repair in the egg (Ch13.5.2). However, much of the knowledge presented here is quite speculative, and our understanding of DNA repair in the haploid part of the germline is still very limited, especially for humans. Across the entire germline, next generation sequencing (NGS) slowly changes this at the level of the individual base, as mutation patterns can be linked to the errors that may occur during DNA repair (this way of thinking and investigating stems from genetic research on cancer). Exogenous and endogenous effects, such as the unwanted side effects of replication and transcription, can be visualised this way. The effects of unwanted demethylation of mCpG (fig 11) are also included in the analysis. Moreover, this method of analysis provides a connection to damage repair in the zygote (Ch13.5.2). The large and expanding databases of gnomAD and TOPMed (Ch3.2) are thereby fundamental to this method of analysis, which seeks to link the mutation patterns to the underlying processes involved in maintaining DNA integrity; it is a specialised task.

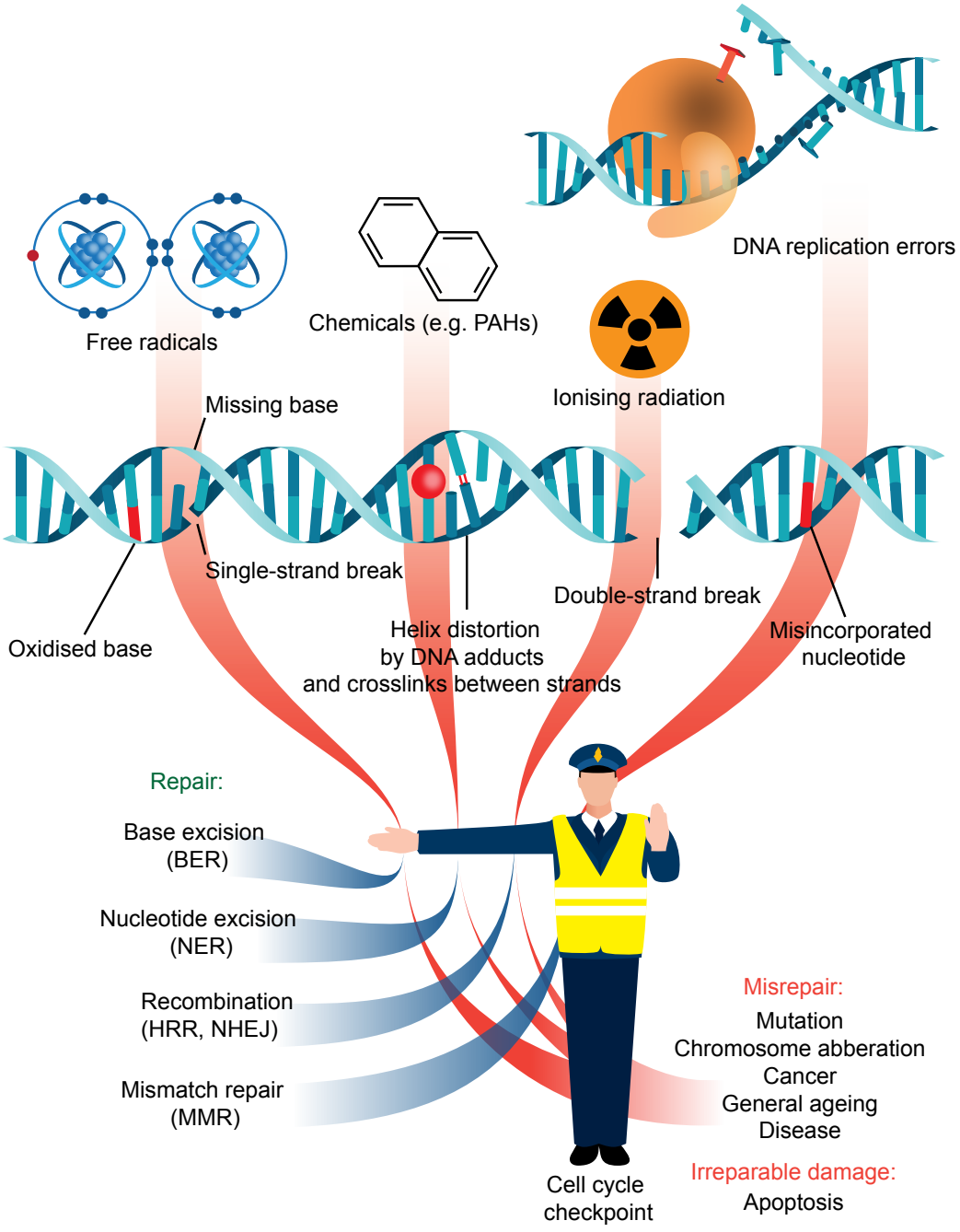

DNA replication errors
Free radicals
Chemicals (e.g. PAHs)
Ionising radiation
Missing base
Single-strand break
Oxidised base
Helix distortion
by DNA adducts
and crosslinks between strands
Double-strand break
Misincorporated
nucleotide
Repair:
Base excision
(BER)
Nucleotide excision
(NER)
Recombination
(HRR, NHEJ)
Mismatch repair
(MMR)
Misrepair:
Mutation
Chromosome abberation
Cancer
General ageing
Disease
Irreparable damage:
Apoptosis
Cell cycle
checkpoint

FIG 33

Highly simplified classical representation of the relationship between **DNA** damage and the classification of DNA repair pathways. They are not as strictly separated from each other as presented here.

Repair can be error-free or it can create errors (**mutation**s). The spectrum of errors made during repair is certainly larger than the eventual **mutation spectrum**, as some errors will be lethal for the cell. When the damage is irreparable, the “police officer” (see also fig 4, the protein p53 is the best known for this function) will more frequently take action and initiate the pathway to **apoptosis**.

Published estimates of DNA damage per nucleus per day provide a good impression of the importance of DNA repair. There are 10 to 50 double-strand breaks and 10 to 50,000 other damages. The loss of an amine group on cytosine (100-500 times per day) is significant: when this cytosine is methylated (fig 11, and that is usually the case), a C changes to a T (the most frequent base change). Without DNA repair, life is not possible. Spontaneous mutations are at the root of **somatic** disease in general, especially cancer and diseases linked to ageing. As depicted on the left, 8-oxoguanine is often used as a marker for oxidative damage due to **oxidative stress** to a base. For abbreviations, refer to the list at the end.

12.8 Numbers of sperm cells, the production

The criteria for assessing sperm production in the ejaculate are initially simple: volume, sperm concentration, total number of sperm cells, motility, and morphology of the head and tail are determined (table 5, fig 34). If you have experience with such tests in mice, rats or farm animals, the different picture in humans immediately stands out. In human sperm, we observe great variation in the appearance and motility of the cells. This coincides with a relatively low sperm production. An overview from 2009, made by Rupert Amann from Colorado, presents a table showing sperm production per gram of testis in humans compared to a whole range of mammals, including rhesus monkeys, bulls, rats, rabbits and hamsters. Humans rank last. Even men with the highest levels still score three times lower than rats, rabbits and hamsters.

In a sperm analysis, most of the attention is still focused on the cell count. The first time I heard something about this was at a meeting of the Flemish Dutch Fertility Study Association

in the 70s. This association, with a strong clinical orientation, still exists. In Belgium, Frank Comhaire, then in Ghent, talked about sperm production. He presented a slide showing a man followed over part of his life. The line usually showed a fairly even course of sperm production, but at times, there could also be large temporary deviations, i.e. significant dips.

There is consensus that (a) the best measure is sperm production per hour, and (b) a single sampling is not sufficient. Regarding the latter, the interval between two consecutive measurements can make a big difference. During ejaculation, the last part of the epididymis (fig 31, 35) is emptied in addition to the content of the vas deferens. Therefore, multiple samples, preferably three, taken at intervals of 42-54 hours will provide the best impression. Variation will, however, still be found. There are different reasons for this. Firstly, the residence of sperm in the epididymis can be quite variable and can last longer when production is lower (fig 35). In mice, you can count on the transport taking about a week; in humans, we see that it can take anywhere from less than two days to as many as 14 days for sperm to pass through the approximately 5-metre-long tube that makes up the epididymis. In chimpanzees, a great ape species with a different reproductive strategy, this process is always short due to high sperm production.

In research and diagnostic settings, sperm samples are usually produced in non-physiological conditions. Regardless of the best efforts of the interior designer of the hospital and the visual stimuli that are present, these conditions are not a natural setting. The literature does not indicate that stimulation by viewing pornography boosts sperm production. Some older research has found that when a measurement is taken from an ejaculate in a condom, used in a situation for which it is intended, more sperm cells are counted, despite the fact that a small fraction sticks to the rubber − it is still science. In bovine artificial insemination (AI), arousal does matter in the relatively sterile environment of the sperm collection centre. A bull needs to mount the artificial cow a few times before ejaculation occurs. Also, the animal should not ejaculate more frequently than two to three times a week.

To a significant extent, reproduction remains a matter of behaviour, as it involves the contraction of involuntary muscles during arousal. Behaviour is variable, and so is the behaviourally determined variation in sperm production, something that is utilised in bovine AI by “stimulating” the bull, as indicated above. There are also genetic mouse models that illustrate this aspect. An article on this subject from the year 2000 addressed a knockout mouse for a gene encoding a protein involved in the transmission of nerve stimuli to the smooth muscle cells in the wall of the efferent duct of the epididymis (the vas deferens, fig 31, 35). These mice had normal testes and

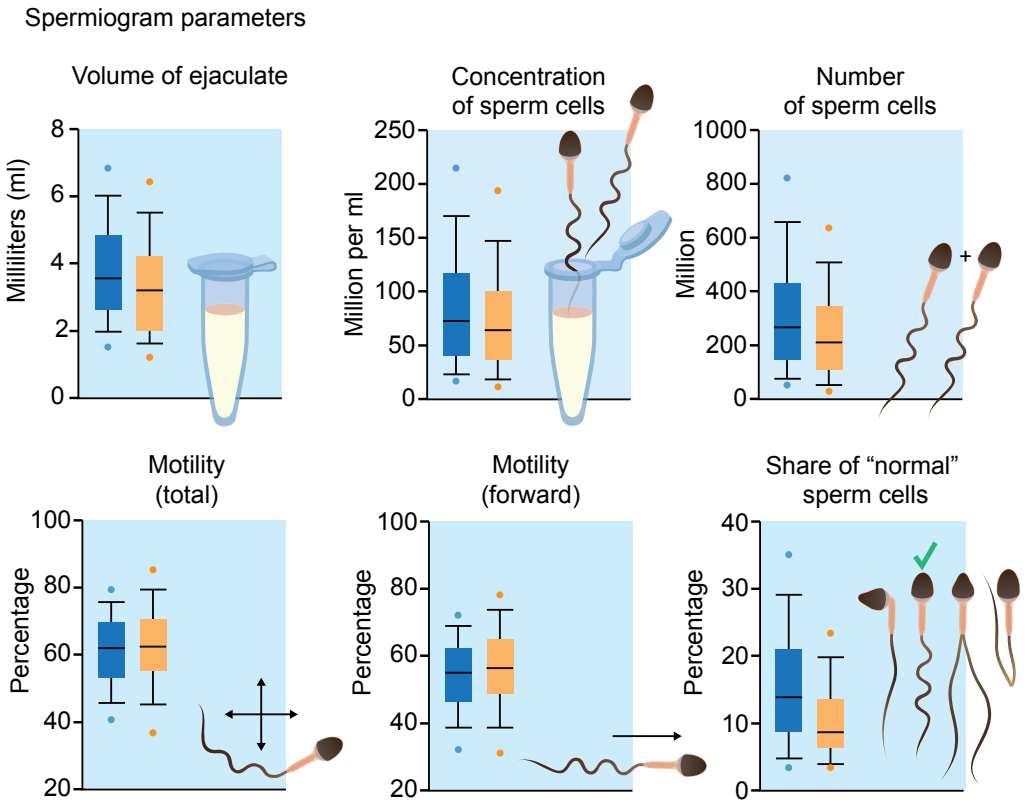


FIG 34

Assessment criteria of a sperm sample, illustrated by the box plots from a WHO publication in the well-known journal *Human Reproduction Update* 2010. The coloured area comprises the percentages 25-75, in an ascending order of values. The vertical line delimits the values from 10-90%, and the values from 5-95% lie between the dots. The horizontal line in the box represents the middle, the median value. Orange: sample from the entire population. Blue: men involved in a pregnancy within 12 months. Forward motility is also described as progressive motility.

copulated without any issues, but they did not impregnate their partners. Not surprisingly, this result made it into the renowned journal *Nature*.

When the number of sperm cells in the ejaculate is counted, the literature usually reports a concentration. Figure 34 provides an idea of the wide variation encountered in men. If you look at the key parameter, the total number of sperm cells, you will find a peak between 100 and 200 x 10^6. By comparison, a billion sperm cells per ml is nothing special for a bull.

12.9 The spermiogram

The results of a semen analysis conducted according to the criteria of the World Health Organisation (WHO) are presented in a **spermiogram** (table 5, fig 34). A general rule is that when the number of sperm cells declines, the entire spermiogram will get worse. Motility decreases, the number of morphologically abnormal sperm increases, as well as the number of dead sperm cells. This phenomenon is so common that it has been given a name as a syndrome: OAT, Oligoasthenoteratozoospermia. OAT is very frequently found in men referred for a semen analysis due to fertility problems (table 5, fig 36). Oligo means few, astheno indicates motionless, and terato stands for deformed. These three parameters can be encountered in various combinations and degrees. Motility is not a black-and-white concept; it is subject to a lot of variation. When there is no movement, there is a high probability of cell death. The expression “lazy sperm” is sometimes used in everyday speech, which refers to cases where lower sperm motility is seen throughout the entire ejaculate. With the major goal of estimating the probability of fertilisation, the spermiogram has been and continues to be the focus of many discussions and publications. Table 5 provides the cut-off values determined for a fertile ejaculate, determined for each parameter, but sperm cells do not readily reveal their secrets. Figure 34 provides a graphical representation of the main spermiogram parameters for both the entire population and for men who were able to impregnate their female partners within 12 months. The large variation in all assessment criteria is evident. This is precisely the variation that is observed in a more pronounced degree in the fertility lab. What happens physiologically with combinations of low values, which are commonly found (fig 36)? Over the past 20 years, the criteria for normal morphology (appearance) have been tightened once again. Now, the term “strict criteria” is used; at least 4% of the sperm cells must meet this stringent rule (table 5). One does wonder about the biological significance of all this. After all, the morphology of the winning sperm is never observed in the natural situation and

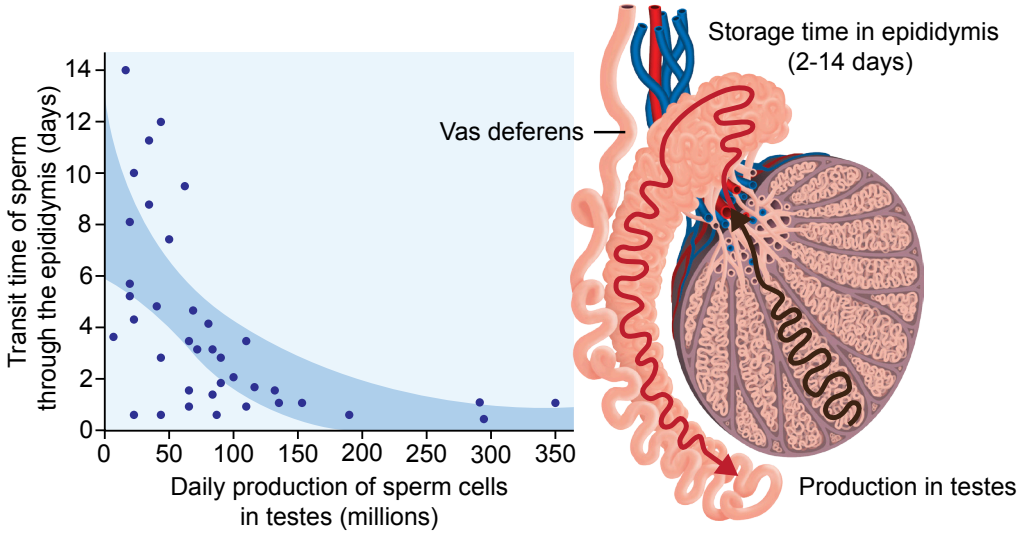


FIG 35

A valuable graph from Rupert Amann's review article from 2009. The relationship between daily sperm production and passage time through the epididymis is clear. With lower production, the variation in passage time increases significantly, thereby increasing the risk of **DNA** damage due to, for example, **oxidative stress**. On the right, the global anatomy of the testis, the epididymis (the red trajectory), and the beginning of the vas deferens. The efferent ducts from the testis to the epididymis are known as the rete testis (see also fig 22).

thus remains unknown. Meanwhile, the IVF field has reported pregnancies following the use of sperm that did not meet the strict criteria. When a TESE-ICSI (Ch16) requires a considerable search for a sperm cell in a testis biopsy, it is unlikely that the prescribed criteria are precisely followed. It is already nice when you can observe some motility, and therefore, determine that the sperm cell is alive. All in all, the results of a semen analysis thus have a limited predictive value for the chance of a spontaneous pregnancy over a prolonged period of time, with the exception of cases involving a complete absence of motility, very "lazy sperm", a consistently very low sperm count, and/or a uniformly occurring morphological abnormality.

12.9.1 Heterogeneity

Usually, heterogeneity increases when the number of sperm cells decreases. In Nijmegen (the Netherlands), at the beginning of this century, we also studied the variation in sperm in men with OAT. This kind of research is actually notoriously challenging. Although the annual number of publications about human sperm is high, progress in our understanding has lagged behind for years. The main reason for this is the compact morphology of the nucleus (Ch12.6, fig 32). However, there are ways to facilitate this research, despite the fact that bias may creep in.

The use of the anticoagulant heparin helps to "relax" the nucleus, even when the cell adheres to a slide. This makes it possible to use **immunofluorescence** microscopy to obtain an impression of the amount of nucleosomal chromatin (fig 32); is it often more than you would expect. In addition, you can make the nuclei light up by using a small fluorescent molecule that can chemically react with the SH group of the amino acid cysteine in the protamine proteins (Ch12.6, fig 32). This way, the change of the chromatin to a more compact structure can be tracked.

In this context, it is actually better to see a lower level of fluorescence. Our study revealed that in OAT patients, nearly 50% of the spermatids in the testis had not completed the protamine-directed transition of chromatin into a more compact structure. However, we also discovered, completely unexpectedly, that something very different occurs in the epididymis of OAT men. The crosslinking of the SH groups of protamine proteins for further stabilisation (Ch12.6, fig 32) primarily takes place there. This is an oxidative process that can also be overly activated, which is what we observed. During the variable stay of sperm in the epididymis, chromatin must be carefully protected against **oxidative stress** caused by **oxygen free radicals** (which actually affects the cell membrane first). This protection is apparently less effective when fewer sperm cells are present. On top of that, there are a couple of sources of cellular turmoil. When formation of a sperm cell is not completed in the testis, there is often a small cytoplasmic droplet attached to its tail (fig 32, 34). This also causes oxidative stress, as does the presence of white blood cells in the prospective ejaculate. Could sperm cells consequently fall victim to a self-selected death, to apoptosis? The consensus is that they can indeed, with a sperm cell-specific variant of this important process. There is a staining method to visualise this, which exploits the fact that the cell cuts up its own DNA during apoptosis. Apparently, the enzyme responsible for this can nibble at the chromatin from the nuclear membrane. Unsurprisingly, the fluorescent staining shows many more positive nuclei in an OAT ejaculate, up to about half of all cells, compared to an ejaculate of "normal" men, where only 10% of cells will be stained.

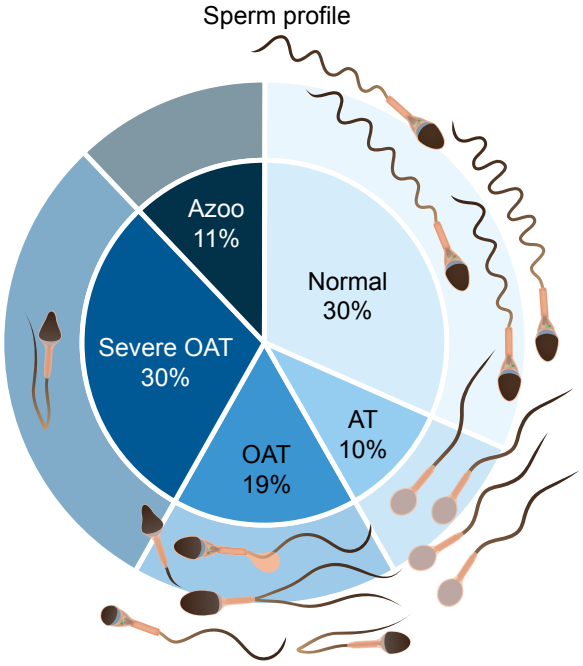


FIG 36

A pie chart of semen analyses of 26,091 men referred to the CeRA in Münster, representing cases where the cause of a couple's fertility problems is not yet known. The O in OAT stands for oligo (few), the A for astheno (no movement) and the T for terato (abnormal shape). AT is characterised by motility and shape problems. In **azoospermia** (Azoo) cases, no sperm cells are found. This condition can be caused by an obstruction in the system of efferent ducts (up to 40% of cases), which is commonly located in the epididymis and most often not congenital, but acquired later in life. However, in most cases (60% or more), azoospermia is related to the state of **spermatogenesis**, with a genetic defect being the most likely cause. The contribution of azoospermia to male infertility is estimated to be 10-15% (see also table 6). This table provides an overview of the current state of knowledge (as of 2021) regarding the contribution of genetics to the explanation of male infertility and subfertility. Among the OAT men, there are also individuals who would have been fertile with a different partner.

Table 5 The cut-off values of the **spermiogram** or semen analysis (with the 95% confidence interval in parentheses) for reaching pregnancy within one year of unprotected intercourse.

Ejaculate volume (ml)	1.4 (1.3-1.5)
Total sperm count (total x 10^6)	39 (35-40)
Sperm concentration (sperm cells x 10^6 per (ml)	16 (15-18)
Total motility (%) (progressive and other)	42 (40-43)
Progressive motility (%)	30 (29-31)
Immotile sperm cells (%)	20 (19-20)
Vitality, live sperm cells (%)	54 (50-56)
Morphology, normal shape (%)	4 (3.9-4.0)

In diagnoses of asthenozoospermia (insufficient motility), the value falls below the cut-off listed in the table, which is also the case for abnormal morphology (teratozoospermia). When numbers are lower, the parameters motility and morphology generally decrease as well. This underlies the commonly diagnosed condition OAT (Oligo Astheno Teratozoospermia, see also fig 36 and table 6).
In 1999, the WHO defined the condition OAT as having a sperm concentration of more than 5 and less than 20 million sperm cells per ml. In cases of severe **oligospermia**, the concentration is lower than 5 million/ml. Some laboratories also mention extreme oligospermia (less than 100,000 sperm cells per ejaculate).

For a long time, practitioners in fertility clinics have been trying to separate sperm cells with a more compact nucleus from those with a loose packaging. This can be done by placing the cells in a tube, using centrifugal force in a centrifuge spinning at high speed, and pushing them through an increasing density gradient of Percoll. The most compact ones will reach the bottom of the tube and can be harvested from there.

Considering that patience is a virtue, that stress increases during extended periods of infertility, and that we live in a world where many things are only a mouse click away, the field of fertility research and counselling is currently witnessing the parallel emergence of many new but also very different trends. For instance, on the free market, devices are offered that can be connected to a mobile phone, which allow individuals to gain insight into sperm production and

quality. On the other hand, a select group of internationally operating researchers is still very actively working within their own sheltered scientific community to unravel the mystery of sperm cell heterogeneity in an attempt to find better predictors of fertility.

12.9.2 DNA methylation in sperm cells

As gonocytes differentiate into spermatogonial stem cells (Ch8), the later stem cells of spermatogenesis, the CpG methylation level of DNA begins to rise. Eventually, this leads to a situation where the DNA in sperm cells is heavily methylated (fig 11, 23). Determining the methylation status (the **methylome**) of DNA in a population of sperm cells could provide information about (a) the history of gene regulation in spermatogenesis and (b) the heterogeneity of sperm generation in the testis. The effects of foreign substances on the methylation status during differentiation are sometimes also considered in this kind of research (Ch17). Jörg Gromoll's group at the CERA (Centrum für Reproductionsmedizin und Andrologie) in Münster, Germany, focused on the fact that sperm cells have "methylation heterogeneity" for selections of CpG positions. They link this to male infertility and to the "**mosaic**" nature of spermatogenesis. This could mean that deviations from the normal methylation pattern can already be established in the stem cells, thereby impacting the reproductive status of the descendants of these stem cells for their entire lives. Douglas Carrell's group from the Andrology Department at the University of Utah, USA, took a different approach and investigated the methylation patterns in sperm cells from men with a normal spermiogram and men who had sought assistance from the fertility lab to become fathers. Both the time needed to conceive with their partners and the "success" in the fertility lab were found to be related to the methylation pattern of sperm DNA. For this, methylation was examined at a number of select positions in the genome.

It was previously known that the methylation status of an ICR (see Ch10 **genomic imprinting**, fig 27) is sensitive to instability during spermatogenesis. For a paternally imprinted ICR, there is a greater chance of incomplete methylation, while a maternally imprinted ICR is more likely to be methylated where it should not be. This thought-provoking observation does not have a clinical application.

Characterisation of the methylome of the sperm cell will probably continue to be of interest to researchers. It can improve our understanding of embryonic development and the influence of the environment on the germline. At this point, how far are we from the moment when a

technique like this can serve as a diagnostic and prognostic indicator of success in assisted reproductive technologies?

12.9.3 The X and Y chromosomes and poor quality sperm

It was mentioned in chapter 12.5.3 that the inactivity of transcription of the sex chromosomes in the sex body during the first meiotic prophase seems to be essential for sperm quantity and quality. In both mice and humans, there is a large-scale breakdown of nucleosomes at this stage, and they are reassembled in a slightly different composition (Ch12.5.2). This occurs considerably less drastically in humans than in mice. When the separation between the sex body and autosomal chromatin blurs (fig 37), and consequently, transcription in the sex body increases, sperm cell production and quality decrease (Ch12.5.3). Therefore, there is a greater chance of encountering Robertsonian and reciprocal translocations (fig 12) in subfertile and infertile men (Ch12.12.1, table 6). Could variation in the manner and degree of inactivation of the sex body be related to the large heterogeneity of the sperm population in humans?

At the Department of Developmental Biology at the Erasmus Medical Centre in Rotterdam, research led by Anton Grootegoed addressed this topic with a molecular approach. In chimpanzees, the selection for a large sperm production is attributed to the mating system; sperm cells from different males in a community compete on their way to the egg that is to be fertilised. This competition has led to large testes, large ejaculates, and perhaps even better performance per sperm cell. In chimpanzees, there is also much less variation in the shape of the sperm head.

The possibility that sperm performance is related to the high or low expression of a select group of genes prompted Anton and his colleagues to examine the expression of genes on the sex chromosomes in testes of humans and apes. What they found was that the downregulation of transcription during spermatogenesis occurred with slightly more control in chimpanzees than in humans. In this regard, you would then also expect less variation in downregulation of gene expression of the sex chromosomes among chimpanzee spermatocytes compared to those of humans, and consequently, more homogeneity in ejaculates. Could the sex chromosomes be involved in this after all?

12.9.4 DNA damage in sperm cells

The DNA of sperm cells can incur damage. This can occur when the DNA is being packaged during the protamine-controlled restructuring of the nucleus of the sperm cell (Ch12.6, fig 32), but more likely during the stay in the epididymis. This is especially true if that stay is prolonged, as is the case when sperm production is reduced (Ch12.8, fig 35). Oxidative stress caused by **free radicals** can play a role in this, although the efferent duct has a defence mechanism for this. The possibilities to assess the structural quality, the integrity of DNA, are rather limited for sperm cells. The extent of damage can therefore never be determined with a high degree of sensitivity. Residual DNA damage, such as double-strand breaks, can be detected by the egg, then a zygote (Ch13.5.2). Available *in vitro* laboratory tests for DNA damage can be used to characterise DNA from the entire cell population or from a purified fraction thereof. Then, attempts are made to predict the "embryo developmental potential" of the patient's sperm. The most reliable technique used for this involves letting the DNA migrate from the nucleus under the influence of an electrically charged field. When DNA breaks are present, chromosomal fragments have been created. Since DNA is negatively charged, these fragments migrate to the positive pole. The DNA that has migrated from the nuclear matrix (fig 32) can then be quantified after staining. The image will be the clearest for apoptotic sperm cells. The value of such an assessment for the molecular integrity of DNA from the entire cell population of the ejaculate has been questioned for years. A meta-analysis published in 2017, based on 41 studies, demonstrated that the instability of the DNA of a sperm sample, coarsely identified in this way, does actually have predictive value for the likelihood of a successful pregnancy. However, it is not the main variable in this context; the quality of the egg is much more important (Ch11.2, Ch11.5, Ch13.5.3, Ch16.3.2).

12.10 Declining sperm counts

In 1992, the world was alarmed by the publication of a study from Denmark that reported a sharp decline in the number of sperm per ejaculate over the previous 50 years. This study generated controversy for a long time. Even in 2013, high-level debates were still being held about it, as the outcomes of the study were not widely trusted. The appearance of a large overview of sperm counts from the period 1973-2011, summarising 185 publications and measurements taken from nearly 43,000 men, settled the debate in 2017. Whether looking at concentration or total quantity, the decline over this period ranges around 50%, with a decline rate of approximately 1.4% per

year. The authors of this paper, which was published in *Human Reproduction Update*, emphasise that the decline is happening throughout the Western world: In Europe, the USA, Australia, and New Zealand. However, the same trend is also observed in China. In 2022, the same consortium published an update, covering data from all continents with even more alarming conclusions. The rate of decline between 2000 and 2018 was 2.6% per year, more than double the rate of decline between 1973 and 2018. The discussion about the decreasing sperm production in the Western world began in 1992, initiated by Niels Skakkebaek, who is now in his eighties. These concerns in Denmark did not find their origin in the IVF lab but came from the fields of oncology and sex development. As a kind of *opus majus*, the group of researchers around Skakkebaek gathered the key facts again in 2016, an effort that was repeated even more extensively in 2021. Several trends do indeed raise concerns. An association is observed with sex differentiation and with the reprogramming of the germline (Ch9). Regarding the latter process, it is noteworthy that the incidence of testicular germ cell cancer in relatively young men (14-44) has shown an increasing pattern over the last three decades. This form of cancer is found in approximately nine out of every 100,000 men in Europe. In these cases, cancer cells originate from prenatal germ cells, where errors were made in the confirmation of their status as prospective stem cells. The **pluripotent** aspect of these cells apparently remains too dominant, which increases the risk of germ cell neoplasia (box 2). One single “driver” mutation that causes this type of tumours and that is shared by all of them has never been found. However, mutations do, in fact, play their well-known role in the development of germ cell tumours. Other known risk factors include improper descent of the testes (cryptorchidism) and infertility.

In his analysis of declining sperm counts, Skakkebaek incorporated abnormalities in the development of the testis and the efferent duct of the male urogenital system, characterised by increasing rates of hypospadias (the opening of the urethra is not located at the tip of the penis), and cryptorchidism. Out of a wide range of findings, all related to the development of the testis, Skakkebaek distilled his life’s work: defining the “testicular dysgenesis syndrome” (TDS), a condition in which the testis is at risk in the embryonic stage and perhaps also thereafter. The concept of “**endocrine disruptor**”, a separate research area as such (and one that is, in fact, too extensive for this book), plays a role in this analysis. Perhaps that is why it reappears in the chapter on **epigenetic** inheritance, chapter 15. In Niels Skakkebaek’s view, the testis (including its prenatal development) is primarily an organ that is markedly impacted by environmental factors, and he is not alone in this belief.

12.11 The general health aspect of low sperm production

Does a low sperm count occur independently of other bodily functions? It can, but that is not necessarily the case. When the advancement of artificial reproductive technologies in principle enabled fertilisation with a minimal yield of testicular sperm cells (TESE, Ch16.1, Ch16.3.1), publications in the literature (around the year 2000) raised questions about whether the observed sterility was entirely isolated or if there was any connection with other medical-physiological issues. This was followed by a period of silence, but in recent years, larger epidemiological studies have been published that demonstrate associations with other clinical problems. Firstly, the finding that a poorly functioning testis poses a risk for testicular cancer reappears. Secondly, it is once again confirmed that large body size (waist circumference) is not a positive predictor of sperm cell production. However, it provides little information about sperm motility.

Right now, at the time of writing this, there are at least three publications showing that sperm production, as well as sperm motility, are good predictors of overall health. In the group of men who were eligible for a semen analysis, a correlation was found with a higher likelihood of hospitalisation and a reduced life expectancy. The apparent poorer health in these men could be correlated with issues in the cardiovascular system and the prevalence of diabetes (types 1 and 2). The key question is, of course, whether this is “nature” or “nurture”, or perhaps a combination of both. In this context, “nature” refers to the idea that genes functioning in the testis are also involved in other aspects of health. This seems to be a reasonable assumption due to the large number of genes involved in each trait, which makes it more than likely that there are overlapping effects occurring in multiple tissues or organs at the same time. “Nurture” will be further addressed in chapter 17, when we will, although briefly, touch upon environmental effects on sperm production from a more toxicological perspective. As mentioned earlier, that area is not the focus of this book. Recently, research from China has also shown an association between the microbiome of the male genital tract and disrupted fertility, once again demonstrating that literally any biological mechanism can have an effect on fertility.

12.12 Genetic causes of male infertility

To date, relatively little progress has been made in identifying the presumed genetic causes of disturbed or suboptimal sperm formation, despite ongoing reports of genes involved in this (table 6). The notion that obviously there must be a genetic aspect became evident in an English

research project published in 1994. In this study, conducted in a relatively small population, they looked at the brothers of men who had visited a fertility clinic, while there were no indications of reduced fertility in their female partners. Of those brothers, 17% also reported needing fertility assistance, while none of the brothers of the men in the control group faced this issue. The observed difference is highly significant. In order to advance genetic research on spermatogenesis, it is essential to first elaborate on the histological manifestations of infertility.

In chapter 12.2, it was described that characterising human spermatogenesis based on histology is not an easy task. It is even more challenging in men whose spermatogenesis follows a less optimal course. After all, there is a sliding scale when it comes to fertility problems, ranging from the complete absence of spermatogenetic cells to a situation where sperm do end up in the epididymis and also in the ejaculate. The first scenario is referred to as “Sertoli cell-only syndrome” (SOS). When spermatogenesis halts at a specific stage (see Ch12.5.3 on the uniform initiation of a checkpoint), this is called a “maturation arrest”. That arrest can be partial or complete, depending on the underlying genetic condition. Spermatogenesis may be locally present (focal), as if there is (or has been) a stem cell problem that has been resolved here and there. Finally, older literature mentions hypospermatogenesis; the stem cells are predominantly doing their job, yet there is loss of cells throughout spermatogenesis. Besides **oligospermia** (table 5, 6, fig 36), two types of azoospermia are found in the fertility clinic (fig 36): obstructive and non-obstructive azoospermia. The obstructive type, which might be due to inflammation in the epididymis, does not really fit in the context of this section, if at all. It is much less likely to have a **genetic background**, so it will not be discussed here.

In the case of non-obstructive azoospermia, a problem that affects at least one in 200 men (see table 6 for an estimate), the histological picture can still be highly variable. In about 50% of these cases, sperm cells are present in the testis. During fertility treatment using a TESE, these cells can then be sought, or it can be conclusively determined that there is really nothing to be found.

12.12.1 Chromosomal aberrations

In about one in 1,000 men, there are no or hardly any spermatogenetic cells in the testis. In that case, we are dealing with the SOS syndrome from the previous paragraph. The most common cause of this is the presence of two X chromosomes. Together with the Y chromosome, this results in the XXY sex chromosome constitution, the karyotype associated with Klinefelter

syndrome (Ch3.4, table 4). Somehow, spermatogonia cannot cope with two X chromosomes. The fact that quite a few genes involved in spermatogonial multiplication are located on the X chromosome could add to this as **dosage compensation** through X chromosome inactivation (box 4) does not seem possible. An excess of protein production from these genes can be as detrimental as a deficiency. So why is it that you can still find small foci of spermatogenesis in the testis of Klinefelter men, even including sperm cells? Apparently, during the divisions of spermatogonial stem cells or their precursors, an X chromosome is occasionally lost. This results in a clone of XY-carrying stem cells with normal chances of differentiation. However small those XY clones may be, sperm cells can be found in up to 50% of Klinefelter men with the use of TESE. The chance of having children is the same as in men with non-obstructive azoospermia due to other causes (Ch16.3.2).

In men who still produce some sperm, and who fall into the OAT category (fig 36), more structural chromosomal aberrations are found, such as Robertsonian and reciprocal translocations (Ch3.4.1, table 6, fig 12, 37). The cause of this lies in problems with meiotic chromosome pairing, resulting in unpaired chromosome segments that interfere with the assembly of the sex body (Ch12.5.3, Ch12.9.3, fig 37). Consequently, the frequency of such chromosomal abnormalities is significantly elevated in men with fertility problems.

12.12.2 Y chromosome (micro)deletions

In 1976, a remarkable paper was published by researchers from Pavia, the old university town south of Milan, Italy. It was not the first time either that an important announcement in the field of reproductive genetics came from the Po Valley: back in 1865, the Sertoli cell was named after Enrico Sertoli, who worked in Milan but was born in Pavia. The scientists from Pavia reported in their publication that a shortening of the long arm of the human Y chromosome (fig 38) that was visible under the light microscope was associated with sterility. To appreciate this finding, some knowledge of the DNA landscape of the Y chromosome is needed. This knowledge has since been acquired, but the Y chromosome proved to be a tough nut to crack, although in the summer of 2023, the telomere-to-telomere DNA sequence of the human Y chromosome was consecutively published by two research groups in the prestigious journal *Nature*. An initial hint emerged in 1996, when it was reported that three regions on the long arm of the Y chromosome were specifically involved in either total azoospermia or variable oligospermia. The group led by Peter Vogt in Heidelberg, Germany, labelled these regions, from the **centromere** to the

Table 6 Genetic diagnosis of **spermatogenesis** in male infertility (excluding hormonal aspects and abnormalities of the efferent ducts).

Type of mutation	Azoospermia (0.7%) * Oligospermia (6.9%) *	*Asthenospermia and teratozoospermia 1% ****	Population incidence rate (table 4)
Y chromosome deletions	3.3-9.1% (6.8%), mostly infertile men, ranging from SOS to mild oligospermia		?
Sex chromosomes: mainly XXY	12.6-16.2% of men with azoospermia		0.16%
Reciprocal translocations	azoospermia 0.9% oligospermia 0.6%		0.14%
Robertsonian translocations	azoospermia 0.09% oligospemia 1.6%		0.12%
Gene defects	Azoospermia ** 1.2% N genes = 31 Oligospermia *** N genes?	50%	?

* The assumptions used to compile this table are the following. One in 7 couples experience difficulty to conceive. In 50% of these couples, a male factor is involved, ranging from reduced fertility to infertility (in approximately 7% of men). The pie chart from figure 36 was taken as a baseline for the distribution between men with **azoospermia**, men with **oligospermia** and men with morphology or motility issues. For men with azoospermia (a frequency of one in 100 is often mentioned in the literature), corrections were made for other causes (for instance, blockage of the epididymis, 30-40%). Among the men with oligospermia, there will surely be some with a partner with reduced fertility. The count may also indicate a temporary reduction.

** In azoospermia, most of the affected genes show a **recessive** pattern of inheritance. To increase the chance of finding the mutation, researchers focus on a clear **histological** image such as a uniformly occurring cessation in spermatogenesis (maturation arrest). By the end of 2020, 12 genes were known to be associated with this condition, and at least half of them are also involved in female

gametogenesis (particularly in **meiosis**). The understanding of the genetic background of azoospermia in particular continues to increase as research progresses, now with the use of methods like trio sequencing (Ch3.2). A recent study from Radboudumc reported 29 mutations that are likely to contribute to infertility, out of a total of 192 in 185 patients (mostly cases of azoospermia, besides that severe oligospermia). **WES** was used for comparison between patients and their parents. This analysis assumes a **dominant** effect of the allele that performs less or not at all due to the mutation.

*** The genetic analysis of men with oligospermia whose histological image of the testis shows great variation, has only just begun.

**** These genetic diagnoses exhibit a uniform pattern: abnormalities in the tail/motility (astheno) or the absence of the **acrosome** (globospermia, a type of teratozoospermia).

telomere, as AZFa, AZFb and AZFc (fig 38), but knowledge about the specific nature of the DNA in these regions was still lacking. The reason why clarity on this did not come easily with the state of DNA research at that time lay in the genetic structure of the Y chromosome, which is full of "**copy number variation**" (Ch3.3). David Page's lab at MIT (Massachusetts Institute of Technology) in Boston was on top of this subject in good time. Consequently, the further unravelling of the role of the human Y chromosome in infertility largely took place at that lab. A breakthrough publication on the research results of a team of 40 authors, with David Page as last author, appeared in the journal *Nature* in 2003. It described that there are eight regions on the long arm of the human Y chromosome where the DNA sequence is mirrored (fig 38). Such a structure is known as a palindrome, meaning that DNA base sequences occur in two orientations (reversed and not reversed) in duplicate or multiple instances. It is justified to expect that this provides an evolutionary advantage for the Y chromosome. The base DNA motifs of these palindromes contain several genes related to spermatogenesis. Due to the palindromic structure of the chromosome, these genes typically appear in multiple copies and are said to be ampliconic, i.e. housed in an amplicon. In humans, as well as in great apes and other primates, the orientation of genes located on the Y chromosome towards male fertility is remarkable. A consequence of this homology within the Y chromosome is the formation of hairpin-like structures in the sex body, between the "legs" where undesired recombination can occur (Ch12.5.1).

As a result, gene copies located within these “hairpin structures” are lost. This is referred to as AZFa, -b and -c deletions. In almost all cases, these deletions are not visible under the microscope, and molecular DNA technology is used to detect them. Table 6 provides an impression of the negative consequences for male fertility. The genetic research on AZFa, -b and -c deletions that is conducted worldwide is a great example of a scientific field that is becoming increasingly complex to explain (fig 38), but that does open the gate for understanding the evolutionary existence of “our Y chromosome”. Modern DNA research now enables us to globally observe that the evolution of these “multicopy” genes on the Y chromosome occurs faster than that of autosomal multicopy genes. Given their unique positioning and the aforementioned mechanisms, this is entirely in line with expectations. The special recombination events that the Y chromosome may undergo can, albeit in a minority of cases, also be advantageous due to their effect on gene dosage. A much finer and intriguing insight into “our Y” is provided by one of the aforementioned telomere-to-telomere (T2T) studies from the summer of 2023. This is because this group utilised a pangenomic (table 3) approach when comparing 43 human Y chromosomes from all over the globe. Apart from reaffirming that the Y chromosome has originated from a once homologous pair (as evidenced by the X-degenerate region of the Y chromosome), this paper pays attention to the selection pressure that must play a role in fertility, as approximately 180,000 years of evolution can now be covered by the geographic distribution of the 43 Y chromosomes. The Y chromosome stands out due to its susceptibility to structural variation (for example Y deletions) on the one hand, and on the other hand, due to the signs of purifying selection that it shows. This leads to loss of genetic variation, as is for instance demonstrated by the low number of SNPs in single-copy protein-coding Y-specific genes. Another possible explanation for this phenomenon is a male-specific bottleneck in the more recent human history.

A clinical picture thus only emerges when very large segments are cleaved from the chromosome as a ring-shaped structure through recombination between related DNA sequences. Deletions of the AZFa and -b regions (fig 38) almost always lead to “Sertoli cell-only” (SOS, the a deletion) or to a complete blockage of spermatogenesis (the b deletion). In the case of a complete AZFc deletion, the picture varies from SOS to severe oligospermia. Determining the status of the AZF regions is, therefore, part of the standard work-up in fertility diagnostics. Due to the somewhat more complicated structure of the AZFc region, small deletions can occur more frequently than in the other two regions, but also as a result of undesired recombination. The consequences of this range from azoospermia to normal spermatogenesis. Among geneticists, this phenomenon is known as the influence of “the residual genotype”, the individual genome

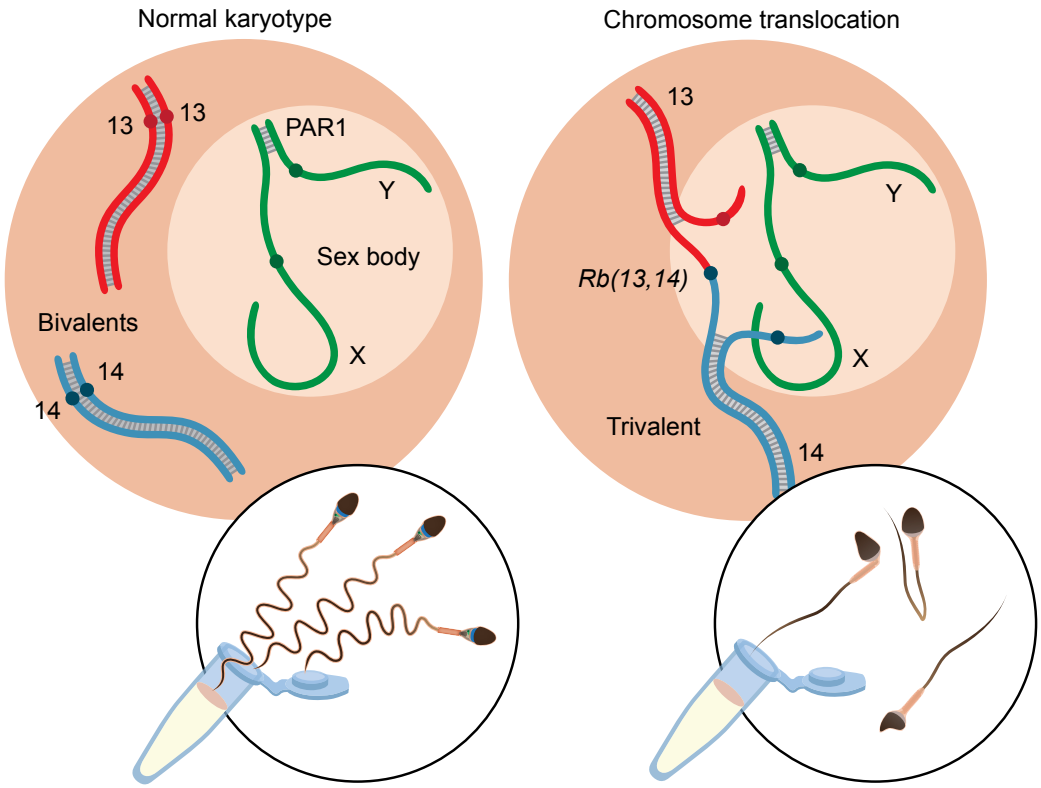


FIG 37

Problems with **meiotic chromosome** pairing are more common in carriers of a chromosomal aberration, such as reciprocal and Robertsonian translocations (fig 12). This results in a reduced sperm count and quality (from **azoospermia** to **oligospermia** (OAT), tables 5 and 6, fig 36). The search by unpaired **autosomal** chromosome segments for the largely unpaired **sex chromosomes** is responsible for this. This disrupts both the silencing of **transcription** of the "**sex body**" and the desired transcription of the unpaired autosomal segments. PAR stands for **pseudoautosomal region**, of which the main one (PAR1) is depicted here (see also fig 38).

that is apparently able to mask the effects of these smaller deletions. Because of genetic variation in the population, this is not possible for everyone.

When an AZFc deletion is detected in the clinic, it is usually caused by a *de novo* mutation that occurred in the meiotic cell that produced the successful sperm cell. However, a fertile father who is a carrier of this mutation can also have an infertile son who has inherited this same mutation but reacts differently to it.

In the early days, after the discovery of the relationship between Y chromosome deletions and sperm cell production, a graph was presented during a lecture on the subject, showing the number of generations on the X-axis and the fraction of men with a complete or partial AZFc Y chromosome deletion on the Y-axis. When nothing is done, that is, one does not search for sperm cells to use in IVF/ICSI, you will see a horizontal straight line in this graph: the creation of new mutations is equal to the elimination of mutations due to infertility; there is equilibrium. However, when carriers are genetically rescued, at least partially, through the assistance of the IVF laboratory, the deletion can spread in the population and will thus increase in frequency. Since not every man with this mutation will have children, a new equilibrium will eventually be established at a higher level after many generations. This is a topic of interest for population geneticists, and also for clinicians (?).

Could those smoothly written newspaper articles about the end of the Y chromosome due to continuous DNA loss actually be true after all? Our evolutionary existence has already lasted long enough to allow such a scenario to come true. However, the Y chromosome has not been lost. On the contrary, the human Y chromosome has strengthened itself by duplicating regions that are crucial for spermatogenesis. The small business risk of illegitimate recombination seems to be accepted in this process. At the population level, you will not notice the slight loss of male fertility; after all, there are more than enough men to sustain the population.

12.12.3 Mutations at the gene level

Along with the brain, where many genes are required for complex functioning, the testis is always mentioned as an organ for which the same applies. Increased expression of 2,274 genes, slightly more than 10% of the total number of protein-coding genes, has been identified in this organ (2020). Roughly 500 of these have been classified as truly testis-specific genes (Ch12.5.2). The so-called “target”, the part of the genome in which mutations that have consequences can occur, is thereby remarkably large for the testis, as it is for the brain.

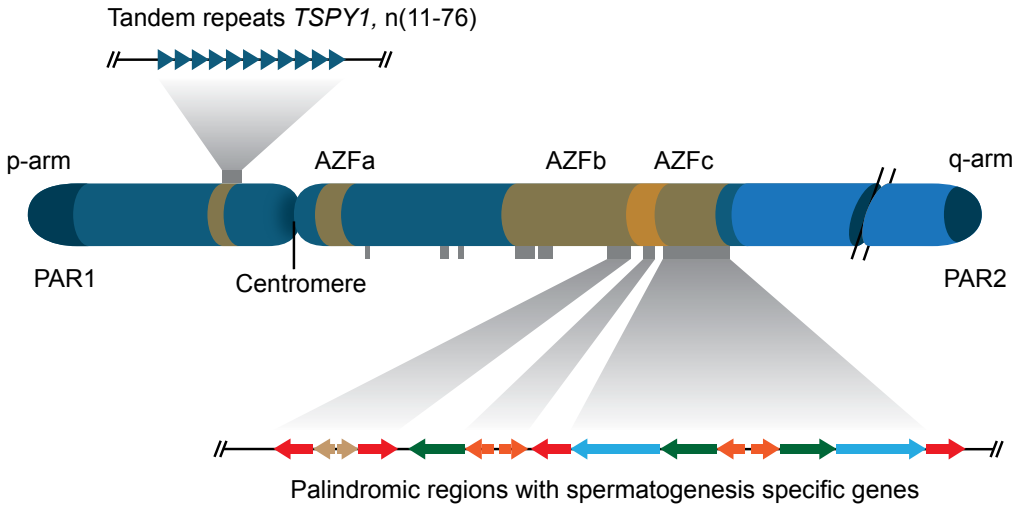


FIG 38

The male-specific part of the Y **chromosome** contains nine "multicopy" **gene**s, mainly located in six regions that are organised as a palindrome, the palindromic or ampliconic regions. Three of them are shown in the figure. The expression of these genes is testis-specific. The total number of protein-coding genes on the Y chromosome is small, with 16 in PAR1, four in PAR2, and 23 in the male-specific region. However, now that the base sequence of the Y chromosome has finally been resolved, some more coding genes have been discovered. Even in the absence of a clinical fertility condition, the number of copies per gene in the palindromic regions shows great variation between men. The organisation of these genes, some of which also have family members on **autosomes**, is a nice example of the infinite complexity of the functioning of our **genome**. NB: the **PAR** regions are **homologous** to the ends of the X chromosome (see also fig 15).

The search for male fertility genes started quite some time ago, stimulated by knockout mouse models with attractive phenotypes. However, compiling a list of candidate genes obtained from such research and using that to screen a poorly defined patient population for mutations in these genes proved to have very limited success. In 2017, a report on such a search was issued from Radboud Radboudumc. In 1,112 men with azoospermia or severe oligospermia without an identifiable cause, genetic causes of infertility were sought in 101 genes known from mouse research. This resulted in only a 1-1.5% yield of genetic diagnoses.

A more efficient way to search for fertility mutations is to look for a uniformly occurring deviation in the morphology of human sperm, such as a missing acrosome (globospermia) or an abnormality of the tail (table 6, fig 36). Then, you check whether there are any genetic models for this phenotype in the mouse. This literature search can yield candidate genes that can be used in further screening of the patient. This approach is effective, and currently, a genetic diagnosis can be reached in about 50% of patients in this way (table 6). The phenotype can be extended to a characteristic histological image for meiotic cells that are trapped at a checkpoint (Ch15.5.3), for example. In Rotterdam, at the Department of Developmental Biology at Erasmus Medical Centre where Willy Baarends works, this approach was applied to a series of 33 men with non-obstructive azoospermia. They screened for mutations in 175 genes found in both humans and mice, which lead to the histological phenotype of interest in homozygous mutant mice. That resulted in two “hits”, both of them for a gene that is active in meiosis and that also affects the female germline. Through global collaboration, more of such gene mutations have now been found. Involved researchers have united under the name International Male Infertility Genomics Consortium (web address http://infertilegenome.org). This illustrates the importance attached to understanding the cause of impaired spermatogenesis in men. To get an idea of the current state of knowledge (up to 2021) on the genetic origins of male infertility, please refer to table 6.

Currently, great progress is being made in the search for the role of mutations in an aberrant phenotype, thanks to a technique called trio sequencing (Ch3.2, Ch11.3, Ch18.1). The assumption here is that the causal mutation arose in the germline of one of the parents, or alternatively in the very early stages of embryonic development, and that it is **dominant**ly inherited. Hence, the mutation found in the patient is not present in DNA from somatic tissues of the parents, and it must have affected a gene that matches the clinical picture from a cell biological point of view. Over the past 10 years, this approach has gained prominence at the Genetics Department of the Radboudumc in Nijmegen, and it has also been successfully used in research on male infertility. Initial findings indicate the involvement of genes that (may) function in spermatogen-

esis. However, it is proving much more difficult to obtain the necessary evidence for this than it is when a similar approach is used focused on the brain. Why is this the case? It turns out that the fertility genes (usually the paternal gene copy) that are identified in this way do not tolerate a missense mutation (also called non-synonymous, box 1) very well: they are often prone to **haploinsufficiency**. However, the genetic basis of infertility can also be explained using the concept employed in quantitative genetics (Ch6). In this model, subfertility and infertility are caused by the convergence of several less favourable alleles for sperm production in the patient. A picture like that also fits into the overall view of the increasing "mutational load". More on this will be covered in chapter 18.1.

12.13 Older fathers

When people write about the effect of age on the chance of reproduction, they usually don't focus on men. One may now teasingly (but truthfully) remark that this is because most researchers in this field are men. Another reason could be that, up until now, sperm cells have not been an easy research subject for those who aim for a more in-depth scientific investigation (this chapter).

Of course, also in men, sperm production and the quality of the product are finite. The moment at which this leads to fertility problems approaches faster if the initial situation is worse. In the literature, it is reported that sperm production declines by two-thirds between the ages of 20-30 years and 60-70 years. An OAT man (table 5) identified in the clinic may actually have had normal fertility earlier in life. Numbers on declining fertility are often epidemiological in nature. The age of a woman and a man is usually linked. Aren't men, on average, three years older than their female partners (?), and what is the distribution of the age difference? The influences of male and female ageing are thus intertwined. That, and other factors that complicate the statistics, are possible reasons for the relative scarcity of review articles in this area. Data from the USA, compiled with statistical care, indicate that just under 80% of men younger than 25 will impregnate their female partner within 6 months of trying to conceive. The comparable (adjusted) rate for men over the age of 45 is 53%. Spontaneous abortions are also more likely to be statistically linked to the father when he is older.

Naturally, the sperm of older men has been studied extensively. The ejaculate is less voluminous, but the concentration of sperm cells is not immediately lower. However, motility is lower, and fewer sperm heads meet the high morphological requirements that are set in the clinic. But does nature impose equally high standards?

Does an older age of the father result in an increased risk of a genetic abnormality that shows up in the phenotype? Some insight into this is quite old. Lionel Penrose, who was already mentioned when the effect of maternal age on the likelihood of having a child with Down syndrome was addressed (Ch11.4.1), also published in *The Lancet* in 1955 that he observed a paternal age effect on the spontaneous appearance of severe growth retardation (achondroplasia). The cause of this is a dominant mutation in the *FGFR3* (fibroblast growth factor receptor type 3) gene. Coincidentally, this gene is expressed in the spermatogonial stem cells, and even more coincidentally, this mutation gives cells a proliferative advantage. As a result, there are clones of rapidly dividing cells that make a greater contribution to the sperm population over time, leading to an increasing likelihood of the appearance of this phenotype in the offspring of fathers with the *FGFR3* mutation in their stem cells. Fortunately, this does not happen often; the risk of achondroplasia is one in 15,000 births. That rises to one in 1,250 when the father is between 50 and 54 years old.

In addition to Penrose's paper, there was another publication in the 50s that established a link between disease in offspring and the age of the father. In 1958, Johanson reported a suspicion, based on a small cohort, that the older age of the father is involved in the development of schizophrenia in offspring. Nowadays, large sections of reviews on the genetics of such conditions, also including bipolar disorder and autism, focus on this association. As per our current understanding, these conditions are all genetically related. The relative risk for fathers over the age of 55 to have a child with schizophrenia is about 6 (compared to fathers around the age of 25). In a large Dutch study from 2011, led by Roel Ophoff, a former student of Wageningen University, it was found that autism is the greatest age-related risk, followed by schizophrenia. The big question at this moment is what might be behind this. Is it all *de novo* mutagenesis at the gene level? We now know that new mutations at the base level mostly emerge in the male germline (table 3), and the male contribution doubles in approximately 20 years (table 3). The theoretical explanations behind this were already mentioned at the end of chapter 12.3 (see also table 3). Out of the approximately 45 base pair mutations that are passed on in each generation from the father (at the age of 30) to the offspring, there will only be a few that occur in an exon, which is not enough to explain the increased risk of schizophrenia. Or could there also be an epigenetic aspect to this, because, for example, in the case of decreasing sperm quality, the DNA is less well packaged and instructed (Ch15)? As far as I know, this has never been thoroughly investigated, while it has, of course, been studied in the rat, with the following predictable outcome: at an older age,

the share of nucleosomal DNA (Ch12.6) increases. Without a doubt, future molecular genetic and chromatin-focused research will shed more light on this.

The overall picture of an increase in risks that are still relatively small is completed when childhood cancer is also considered. For fathers older than 36, the risk for their child increases by a factor of 1.63 compared to fathers younger than 25.

Finally, there is a bonus: since **telomerase** is involved in the greater number of chromosome duplications of the spermatogonial stem cell, the telomeres (Ch2.3.3) of the chromosomes passed on by older fathers are longer than those transmitted by younger fathers. Whether this makes the offspring live longer is a different discussion. At fertilisation, paternal telomeres are longer than the telomeres of maternal chromosomes anyway (Ch13.6).

CHAPTER 13

Fertilisation and cleavage

13.1 Introduction

At ovulation, the egg (or oocyte) is in the metaphase of the second **meiotic** division. The **spindle apparatus** lies beneath the cell membrane near the first polar body (fig 39). Many pages have been written about how sperm cells reach the oviduct after mating, intercourse, coitus or (artificial) insemination, whether or not preceded by solid research. Much of this research is old, and part of it can stand the test of time very well.

Fertilisation takes place in the ampulla of the oviduct, on the side of the ovary. Naturally, the sperm cells arrive from the other side, through the isthmus of the oviduct (fig 39). If all has gone well, sperm are present at the time of ovulation, albeit in small numbers. Fertilisation then proceeds as a fascinating sequence of extraordinary events. It starts with the union of the **gametes** through cell fusion, resulting in the formation of the **zygote** (fig 42). At the end of the zygote's cell cycle, the **chromosome**s of both female and male origin are in the metaphase of the first cleavage division (fig 42). After this, the two sets of chromosomes convene in the nuclei of the 2-cell **embryo** (fig 43). The dynamics in chromosome behaviour, and the changes in **chromatin** in particular, are especially enormous on the male side. Here follows a concise overview of our current insights into the entire process.

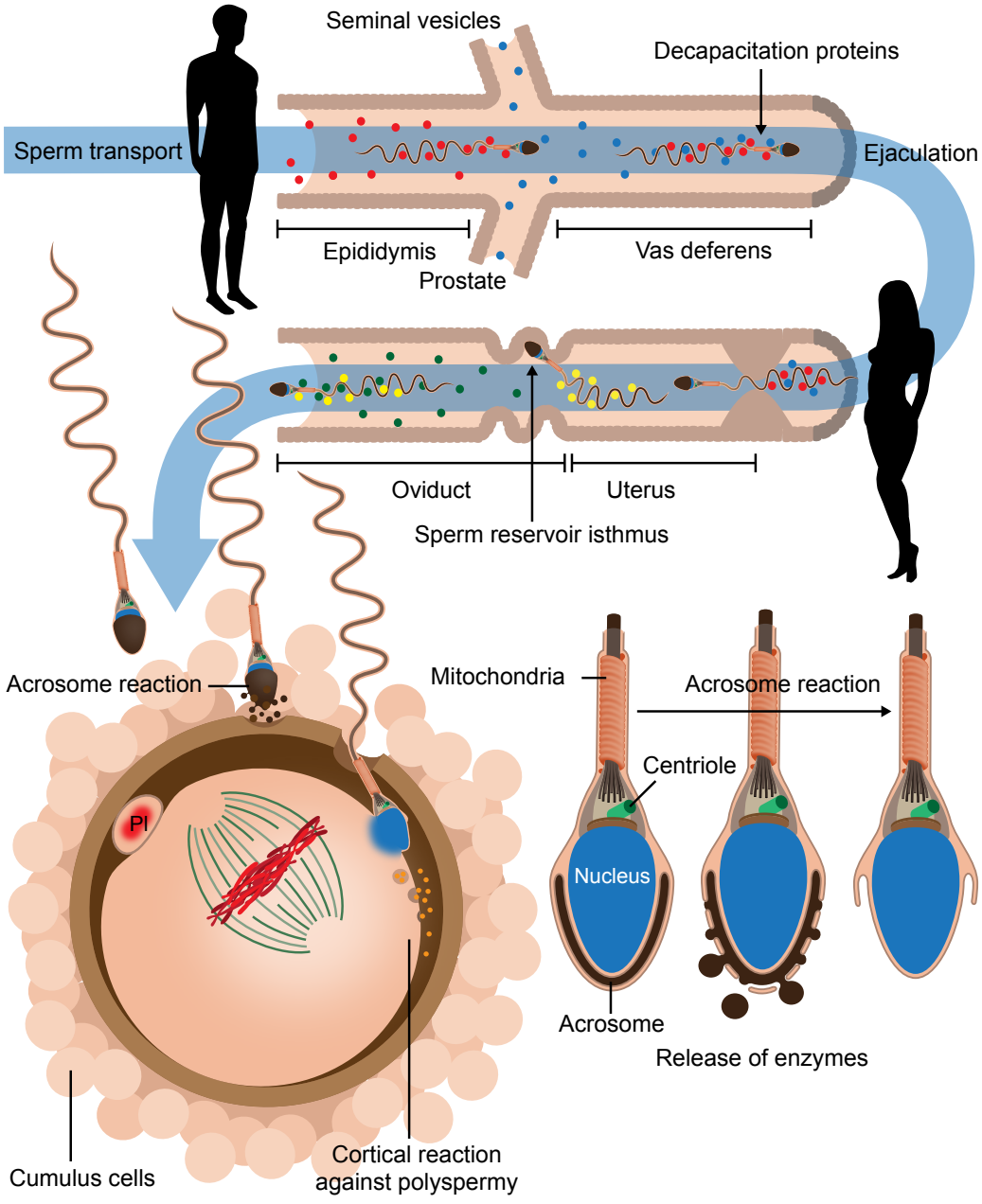

Seminal vesicles
Decapacitation proteins
Sperm transport
Ejaculation
Epididymis
Vas deferens
Prostate
Oviduct
Uterus
Sperm reservoir isthmus
Acrosome reaction
Mitochondria
Acrosome reaction
Centriole
PI
Nucleus
Acrosome
Release of enzymes
Cumulus cells
Cortical reaction
against polyspermy

FIG 39

Overview of sperm transport from ejaculation until fusion of **gametes** in the ampulla of the oviduct, which forms its beginning portion. The longer sperm keep their fertilising capacity in the female genital tract, the greater the chance of pregnancy per cycle. The main way to achieve this is to delay **capacitation** in the sperm reservoir, the existence of which in humans is assumed here. The red spheres, proteins secreted by the wall of the tube in the epididymis, contribute to sperm maturation. The blue spheres represent proteins that accompany the secretions of the prostate and seminal vesicles during ejaculation, and they act as capacitation inhibitors (the decapacitation proteins). These are lost during passage through the cervix, and are replaced by the capacitation-regulating proteins of the female genital tract (the yellow spheres). Finally, as ovulation approaches, proteins that stimulate capacitation are released from the **epithelium** of the oviduct by the isthmus (the green spheres). When ovulation is not yet imminent, proteins in the isthmus portion of the oviduct contribute to delaying capacitation.

Cumulus cell expansion increases the probability of encounter in the ampulla (see also fig 28, 29). The **acrosome reaction** is also shown in this figure. In this reaction, the acrosome membrane fuses with the cell membrane, thereby releasing its contents. In humans, the sperm carries the centriole (fig 3), shown here in green and blue. A second centriole has recently been discovered, hidden in the base of the tail, although its function remains to be clarified. After membrane fusion of the cortical granules with the membrane of the egg, their contents are released, which changes the **zona pellucida** (light brown) in a way that greatly reduces the chance of polyspermy. This cortical reaction starts at the site of cell fusion. The space between the zona and the wall of the oocyte is known as the perivitelline space. PI, first polar body that rapidly goes into regression (see also fig 13). Until egg/oocyte activation, the cell is in the metaphase of the second meiotic division (the secondary oocyte of fig 13).

13.2 Sperm transport

When I began delving into this subject in the late 1960s, it was already a well-established fact that the transport of sperm from the cervix to the beginning of the oviduct is primarily facilitated by action of the uterus. In pigs, this transport can be completed as quickly as 15 minutes after insemination, and the pig uterus is long! In rabbits, the distance to the oviduct is covered in about 5 minutes. The process is also swift in laboratory rodents. In his recent overview titled *Copulatory and Postcopulatory Sexual Selection in Primates*, Alan Dixon notes that human sperm moving at full speed, reaching a top velocity of 7 inches (approximately 18 cm) per hour, would still require at least 45 minutes to reach the beginning of the oviduct under their own power. However, as in other mammals, uterine contractions reduce this time to several minutes in humans, according to the literature. The small volume of the uterine cavity (100 µl) aids in this process. In farm animals, much is understood about the formation of the sperm reservoir that will supply the oviduct in the direction of the ampulla: around a million sperm cells are present at the transition from the uterus to the oviduct in pigs and within the folds of the surface at the beginning of the oviduct in cows (in the isthmus, fig 39). Hereafter, it will become apparent that there are strong indications of the existence of such a reservoir in humans as well.

13.3 Sperm heterogeneity

The population of sperm cells in an ejaculate is fundamentally heterogeneous. This is particularly true in humans (Ch12.9.1). The heterogeneity extends to the likelihood of fertilisation. In the mid-70s, J Cohen's group from Birmingham reported on an experiment with rabbits, thereby bringing attention to this phenomenon. A small amount of sperm that had reached the oviduct was harvested and mixed with an excess of fresh sperm, after which artificial insemination was performed. The sperm that had already "seen" the oviduct came from an agouti (wildtype) ram, and the excess came from a white albino ram. The doe was also albino. To their surprise, the sperm that embarked on a second journey was much more effective. As a matter of fact, the percentage of agouti offspring was substantially higher than expected based on the fraction of "agouti sperm" in the mixed sperm used for artificial insemination.

The extent to which sperm heterogeneity is manifested is strongly determined by the female reproductive system, which can be viewed as a sorting machine. When we discovered a mouse with a chromosomal anomaly in Wageningen in the 1970s, which produced few and poor-quality

sperm, we were among the first to study this phenomenon. After mating, these mice had lower litter sizes, but they were rarely completely sterile. Sperm from normal mice show little variation in head and tail, but morphological heterogeneity was considerable in males with the chromosomal anomaly (compare with fig 37). When we flushed the sperm from the female's oviduct after mating, the variation observed in this population was much lower than in the uterus or in the male reproductive system itself. The motility of these sperm was also higher. As is usually the case in scientific research, these observations raised further questions. Is there selection against head abnormalities of sperm during the transition from the uterus to the oviduct? Is there competition based on motility, or is the phenomenon a combination of both?

13.4 Gametes and the "fertile window"

For years, it has been recognised in both laboratory and farm animals that it is reproductively inconvenient to have the egg wait for the sperm. During my studies in animal husbandry, I was greatly impressed by the work of RHF Hunter. This researcher worked at the ARC Reproductive Physiology Unit, Huntington Road in Cambridge, a place where advancements in reproductive research happened years earlier.

In his research on gilts (Ch11.5), he induced ovulation with HCG (Ch11.3). Since **oestrus** (the period during which female animals allow males to mate) is relatively long in pigs, the injecting, which is done at the start of this period, precedes the LH peak (the ovulation-inducing hormone, fig 29). The oocytes are then released from the Graafian follicles after about 40 hours (fig 29). Around 6 hours after insemination, a reservoir of approximately one million sperm cells has formed at the junction of the uterus to the oviduct. This reservoir remains in place for 24 hours. Ronald Hunter studied the effect of the timing of insemination relative to the timing of ovulation. When inseminating between 6 and 10 hours after ovulation, he observed that the percentage of fertilised eggs began to decrease from 95% to 70%. The fraction of fertilised eggs that did not behave normally during cleavage (fig 43) began to increase, from a few to as much as 25%. The egg's ability to start normal development after fertilisation already decreases about 8 hours after ovulation.

Also in humans, reproduction is much more successful when sperm cells are present around the entrance of the oviduct at the time of ovulation. A few large studies from around the turn of the century clearly demonstrate this. For a woman, there are two ways to approximate the timing of

ovulation: by determining a temperature curve or by measuring the waste products of **oestrogens** and progesterone in the urine. The oestrogens produced by the follicle peak at the time of the LH peak. After that, the follicle also starts producing progesterone, which is taken over by the corpus luteum after ovulation (fig 28). Ovulation occurs about 36 hours after the LH peak, and this moment can be estimated using the ratio of metabolites of oestrogens and progesterone in urine (fig 29). With this method, it was found that there is a wide margin in the timing of ovulation within the menstrual cycle and thus in the likelihood of pregnancy (fig 40). The temperature method, which is slightly less accurate but can also be used to estimate the timing of ovulation, is based on the fact that the basal body temperature is slightly lower around the LH peak. After this peak, it rises to a plateau that is reached 2-3 days after ovulation. Ovulation occurs around the halfway point of the temperature rise to this plateau level.

In 2002, researchers from Padua, in collaboration with a skilled biostatistician from Research Triangle Park, North Carolina, USA, published the results of a large study on the relationship between ovulation and conception. It involved 782 allegedly healthy couples who used only natural methods to regulate their fertility. Occasionally, a pregnancy did occur.

The time of ovulation was estimated using the temperature method, and the participants had to specify on which day they had had sexual intercourse. The results were exceptionally consistent (fig 41). Fertility increases from days before ovulation and peaks 2 days before. After that, and also on the day of ovulation, the chance of pregnancy decreases substantially. The effect of ageing of the oocyte is clearly visible in figure 41. The “fertile window” (fig 40) is determined by two variables: first, the chance of ovulation, and second, the maximum survival time of sperm cells at the beginning of the oviduct (the most likely location for survival). Additional variation in the maximum survival time will emerge as an outcome of the interaction between man and woman.

Fertility is highest in women between 19 and 26 years old. In women aged 35 and over, fertility is roughly halved (fig 41). If one looks at fertility as a parameter for a couple, we know that at older ages, a larger age difference with the man (for example, 5 years older or more) negatively affects the prognosis.

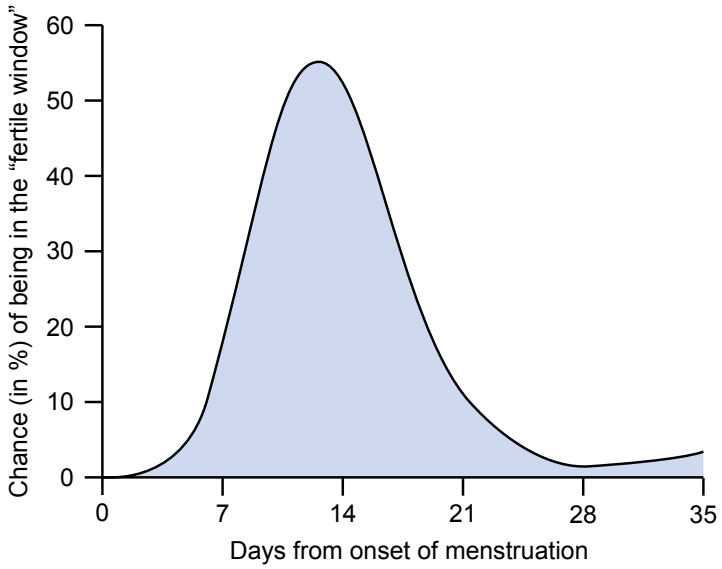


FIG 40

The ratio between **metabolites** of **oestrogens** and progesterone in morning urine was used to predict the LH peak for this graph. Ovulation occurs about 36 hours later (at the end of the follicular phase of the cycle). Based on this, it is possible to determine the probability that the woman is in the fertile part of her cycle, the so-called “fertile window” of about 6 days, covering the 5 days before ovulation and the day of ovulation itself (the Y axis). In the period prior to ovulation, sperm cells are able to survive in the female genital tract (fig 39, 41).

In all situations, it is advantageous if sperm cells are present around 2 days prior to ovulation. An egg will naturally not distinguish between sperm cells introduced through intercourse or intrauterine insemination (IUI), a fertility treatment. Could an insufficiently long interval between insemination and ovulation have contributed to the low success rate (around 10% per cycle) of IUI over the years?

13.5 Fertilisation

When sperm cells reach the sperm reservoir in the isthmus of the oviduct, they are not yet capable of fertilisation; they must first **capacitate** (fig 39). By definition, this means that the sperm are now capable of undergoing the **acrosome reaction** (fig 39). This reaction normally occurs when the sperm cell makes contact with the **zona pellucida** and it is necessary for fertilisation (fig 39, 42).

Capacitated sperm cells also exhibit a different tail movement, with a larger and faster amplitude. This is why the tail of mouse sperm, when the process of **capacitation** is followed under a microscope, appears in the shape of the number 8. The triggered hyperactivity increases the power of motility as well as the swimming speed. Once capacitated, there is no turning back for the sperm and its lifespan is limited. Capacitation with hyperactivity is essential for the sperm to reach and attach to the egg, which is surrounded by a layer of now loosely linked cumulus cells (fig 39). The transition from granulosa to cumulus cells begins as early as in the tertiary follicle (fig 28). The egg plays an active role in this process.

A matrix is created between the cumulus cells, consisting of proteins enriched with sugar molecules, which can be penetrated by the sperm cells (fig 39, 42). This essentially enlarges the target for the sperm. The penetration of this layer of cells can be seen as a kind of third-to-last obstacle test. Overcoming the zona pellucida is the second-to-last step. This occurs at a slight angle, and from the deformations in the zona that visibly occur in the process, researchers deduce that the sperm uses force to penetrate it with its head (fig 39, 42). Once underneath the zona, the cell membrane of the sperm meets the cell membrane of the egg, allowing the membranes to fuse (fig 42). There were already electron microscope images showing this effect around 1970. Gamete fusion takes place in the ampulla of the oviduct (legend to fig 39). This marks the first step in the fertilisation process, which is completed only when the maternal and paternal chromosomes are united in the nuclei of the 2-cell embryo after the first cleavage division (fig 42, 43).

Do the proteins that came with the ejaculate play any role in the fertilisation process? It is believed that they do, as they contain proteins that hinder capacitation. After all, when that happens, the final stage of the sperm's lifespan has been reached. Also, the female reproductive system is primarily equipped to regulate capacitation (fig 39) and usually to delay it. However, there are no coherent insights yet as to whether this forms a truly selective barrier during the "waiting phase"

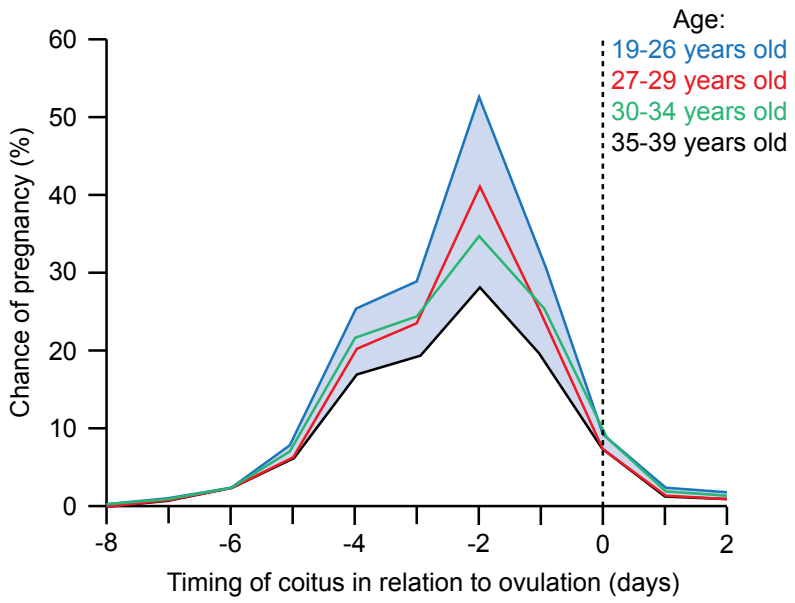


FIG 41

The relationship between sexual intercourse, ovulation and fertility. The indicated age is that of the woman.

of sperm, and, if so, how this works. However, when ovulation approaches, the female inhibitors of capacitation disappear and sperm cells are stimulated (fig 39). Motility is increased and the search for the egg begins. In 2009, a group of German researchers published a more complete picture of gamete interaction. For this, they studied the cow's oviduct *in vitro* and they were able to film the process in an oviduct that had been cut open. All steps of fertilisation, including the ciliary beating of epithelial cells of the wall of the oviduct, could be clearly followed this way. The cilia help move the egg(s) with the cumulus cells all the way to the isthmus if fertilisation did not occur (fig 39). Besides their own motility, sperm cells also utilise the peristalsis of the muscle cells in the wall of the oviduct.

In the German *in vitro* study, it was observed that hyperactivation and, thus, capacitation depended on the presence of the cumulus-oocyte complex: the increased motility caused the

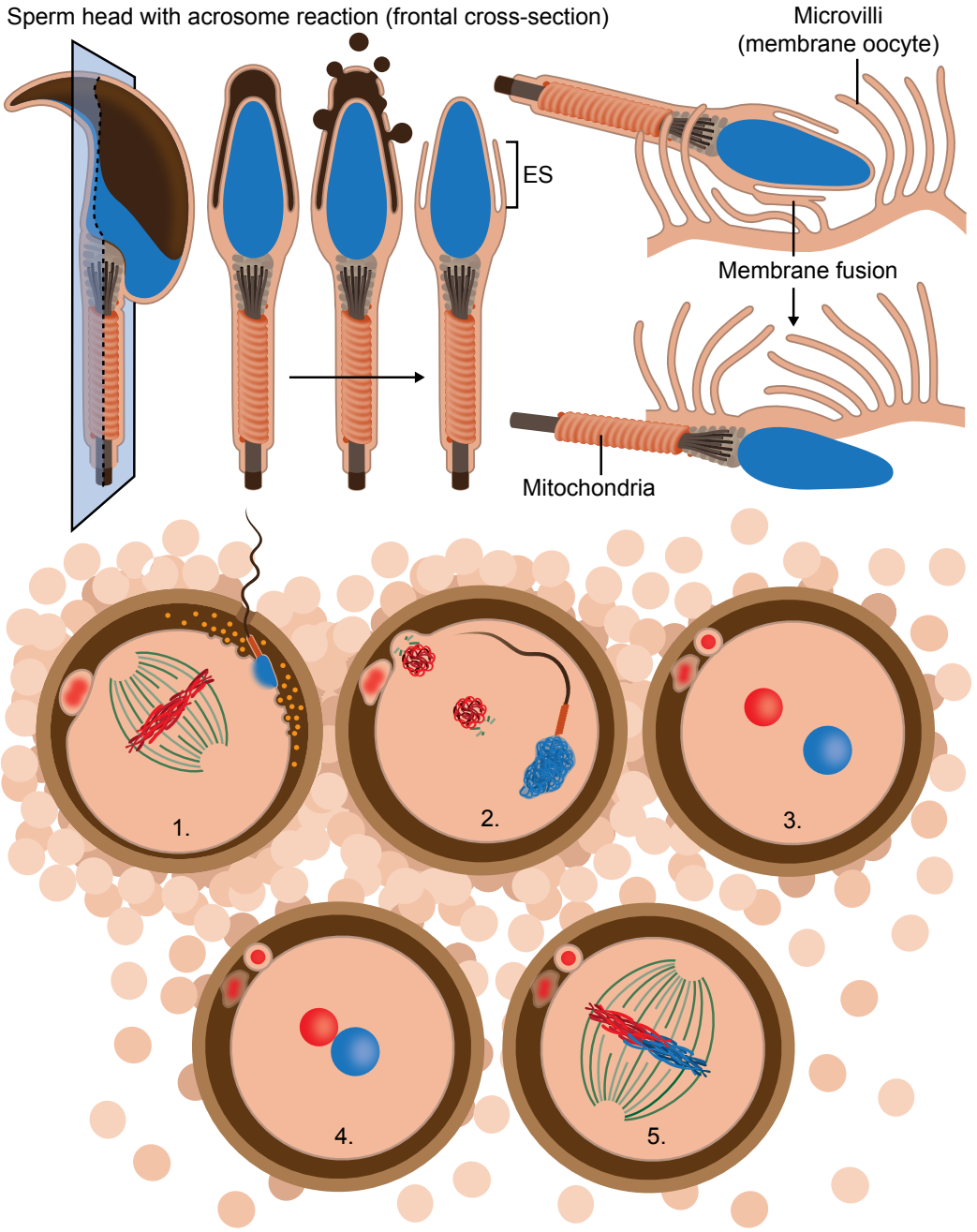

Sperm head with acrosome reaction (frontal cross-section)
ES
Microvilli
(membrane oocyte)
Membrane fusion
Mitochondria
1.
2.
3.
4.
5.

FIG 42

Sequence of developments in maternal and paternal **chromatin** during fertilisation in mice. Fundamentally, the events in the **zygote** in humans proceed in an identical manner but at a slower pace (see the text of Ch13.5.2; mammals differ in specific aspects but not in basic principles). During stage 1, the egg is activated and the second **meiotic** division ends. In stage 2, 30 minutes after entry of the sperm nucleus, the transition from a protamine structure (fig 32) to a **nucleosomal** structure is in progress. Three to at most 4 hours after entry, the maternal and paternal chromatin is packed into the pronuclei (3). In stage 4, the pronuclei lie adjacent to each other and **DNA replication** takes place. Stage 5 illustrates the metaphase of the first (**mitotic**) cleavage division in which the maternal and paternal **chromosome**s are not (yet) intermingled. The panel at the top right shows a magnified view of the membrane fusion between sperm and egg. The fusion, of which the molecular details are increasingly unravelled, is initiated from the so-called equatorial segment (ES) of the sperm cell membrane. In humans, the supply of the centriole (fig 3) is attributed to the sperm cell (fig 39). The few mitochondria of the sperm and its tail also enter the egg. Co-injection of the tail is necessary in human ICSI. In the top left panel, the **acrosome reaction** is visualised once more (see also fig 39) in the distinctive mouse sperm head. The spiral depicts the mitochondria (Ch14) in the so-called “midpiece” of the sperm.

sperm cells to detach from the crypts in the wall of the oviduct. In summary, there is accumulating evidence that the cycle, and in this case ovulation, strongly influences the fate of the male gamete. As intriguing as this research is, it continues to progress at a relatively modest level with an increasing role for molecular observations in farm animals.

Why would it be beneficial for fertilisation to put sperm on hold for 2 days (fig 41)? One might think this could aid in sperm cell selection, but another reason might also play a role. If, as in humans, oestrous behaviour was to evolutionarily disappear, thereby making ovulation “hidden”, men whose sperm survived longer in the female genital tract would be at an advantage. This may be a reproductive trait that natural selection /”hidden” sexual selection taps into. Meanwhile, women who want to conceive naturally benefit from a man whose sperm cells are long-lived in

her body. Thus, it is possible that selection pressures also act on characteristics of the female reproductive system that promote sperm survival.

For many decades, literature has covered attractants that are allegedly secreted by the cumulus-oocyte complex (fig 29) to guide sperm. It is well known that very few sperm cells are present around this complex at the time of fertilisation. The egg may only be fertilised by a single sperm cell, and a high concentration of sperm does not support that. Fortunately, there are mechanisms to avoid **triploidy** caused by fertilisation with two sperm cells. One of them may be the selection of a very limited number of sperm by a chemical attractant. Chemotaxis, movement based on an attractant, is very well known in life forms that use external fertilisation (outside the body). However, evidence for chemotaxis also exists in mammals, where internal fertilisation takes place. The literature focuses on olfactory and taste receptors, also in humans. This field of research is still active. For an insight into the state of this research and the proposed molecular biological mechanisms, see the bibliography.

13.5.1 Oocyte activation and the zona reaction

What happens in the egg (or secondary oocyte, fig 13) when the fusion between sperm cell and oocyte is complete may be regarded as one of the most miraculous biological phenomena. The original cell cycle of the secondary oocyte continues because the second meiotic division has not been completed (fig 39, 42). This division takes place only when the sperm has activated the arrested cell cycle in the oocyte. Around the late 1990s, it became clear that the oocyte produces a few spikes of calcium ions (Ca^{2+}) for this purpose at the signal of the sperm cell. These spikes have been imaged and they can also be artificially manipulated. In 2001, Jean-Pierre Ozil of the INRA institute at Jouy-en-Josas, south of Paris, published intriguing observations on rabbit embryos which, guided by administered Ca^{2+} spikes, had developed **parthenogenetically** until day 11.5 of the gestation period of about 30 days. The researchers observed that the administration regimen after transplantation into a surrogate mother had an effect on the development of these immaculately conceived embryos. Publications on this topic, the embryological implications of egg cell activation, have never abounded, but in 2017, the INRA researchers reported a very intriguing finding. The key question, of course, is whether events that occur in the very early stages of the zygote can have repercussions throughout the rest of life. In Jouy-en-Josas, they monitored the pattern of Ca^{2+} spikes after ICSI (intracytoplasmic sperm injection)

in mice, a technological masterpiece. The spikes cease around 3-4 hours after sperm entry, when the pronuclei have just formed (fig 42). During the *in vitro* period leading up to the embryo transfer, which was performed no later than the 2-cell stage, two different culture media that are accepted in the research community were used (Ch13.7). This variation in culture conditions had a large effect on the number of spikes but not on the number of young mice born after embryo transfer. However, the researchers did observe long-term effects of the different compositions of the culture medium and the resulting spikes on adipose tissue (significant in females) and on brain development (see also Ch16.6). One observation, which is repeatedly encountered in the many articles on early embryonic development, is that the period from zygote to implantation is among the most sensitive found in the **germline**. The findings in the 2017 mouse article thus fully confirm this perspective.

Another early response of the egg cell to fusion with the sperm cell occurs in its cell membrane. Among other things, this change prevents successive spermatozoa from penetrating the zona pellucida. Small vesicles located beneath the cell membrane, called cortical granules, fuse with it and empty their contents into the space between the membrane and zona (fig 39, 42), making the latter impenetrable. This process does not always work perfectly. About 1% of non-aged egg cells of the pig are fertilised by more than one (usually two) sperm. Things occasionally go wrong in IVF dishes as well, resulting in around 3% of dispermic fertilisations, hence triploidy. In the spontaneous abortions from the Boués material in Paris (Ch3.4.2), there were still quite a few triploid spontaneous abortions, with around two-thirds of them being of paternal origin.

13.5.2 Chromatin changes and DNA repair

Simultaneously with the completion of the second meiotic division, the nucleus of the sperm cell is unpacked. It loses its shape in this process. In the mouse, the sperm nucleus initially begins to transform into a sort of flying carpet (fig 42), after which it contracts again. The protamine nuclear proteins (Ch12.6, fig 32) are rapidly removed from the **DNA**. Locally, small residues with a nucleosomal structure will remain, as they were skipped during chromatin remodelling in **spermiogenesis** (Ch12.6, Ch15.2.4, fig 8, 32). Their role in the transmission of information through an **epigenetic** mechanism is disputed, as their general inheritance potential is considered to be low, at least in mice. The rate of **nucleosome** remodelling is also spectacular. The whole process already starts when the sperm nucleus starts to decondense, and takes only a few hours in the mouse. Both the paternal and maternal chromosomes must also be neatly enclosed by

a nuclear membrane again (fig 42). As mentioned in the previous section, only a few hours will have elapsed until the formation of these pronuclei. After 5 hours in mice and 8 hours in humans, DNA synthesis begins in the cellular S phase (fig 4) to copy the chromosomal DNA. Shortly after this (fig 4), the nuclear membranes dissolve and the first mitotic division can begin. The time frame from sperm entry until the first cleavage division takes around 17 hours in mice and around 23 hours (+/- 1 hour) in humans. Knowledge derived from IVF experience (humans) was used for this latter estimate. For the mouse, we know that the timing and outcome of the entire aforementioned process are almost identical *in vivo* and *in vitro*.

As mentioned in chapter 9, the major remodelling of the paternal chromatin is accompanied by the resetting of chromatin memory. The finest illustration of this is the active loss of cytosine methylation (fig 11, 23). DNA repair is involved in this (BER, fig 33). The oocyte's ability to repair DNA damage (Ch11.3) is still present after ovulation. In the early years of the 21st century, we had the opportunity in Nijmegen to take a closer look at the stability of DNA in the mouse zygote. From a *Nature Reviews Genetics* publication from 2000 by the well-known American geneticist James Crow, and from earlier reports from 1988, we already had evidence of a much greater contribution of the father to the *de novo* occurrence of structural chromosomal abnormalities in humans (table 4). In the Nijmegen study, the signalling function of gammaH2AX for a double-strand DNA break came in handy (Ch5.1, table 1). For the detection of these breaks, we made use of **immunofluorescence**. Furthermore, we had access to mouse strains with genetic defects for the repair of double-strand DNA breaks, and we could follow the whole process from damage induction to damage repair over time using mouse IVF. In order to induce double-strand DNA breaks, we used routine techniques consisting of irradiating the sperm prior to IVF, applying low doses of specific chemotherapeutics or other substances that damage DNA, and finally irradiating the zygotes at special times, such as at the beginning of S phase. We knew how many zygotes divided into 2-cell embryos after treatment, and also how many chromosomal aberrations could be seen using the techniques from **cytogenetics** in the metaphase of the first cleavage division (Ch3.4, fig 3, 42).

Following the chromatin change in the head of the sperm shortly after fusion with the egg, we observed between zero and four double-strand DNA breaks, about 18 times more than what is seen in a fibroblast, a connective tissue cell. Were these breaks already present, or are they the result of the chromatin revolution, that has just taken place? We believe the latter. Compared to fibroblasts, sperm nuclei are up to 15 times more resistant to double-strand DNA breaks caused

by **ionising radiation** because of their compact structure. So you can really bombard sperm cells with radiation, and they will still fertilise, but their insensitivity changes completely once they are in the egg cell, even though that is where DNA repair happens again (fig 33). Already in 1974, Tony Searle and his colleagues reported from Harwell that the way to create mouse strains with reciprocal translocations (fig 12) was to irradiate the male genital tract shortly before mating. The subject resurfaced a few years later, when Waldo Generoso of the National Laboratory in Oak Ridge, Tennessee discovered in 1979 that when he subjected male mice to **mutagenic** substances and placed them with females immediately afterwards, the amount of damage caused by those substances depended on the **genotype** (different inbred lines) of those females. His conclusion was that eggs from one inbred line were much better DNA repairers than eggs from another.

In Nijmegen, since we had eggs from mother mice with a defect in the repair mechanism of double-strand DNA breaks (fig 33), we could see how and when during the functioning of the zygote these systems were active. Roughly speaking, two mechanisms can be distinguished. One mechanism, non-homologous end joining (NHEJ), is quick and dirty (fig 33). The other mechanism works very precisely and is based on the mechanism of **homologous recombination** (Ch5.1 fig 14, HRR fig 33). The NHEJ system tackles the double-strand breaks that are brought in by the sperm. The zygote has recovered about 80% of the breaks after 3-4 hours, but is sloppy in the process. The NHEJ system in the egg is also not very diligent compared to that of white blood cells *in vitro*, for example. There are breaks that, probably because they are too complex, simply are not repaired. The HRR system becomes active alongside the NHEJ system when DNA **replication** starts in the S phase (fig 33, 4).

No matter how we looked, the situation in the male nucleus was always somewhat more concerning than in the female nucleus, although the variation between the eggs obtained via superovulation (Ch11.3, Ch16.3.2) could be quite large. Some eggs (now zygotes) were clearly better able to cope with the “male problems” than others. The importance of repairing sperm DNA (Ch12.9.4) in the zygote’s cytoplasm in both the natural situation and also for the practice of IVF (Ch16.5) was already realised in the last century. Together with a collaborator, Robert Edwards wrote the first review article on this in 1996.

Whenever you enter largely unexplored territory with a new approach, you discover extraordinary things. One of the compounds we use to put a strain on the zygote’s DNA (4-Nitroquinoline 1-oxyde, 4NQO) mimicked, as well as possible, the effect of ultraviolet (UV) radiation. When

this compound was applied to zygotes generated from mothers with a defect in HHR, reciprocal translocations already occurred abundantly at very low concentrations (fig 12). This confirms the idea that the zygote is a source of new structural chromosomal abnormalities in the population (Ch16.5, table 4).

Referring to chapter 16.3.1, it is valuable to introduce the latest developments (2022) in the field of study of the zygotic S phase in humans and mice here. Quite astonishingly, but not in contradiction with the above observations, is that the S phase in humans has been found to exhibit marked variation in the synchrony of the initiation and execution of DNA replication *in vitro*. This, as such, is mutagenic, as it predisposes the chromatid to breakage and, hence, to nondisjunction for parts of chromosomes during the first cleavage division. In addition, structural chromosome aberrations may also emerge. Eggs aged before fertilisation, where a shortened G1 phase is expected, are likely more vulnerable to these asynchronies (entry and exit of S phase). S phase can be prolonged into G2 and can even be incomplete (compare with fig 4). The late replicating gene-poor regions (Ch2.3.3, Ch3.4) are the most vulnerable to this. Searching for a biological rationale, a mouse study from the same year strongly suggests that the slow replication rate (fig 5) is associated with reprogramming to **totipotency** (Ch9, box 2).

Currently, double-strand DNA breaks in the zygote have received renewed attention from scientists because of the new possibilities for **gene** editing in the germline using the CRISPR-Cas system. The Cas enzyme/guide **RNA** complex used in this technique also creates a double-strand DNA break in a targeted site-specific manner (box 3). The correct repair of this is, of course, essential for the success of the desired DNA change in the gene to be edited. Therefore, the CRISPR-Cas system has further evolved, now relying on a single-strand break and additional base editing and prime editing (in short in box 3, Ch18.2).

Looking back, it seems that in evolution, repairing double-strand DNA breaks in the zygote was not very high on the priority list. However, this is not true for every type of damage.

Earlier in this book, I discussed that metabolic activity can trigger the release of **oxygen free radicals** that damage a variety of cellular structures, including DNA (Ch12.9.1, Ch12.9.4, Ch14, fig 33). Throughout evolution, cells have armed themselves against this with their availability of protective antioxidants and radical-neutralising enzyme systems. A simple molecule that plays a major role in the protection against oxygen-derived free radicals and the H_2O_2 (hydrogen peroxide) often formed in the process, consists of three amino acids: glutamic acid, cysteine and

glycine. It is referred to as glutathione (abbreviated GSH; the S is the sulphur on the cysteine, which, in its reduced form, contains a hydrogen atom, the H). The SH group of the cysteine plays the key role in its antioxidant action. After contact with e.g. H_2O_2, the oxidised form of GSH, GSSG, is temporarily formed, which can then be converted back to the reduced form by other systems in the cell. The to-be-fertilised egg in the metaphase of the second meiotic division already naturally has a high level of GSH, as this also plays a role in the transformation of the sperm nucleus into the male pronucleus. This high concentration thus also appears to be important for controlling the balance between oxidation and reduction, thereby controlling damage in the cell. An additional significant factor is that the oxygen content in the female reproductive tract is relatively low. When oxidative damage does occur and induces base changes in DNA, it is primarily tackled by the BER system in the zygote (fig 33). The group led by John Aitken, originally from Edinburgh and now affiliated with the University of Newcastle in Australia, provided evidence for this some years ago. So as an egg, you'd better be prepared for oxidative DNA damage, which may have originated in the epididymis (Ch12.9.1) and is thus introduced via the sperm. Apparently, there has been evolutionary selection pressure on the egg to both (a) control **oxidative stress** using the glutathione system and (b) repair oxidative DNA damage of mainly paternal origin. As introduced earlier in this section, the BER system used in the latter process also appears to be involved in removing the methyl group at the CpG **dimer**s in paternal DNA after sperm entry (Ch9, fig 11, 23). A nice example of multitasking in the cellular stages of the female germline.

Once our own research in the mouse had given us a good impression of the amount of double-strand breaks in the DNA of the sperm nucleus, as observed after entry into the egg, the road was open to take a look in human sperm. Some sort of insight had already been obtained in Japan in the 1990s, when a human sperm cell was introduced into a mouse egg via ICSI, upon which the egg became activated (Ch13.5.1). At that point, nothing stops the development of the pronuclei in the zygote (fig 42). Before the first cleavage division began, the Japanese researchers prevented the formation of the spindle apparatus by using colchicine (Ch3.4). The fusion of the gametes had already taken place, but the fusion of their nuclei had not. In this incomplete fertilisation, the mouse chromosomes could be observed and studied alongside human chromosomes, and an ethically unacceptable experiment (a human-mouse embryonic **hybrid**) had just been avoided.

For our purpose, it was sufficient to see that the mouse secondary oocyte extracted the chromatin of the human sperm cell after ICSI. Due to the change into a nucleosomal structure, it was possible to use immunofluorescence on gammaH2AX to see how many double-strand DNA breaks were introduced. In this pilot study, which was just enough to get a clue, we examined a couple of men with poor semen, and two men with a normal **spermiogram** (Ch12.9).

Here we also saw the double-strand DNA breaks, and in roughly the same order of magnitude as in the mouse. However, the number of sperm nuclei without breaks was somewhat lower in the three donors with poor semen than in the normal controls. This highly focused approach to determining the number of double-strand DNA breaks of the sperm in the context of the zygote is very intensive in terms of labour and experimental animals. It is, therefore, not suitable for a practical application such as characterising DNA damage in human sperm.

13.5.3 The maternal to zygotic transition

This transition, which is the essential concept to understand the start of embryonic development, first came up in chapter 11.2 under the heading “developments in the oocyte”. A select group of genes, of which the exact number is not yet known, is at least partially responsible for this. The accumulation of a stock of **mRNA transcripts** from these so-called “maternal effect” genes takes place relatively late in the oocyte’s maturation process. The principle of the role of the oocyte in initiating embryonic development has been known in more detail for a longer time for the fruit fly (*Drosophila*) and the clawed toad (*Xenopus*). For all mentioned organisms, the production of mRNA from the maternal effect genes is essential. While the egg’s mRNA supply diminishes during cleavage, **transcription** on the embryonic **genome** starts and mRNA molecules are produced which, overlapping with the activity of the maternal effect proteins, further advance embryonic development. Think of it as a kind of changing of the guard, one in which these genes play a large role. A short publication in *Nature Genetics* from the year 2000 first mentioned their existence in mice. The female mice described were **homozygous** for a **mutation** in the gene *Mater* (-/-). They have a normal cycle of about 5 days. Their ovaries look completely normal, and so do their eggs. They mate normally, but pregnancy does not occur. The mutation is **recessive**, the +/- females are completely normal. *Mater* has nothing to do with male fertility, the gene is only expressed in oocytes in the ovary. The embryos of the -/- mothers die in the 2-cell stage. By 2016, a total of 44 genes were already known to conform to the same pattern. They have no role in the male germline, but they do have a role from the zygote stage

onwards as a result of transcription during **oogenesis**. There are certainly more of these genes; research-wise, this is not an easy area, but it is developing. From a cell biological point of view, the change from an egg that is about to complete the second meiotic division to a zygote that will split into two equal halves is tremendous. We now know that maternal effect proteins, like so many proteins, work in larger complexes (modules). The module that is currently best described has the name Subcortical Maternal Complex (SCMC) and was discovered by a group of Chinese researchers from Beijing in 2008. The protein of the archetypal maternal effect gene *Mater* is also part of it. This protein complex, which consists of at least six or more members and resides beneath the cell membrane, is involved in various processes. It controls the changes of the cytoskeleton, helps to reorganise cell **organelle**s (fig 2) and to prepare paternal and maternal chromatin in order to express genes. Other maternal effect genes, such as *Trim28* (fig 24), are involved in suppressing retrotransposon activity, maintaining DNA methylation in the "imprint control regions" (ICRs) of clusters of imprinted genes (**genomic imprinting**, Ch9, Ch10, fig 24, 27) and in repairing DNA. TRIM28 is an extreme example of a protein that acts as a connector of functions in a large complex (fig 24). In this case, these functions are all related to regulating chromatin (fig 24).

Yet another and final example of this fascinating group of genes, of how they act as integrators of cell biological processes and are needed to facilitate embryonic development, is *ATRX*. *ATRX* is a gene located on the X chromosome in both mice (*Atrx*) and humans, that has been picked up for causing a serious genetic disorder with the complicated name "alpha thalassemia/intellectual disability syndrome X-linked" (a thalassemia is an abnormality of haemoglobin). It can be found in the OMIM (Ch4.2) under the number 300032. The product encoded by this gene functions as a DNA helicase, a protein that can unwind the double helix, for the purpose of transcription, for example. Investigation of a mouse model for this disease, the results of which were published in 2015, revealed expression of the *Atrx* gene in oogenesis. Mutation led to a halt in development after fertilisation during cleavage: it is a maternal effect gene. The researchers observed signs on the **heterochromatin** near the **centromeres**, when synchronising the chromatin of the paternal and maternal chromosomes in the zygote. Something was also wrong in "managing" double-strand DNA breaks. Thus, ATRX appears to be involved in preventing chromosomal instability (CIN). This phenomenon is a nightmare in the IVF lab, as we will see in chapter 16.3.1. The three-part functions of *Atrx*: its involvement in (a) chromatin modelling, (b), the participation of maternal chromosomes in the spindle apparatus, and (c) keeping chromosomes intact during

the first cell cycles of the early embryo, beautifully illustrate that the alignment of cellular processes is a life skill. Tampering with this damages early life.

13.5.4 Abnormalities during fertilisation

During **meiosis**, fertilisation and the cleavage divisions, Murphy's law applies: anything that can go wrong will go wrong. A relatively frequent error is triploidy. As we have seen, this can result from fertilisation by two spermatozoa (Ch13.5 and Ch13.5.1). Another pathway to triploidy and the appearance of three pronuclei is through an aberrant second meiotic division of the oocyte. In this case, a second polar body is not formed (fig 42), and the chromosomes designated for this appear in a second female pronucleus. This occasionally happens in the IVF lab during ICSI fertilisation (the incidence rate of this is about 1%). Other abnormalities are much rarer and scientifically more spectacular.

My first doctoral student, in the early 70s, looked at egg fertilisation in mice that had been selected for fertility. She witnessed something that we have never observed again afterwards. The second meiotic division had adopted characteristics of the first cleavage division. Two eggs, each with half of the cytoplasm, had now developed, and both were fertilised. The individual that could develop from this is a **chimera**, because the embryo that is created consists of a mixture of cells with two different genotypes, that did not originate from one zygote. Most chimeras will result from the fusion of two zygotes or cleavage stage embryos (fig 43), which is also a very rare event. Nevertheless, it does occur in humans. When there are genetic differences between cells in the body that have originated from one zygote, it is referred to as **mosaic**ism. This is very common at the beginning of embryonic development (referring to cleavage divisions, fig 43), which we now know more about through the use of IVF (Ch16.3.1). Moreover, all women are actually mosaics as a result of X chromosome inactivation (box 4).

Through DNA analysis on single bovine and human **blastomeres**, another remarkable abnormality was found. During the first cleavage division, possibly preceded by abnormalities during fertilisation, all the female chromosomes ended up in one blastomere and all the male chromosomes in the other. This results in a mosaic with aspects of a **gynogenote** and an **androgenote** (Ch10). The blastomeres can also be not only **diploid**, but triploid and **haploid** as well. Further development of such embryos is naturally terminated (Ch10). Abnormalities

during fertilisation usually do not result in offspring, but in very rare instances, as in the case of a chimera, it can occur.

13.6 The cleavage divisions

With the development of IVF in humans, interest in the progression of cleavage stages and the division pattern of blastomeres has, of course, grown tremendously (fig 43). As this is addressed in chapter 16, it will receive less attention here (fig 43). In all well-studied mammals, such as mice, rats, and livestock animals, a delay in the activity of cell division in preimplantation embryos is not a good sign: there is a timetable that needs to be followed (fig 43). Even during early pregnancy, the embryo's environment is highly hormonally regulated.

A delay caused by problems during DNA replication or mitotic cell divisions is undesirable. As we saw in chapter 11.5, early embryonic mortality is a common event. The selection of fit embryos thereby takes priority over saving embryos that lag behind. So do the cell cycle checkpoints (fig 4, and thus the **DDR**, which takes time to be executed) not work properly at all in the early embryo? Research in mice has shown that the tools needed for this are present, but as in meiosis, they are initially not very finely tuned. There are also differences between species; mice perform better in this regard during cleavage divisions (Ch13.6.1). In chapter 16.3.1, we will revisit the chromosomal chaos that can arise in the human embryo due to disruption of cleavage divisions: it is closely related to the practice of IVF, as there is a direct link to preimplantation and early postimplantation mortality.

From the zygote stage onwards, it is not only the correct distribution of chromosomes among the daughter blastomeres that is important. The **telomere** length of individual chromosomes must also be carefully regulated as early as in the preimplantation stage, because paternal and maternal chromosomes differ greatly in this regard. In preimplantation embryos, **telomerase** activity (Ch2.3.3) rises slowly and it peaks at the **blastocyst** stage (fig 43). After that, it slowly decreases until the telomerase gene is no longer transcribed after birth.

During **spermatogenesis**, telomerase is abundantly present and active (Ch12.13), but during oogenesis, much less to no activity is observed. Consequently, the resulting difference in telomere length between the maternal and paternal chromosomes is large in the early embryo.

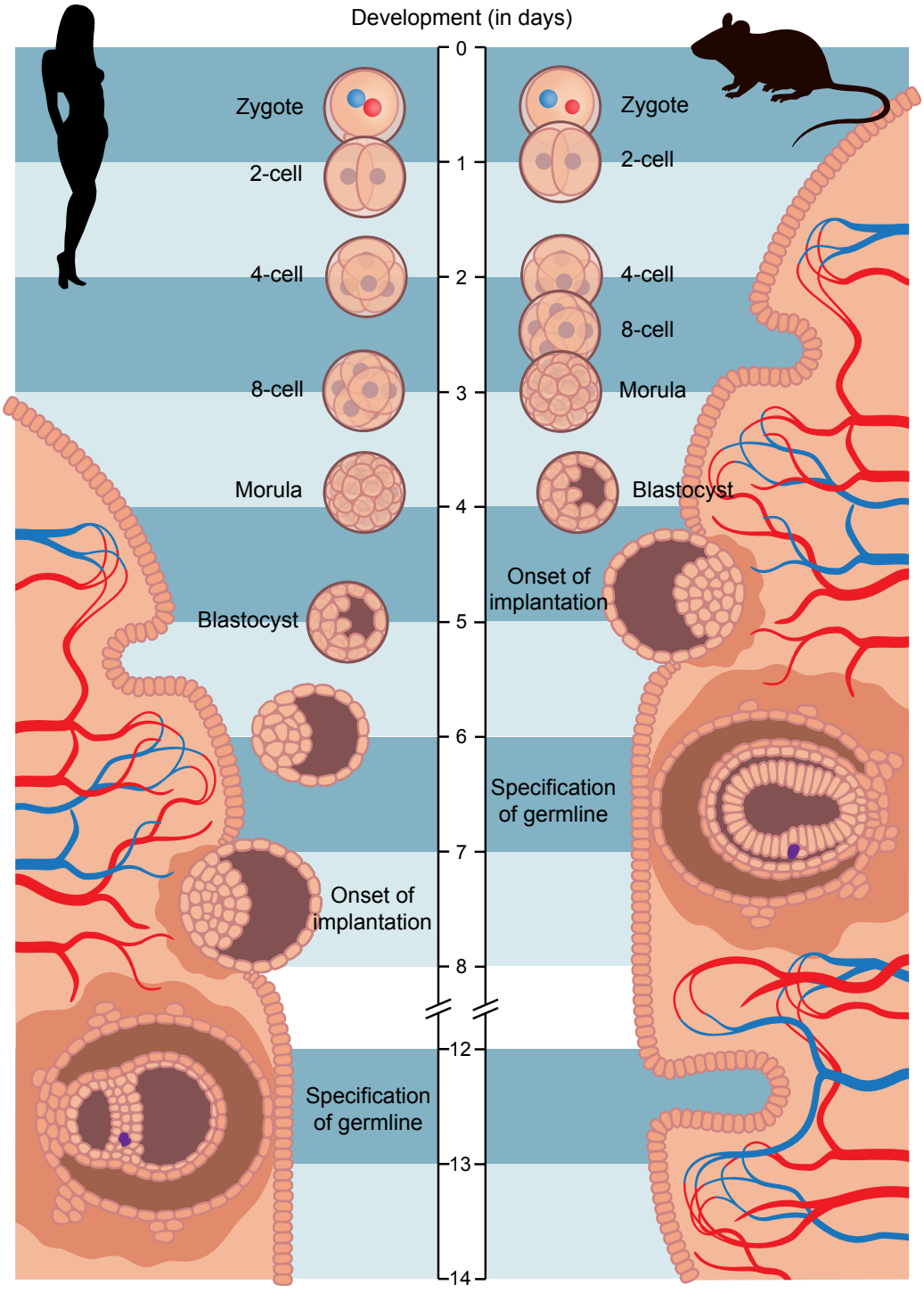

Development (in days)
0
1
2
3
4
5
6
7
8
12
13
14
Zygote
2-cell
4-cell
8-cell
Morula
Blastocyst
Onset of implantation
Specification of germline
Zygote
2-cell
4-cell
8-cell
Morula
Blastocyst
Onset of implantation
Specification of germline

FIG 43

A comparison of the timeframes of cleavage divisions in mice and humans, via the **blastocyst** stage and implantation to the development of the **germ layers** (gastrulation) and the appearance of the first **primordial germ cell**s (in dark purple). During the 8-cell stage, the outlines of the **blastomeres** are at some point no longer individually discernible. They move closer together, their mutual contact intensifies: this is referred to as compaction. In the early blastocyst, one can begin to see a distinction between the "**inner cell mass**", from which the **epiblast** and thus the **embryo** will later develop, and the surrounding cells of the trophoblast. The **zona pellucida** is shown as a thick line.

This is perhaps why another recombination-based mechanism is also active in the preimplantation embryo to lengthen telomeres, a system that ultimately eliminates the difference in length.

13.6.1 Is the egg oversized?

Yes, the egg is oversized in both its size and content. There is room for some things to go wrong up to the development into a blastocyst (fig 43). Only after that, growth in mass occurs again (besides an increase in the number of cells).

Already around 50 years ago, the question of how many embryos could be obtained from a zygote was raised in the ARC Reproductive Physiology Unit in Cambridge. The maximum achievable number in sheep turned out to be quadruplets, created from a whole series of 8-cell embryos that had been "manipulated" apart into four sets of 2-cell embryos after removal of the zona pellucida. The researchers then placed these 2-cell embryos back into empty zona pellucidae from ovulated oocytes. Since good culture media for livestock embryos were lacking in those years, they further cultured these 2-cell embryos in the ligated oviducts of a sheep in the appropriate stage of the cycle. The early blastocysts obtained in this way were then transplanted into the uterus of a third ewe. Such an approach, searching for the most advanced blastomere stage combined with the smallest amount of cytoplasm that still allows complete development of life, served as a proof of principle. A little later, the test was repeated in cattle in the same institute. In that experiment, the creation of these monozygous multiples stopped at **triplet**s, which, however, does not prove that quadruplets would not have been possible. In rhesus monkeys, a

quarter embryo has once led to a birth, and in the IVF lab at the Free University of Brussels, led by professors Liebaers, Van Steirteghem, Tournaye and Devroey, four normally appearing blastocysts could be grown from a fertilised human egg. This is not entirely incomprehensible. In humans, identical twins occur in 0.3% of spontaneous births, and identical quadruplets are also very incidentally born.

In mice, there has also been extensive research on the capacity of the production of monozygous multiples. It remains limited to twins, perhaps because the egg is slightly smaller and also contains less yolk material (it is clearer). For a very long time, it was assumed that the eight cells prior to compaction (fig 43) were equal and thus fully totipotent (box 2). A recent publication from Michele Boiani's group in Münster described that "each for themselves development" was not equally successful in every 2-cell embryo. Only one-third of the large series of embryos showed that success. In the others, only one of the two could develop into a blastocyst. The true embryo forms when the **inner cell mass** in late blastocysts divides into a cell layer, the primitive endoderm, and a group of cells known as **epiblast**s, out of which the embryo eventually develops (fig 43). In the mouse, the epiblast initially normally consists of about nine cells, but how many of those are minimally needed? In 2016, additional scientific insight once again emerged from Cambridge. By inhibiting the checkpoint for the proper attachment of chromosomes to the spindle apparatus (fig 4) in the division from the 4- to 8-cell stage, deviations in the diploid number emerged in about two-thirds of blastomeres. That is too much for normal embryonic development. However, if you reduce that fraction by 50% by creating a new 8-cell embryo (a chimera) that includes four untreated blastomeres from a control embryo, it succeeds. It is estimated that in the mouse, two to three cells in the 8-cell embryo can be chromosomally abnormal without compromising its development. This means that some things can go wrong in the first few days without jeopardising the chances of a continuing pregnancy (see also Ch16.3.1).

13.7 *In vitro* techniques

In vitro research on the **metabolism** of mouse embryos started in the 60s with the work of Ralph Brinster from the veterinary faculty of the University of Philadelphia, USA. This group discovered that up to the 8-cell stage, a mouse embryo can only use pyruvate (a product of glycolysis) as an energy source. It is only from the 8-cell stage onwards that glucose can also be metabolised.

In those days, two “recipes”, which did not substantially differ from each other, were used for the composition of the media used to grow blastocysts from zygotes or from 2-cell embryos. In the USA, researcher Wes Whitten from the Jackson lab in Maine had given his name to the Whitten medium, while the British variant in the UK was called Whittingham’s medium. Both media consisted of relatively simple salt solutions with a pH buffer (bicarbonate, the effervescent tablets from the drugstore) and CO_2 in the air above it for equilibrium. Albumin from bovine blood plasma, available in a range of purities, was used as a protein additive. These media have been further developed since, which included the addition of amino acids. These more complexly composed media were, for instance, used by Jean-Pierre Ozil, among others, in his research on egg activation (Ch13.5.1).

Henry Leese from the University of Hull in the UK devoted many years to studying the metabolism of preimplantation embryos up to the blastocyst stage *in vitro*. He is the one who proposed the “quiet embryo hypothesis”. The basis for this hypothesis lies in the compelling evidence of an almost dormant existence of the embryo in the first days of its development. As noted earlier in chapter 13.6.1, there is no increase in cell mass up to the early blastocyst stage. During the initial cell divisions, the stock of mitochondria, as well as that of other cell organelles, is repeatedly divided. Consequently, the cellular energy requirement is relatively low in the early embryo. Moreover, no energy is lost to communication between individual blastomeres, as they are autonomous until the 8-cell stage. That being said, the energy supply regulation is a complex process that is too specialised for further discussion here.

All experts agree that the cleavage stage embryos are sensitive to external influences. In apparent contradiction to this, there is a widely shared observation that no cell selection takes place during this period. **Apoptosis** does not really happen yet, this becomes effective only at the blastocyst stage (fig 43).

In 2003, a former PhD candidate in Wageningen, Marleen Boerjan, published a study in collaboration with Ann Van Soom’s group in Ghent on the effect of oxygen concentration on the *in vitro* development of bovine embryos. At that time, it was not yet widely accepted that it would be wiser not to let embryos develop under atmospheric conditions; rather than a level of 20% oxygen, a level of up to 5% should be used. In 2012, the IVF lab in the former AMC hospital in Amsterdam published a Cochrane database review on the effect of using 5% oxygen versus 20% in global IVF practice in humans. The conclusion is clear, low oxygen improves the “take home baby rate”. Henry Leese stated that in 2013, the conditions in the oviduct at the time of

transport of the cleavage stage embryo to the uterus (fig 43) were still not well understood. One thing that is known, however, is that oxygen is largely kept at bay here. This aligns with what we already knew about tertiary follicles growing into Graafian follicles in the ovary (fig 29).

CHAPTER 14

Mitochondrial inheritance

Mitochondria (fig 2) are cell **organelle**s with a rather variable shape and a size ranging from 0.5-10 µm. Virtually every cell has mitochondria. They play a very central role in the cell, involving a variety of aspects. Their most well-known function is that of being the powerhouse of the cell, indispensable as they are in utilising oxygen in the **metabolism** of pyruvate and fatty acids to produce ATP (adenosine triphosphate). ATP is the currency for all processes in the cell that cost energy. However, the cellular significance of mitochondria is much broader. They are involved in steroid synthesis, assist in the storage of calcium ions, play a role in the regulation of **apoptosis**, and also regulate heat production in warm-blooded animals.

The now widely accepted explanation for the origin of mitochondria is associated with the name of the American biologist Lynn Margulis. Around 1970, she made it plausible that during the times of the earliest life forms, fusions occurred between a (primitive) unicellular organism and a simple bacterium. This hypothetical fusion event is now known by the term endosymbiosis. While it was initially strongly criticised, much additional evidence has been provided for this origin story, and it is now widely accepted.

What is the current state of the descendants of this endosymbiotic form of coexistence? Every mitochondrion in the cell is now enveloped by a double membrane (fig 2). The inner membrane originates from what was once the membrane of the integrated bacteria. The outer membrane must have formed during the fusion because it has characteristics of the primal membrane of the unicellular organism that acted as a kind of host for the bacteria. Finally: a mitochondrion possesses its own **genome**, formed by **DNA** that still carries bacterial features from ancient times.

There is a clear division of tasks between mitochondrial DNA (mtDNA) and nuclear DNA. The mitochondrion usually contains several (two to 10) circular DNA molecules that are supported

by proteins which provide structure and function. These DNA-binding proteins are encoded by genes in the nuclear genome and, after their synthesis, they are imported from the **cytoplasm** into the mitochondrion. In total, a mitochondrion is composed of as many as 1,500 different proteins. All mitochondria together sometimes make up 10% or more of the total protein mass of a cell. The regulation of the number of mitochondria per cell occurs through division of the organelle and is controlled by the nucleus. This process happens independently of the cell cycle progression. Through constriction, a mitochondrion is divided into two parts along its longitudinal axis ("mitochondrial fission"), and its mtDNA is divided over the two new mitochondria. Of course, this is preceded by the **replication** of mitochondrial DNA into multiple copies per mitochondrion. This is done with mitochondria-specific tools, with the nuclear genome providing the recipe and controlling the process.

The circular mitochondrial DNA, the mt genome of the different species discussed here, varies very little in size. In humans, this DNA has a length of 16.6 kb, most of which is coding DNA. It contains 13 genes for proteins and 24 genes for **RNA** molecules (22 tRNAs and two rRNAs) with a function in the **translation** of those 13 proteins. Mitochondria even have their own ribosomes for this purpose. In this sense, the mitochondrion is still self-sufficient. The 13 proteins are subunits of five much larger protein complexes that lie against the inner membrane of the mitochondrion, where they are responsible for a cascade of reactions, ultimately producing the cellular energy fuel ATP. This process is usually simply referred to as oxidative phosphorylation (OXPHOS). Most of the proteins in the OXPHOS protein complexes, as well as all the other 1,500 proteins (mentioned above) with a role in processes like the import of ions, pyruvate and fatty acids, metabolic conversions, the formation of mitochondria and ribosomes, and the mtDNA replication (just mentioned), are encoded in the nucleus. They are synthesised in the cytoplasm of the cell, after which they are imported into the mitochondrion and correctly incorporated into the organelle. In cell types that are highly energy-demanding, such as neurons and muscle cells, but also in cells involved in hormonal systems, mitochondria are found in relatively high numbers and they are also more active than in cells that are in a state of metabolic rest.

The regulation of the quantity of mitochondria and their activity is an extremely complex process in the **germline** (and during early embryogenesis after the formation of a **zygote**). During **oogenesis**, a huge population consisting of several hundred thousand mitochondria is established. In an egg cell, which already has an unusually large size (Ch1), this population can occupy up to 30% of the volume. In the oocyte, mitochondria remain responsible for ATP supply via OXPHOS, but they receive the fuel for this (pyruvate) from the granulosa and cumulus cells.

During **spermatogenesis**, a sort of reversed process occurs. Due to complex morphological changes, the volume of the eventual sperm cell decreases by about 40-50 times, and 80-90% of the original quantity of mitochondria is lost. In this process, the mitochondria undergo gradual changes in both shape and position in the cell. Finally, when only a few dozen to around 70 mitochondria remain (human), they accumulate behind the head at the beginning of the tail (in the so-called midpiece, fig 42). Throughout the entire maturation process, the OXPHOS-mediated ATP supply, for which fuel is received from the Sertoli cells in the form of lactate (which is converted into pyruvate), is active.

The life cycle and population dynamics of mitochondria are thus very different during the formation of oocytes and sperm. At the onset of fertilisation, almost all of the mitochondria originate from the egg cell. What happens next to the paternal mitochondria was discovered by Peter Sutovsky from the USA: the egg labels the paternal mitochondria by attaching a small, very common protein called ubiquitin, causing them to fall prey to the protein degradation machinery of the cell. As a result, the small number of paternal mitochondria has been cleared out long before **blastocysts** are formed. Ubiquitination is only the last step; very recently, it has been discovered that those paternal mitochondria are incomplete, as they are devoid of mtDNA.

After zygote formation, the vast population of maternal mitochondria is distributed among the **blastomeres** during the subsequent cleavage divisions (Ch13.7). Since there is no synthesis of new mtDNA and thus no biogenesis of new mitochondria during this early period, the number will gradually decrease per cell division. Of course, the cellular tasks must still be fulfilled. The mitochondria continue to supply energy through OXPHOS, and they are also involved in some of the other aforementioned physiological processes required for the regulation of cell division. These include the extensive **epigenetic** changes in **chromatin** (Ch9) and the communication between cells in the early **embryo**. It is therefore not surprising that when their number and state of functioning fall below a certain threshold during oogenesis, due to ageing and/or other internal or external causes, it compromises embryonic development.

During later stages of early embryonic development, the production of new mitochondria slowly resumes. In mice, the number of mitochondria per cell increases again after implantation, when embryonic differentiation accelerates (fig 43).

Together with the fact that mitochondria carry their own genome, the enormous variation in metabolic functions and life stages of mitochondria in the different cells that constitute the germline (including early embryonic development) makes the story of their inheritance and genome evolution very complex. After all, the mitochondrial genome also experiences **mutation**

pressure. One of the main causes of mutations in mtDNA comes from base damage that can be inflicted by free radicals (H_2O_2, O_2•-). As particular OXPHOS protein complexes in the mitochondrion are the primary source of these free radicals, the mitochondrial genome is always in a kind of danger zone. However, the risk of DNA damage is highly dependent on the metabolic specialisation and the metabolic activity of the cell type in which they occur. One of the modifications that can occur due to the action of free radicals on DNA is the formation of 8-oxoguanine (fig 33). The concentration of this found in DNA (mtDNA or nuclear DNA) can be used in research as a measure of the extent of damage that has been inflicted. Among other things, antibodies that can recognise 8-oxoguanine in DNA are used in this type of analysis (fig 33). Other types of DNA damage, which are caused by replication errors or toxic substances, can also occur in mitochondria. The mitochondrion has its own tools for repairing these forms of damage. While the machinery for this is composed of nucleus-encoded proteins, its operation is clearly different from that for the repair of DNA in the cell nucleus. There are functional similarities since well-known DNA repair pathways such as BER and NER (see fig 33) are also active in mitochondria, but they are seemingly less effective. Mitochondrial DNA is, in fact, quite mutation-prone, approximately 20 times more than nuclear DNA in mammals.

So the mtDNA in early embryos is of maternal origin. If you have an idea about the multiplication of mitochondria during oogenesis up to the stage of the mature egg, and you know the mutation rate per mtDNA base pair per replication event, you could estimate the influx of new mutations per generation. However, that is not the approach taken in recent review articles on the genetic variation of mitochondria. The starting point that is usually taken, at least for humans, involves the observation that a disease resulting from a distinct mutation in mitochondrial DNA is present in one in 5,000 births. If this mutation is present in (more than) 60-80% of the mitochondria in the egg cell, and this percentage is maintained in the body cells (**somatic** cells) of the developing individual over the course of development and growth, the **phenotype** associated with this mutation emerges, characterised by a severe metabolic or physiological impairment. When the population of mitochondria in a cell contains both mutant and non-mutant mtDNA, it is referred to as heteroplasmy (as opposed to homoplasmy, where either only the non-mutated or only the mutated form is present). New technologies used to determine the DNA base sequence have revealed that homoplasmy is much less common than previously thought; it is actually rare. In fact, no one appears to be purely “homoplasmic”: cells with some mutated mitochondria are always found in tissues. It is even more likely that there are often small subpopulations which together contain a spectrum of different mutations.

Something that has greatly surprised researchers studying the transmission of mitochondria to the next generation, is how few generations it can take to transition from one **dominant** mtDNA **genotype** to another. The presence of a few mitochondria with a DNA alteration can already be enough. These are findings from around 1980, discovered in sheep and Holstein-Friesian black-and-white cattle. To explain how a new mutation can rapidly assume a dominant position, scientists introduced the "bottleneck hypothesis". Somewhere in the female germline, there is a bottleneck where the number of mitochondria per cell is very low. When these cells divide, chance can play a role in the distribution of different mitochondrial genotypes among the daughter cells. If that results in the emergence of a population with a majority of one particular mitochondrial DNA type, and this "coincidence" repeats itself a few more times, you can understand why a change can occur so quickly. What is a realistic estimate for the number of mitochondria during the bottleneck, and at which stage of the germline does this occur? It was not until 2008 that an article on this was published out of Cambridge, providing these estimates for **primordial germ cell**s in mice. The technique they used, determining estimates based on the strength of a DNA signal, had only recently been developed with sufficient accuracy at that time. Counting mitochondria using the electron microscope is far too laborious. The bottleneck appears to be in the early stages of the primordial germ cells, in which an average number of 450 mitochondria per cell was observed, a number that continuously increases thereafter. Assisted by further technical advancements, several groups, mostly centred in Cambridge, published an even more detailed picture for humans in 2018: their estimate now ranged from 200 to 400 mitochondria per primordial germ cell. In these mitochondria, the researchers observed the emergence and proliferation of new mutations, and they also found that selection against unfavourable mutations takes place. Both findings contribute to the attempt to understand the dynamic landscape of the transmission of mitochondrial genetic information from mother to offspring.

What remains most remarkable is that the characteristic mutations in the mtDNA, which serve as indicators of the evolutionary history of humanity (search for "mitochondrial Eve"), must have occurred in a single copy or repair event. After that, the mtDNA variants must have multiplied into a state of dominance in human populations. This enables us to track the migratory behaviour of individuals from this population and their descendants. It is likely that preferential replication or selective destruction of certain mitochondria also plays a role in the rise or disappearance of distinctive mtDNA variants within populations.

The number of serious diseases caused exclusively by mutations in mitochondrial DNA is not very high. These diseases can really appear out of nowhere. They can manifest after only

a few generations of “enrichment” of mitochondria with the causative mtDNA mutation once a sufficiently high level of heteroplasmy is reached (see above).

CHAPTER 15

Epigenetic aspects of inheritance

15.1 Introduction

An old doctrine in biology revolves around the separation between **phenotype** and **genotype**. This implies that the phenotype cannot interfere with the genotype: information transfer always flows from genotype to phenotype. This proposition is known as the Weismann barrier and dates back to 1893. However, we have known for some time that this proposition presents an overly simple picture of how inheritance can take place. Therefore, I will make an attempt to explain new knowledge on this matter in more detail in this chapter.

In quantitative genetics, the formula P = G + E is used (Ch6.1). There are no influences on phenotype P other than (a) the **allele**s of the **gene** set that segregate and recombine according to Mendel's laws (summed up as G) and (b) environmental factors (wrapped up in the term E; the formula can be further extended to include the interaction between G and E, known as genotype-environment interaction). These two influences are expressed as contributions to deviations from the mean (Ch6.1, fig 17). However, in this formula, the mathematicians do not take into account any other ways of transmitting information from one generation to the next. Their approach is reinforced by the applicability, for example in selection experiments, of formulas from quantitative genetics for traits based on many genes (Ch6). This method is related to what is referred to in genetics as the "modern synthesis" (which, incidentally, dates back to 1942).

In its simplest form, this modern synthesis combined Darwin's ideas (the importance of natural selection) with Mendel's laws (explaining how genetic (allele) variation within parents is expressed in offspring, Ch5.3). The genetic variation that exists between individuals is a part of it. It also implies that the genotype interacts with the environment to reach phenotypic variation.

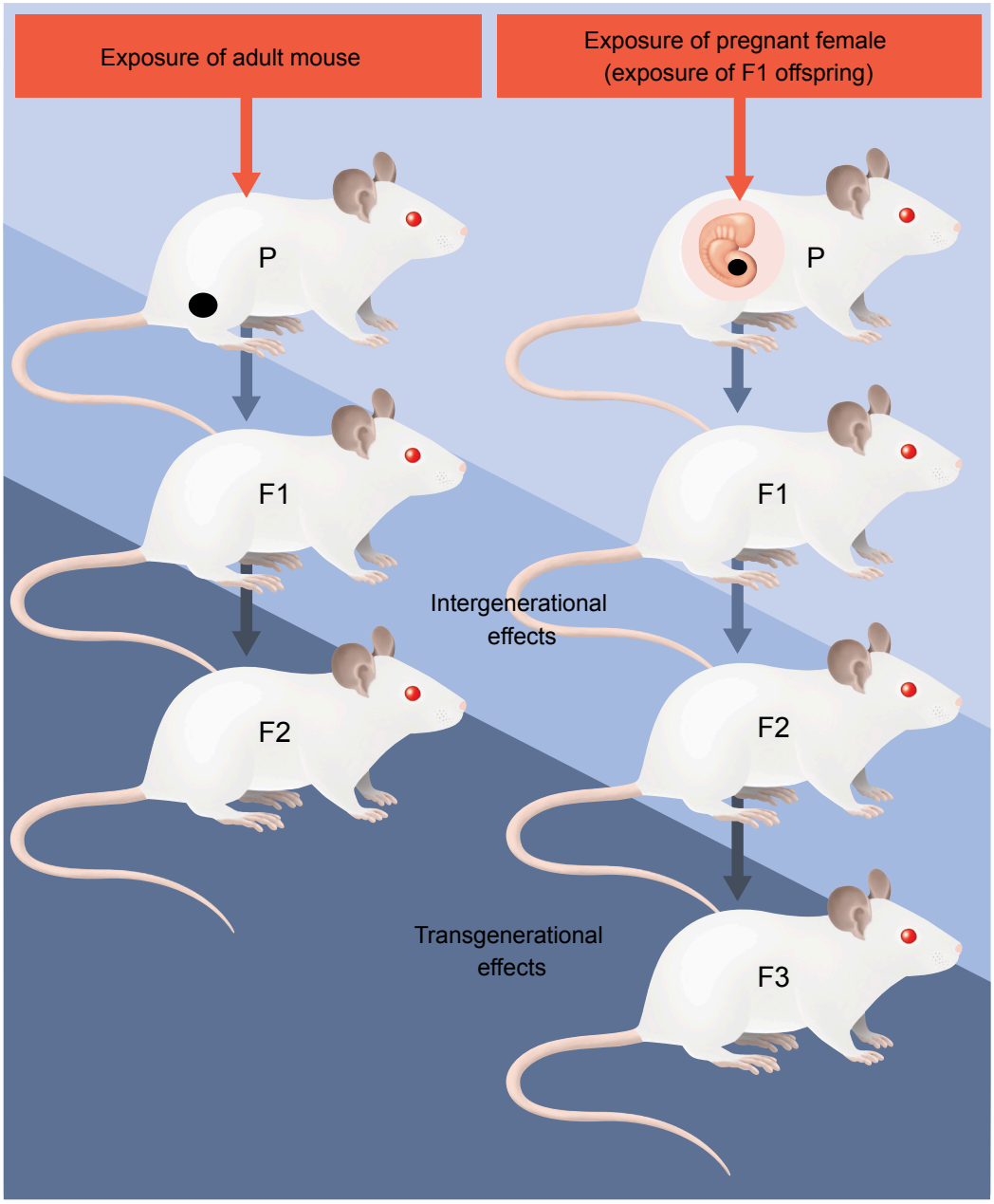

Exposure of adult mouse
Exposure of pregnant female
(exposure of F1 offspring)
P
P
F1
F1
Intergenerational
effects
F2
F2
Transgenerational
effects
F3

FIG 44

The P (for parental) represents the treated generation. In adult males, this concerns the exposure of cells involved in **spermatogenesis** (black), in adult females it concerns **oogenesis** in the follicles (idem). When the female is pregnant, these experiments involve the exposure of **primordial germ cell**s until the end of the sex differentiation of the early gonad (Ch8, black). In the case of an intergenerational effect, the **germline** of the parental generation (P) has "experienced" the stressor (i.e. has been irradiated or chemically treated, irrespective of the stage of life, prenatal or postnatal). In the case of a transgenerational effect, the germline of the subsequent parental generation has not experienced the direct influence of the stressor. Generation P is also referred to as F0.

In natural selection, the selection is based on phenotypes rather than **DNA** parameters. The better the phenotype predicts the genotype, the more effective that selection is. However, that accumulation of genetic variation, the emergence of new alleles, occurs slowly (table 3). This also means that when conditions change rapidly, it will be challenging for organisms with a longer **generation interval** to adapt to the new conditions through natural selection (according to Darwin). It is even more difficult when the population is small and/or fragmented. Experiments with various life forms (for example, the nematode *Caenorhabditis elegans*, which is often used in genetic research) provided evidence that there are additional ways to transmit information to the offspring, namely through the **chromatin** of the **gametes** and/or via the "instruction manual" of that chromatin. This is referred to as **epigenetic** intergenerational or, when it persists for more than one generation, transgenerational inheritance (fig 44). Especially Eva Jablonka from Tel Aviv University has been actively spreading this line of thought for years. In her publications, she provides examples of this across different life forms.

A mode of inheritance like this is known as "soft inheritance", and mechanistically, it is completely different from the DNA sequence-based "hard inheritance". In soft inheritance, there is no segregation à la Mendel, but there is a shift of the phenotypic variation in the offspring in a particular direction, namely that of the parent. The effect can also be reversed within a few generations; it fades away and is thus reversible. Since special and often more extreme environmental conditions are used in experiments on this in rats and mice (mostly inbred lines), these animals appear to be sensitive to soft inheritance. This seemingly confirms the old ideas about heredity put forward by the French biologist Jean-Baptiste de Lamarck. Before Darwin published

his theory of natural selection, Lamarck proposed that traits acquired during life due to environmental influences are heritable. That way, a bodybuilder could have very muscular offspring.

This way of thinking, which effectively represented a shift from “nature” to “nurture”, was turned into a communist ideology by biologist Trofim Lysenko under the Stalin regime in the Soviet Union, and it essentially denied the existence of genes. Only environmental factors were acknowledged as determinants of the traits of plants and animals. Attempting to modernise food production in the Soviet Union through adjustments in crop growth, conforming to these delusional ideas, was one of the factors that led to an unprecedented food crisis and the deaths of millions of inhabitants. Perhaps, this is partly the reason why, as long as we lack a better understanding of the mechanisms of epigenetic inheritance, there may be a somewhat heavy atmosphere around this topic when it comes up in the literature and in conversations among colleagues. Nevertheless, the results obtained in experimental animals have thus far given the subject momentum. This is mainly because the possibilities for analysis of DNA methylation (fig 11), **RNA** base sequence determinations, and chromatin have advanced so much. These analytical methods, which are becoming increasingly sensitive, can be combined with (conditional) **knock-out**s and **knock-in**s in mice (box 3) to unravel the mechanisms of epigenetic inheritance. Environmental conditions that seem to lend themselves to the identification of epigenetic effects in inheritance, such as chemical pollution, mental stress, and changes in diet (and their associated phenotypes, such as obesity), also contribute to the growing interest.

Even now, we do not know all the mechanisms and details of the epigenetic phenomena that occur between generations, and there are plenty of questions that remain unanswered. How can the recipe for chromatin changes become anchored in the **germline** of the next generation and the one thereafter? Nonetheless, numerous indications of the existence of transgenerational epigenetic inheritance have been found in mice and rats (fig 44).

15.2 Intergenerational and transgenerational epigenetic inheritance in experimental animals: the experimental design

The fact that the **embryo**nic chromatin is reset in the germline (Ch9, fig 23), first in the period before implantation and then again after the determination of the **primordial germ cell**s, actually serves as a barrier against excessive environmental interference with the “instruction manual” of the chromatin of the gametes (before and after it has been transmitted to the next generation).

Nevertheless, there are experimental approaches that can be used to induce and study epigenetic inheritance. Often, the use of experimental animals is central to this, and the emergence of a pattern of epigenetic inheritance in, for example, mice, rats, or model organisms such as the nematode *C. elegans* or the fruit fly *Drosophila* is related to exposing the germline to stress. To label something as epigenetic inheritance (and not as a **mutation**), you must be certain that the germline stage, and thus the cell from which the individual under study develops, has not been in direct contact with the stressor (fig 44). If the treatment takes place during the embryonic stage and the mother is the F0 (or P for parental), then you will only have evidence of a transgenerational epigenetic effect in the **F3** generation. If you treat the germline of an adult animal, this effect can already be observed in the **F2** (fig 44).

15.2.1 Initial effects in outbred rats

Since 2005, an American group that uses outbred rats has been working on this topic through a steady series of publications. Pregnant females are treated between day 8 and day 14 of the embryonic period. Experiments often extend into **F4** (fig 44, here meaning crossing among descendants). The output initially covered overall pathology. The stressors that produce a transgenerational epigenetic effect are often related to endocrine influences. Examples include vinclozolin, a fungicide (and **endocrine disruptor**, Ch12.10) as the "founding father" of this type of research, along with BPA (also an endocrine disruptor, Ch11.4.3, Ch12.10), and plasticisers in plastics (phthalates, likewise). However, this transgenerational and probably epigenetic effect can also be induced by a series of other chemicals and treatments (diet, stress). Recently, in 2019 and 2021, glyphosate (known from "Roundup", a herbicide) also entered the scientific literature this way. An important variable is the period in the development of the germline during which the stress is applied. It seems that the periods associated with changes, including the resetting of chromatin (Ch9, fig 23), are the most sensitive to stress.

15.2.2 The first mouse model: agouti yellow A^{vy}

The first mouse model discovered for the detection of transgenerational epigenetic effects carries a specific **dominant** allele for the gene that causes wildtype coat colour (fig 45). This gene, with the fancy name "agouti", causes a typical grey-brown colour in mice and other rodents such as rabbits. These animals are almost always **homozygous** for the wildtype allele *A*. The most

well-known alternative allele is labelled with the symbol *a*. This **recessive** allele will result in a black coat. In the past, you would commonly see black *a/a* rabbits in the forest, and hopefully they still exist in the wild. In 1962, a new dominant allele was described, known as A^{vy}; vy stands for "viable yellow". This allele was formed through the action of a **retrotransposon** that ended up just before the **promoter** region of the *A* **locus**, via an RNA intermediate (Ch2.3.3). This type of retrotransposon, which is originally of viral origin, can take over the regulation of a nearby gene through a DNA sequence that is present in it; it has its own promoter activity and can thus control the gene differently than normal. This explains the change in coat colour to yellow (fig 45). The retrotransposon insertion at the start of the *A* locus owes its promoter activity to a clustering of CpG motifs in the DNA base sequence, which is something that occurs in many promoter regions and often affects their "strength" (Ch2.3.3).

When you blow into the fur of a grey mouse (or rabbit), you will see characteristic ring-like zones from the base outward, going from black to yellow and then back to black at the tip: every single hair has these three zones. In the presence of A^{vy}, the black is not expressed. The genotype of the "readout mouse" in experiments to unravel the "soft side" of inheritance patterns that, so to speak, evade Mendel's gaze, is A^{vy}/a (fig 45). This genotype is introduced in an inbred line: after all, you want as little disturbing genetic variation as possible, as that might obscure what you are looking for. The first description of the "utility value" of the A^{vy}/a mouse line for epigenetic research came out of Emma Whitelaw's laboratory in Brisbane, Australia, in 1999. It is not the case that every carrier is yellow due to the dominant A^{vy} allele. The coat colour ranges from yellow to almost indistinguishable from wildtype agouti-grey. For statistical analysis, three phenotypes are defined: yellow, mottled, and pseudo-agouti (fig 45). Fathers of each type produce offspring with the same spectrum of colour variations. However, this was different for mothers. Although each mother still produced all variants, grey mothers tended to have more grey offspring, while yellow mothers tended to have more yellow offspring. Moreover, this tendency is transmitted via the mother, while this transmission does not occur via the father. It is unclear why that is the case.

It soon became apparent that coat colour was associated with the methylation level of the CpG-rich promoter region of the retrotransposon. With a higher degree of methylation, the promoter of this DNA segment was switched off, and the promoter of the *A* allele that was still intact took over the regulation of expression (fig 11, 45). This resulted in offspring with a grey coat. When the methylation level is low, the retroviral promoter is dominantly active, and the mice will be yellow. In situations where both the retroviral and the natural promoter are active in different

cells, a variable outcome is possible, and the mice look mottled (fig 45). The expression level of the *A* sequence will ultimately be determined by local chromatin structure changes characterised by the **histone** code present at that site (Ch2.3.2, table 1), together with the methylation level of the CpG motifs in each of the promoters. Hence, the degree of overactivation of agouti in the A^{vy} allele can vary between cells in the hair follicles of the mouse. The **mosaic** pattern of gene expression across the coat, characteristic of the mottled phenotype, resembles the outcome of X chromosome inactivation (box 4).

But why is the A^{vy} retrotransposon still so "unstable" regarding its effect on the expression of nearby genes? Previously (Ch2.3.3, Ch9, Ch11.1, Ch12.5.2.), it was discussed that from early times in evolutionary history, when **retrotransposons** were developing the blueprint of chromatin/DNA, a defence mechanism also had to be set up to keep their genetic activity under control. This was necessary to prevent the whole system from succumbing to instability. The A^{vy} retrotransposon that "hijacked" the agouti locus is still relatively young by evolutionary standards. As a result, the defence machinery of the cell has apparently not had enough opportunities yet to stably inactivate the new integration position. Consequently, somewhere in the maternal germline and/or in the early embryo, the chromatin around the retrotransposon region is not fully "reset" (Ch9). In the literature, alleles of which the expression can vary through an epigenetic mechanism are referred to as "metastable epialleles".

The A^{vy} system has been used to test environmental factors for their ability to shift the spectrum of coat colours among the offspring when applied during pregnancy (fig 45). First, a diet that increases the availability of methyl groups was tested (DNA methylation, fig 11; histone methylation, Ch2.3.2, table 1, fig 8). This diet contains extra folic acid, vitamin B12, choline, and betaine. When the mother mouse (the F0) was fed this diet during gestation (from at least day 8 to day 15), the offspring (the **F1**) would be darker on average (fig 45). This effect persisted into F2, probably because the **germ cell**s for that generation were already in the sensitive phase of chromatin change during the period that the diet was fed (fig 45). When a similar protocol was used to test for the effect of the pseudo-oestrogen BPA, the offspring would be lighter coloured. However, when genistein, the oestrogen-like substance found in soy, was administered, the young mice would be darker. The molecular backgrounds of these different effects are not known. No matter how relatively clear the image of the variably methylated retroviral CpGs appears, it cannot explain the transmission of the phenotype to the next generation. This methylation is roughly preserved in the oocytes and is then gone in the **blastocyst**s (Ch9, fig 23), only to return later.

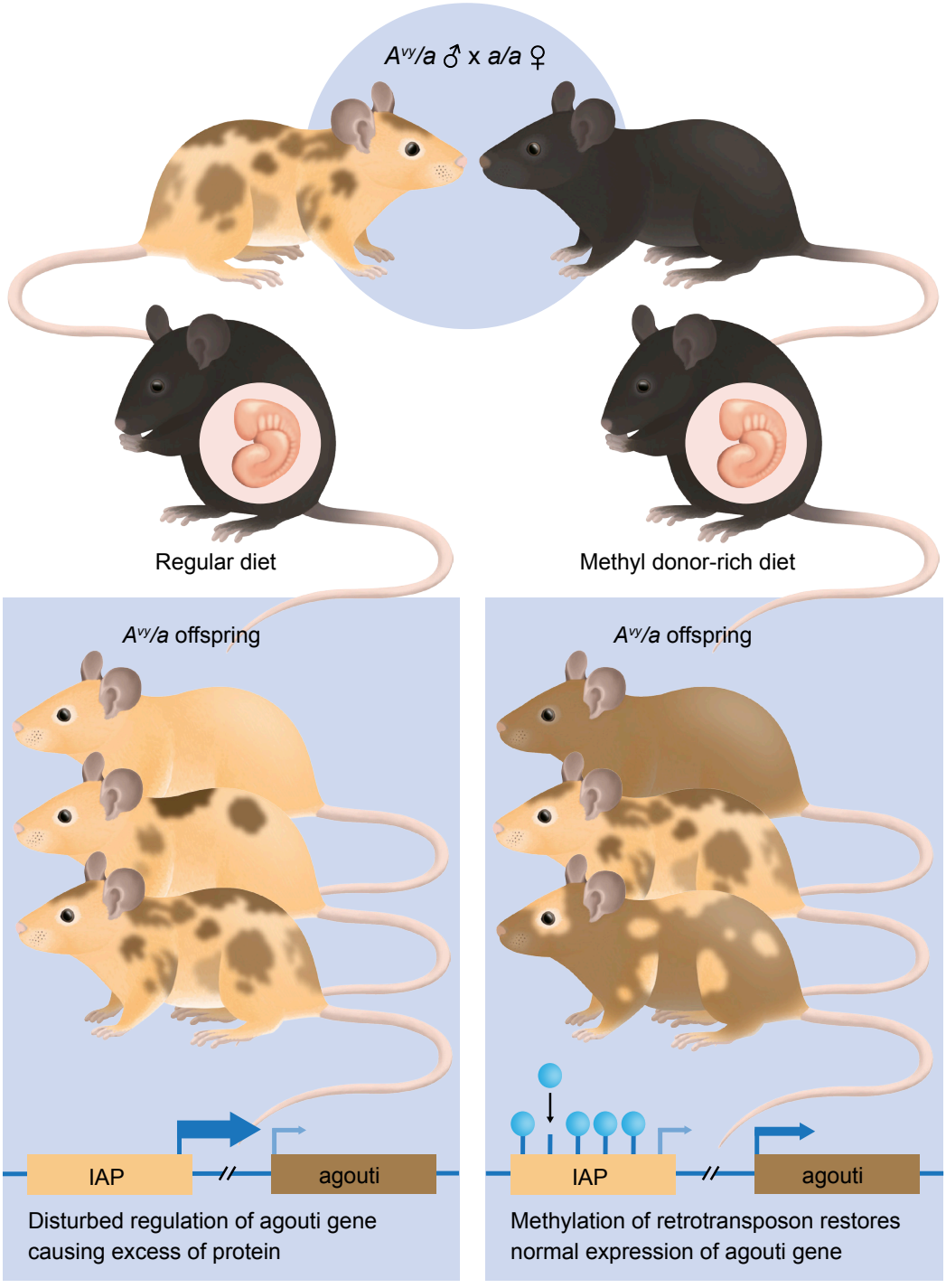

Avy/a ♂ x a/a ♀
Regular diet
Methyl donor-rich diet
Avy/a offspring
IAP
agouti
Disturbed regulation of agouti gene causing excess of protein
Avy/a offspring
IAP
agouti
Methylation of retrotransposon restores normal expression of agouti gene

FIG 45

The effects of a methylation-promoting maternal diet on offspring. IAP stands for "intracisternal A particle", a **retrotransposon** that has not yet evolved to be completely non-functional or neutral, and that can have a **promoter** function. In this experiment, the *A* "viable yellow" (A^{vy}) **allele** originates from the father, who is **heterozygous**. The mother is **homozygous** *a/a* and is thus black. Only the offspring that have inherited the **dominant** A^{vy} from the father are shown here. The variation in methyl donor availability caused by the maternal diet during gestation becomes evident as a change in the spectrum of colour variants. Reduced levels of CpG methylation (fig 11) of the retrotransposon render it more or less active, thereby disrupting the regulation of the agouti **gene**. The promoter activity that determines the **transcription** level of the A^{vy} agouti allele is increased, leading to overproduction of the agouti gene products (first RNA, then protein). In the normal situation, the yellow is only temporarily expressed, resulting in the formation of the yellow band in the hairs of the wildtype coat. As the A^{vy} allele is more active, the offspring are yellower, and the promoter is less methylated (thus overriding the promoter of the *A* sequence (wildtype agouti)). The effect remains visible a generation later, probably because the **germ cell**s that will produce this generation are already present in the **embryo**s that were subjected to the variable nutrition (there is an intergenerational effect, see fig 44).

15.2.3 A role for small non-coding RNAs (sncRNA) in transgenerational inheritance

From the preceding, it is evident that CpG methylation cannot explain the transmission of A^{vy} promoter activity, however appealing that idea may be. More recent literature indicates that entirely different mechanisms may also be at play. The first publication that shed light on the role of sncRNA in mice appeared in 2006, but it was regarded as somewhat peculiar at the time because of its revolutionary nature. The *Kit* gene (encoding a so-called tyrosine **kinase** receptor) is involved in signal transduction from the external to the internal environment of the cell. As such, it has a role in many functions. When only one locus is active, a phenotype appears of which a particular aspect is relevant: in an otherwise grey mouse, the feet and the tip of the tail are white. In research conducted by Minoo Rassoulzadegan's group in Nice, France, a **transgenic** mouse model was used in which one *Kit* allele was inactivated by inserting a piece of

bacterial DNA at the start of the gene using genetic manipulation (see box 3). In these mice, an (overly) long stretch of **mRNA** was produced from this allele, and although it also contained the base sequence of the KIT protein, this mRNA was not translated. Just like the mice with only one *Kit* allele, these mice also had white feet and a white tail tip. When these transgenic mice were backcrossed with normal wildtype mice, and the mutation was passed on, a large proportion of the offspring also had white feet and a white tail tip, as expected. However, this also happened, albeit to a lesser extent, when the mutation was not transmitted and the parent in question (which could be either male or female) had two normal *Kit* alleles. After several generations of back-crossing, the white feet and white tail tip disappeared, which happened slightly faster through the male line than through the female line.

Eventually, the study focused on the **spermatogenesis** of males that were carriers of the abnormal *Kit* allele. In the sperm cells, they found a highly increased amount of *Kit* RNA. In a follow-up step, they discovered that two waste products of this RNA would already lead to the phenotype (the white tail tip and feet) when injected into the **zygote**. These small pieces of RNA were identified as miRNAs. They have a characteristic short length of 19-23 b. There are many of these types of RNAs, they belong to the “small non-coding RNAs” (sncRNAs), which also include the 26-32 bp long piRNAs (Ch12.5.2). The normal biological role of these small RNA molecules is that they bind to a target mRNA molecule through base pair homology and thereby label it for degradation. The remarkable aspect of the discovery of the *Kit*-specific miRNAs is that when they are introduced into the zygote at fertilisation, they can obstruct the action of a *Kit* gene throughout development and for the rest of life. This effect can also be passed on to the subsequent generation but will ultimately become diluted. This is known as a transgenerational effect (fig 44).

After this first publication on small RNA molecules, which act as a kind of transferable regulatory factor between generations, more papers followed. A more recent example (2014), published by a Zurich-based research group led by Isabelle Mansuy, concerns epigenetic effects caused by mental distress, anxiety and trauma. During the period between birth and the age of 3 weeks, when mice are usually weaned and thus taken away from their mothers, the young animals were exposed to stress by regularly being separated from their mothers. This was combined with subjecting the mums to stressful treatments. Exposing laboratory animals to this form of stress is a standard protocol in brain research. The study focused on both the traumatised male offspring (the F1) and their sons (the F2). Males from both generations were found to exhibit signs of depression and anxiety in a variety of standard tests. This indicates that we are dealing

with intergenerational epigenetic effects. Through further molecular research, specific sncRNA molecules (miRNAs and piRNAs) were found in the sperm of the previously stress-exposed male offspring, while these are not normally present. When RNA was isolated and purified from the sperm of these males and subsequently microinjected into the male pronucleus (fig 42) of control zygotes, the offspring developing out of these zygotes exhibited signs of trauma. This also suggests a role of sncRNA molecules in the transmission of the behavioural phenotype.

Anxiety and depression are often mentioned in the same breath. A thorough recent study from China (2021) included a standard chronic “mild stress” protocol (noise, light, and reversing day and night, carried out daily for 5 weeks in an unpredictable manner). After that, the researchers checked whether the depression phenotype that had been induced in male mice could be transmitted to their offspring. This was indeed found to be the case, but the transmission stopped there: it was an intergenerational effect. The researchers could provide evidence for the involvement of miRNAs in the sperm in this effect.

These observations support the idea that certain small RNA molecules, which arrive with the sperm cell, are candidates for a mechanism of epigenetic information transfer. Here is one final example. In 2012, it was discovered that the mouse sperm cell is rich in a special class of tRNA molecules (fig 9), a shortened form called tsRNA (30-40 bases). If this component is obtained from males that were fed a high-fat diet, and if it is then injected into the zygote, the metabolic disorder from which the father suffered is transmitted to the offspring. Even more intriguing is an observation from 2018, which revealed that the effectiveness of this depends on an enzyme that specifically places a methyl group on tRNA. However, the exact molecular details of this information transfer to the next generation are not well understood to date. Reviews on the sperm RNA code are already emerging, with speculations on the inter- and transgenerational effects of sncRNAs derived from sperm on embryonic development. Besides tsRNAs, rsRNA (molecules related to ribosomal RNA, box 1, fig 9) is also mentioned as a potential messenger molecule. Human sperm cells are also rich in such classes of small RNAs; similar patterns of tsRNA molecules are found. They appear to respond to diet, in this case, a week of high sugar intake. This Swedish research group of concern (2019) incidentally discovered that small tRNA molecules of mitochondrial origin also respond to diet, thus adding another layer of potential information transfer from the sperm to the embryo. It will not be long before these micro RNA molecules become part of a sperm assessment in a research setting.

15.2.4 Do histones play a role in epigenetic inheritance?

The most comprehensive research on the involvement of **post-translational modification (PTM)** of histones in epigenetic inheritance has been conducted in a collaboration between the Faculty of Medicine at McGill University in Montreal, Canada, and the Friedrich Miescher Institute (FMI) in Basel, Switzerland. The work was based on a well-known experiment with the nematode *C. elegans*, in which a genetic intervention in a gene involved in the post-translational modification of histone H3 at position K4 (H3K4, table 1, fig 8) led to an epigenetic transgenerational effect (fig 44). The protein encoded by this gene recognises H3K4me2 and removes methyl groups (it is a demethylase). Subsequently, the researchers focused on the situation in mice, using the **orthologous** human demethylase gene for this purpose. The gene (the lysine demethylase called *KDM1A*) was first equipped with a promoter that only works during spermatogenesis. After that, this construct was injected into the male pronucleus of the zygote (box 3, fig 42). This resulted in the birth of **transgenic** mice with a genomically integrated construct (the P (F0) of fig 44), and the male offspring were further used in the study. The endogenous mouse-specific H3K4me2 (lysine) demethylase is active as usual during spermatogenesis. As is common in such experiments, multiple generations were followed through the paternal line, until the F4. Skeletal abnormalities and increased mortality were seen up to the F3 (fig 44), also when the males were no longer transgenic in these subsequent generations. Of course, the researchers investigated the sperm cells in an attempt to find an explanation for these large effects. In mouse sperm, around 1% of the genomic DNA is wrapped around **nucleosome**s (Ch12.6, fig 8, 32). Earlier, researchers had found an association between this nucleosomal DNA and the promoter regions of genes, many of which are involved in housekeeping functions, commonly expressed in, for instance, embryonic cells. These genes are thus associated with histones, and they are, on average, rich in CpGs (fig 11) in their promoter regions, which are hypomethylated (Ch2.3.3), despite the generally high methylation level of DNA in sperm cells (Ch9, fig 23). Together with the methylated H3K4 (me2 and 3), these are indicators of an open state of the chromatin in the promoter, which enables activation of the gene (table 1). All of this indicates a “readiness” of these genes for expression during the prospective embryonic development, however, see the last sentence of this paragraph.

Researcher Antoine Peters (FMI, Basel) estimates that due to the intervention of the human transgenic H3K4me2 demethylase, a modification in the chromatin structure of the promoter has occurred in at least 2,300 genes within this group of genes. However, a direct effect on the level

of CpG methylation of these promoter regions was not found. Nevertheless, this is not yet evidence of a direct involvement of histones in the transgenerational transmission demonstrated by this experiment. We now know that the effects of a single change in chromatin regulation during spermatogenesis can persist for multiple generations, resulting in phenotypic consequences extending to the F3. However, this detailed research does not provide insight into the molecular mechanisms underlying these effects. Of note here is that there is no good evidence yet pointing at the preservation of sperm-born histones at the time of paternal chromatin remodelling, from the moment of entry of the sperm in the oocyte onwards (Ch13.5.2).

15.2.5 An unexpected interaction between the egg and the sperm in the zygote

One of the previous sections (Ch15.2.3) included an example of sugar intake and a short-term effect of this on tsRNA molecules in human sperm. Parental energy **metabolism** has been associated with at least intergenerational effects for a longer time. In this area, scientific interest in obesity has grown considerably. While obesity is naturally associated with changing dietary patterns, could an epigenetic intergenerational mechanism also play a role? This idea becomes more plausible as obesity rates increase sharply, even among children.

Recently, a study was published on epigenetic inheritance, this time of high blood sugar levels, which sheds a completely different light on the range of mechanisms that may be involved. This study commenced with the observation that the maturation of oocytes before ovulation does not occur entirely normally in female mice with high blood sugar levels.

The researchers demonstrated this using the enzyme TET3. This enzyme ensures that the paternal DNA in the zygote is demethylated (Ch9, fig 23). When this enzyme is not active enough in the zygote, it affects the demethylation process, which in turn affects the activity of paternal alleles in the offspring. In this way, the functioning of several genes involved in insulin secretion by the pancreas was affected: the excessively high methylation level inhibited their activity. In the offspring, this led to lower insulin secretion and reduced glucose tolerance, precursors of diabetes. This is an example of a publication for which the researchers (from Hangzhou and Shanghai, China) had to undertake a large amount of work. The potentially considerable social implications led to the publication of this work in *Nature* (2022).

15.2.6 More mouse models: anxiety, olfaction, and disease

The following experiment provides another example of how remarkable transgenerational epigenetic inheritance is in mice. In the roof of the nasal cavity, mammals have receptors that are responsible for the recognition of odour molecules (MOE, the main olfactory **epithelium**). The many different protein receptors responsible for the sense of smell are encoded by an extensive series of genes collectively known as *Olfr*. The receptor from this gene family that is encoded by the *Olfr151* gene can perceive the substance acetophenone. This substance is found in various natural food products, including cheese, meat, apricot, and cauliflower. In 2014, a group from Atlanta, Georgia, USA, published a study conducted in adult male mice, in which the following protocol was used. The odour sensation of acetophenone was paired with punishment in the form of an irritating sound. The animals thus underwent "Pavlovian" conditioning, meaning that they became stressed simply by smelling that specific odour. Next, a breeding programme through the male line was initiated with the use of IVF. After testing animals from different generations, it was found that the stress response triggered by acetophenone odour in the first generation of animals persisted in the two subsequent generations: in the sons and grandsons that had never been confronted with the unpleasant sound (fig 44). The researchers then looked at the CpG methylation of DNA of the *Olfr151* gene in the sperm cells of the stress-exposed generation and their sons. They observed a loss of methylated CpGs. However, when they examined the DNA of the same gene in the olfactory epithelium, no abnormalities were found in the sons and grandsons. This is yet another example of the mysteries that still abound in this area.

The step from stress to pathology is not large. In male mice, a chronic toxoplasmosis infection reduces sperm production, which is combined with morphological deterioration. When two successive generations were produced, work conducted in 2020, it was noted that offspring in the male lineage exhibited abnormal results in a number of standard behavioural tests. While searching for an explanation in the sperm of the infected fathers, these researchers also encountered the sncRNA population as a potential cause of the transgenerational effect.

Lastly, let me present an even more recent example. This study links paternal experience of a sublethal infection, triggered by intravenous administration of the fungus *Candida* to the father, to an increased reactivity of acute immune responses in the next generation and in the generation thereafter, transmitted through the male lineage. In the generations thereafter, the heightened immune response subsided. Changes were found in the sperm **methylome** of the affected father. Some of these could be linked to genes involved in the immune response. This suggests

a direct relationship between the methylation status of genes in the sperm and the functioning of the immune system in the next generation (in which these genes are involved), although direct evidence for this has not yet been provided.

15.3 Stress-related epigenetic phenomena in humans

In the previous sections, I have tried to broadly illustrate developments in research on intergenerational and transgenerational epigenetic inheritance in laboratory animals. In recent years, several review articles have appeared in which the question is raised whether there are clear indications of heritable epigenetic effects (not based on a change in DNA base sequence) in humans as well. Due to the long generation interval, it is difficult to conduct research on this in humans. To prepare for this question, one could examine observations within one generation to determine whether stress-related **somatic** epigenetic variation also occurs in humans.

In 1995, an article by JPM Barker was published in the *British Medical Journal*, which focused on the relationship between inadequate conditions during pregnancy and health risks later in life. The knowledge outlined in this article became known as the "Barker hypothesis" and was later supplemented with additional studies. This hypothesis presumes that the **transcription** activity of susceptible genes adapts to the poorer conditions in the early stages of pregnancy (the embryonic period, from fertilisation to the onset of organ formation, which is completed in the 8th week). Subsequently, the level of gene regulation becomes fixed. Then, even when there is no longer a shortage later on, the **genome** remains set in a state as if there is scarcity, leading to diseases later in life, including lifestyle diseases such as heart disease, increased BMI (obesity), and type 2 diabetes, but schizophrenia is also mentioned.

A particularly well-studied group of people in this regard consists of a sample of people from the Dutch population who were conceived during the Dutch famine between December 1944 and May 1945. By the end of this period, severe stress due to food scarcity was experienced, with individuals having only 500 kcal to eat each day (25% of the normal daily intake). As expected, this did affect the future health of children conceived during the so-called Hunger Winter. This manifests as an altered metabolic profile (reduced glucose tolerance, increased BMI, elevated cholesterol). A study published in 2014 compared two children of the same sex within 24 families. One child was conceived during the Dutch famine, while the other one was conceived outside the December 1944-May 1945 period. In order to detect any potential epigenetic changes, the researchers chose to look at CpG methylation (fig 11). After DNA extraction

from white blood cells, they examined more than 90 thousand regions, which contained nearly seven CpG dinucleotides on average. This represents a selection from the entire human methylome, which is, of course, much larger (29 x 10^6 CpGs). It was determined that there was a difference in methylation levels between the two siblings in 181 chromosomal regions. Since the whole genome has now been mapped (Ch3.1), further investigation was possible. After all, we know exactly where those regions are located in the genomic DNA and whether that can predict or explain any changes in gene activity. The group of researchers, with Elmar Tobi as the first author and Bastiaan Heijmans as the last author, found an increase in CpG methylation levels in 60% of these "methylation difference regions" in adults conceived during the Dutch famine, and a decrease in the remaining 40%. The genes in which these changes had occurred had functions in early embryonic development, particularly in growth and metabolism. The changes were not so much related to CpG-rich promoters (fig 11), but rather to CpG-poor promoters and more distant CpG motifs, known as "enhancers," which influence the level of transcription. The take-home message from this is that when the early prenatal environment becomes more extreme, it has a lasting effect in the first generation, as reflected in the methylation pattern of CpGs near and within genes. Unfortunately, this output has not been used for further research in the generation following the war generation due to privacy reasons. That could have provided valuable data on intergenerational epigenetic inheritance effects (fig 44) after food scarcity. However, there is still some evidence for this. An increased BMI was found in the offspring of fathers (but not mothers) with a prenatal history from the Dutch famine. The children of women with the same history were 1.8 times more likely to have health issues later in life.

Due to the long generation interval in humans, indications of epigenetic inheritance largely rely on long-term epidemiological research. The first datasets used for a comprehensive study for this purpose originated from Överkalix, a village in northern Sweden. The grandparents were born in the second half of the 19th century, and their grandchildren were born in this century. For the grandparents, there are harvest records and food prices, from which it can be inferred whether there was abundance or scarcity (and to what extent). The strongest argument for the existence of transgenerational epigenetic inheritance that emerged from this study was that if the grandfather was a pre-adolescent (with slowly dividing spermatogonial stem cells) during a period of food abundance or food scarcity, this increased the mortality risk in the grandsons. Not all findings were confirmed in an even larger second Swedish study from 2018. However, findings on the effect of an excess of food remained intact. Interestingly, the researchers specifically mention cancer as a contributing factor to the shortened life expectancy of the grandsons.

15.4 From epigenetics to genetics

In 2000, Yuri Dubrova, who worked in the laboratory of the well-known researcher Alec Jeffreys in Leicester, UK (Ch2.3.3), published a "brief communication" in *Nature*, in which it was shown that **ionising radiation** has another heritable effect on the genome, in addition to the already known **mutagenic** DNA breakage and repair effects (Ch11.3, Ch12.7, Ch13.5.2). The study focused on the so-called microsatellite repeats in DNA (Ch2.3.3, table 2). This concerned the approximately 600 copies of the basic motif GGGCA in the mouse genome. The high spontaneous **mutation frequency** of such DNA repeats is always attributed to problems that the DNA **replication** machinery (fig 5) has with repetitions of a short base motif. For this type of repeats, this applies to both somatic cells and germ cells. In the archetypal experiment conducted in Leicester, the spermatogonial stem cells of F0 male mice were irradiated with fast neutrons at a dose of 0.5 **Gy**. These fast neutrons are effective breakers of the DNA double helix, and, like X-rays, they generate a wave of **oxygen free radicals**. Hence, this type of radiation causes widespread cellular stress. Analysis of the copy number of the GGGCA repeat showed that new mutations had appeared in the germline of F1 offspring; the mutation frequency had approximately doubled.

Remarkably, mutations were also found in the repeat locus that originated from the mother, which could thus not have any "memory" of exposure to radiation. Apparently, a reaction had occurred in the irradiated F0 spermatogonial stem cells that genotypically manifested in the spermatogenesis of the next F1 generation as a reduced ability to accurately replicate these repetitive DNA sequences which are already prone to error themselves. Dubrova and Jeffreys' findings demonstrate that there is not always a sharp distinction between genetic and epigenetic inheritance. For now, the mechanism that translates the memory of cellular radiation stress into less effective functioning of the DNA replication machinery in the germline (which is challenged by simple repeats), even affecting the subsequent generation (the "DNA replication mutation genotype"), remains a complete mystery.

In summary, you could say that we are slowly getting more glimpses of the molecular aspects of transgenerational epigenetic inheritance, but most of it is almost certainly still unknown. Is epigenetic adaptation an emergency mechanism to deal with harsh times, and does it always play a role to some extent?

15.5 Clues for a balance between genetic and epigenetic variation around methyl-CpG

In the previous chapters of this book, CpG methylation has already been mentioned several times (Ch9, box 4, fig 11, 23). In the current chapter on environmentally induced epigenetic changes across generations, this mechanism for DNA modification has received extra attention, and it will be further discussed here. In the 1980s, it was still a laborious task to determine for a position in the base sequence of the chromosomal DNA whether the C was methylated or not; now, with a modified **NGS** procedure, it is possible to determine the methylome of the whole genome. Special **chip array**s for CpG methylation analysis are also available.

CpG methylation in genomic DNA is widespread in the realm of living organisms. This phenomenon is evolutionarily old (900 million years) and robust, although it can also be missed in certain organisms. The mechanism that ensures retention of the methyl group at the C position in a newly synthesised DNA strand during DNA replication has been described in detail. While this copying process involves many more errors than the replication of the DNA base sequence itself (Ch2.2.2), it is still estimated to be biologically reliable, with less than 0.3% of Cs being missed. With ageing, more errors occur in the preservation of the methylation status, which leads to increased variation in the methylation level of CpGs in somatic cells, which is generally high (fig 23). For many positions, this is also sex-dependent. Much less is understood about how other chromatin parameters involved in the regulation of transcription, such as the histone PTMs (table 1), are conserved during DNA replication in the S phase of the cell cycle. It is, therefore, attractive to assume that CpG methylation plays a key role in regulating the highly detailed three-dimensional structure of chromatin.

Everything comes together in CpG methylation: genetics, environmental influence, chance and the unknown. Unlike the actual CpG methylation steps, the biological role of CpG methylation and the details of its regulation are still subject to many questions. For example, how fixed are the positions of cytosine methylation, and can any genetic variation occur there? How is the CpG methylation pattern established during embryonic development (fig 23)? Does chance play a role in this, and/or how can environmental factors affect this process? The next chapter (Ch16, on artificial reproduction) will also address CpG methylation as an explanatory and predictive parameter. Here, I will therefore attempt to outline our growing understanding of CpG methylation as a molecular guiding principle in "**nature versus nurture**" interactions.

When it comes to conducting genome-wide studies on the methylome, monozygotic twins who differ for the trait under study (read: a disease) are ideal subjects. The phenomenon of one twin having a certain disease while the other does not (disease discordant monozygotic twins) is not rare at all; on the contrary, sometimes this is the norm. By analysing the individual methylation status of the DNA of such twins, gene-associated regions whose methylation levels are associated with traits such as aggression and obesity have been found in the genome. Another approach is to use population screening to search for areas in the genome related to disease and to include CpG methylation in the analysis. In a screening like this conducted in Scotland, citizens born in 1936 were approached approximately 60 years later with the intention of having them participate in the following study. While it was conducted, these people were 70 years old and healthy. This research focused on the levels of 96 proteins in blood plasma, which all have significance in the context of neurological disorders. In addition to 41 regions which could be linked to the level of 33 proteins based on variation in the DNA base sequence (**GWAS**, Ch6.2), the researchers also found 26 regions that stood out due to the degree of CpG methylation, and those could be linked to the level of nine proteins. This approach thus yielded candidate genes for neurological disorders, even if these may derail due to an epigenetic aspect: after all, the methylation profile may be linked to the regulation of gene expression levels.

The high level of methylation implies that it has an important function in suppressing unwanted transcriptional activity of DNA, which is most notable for retrotransposons (fig 11), the regulation of imprinting (Ch10), and the inactivation of the X chromosome (box 4). That partially rules out a pattern of variation, but variation certainly exists beyond these aspects. The genomic methylation profiles of identical twins (compared to fraternal twins) can also be used to determine the **heritability** (the h^2, Ch6.1) of methylation at a single CpG position, which is an indication of the degree of variability. This appears to differ tremendously for positions that show variation per position, and it is not high on average; higher h^2 values are found only for a small fraction of these CpG positions. A high h^2 value indicates that the situation is fairly stable from generation to generation, with DNA being either predominantly methylated or unmethylated. The CpG-rich regions near genes (the CpG islands; Ch2.3.3) that have a promoter function are commonly unmethylated (see also the legend of fig 11). Gene activity regulation does not depend on this. Genes that are specific to the germline form a remarkable exception to this rule; in somatic cells, the promoter region is methylated. According to the current state of research, the situation is different for CpG positions that are located further from the gene and that are also involved in the regulation of gene expression (enhancer or silencer regions). Variation in these regions is

linked to fluctuations in transcription status (see also the section on the Dutch famine, Ch15.3). Another method to establish that there is meaningful variation stems from quantitative genetics (Ch6). Instead of a GWAS, an EWAS can be performed (an “epigenome-wide association study” based on a genome-wide (m)CpG profile). This method leads to the identification of areas in the genome that can be linked to the phenotype under study through variation in this epigenetic modification.

The influence of prenatal conditions, nutrition, and environment on variation in the methylome was already evident in the analysis of DNA from children conceived during the Dutch famine (Ch15.3). This impact was also seen in other studies, with two more examples following here. In a study conducted in Singapore in 2014, umbilical cord blood was collected from 237 newborns. Using a chip specifically designed for this purpose (Ch3.2, Illumina), the genome was screened to determine the methylation level of 450,000 CpG positions. The babies exhibited large differences in over 1,400 genomic regions, in contrast to a mostly homogeneous picture for the rest of the genome. The researchers monitored several factors, including maternal smoking during pregnancy, depression, but also BMI, and the baby’s weight. The quantitative genetic statistical models they applied primarily revealed that the “overall genotype” (the personal genome) was the best predictor of how the baby had responded to various conditions during pregnancy by methylation of the CpG motif. This work illustrates that the value of CpG methylation as a readout of environmental conditions is greater when that environment exhibits more extreme deviations for an extended period of time during prenatal development. With a different approach studying the influence of the early maternal environment on CpG methylation, researchers found 687 positions that showed a “variable” (or “metastable”, as termed by the researchers (Ch15.2.2)) response in the first weeks of embryonic development, including gastrulation (fig 43). They were also able to establish a link to the regulation of the activity of specific genes.

From this work and several similar studies, it has become clear that the genome responds to environmental conditions via methylation, already from the earliest possible embryonic stage (and throughout subsequent development). Several “players” (gene products such as proteins and certain types of RNA molecules) are involved in generating the characteristic features of the methylome. Each of these “players” is, in turn, also subject to genetic variation. This illustrates the influence of genotype on the response of the methylome to environmental conditions. It thus seems that variation in CpG methylation partly has a genetic basis, and that this basis also exists for the way in which our genome responds to the environment via methylation from conception

onwards. However, twin studies teach us that “chance”, including aspects that may have yet to be discovered, also plays a major role.

Lastly, I will provide a micro-level illustration of the genetic basis of CpG methylation. In a study published in 2016, a female mouse was fed a low-protein diet, 8% as opposed to the normal 20%, for a period of several weeks from the second day of cohabitation with a male mouse. The researchers used the well-known inbred strain B6. Mating occurred at the beginning of this period, and the diet regimen was maintained until the end of the nursing period, 3 weeks after birth. The offspring exhibited growth restriction and reduced physical activity, and they also produced less insulin when subjected to a glucose challenge test. While investigating the effects of protein restriction in these animals, the researchers stumbled upon a (methylated or unmethylated) CpG motif in DNA from liver cells and sperm cells of the male offspring. This motif is located 133 bases upstream of the transcription start of the gene for ribosomal RNA (position indicated as -133). This gene is found in many copies, head-to-tail oriented, in multiple clusters in the genome. In humans, these polymorphic gene clusters are found in the genomic DNA on the short arms of chromosomes 13, 14, 15, and 20, 21 (fig 12). Methylation of the CpG at position -133 in rRNA genes negatively affects transcription. This in itself is remarkable. However, the point that these researchers make is that a DNA base located 29 base positions closer to the start of transcription of rRNA genes (at position -104) has a great impact on the methylation level of this particular CpG. If there is an A at that position, the methylation level of the CpG at position -133 is much higher than when there is a C (so there is a DNA polymorphism (SNP) within the B6 inbred strain). In male offspring with the A base that had experienced the diet restriction through their mother, the growth restriction could be estimated from the percentage of methylated CpGs at position -133. The higher the proportion of cells that have the methylated CpG, the lower the weight of the sons. In this case, a seemingly minuscule detail in the genotype has an effect on the epigenetic response to protein restriction that was already present in utero.

This is a great example of how genotype and methylome are closely intertwined, just as nature and nurture are intertwined.

CHAPTER 16

What artificial reproductive technologies teach us about human reproductive genetics

16.1 Introduction

Without a doubt, the development of *in vitro* fertilisation (IVF) in humans has stimulated reproductive research in our species. Scientific journals with titles like *Fertility and Sterility* and *Human Reproduction* are flourishing and the prestigious *Human Reproduction Update* is coveted to publish in. Long before the emergence of IVF, the medical science community was already looking at the earliest stages of human **embryo**nic development.

From 1938 to 1959, gynaecologist John Rock (Baltimore) and pathologist Arthur Hertig (Boston) reported their findings, which were outcomes of hysterectomies, surgical removal of the uterus. During the months that patients awaited their surgery, they were asked to track their cycles and to record the days of sexual intercourse as well. In most cases, only the uterus was flushed out after the removal, but sometimes this was also done with the oviducts (or fallopian tubes). Oocytes were harvested from the ovaries, and this led to what I believe was the first known IVF attempt in humans (with an indication of success). The patients were between 25 and 43 years old. Based on the collected embryos, Hertig and Rock were the first to construct the human timeline of early embryonic development up to day 17 (an implanted embryo with placental development). This research has also been subject to some criticism. For instance, the patient population used in their studies seems heterogeneous, since the reason for the

hysterectomy procedure is not really provided. Also, Rock and Hertig did not know that sperm can survive for a long time (likely around the isthmus fig 39) and that it is not a good idea to let the egg age before fertilisation (Ch13.4, fig 41). During the period around World War II until 1947, they collected 26 embryos from 122 women, and another eight embryos from 88 women were added over the next 9 years. Ultimately, they rated 24 of the 34 embryos as normal based on appearance. They were the first to observe many aspects of human early embryonic development, which we are now so familiar with through routine work in the IVF lab.

With all the knowledge that IVF research and the practice of IVF have provided about the early stages of embryonic development, can we now better understand human reproduction? It has certainly sparked interest in it. Communities of researchers who are actively engaged in assisted reproduction have suggested that natural human reproduction is not always problem-free either. Therefore, the differences between the laboratory route and the evolutionarily proven method may not even be that great. But is that really the case? What plays a role here is that the reproductive capacity of our species is estimated to be low from the laboratory perspective, as if reproduction in humans is always a matter of chance. Nonetheless, the species has managed quite well over the course of recent evolution. Hence, it seems more likely that there is a high degree of variability in reproductive capacity. This is especially true for a woman, since good eggs are precious and many of the conditions in which the oocyte/egg and embryo reside, such as the environment in the ovary, oviduct, and uterus, depend on her. The variation in sperm production (Ch12.8, Ch12.10, table 5, fig 34) can never have greatly disadvantaged the delivery of the next generation over the centuries: oocytes are many times scarcer than sperm.

Here are some known statistics from contemporary research: approximately one in six to seven couples encounter difficulties in fulfilling their desire for parenthood. What does each partner contribute to this? Most publications state that the male-female ratio in this is approximately 50-50. A smaller proportion will be caused by ineffective interaction within the couple. This can also include couples where no clinical abnormalities can be found (around 20% of cases). From figure 36, you can deduce that in couples who seek reproductive assistance, the male contribution to fertility issues may actually be more than 50%; a clear-cut line cannot be drawn. In the Netherlands, a request for reproductive assistance can be honoured after one year of trying without success. A fertility test at that point can roughly distinguish between couples for whom trying to conceive naturally is still worthwhile, with an estimated chance of a successful pregnancy of 15% in the following year, and those for whom this is not the case. The first group will probably largely coincide with the group that, once the stress of the entire IVF treatment pro-

cess is over, still conceives spontaneously (which can also occur during the IVF trajectory). In a follow-up study conducted at the University Medical Centre of Maastricht (including participants up to 7 years after the initial visit for fertility treatment), it was found that about 10% of children born after the time of intake were the result of natural conception.

A spontaneous pregnancy like that can occur while people are on the waiting list, during treatment, or after the completion of a fertility procedure. The link between the number of cycles and the cumulative chance of spontaneous pregnancy is shown in figure 46. After 2 months, there is already a rise in couples who could benefit from some form of assistance. Imagine that you could replicate this graph using the same couples while correcting for age, this time reflecting their attempt to have another child; would the more fertile couples from the first round be the more fertile ones again? And would the less fertile ones still turn out to be less fertile? There are various reasons to assume that this is true, but to my knowledge, there is no measure for this (the repeatability of the time to reach pregnancy) that has been consistently acknowledged in the literature.

Shortly after Robert Edwards' first publications in 1965, about human oocytes that had been extracted from mature follicles which spontaneously resumed their journey through the first meiotic division (Ch13.7), his group reported in 1969 and 1970 that they had succeeded in the *in vitro* fertilisation of eggs in the second meiotic division (fig 39). It then took another eight years until Louise Brown, the first IVF baby, was born in 1978. The reason for developing IVF at that time was that it offered a solution when the oviducts were blocked, for example, caused by a *Chlamydia* infection. In fact, it was a period in which everything was still being tried out, and the approaches were also rather primitive by current standards. I have no idea whether the founding fathers of this technology foresaw the extent to which IVF, now part of ART (artificial reproduction techniques), would become used among the general population. For a long time already, pathology of the oviduct has not been the main reason for the involvement of the IVF lab in the treatment of infertility (this is now only the case in about 10% of treatments). In a recent representative Dutch multicentre study on the effect of the chemical composition of the culture medium on the various outcomes of IVF treatment, male subfertility or infertility was present in 50-60% of cases. In around 20%, no medical abnormalities were detected in the partners (incidentally, this is the group in which most spontaneous pregnancies occur during the waiting period for treatment).

The fact that IVF requires far fewer sperm cells than natural conception has actually linked urology to gynaecology. The common thread running through many research efforts around ART

is that the international medical profession goes to great lengths to fulfil a desire for parenthood. The feelings of despair and profound sadness evoked by an unfulfilled desire to have children serve as a stimulus in the consulting room to explore the boundaries of the possibilities of artificial reproduction. On the other hand, there is a societal issue, namely that it seems increasingly difficult to accept fate, which is in line with the general idea of a society in which anything should be achievable. The experiment that led to the discovery of the effectiveness of ICSI (intracytoplasmic sperm injection) in humans can be explained against this background. In a small fraction of classical IVF situations, in which the eggs are surrounded by an excess of sperm cells in a small drop of culture medium under a layer of oil, the sperm fail to penetrate the **zona pellucida** (Ch13.5 fig 39). In order to solve this, attempts were made to use a pipette to assist the sperm cell a bit by depositing it in the so-called perivitelline space (fig 39). In 1992, at the IVF lab of the Free University of Brussels, Belgium, a sort of workplace accident occurred while Hubert Joris operated the micromanipulator; one arm of the machine moved too far, thereby depositing the sperm in the **cytoplasm** of the egg cell. If this had been attempted with a mouse egg, it would immediately have been punctured and deflated. In humans, however, this technique proved to be useful. When the researchers were confident enough about the reliability of this method of creating a **zygote**, they published about it in *The Lancet*, describing that healthy babies were born from eggs that had been fertilised in this way.

This made Gianpiero Palermo and his boss, Professor André van Steirteghem, instantly famous. ICSI had emerged as a modified version of IVF treatment, and this greatly expanded the population of prospective parents who could seek treatment. Initially, for men with a very low sperm count, the sperm cell selected by the laboratory technician (of normal size and visibly alive based on observation of tail movement) was obtained from the ejaculate (table 5). However, this was no longer necessary in the next steps. Men can become sterile due to an obstruction in the epididymis, for example as a result of inflammation. In that case, it makes sense to search for usable sperm in the duct upstream of the inflammation (a PESA, percutaneous epididymal sperm extraction). If nothing can be found there and there is non-obstructive **azoospermia** (Ch12.12), then the search can be extended to the testis. This technique is referred to by the abbreviation TESE (Ch12.9, Ch12.11, Ch12.12). I still clearly remember the pictures that a doctor from the USA presented in Amsterdam, which showed a testicular wall cut open longitudinally with a long incision to facilitate the search for thicker segments in the testicular tubules. It is not rocket science to assume that the chance of finding a focal cluster of spermatogenetic cells is greatest wherever the tubule bulges. In 40-50% of men who have no obstruction in the

epididymis but whose sperm production is too low for sperm cells to reach the ejaculate (via the seminiferous tubules, the rete testis, the long duct through the epididymis, and eventually the vas deferens (fig 31)), sperm cells can still be found using this method. IVF analysts can sometimes spend hours searching for one or a few sperm cells in order to still keep the possibility of genetic parenthood open.

The live-dead selection test (a slight movement of the tail, visible after touching the sperm with a pipette) is difficult to apply here because motility normally only arises in the epididymis (Ch12.12). Initially, publications appeared that aimed to nuance the importance of the sperm "being alive" when it is used for ICSI. We now know that this is a prerequisite (Ch13.5.2). In the first years after its publication in *The Lancet* in 1992, ICSI was introduced in Dutch IVF centres. However, as an exception to the widespread implementation of this technique in Europe and around the world, the Dutch Society of Obstetrics and Gynaecology (NVOG) imposed a moratorium on the use of ICSI in 1996, which lasted until 2000. They waited for the release of findings from abroad about the efficiency of ICSI and the risk of **congenital disorders** in children. At that time, it was already known that the conventional IVF method slightly increased the risk of congenital abnormalities in children, which formed a real concern for the professionals in the field. We now know that this is not quantitatively different for ICSI children.

There is also a positive early embryological aspect to ICSI: there are easily about 2,500 sperm cells per small drop of 50 µl culture medium in an IVF dish. This number is many times higher than in the *in vivo* situation (Ch13.5). The sperm cells are metabolically active, and due to the excessive formation of **oxygen free radicals** (Ch13.5.2), they almost certainly cause a physiological effect that is potentially **mutagenic**. Under these conditions, it is already remarkable that polyspermy only occurs in around 3% of cases. There are human IVF laboratories that may predominantly use ICSI for this reason and/or for operational considerations. This will then also apply to a prospective father with a normal **spermiogram** (Ch12.9), thereby eliminating any form of sperm competition.

It is not an overstatement to say that ART has conquered the world and that the contribution of these techniques to bringing the next generation into the world is increasing everywhere. It is also true that in most countries, except for the Netherlands and Denmark, the procedure has been commercialised, with an expanding business sector and venture capital in the background. The well-known statistic for ART in the Netherlands stands at one in 30 births. This means that, so to speak, one child in every classroom has been brought into the world with the use of ART.

16.2 *In vitro*

IVF would not have been possible without the discovery that the zygote can undergo cleavage divisions *in vitro*. In mice, most of the groundwork was done from the 1950s onwards by the aforementioned Wes Whitten of the Jackson Laboratory in Maine (Ch13.7). Two lines of thought can be distinguished in research into the optimal *in vitro* environment: (a) mimicking nature or (b) trial and error (letting the embryo itself indicate what it "wants"). Culture media were initially made in-house and the formulas for their composition were public, but with the growth and globalisation of ART, the supply of media and attributes for sperm and egg isolation and embryo culture has also become a fully commercial affair. It is estimated that as many as 20 different medium formulations are currently in use for humans, probably all different from the culture medium used to create Louise Brown. As the *in vitro* residence time of human embryos has increased due to technical developments, so have the requirements. A "two-step" medium is often used in this process, where the composition during the second step is tailored to the development of 8-cell embryos from day 3 to a **blastocyst** on day 5/6 (fig 43). Developing and testing suitable culture media is an extremely complex matter because the optimal concentrations of the dozens of components depend on each other. A medium that was widely used in the Netherlands during the rise of IVF is HTF (human tubal fluid), based on the composition of fluid in the oviduct, but it has now fallen into oblivion. It is an illusion to think that it will ever be possible to perfectly replicate the *in vivo* environment *in vitro*. The physiological conditions in which cells, especially oocytes and embryos, thrive are too complex and variable for that. Scientific review articles emphasise that the *in vitro* environment represents a form of stress to which the egg and early embryo must adapt. In 2010, it was discovered at the Department of Obstetrics and Gynaecology at the University Medical Centre in Maastricht, under the guidance of John Dumoulin, that the medium in which fertilisation and the initial cleavage divisions had taken place has an effect on the birth weight of babies. This finding was so sensational that it sparked a large study involving six Dutch IVF clinics. Together, they conducted a particularly informative experiment comparing two media. One culture medium was slightly better at producing day 3 embryos with a sufficiently good morphology for transplantation. Its use led to a slightly higher number of successful births. With the other medium, the average birth weight of the babies was about 5% higher (I use this study in Ch16.3 to introduce key statistics of IVF treatment). In three of the six labs, the preimplantation development took place at an oxygen level of 5%, because the oviduct is a relatively oxygen-poor environment. The other three laboratories used a gas phase in the

incubator with normal air, containing 20-21% oxygen. In chapter 13.7, it was already mentioned that for cattle and humans, it is truly better to culture preimplantation embryos at an oxygen level of 5%, which is a more normal condition for tissues. This seems logical when you consider that **gametes** are sensitive to **oxidative stress** (Ch12.7, Ch13.5.2, Ch13.7). So from a precautionary principle, you would expect that all IVF clinics use the more expensive 5%O_2 incubators, and this is more or less the case at the moment.

16.3 Superovulation

When I started to scientifically focus on biological issues related to the use of artificial reproductive techniques in 2001, an enthusiastic gynaecologist told me that the success of IVF is related to the fact that there is still a choice at the time of embryo transfer. In the Dutch study cited earlier in chapter 16.2, which investigated the effect of the culture medium on the outcomes of IVF, the average yield of a superovulation procedure was nine oocytes, with the average age of the woman being 34. The variation is wide for both age and number of oocytes. On average, the superovulation response naturally decreases with age (Ch11.1). From the nine oocytes mentioned in the study, an average of 2.5 developed into 4- to 8-cell embryos (fig 43) that were suitable for transplantation (the *in vitro* endpoint in this study). All these laboratories use both conventional IVF and ICSI (at a ratio of about 2:3). The transfer of usually one but sometimes two embryos at the appropriate time resulted in implantation in 15-20% of cases. After a year of trying, the chance of giving birth was around 40% (and not really different for the media used here, including the now discontinued HFT, see Ch16.2). Measured over the entire study, the ratio of singletons to twins was 7:1. The ratio of fresh embryos to frozen embryos that were transferred was 2:1. For comparison, here are the statistics from the Netherlands for 2020. The first number concerns classical IVF, the number in parentheses refers to ICSI. IVF and ICSI were used in approximately equal proportions. All cycles were included, leading to the supposition that when only first cycles are considered, a more successful picture will be obtained. When the number of cycles started with superovulation is set at 100%, this leads to follicle puncture in 86.3% (87.7%), embryo transfer in 74.6% (76.7%), pregnancy in 27.2% (29.5%) and an ongoing pregnancy in 19.3% (21.7%). Of the births, 3-4% are twins.

Of all the factors that play a role in the success rate of an IVF procedure, the age of the woman who wishes to become pregnant is the most important one. A professional and advocacy organisation like ESHRE sets an implantation success rate of 25% for an 8-cell embryo and 35%

for a blastocyst as the minimum norm. However, there are biological limits and many variables that affect the chance of success.

How do you select the most promising embryos for transfer? Despite a diligent search for answers to this question, there appear to be no reliable indicators either in the embryo or in the external environment (the culture medium), apart from clear morphological features. The indicators that do have significance include abnormalities during fertilisation and a strongly accelerated or delayed rate of division (more common). Fragmentation, the appearance of small clumps of cellular material during cell division, often occurs in such situations. This phenomenon, which is frequently seen in human embryos, can be reversible. It hardly ever occurs in mouse embryos. IVF embryologists assess embryos once a day under the microscope to determine the degree of fragmentation (increase is negative) and cell count (increase is usually positive, fig 43). This procedural approach has led to a kind of semi-objective, semi-subjective rating scale. With current technological possibilities, could this not be made more objective and informative, you may wonder? The answer to this question is yes, because the embryoscope, an automated microscopic imaging system that takes a picture of each embryo in the incubator every 20 minutes, was developed several years ago. More modern systems currently do this every 5 minutes. When viewing these images like a video, the development of the embryo flies by. Even before the data collected using the embryoscope appeared in the literature, a story had already been published about the predictive power of the timing of the first cleavage division: the progression of the zygote stage (Ch13.5, fig 42). Now, thanks to time-lapse monitoring, the progress of the first cleavage division can be followed in detail. A French study published in 2019 used 22 parameters to describe events occurring between gamete fusion and the appearance of four **blastomeres**. Two of these were related to the likelihood of a full-term pregnancy and, thus, the birth of a baby: the way in which the two pronuclei (fig 42) aligned in the centre of the zygote (a positive sign) and the appearance of multiple nuclei in at least one of the two blastomeres after the first cleavage division (a negative sign). The importance of control around the first cleavage division is demonstrated once again with this new technology. To further assess the potential of time-lapse monitoring for IVF practice, a multicentre study has been launched in the Netherlands, the results of which were released in 2023. Compared to the routine practice of following preimplantation development, time-lapse monitoring did not result in any improvement in the clinical outcome. Even with this sophisticated tool at hand, the biological events that occur during early embryogenesis will remain mysterious territory for some time to come.

Up to this point, the environment (the E from the formula P = G + E of Ch6) has not been taken into account. There are observations indicating that the environment of couples opting for assisted reproduction, more specifically those who need the services of an IVF laboratory, unmistakably influences the chances of success, relating to both ongoing pregnancy as well as live birth. Data from a cohort study conducted at the Rotterdam Erasmus Medical Centre were stratified based on socioeconomic status (SES), linked to the IVF laboratory registry of the Erasmus Medical Centre, and analysed for the period from 2006-2020. SES characteristics clearly vary across different neighbourhoods in Rotterdam, so this city provides a suitable population to study environmental effects, including factors that influence age expectancy and the number of healthy life years. An elaborate statistical analysis revealed odd ratios of 0.66 for a successful pregnancy and 0.63 for childbirth. CpG methylation (fig 11) as another aspect of interaction with the environment has been described in chapter 15.5.

16.3.1 The emergence of chromosomal aberrations during cleavage divisions

Persistent fragmentation of cytoplasm but also of the nucleus; this situation seems to represent two sides of the same negative coin. During the formation of blastomeres, the division apparatus is apparently not entirely reliable. What is responsible for this, and does this only occur in humans?

In 2000, an Australian paper was published in which a chromosomal portrait of individual blastomeres was created from 12 human day 3 embryos using a novel **DNA hybridisation** technique. This work provided a strong indication of incorrect **chromosome** numbers per cell. When this result became evident, the question arose whether a chromosome diagnosis for the whole embryo could already be made at the 8-cell stage, prior to compaction (fig 43). This could be done by determining the number of chromosomes from one or sometimes two blastomeres that would be biopsied. The idea that a diagnosis based on a single cell could be reliable was, of course, strongly influenced by what was already known about the meiotic origin of numerical abnormalities and the strong effect of maternal age on this (Ch11.4, fig 30): meiotic errors are found in every cell of the embryo.

However, further work on this yielded unexpected results. When Esther Baart began her PhD research in Rotterdam at the beginning of this century, she had access to fluorescent pieces of search DNA (probes) from 10 chromosomes, with which she could identify the chromosomes using DNA hybridisation and microscopic analysis. The research material consisted

of 196 embryos. She collected a single cell from 74 of them and two cells from the other 122. In half of those 122, the fluorescence pattern differed between the blastomeres: the embryo turned out to be a **mosaic**. In retrospect, this was the first somewhat stronger indication of a phenomenon called CIN, chromosomal instability. Early discussions about CIN naturally addressed that due to this phenomenon, it is difficult to make predictions about the chances of implantation in the uterine wall and subsequent embryonic development based on examination of a single cell.

Meanwhile, technological advancements had continued, and by 2009, it became possible to amplify the DNA from a single **blastomere** and analyse it using a **chip array** (a glass slide containing a representation of the entire **genome**, by using small fragments that are arranged in the correct order) and DNA hybridisation. In this way, copy numbers that deviate from two can be visualised with much greater resolution. For this rather iconic publication from 2009 by Joris Vermeesch's group in Leuven, Belgium, the researchers sought cooperation from couples who did not require IVF due to reduced fertility, but to prevent transmission of a genetic burden, such as an X-linked disease or the presence of a *BRCA2* **mutation** (one of the two well-known breast cancer genes), to the next generation. In such a procedure, embryos created using IVF are screened for the relevant mutation (preimplantation genetic diagnosis, PGD, now PGT-M, see also Ch18.2), and only genetically healthy embryos are transferred into the uterus. The researchers had access to embryos that still looked reasonably good but were no longer needed to fulfil the couples' desire to have children. Their conclusion was that between 70-90% of these on average 8-cell embryos were chromosomally abnormal in one or more (up to all) cells. By using the chip array, they could also see if chromosomes were broken, a fatal obstacle when genetic information is transferred to the daughter cells. It turned out that this was also far from rare. Collectively, the newly acquired knowledge contributed to the rather sad picture of poor "chromosomal robustness" during cleavage divisions.

Within the development of IVF/ART, a related discussion is currently taking place. The main question in this context concerns the optimal timing of embryo transfer. This question, in turn, is related to a possible influence of the culture media on the correct progression of cleavage divisions and the development of the fertilised egg into a blastocyst. Embryo transfers are currently performed on day 3, when an 8-cell embryo is preferably used, or on day 5/6, when an early or slightly later blastocyst is used. There is strong evidence that a prolonged *in vitro* embryo culture time already leads to selection against embryonic cells (blastomeres) that are genetically highly abnormal and that this selection occurs in both the **ICM** and the trophoblast (fig 43). You can actually already see this simply by looking closely at the embryos. Even before, but certainly

after the 8-cell stage, more poorly developing embryos will indeed appear. However, there does not seem to be a big difference in the eventual "take home baby rate" (including potential subsequent transfers of frozen embryos obtained during a previous stimulation cycle, see Ch16.3.2) when the results of multiple studies of day 3 and day 5/6 transfers are combined. Nevertheless, the duration of embryo culture (and thus the timing of transfer), including quality control through chromosome and genome analysis, remains a subject of ongoing scientific debate. For example, there are outspoken proponents of genetic analysis of blastocysts based on a trophoblast biopsy. This is justified given the effect of age on meiotic **aneuploidy** (Ch11.4, fig 30). As noted before, chromosomal instability (CIN) that does not occur until cleavage divisions take place is a factor that limits IVF success. Recent research indicates that when *in vitro* culturing is continued up to the blastocyst stage, the medium used for this also has an impact. One may wonder whether chromosomal instability occurs to the same extent in natural reproduction that is not guided by hormone stimulation. But how do you find evidence for this and how do you search for possible causes?

Despite the frequent occurrence of CIN, IVF success rates of 25-35% per transfer of 8-cell embryos are still reported in the literature. This is even higher for blastocysts (35-60%). It is clear, however, that these rates are affected by the presence of mosaic embryos (Ch13.6, Ch13.6.1): there must be enough **diploid** cells in the **epiblast** of the early embryo to ensure viability (fig 43). Earlier (Ch13.6.1), it was noted that the egg cell is oversized in terms of cellular content. The egg is large enough for the development of more than the minimum number of cells that is needed to give the epiblast a chance to continue embryonic development. The formation and subsequent loss of some blastomeres due to chromosomal aberrations (CIN) can be tolerated during the cleavage divisions (and even thereafter). Recently, a USA-European consortium examined the implantation chances of aneuploid-mosaic embryos and their chances of eventual birth, based on an **NGS**-based evaluation of 5 trophoblast cells per blastocyst. The chances were less than half compared to those of birth of embryos that were declared **euploid**. It should be noted, however, that aneuploidy in trophoblast cells (the cells used for the genome analyses) is much better tolerated at this stage than in the **ICM**/epiblast cells from which the embryo further develops.

It is also worth mentioning that, even based on this small number of five trophoblast cells, the degree of aneuploidy is indicative of the likelihood of development after implantation. This probability is lower when the percentage of aneuploid cells is higher. Extrapolating this to the *in vivo* situation, it is a rather strange notion to wonder how many people this applies to. Who among us has already passed through the eye of a needle at the blastocyst stage? Where does the lower

limit of the fraction of chromosomally normal cells necessary for the development of the epiblast and later stages (fig 43, Ch13.6.1) lie in humans? Entirely as expected, there will also be embryos that are aneuploid in all their cells, as determined from the biopsy. In such cases, a meiotic error is the most likely cause (see also Ch11.4.1). Even when a woman is younger, 20% of the eggs obtained after superovulation will be aneuploid. After the age of 36, this percentage increases sharply. It then seems sensible to select against this. A question that remains unanswered is whether there is also an age-dependent (increasing) effect for the phenomenon of chromosomal instability (CIN). All in all, it is remarkable that the practice of ART does not indicate that taking biopsies and checking for aneuploidy (this form of PGD is now called PGT-A) improves the chance of successful pregnancy. Currently, the success of ART is simply highly dependent on the woman's age. The condition of her eggs plays a key role in this.

After superovulation and intrauterine insemination (IUI), would it not be possible to flush the embryos from the uterus in the blastocyst stage, take a biopsy for genetic analysis, select those without severe genetic defects and keep them frozen until they can be transferred back into the uterus in the next cycle? With those questions in mind, contact was established between several USA laboratories and the Punta Mita hospital in Punta de Mita, a coastal resort in Mexico. The female volunteers and sperm donors (not always their partners) were recruited in Mexico; the women received $1,400. That amount covers living expenses for more than 2 months. They were informed about the risks: some participants became pregnant and got an abortion. A superovulation procedure was initiated, and IUI was performed at the predicted time of ovulation (Ch13.4, see fig 41). The 134 uterine flushing procedures performed in 81 women yielded 136 embryos, including 72 blastocysts, almost all on day 5 and 6 (fig 43). The ultrasound had predicted an average of 8.5 ovulations and therefore, there were pregnancies. The researchers noticed that the *in vivo* morphology of the blastocysts, which had probably undergone some kind of selection due to the procedure, was better than what is common in IVF practice. IVF data from some women in this study were also available, which confirmed this initial impression. And what was the result of the genetic analysis of the biopsied cells from the trophoblast (fig 43)? Normal diploid embryos were predicted from around 50% of the blastocyst biopsies, consistent with what is found in IVF control blastocysts. Almost 20% of the blastocysts were fully aneuploid, and the rest ranged from low mosaic to highly abnormal.

Another pathway in this difficult research field makes use of cattle as an animal model, in this case Holstein-Friesians, the well-known global dairy producers. The animals had never calved before and a superovulation was induced with hormones. One group of animals was normally

inseminated, while mature eggs were retrieved from another group for an IVF procedure. In a third group, immature oocytes were retrieved from the ovary. These matured *in vitro* and were subsequently fertilised using IVF. Genetic analysis was conducted on individual blastomeres of the 4- to 8-cell embryos. After amplification, the DNA was analysed for any deviations from diploidy using an **SNP** array (Ch3.2). In 19% of the *in vivo* fertilised embryos, one or more genetically affected blastomeres were found, but in the IVF embryos, this was observed in as many as 69%. When oocyte maturation had also taken place *in vitro*, this percentage increased to 85%. This was a pretty shocking finding.

The major problem encountered when analysing the results of IVF/ART is the statistical phenomenon called confounding. When you include (a) age effects, (b) superovulation, (c) *in vitro* culture, (d-e) quality of egg and sperm, and (f) the environment of the uterus as variables affecting the chance of having a healthy baby, these factors generally coincide, they are linked in some way. There may also be interactions between variables, another statistical concept. This makes statistical analysis difficult, and it is almost impossible to reliably investigate the individual influence of each of these variables in the context of IVF. It is also not straightforward to compare results from different IVF laboratories. Finally, it remains very difficult to gain insight into the reality of human artificial reproduction through research on experimental animals. For instance, is the mouse (reproductively basically an all-rounder) actually a suitable animal model with its large litters and low early embryonic mortality? In this respect, *Bos taurus* (cattle) seems to be a better model. Each species will have its own (in)sensitivities to the various technical steps that an IVF procedure consists of. And the reality is that the human population seeking ART is very heterogeneous. Essentially, we are our own test subjects. Further questions about this and additional genetic and **epigenetic** aspects of ART will be addressed hereafter (Ch16.5 and Ch16.6).

16.3.2 Superovulation and the maternal to zygotic transition

In chapter 11.5, it was argued that there may be variation in the quality of egg cells between consecutive cycles of a woman and that this could be the reason why fertilisation may not lead to implantation, followed by sustained embryonic development, every month. When fertility is better, this variation is lower. Apparently, some eggs are more capable than others at making the transition from the “oocyte programme” installed during follicle development (of which one important aspect is to prepare for embryonic development) to the programme for the initiation of embryonic development. Regarding the first phase, epigenetic preparation of the relevant

genes is achieved when the oocyte is fully grown, while the latter program is only “installed” at a late stage of oocyte maturation (Ch11.2). In chapter 11.5, it was suggested that in natural reproduction, there is apparently a relationship between the extent of embryonic mortality (most of which is early) and the success of the **maternal to zygotic** transition (MtZ). Chapter 13.5.3 further elaborated on this concept by addressing the exemplary role of the maternal effect **gene** *Atrx* in mice, which leads to CIN in the early embryo when it is functionally absent in the egg. The maternal to zygotic transition is a gradual process that occurs from the 4- to the 8-cell stage in humans (fig 43). An obvious thought is that the hormonal induction of superovulation potentially contributes to the release of immature oocytes that are not well equipped to support early embryonic development. Variation in eggs plays a much bigger role in the early period, including the MtZ up to the 8-cell stage, than variation in sperm cells. The lesser importance of selection for the sperm is perfectly demonstrated by the success of ICSI, and especially TESE, which is achieved with artificially obtained sperm cells: the chances of conception with a first transfer are not much lower than with the use of ejaculated sperm.

At a scientific meeting in September 2019 in the Netherlands, Tom Stout from the national veterinary faculty in Utrecht presented his results on IVF technology in horses. I had never realised the scale of this application field of IVF, but cost considerations are apparently not an issue here, especially when the offspring that will be produced is worth a lot of money.

Superovulation is not used in horses. In the late follicular phase of the cycle, there are enough follicles larger than 8 mm on the surface of the ovary, from which the cumulus-oocyte complexes (Ch11.1, Ch13.7, fig 29) can be retrieved using a kind of scraping technique (this is called ovum pick up). The eggs obtained in this way are in the specific female diplotene stage (Ch11.1, Ch11.2, fig 28). After maturation, they are not suitable for classical IVF protocols in which they are brought into contact with an excess of motile sperm cells. In horses, the use of ICSI is the rule. Tom presented the results of 1,337 cycles spanning the years 2014-2019. In the last three of this period, his group achieved a 75% chance of foaling after the transfer of a blastocyst (but obtaining a good blastocyst is a lot harder than in humans). The yield of eggs steadily decreased from the age of 16 years, but growing follicles containing oocytes were also present in mares older than 20 years. The population was diverse and consisted of animals with fertility problems as well as horses intended for breeding after a sports career.

Biologically, the most significant aspect of ART in horses is the emergence of a more statistical approach to the variation in blastocyst yield, since clinical aspects do not dominate everything in the population that is used. The statistical value particularly relates to the parental effects of

stallions and mares. This advantage allowed Tom to convincingly demonstrate that stallions contribute less to the variation in blastocyst yield than mares. This was, for instance, related to the yield of eggs during the "pick up" procedure. Mares that were successful in producing an embryo in a first cycle, showed repeatability over subsequent cycles.

An indication of an interaction effect between mare and stallion was also found in this research in Utrecht, in the sense that changing the stallion could be beneficial for the mare's breeding.

Back in the day in the countryside, I was amazed about how some farms and some horses would have a foal every spring. Gestation is 11 months in horses, and mares are quickly mated again in the first period of heat after the birth of a foal. The good IVF outcomes from the Utrecht study could actually be related to this relatively short "time to pregnancy" and the repeatability that can be seen in natural reproduction.

16.4 Comparisons with spontaneous fertility

From the previous sections, it can be inferred that the line between spontaneous fertility and the use of ART is not always sharply defined. This is true, for example, in cases of undiagnosed infertility and fluctuating sperm counts. And there is also great variation in spontaneous fertility. In the field of gynaecology, the maximum probability of a successful pregnancy is considered to be 60% per cycle, a number that is only available to super-fertile couples. A figure showing the probability of pregnancy per consecutive cycle for the entire population was printed in 2003 as part of a clinical review in the *British Medical Journal*. The corresponding graph (fig 46) was discussed earlier in chapter 16.1. About 30% of couples will have conceived after one cycle, and this number rises to around 50% after two cycles; after that, it becomes harder. Again, the same question comes to mind: imagine you could create this same graph for the first and second child, how likely would it be that couples occupy the same successive positions in both graphs? Does a quick first conception serve as a predictor? Let's assume that partners choose each other without any intuition about each other's reproductive capacity. The percentage of less fertile men (with a fertility spectrum ranging from azoospermia to reduced fertility with a normal spermiogram) can be as high as about 10% of the population (fig 36). In that case, the fraction of women who can get pregnant in one attempt is already underestimated by that percentage. Following this, corrections for age and other predictors, such as BMI, can be applied to the entire population. Will you then reach the estimate of maximum fertility of 60% per cycle, the number that is generally accepted in gynaecology? And what is the statistical distribution of that estimate? In

chapter 11.5 on oocyte quality, I introduced the concept of maternal to zygotic transition to start exploring this topic. This was then elaborated on in chapter 13.5.3 and in this chapter (Ch16.3.2). Returning to the central equation in quantitative genetics (P = G + E, Ch6.1), the first assumption that can be made is that genetic variation also applies to the concept of oocyte quality and, therefore, to fertility per cycle (Ch11.5). Every cell biological function has its own genetics. The number of genes involved in female fertility is large (around 8,500 genes with clear expression in the ovary), but the proportion that is more exclusively involved is low compared to the testis (Ch12.5.2). Consequently, there are many possibilities for genetic variation, which is considerable altogether (Ch3, table 3). In (other) mammals, this is also evident in selection experiments on litter size (the mouse is an example). Another demonstration of the genetic component is that despite the low **heritability** of fertility (the h^2, Ch6.1), selection for this trait is possible in cattle breeding (also due to the large phenotypic variation).

In addition to the cell biological aspects of the egg, many other biological functions come together in fertility, such as the behaviour that enables reproduction, hormonal functions, and the environment. We devote little or no attention to the hormonal and environmental aspects in this book, but I would like to make an anecdotal exception for behavioural aspects, also because behaviour plays a role in seeking access to ART. Otherwise, this is the domain of specialists like the sexologist.

Here is an initial example that was published in 2004 by a group of American researchers affiliated with the National Institute of Environmental Health Science in Durham, North Carolina. The 285 women involved in this study, many of whom had a university background and were around 30 years old, tracked their sexual behaviour over 867 cycles. By measuring the metabolites of **oestrogens** and progesterone in urine, the researchers could estimate the time of ovulation (Ch13.4). The frequency of sexual intercourse increased in the follicular phase of the cycle (fig 40) and then sharply decreased on the day of ovulation. That aligns nicely with a goal of conception. However, the analysis also revealed something else. The data suggested that intercourse could induce ovulation. To further investigate that, the data were analysed in a different way, based on the level of sexual activity during weekends: the higher the frequency of sexual activity during that time, the greater the likelihood of ovulation taking place between Sunday and Tuesday.

Next, the second example, which is really old. Phenomena regarding reproduction can be rare, but many things are possible in the natural condition of humans. Such a situation was published by W. Pryll in the *Muenchener Medizinische Wochenschrift* in 1916. Married couples

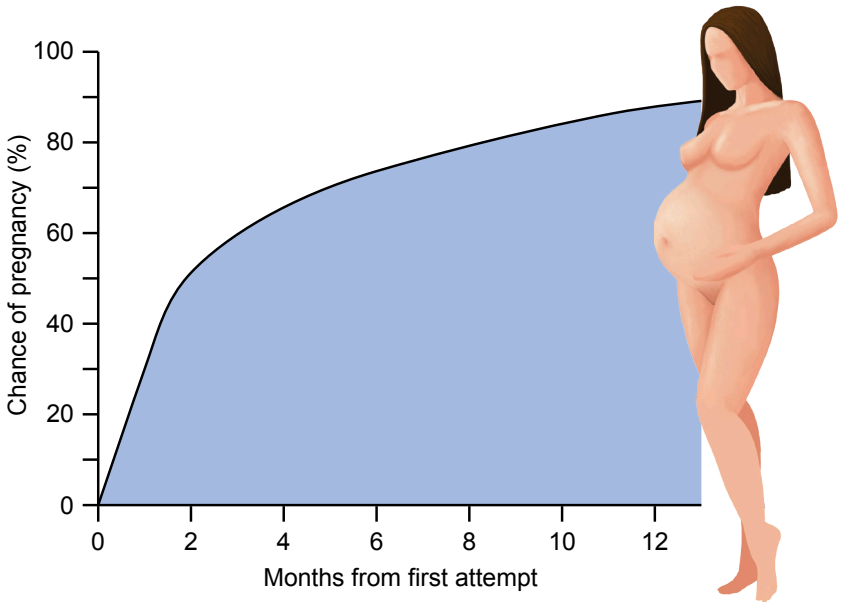


FIG 46

The cumulative chance of pregnancy during the first year of trying to conceive, as published by the *British Medical Journal* in a review article in 2003.

with a husband who was serving at the front (during World War I) and returned home on furlough, and with a wife who tracked her menstruation, were asked about their sexual interactions. Having coitus at least once led to 713 pregnancies in this population, as was analysed by this "Assistent der Kliniek" of the Universitäts-Frauenklinik in Erlangen, Germany. Conception, which was found to be possible from day 2 of the cycle, peaked over the days 6-12 after the onset of menstruation but could still occur up to 30 days later, although the likelihood gradually levelled off during this time. The remark that humans are not rabbits is, of course, obvious, but one can also imagine how much these people longed for each other, with passion being part of this. Such natural behaviour and the stimulating aspects of it are often no longer present in couples starting an IVF procedure.

Genetic predisposition plays a role in women who repeatedly quickly get pregnant, and that genetic predisposition also relates to the maturation process of oocytes prior to ovulation. So

when the quality of that process is an important limiting factor of fertility, I do not believe in that maximum fertility rate of 60% per cycle.

16.5 *De novo* mutation in ART offspring

IVF was introduced and quickly accepted without a solid foundation based on animal experiments. In hindsight, one could argue that this was not really possible either, given the rather large biological differences between reproduction in mice or other experimental animal species and humans. Moreover, in the period from 1970-1980 until the beginning of this century, cell biology techniques and DNA technologies were not sufficiently developed to be very helpful in the evaluation of eggs, sperm and early embryos. As mentioned earlier, the increase in referrals for ART mainly has a urological background. This is partly due to the more extensive search for sperm cells that are trapped behind an obstruction in the efferent duct system or unable to exit the epididymis or testis (fig 31), leading to PESA and TESE. As a consequence of this situation, it was important to quickly conduct thorough research on the course of pregnancy, birth, and the possible presence of congenital anomalies. All reviews on the characteristics of reproduction using IVF/ICSI report a slightly lower average birth weight and a slightly increased frequency of earlier births after fresh embryo transfer. Although these phenomena are most evident in the case of multiple births (of which there are not that many anymore, Ch16.3), they are also apparent in singleton pregnancies. From the very beginning, attention was paid to a possible increase in the percentage of children with congenital anomalies. In the literature, a statistic of around 2% can be found for cases of natural reproduction, but when a closer professional look is taken, this percentage seems to increase (see Ch18.1). Some indications appeared in professional journals that this percentage was somewhat elevated when ART was used, but other researchers contested this. The latest and largest meta-analysis to date includes a collection of 167 publications, and 48 of those, published between 1995 and 2012, met the quality requirements set by the researchers. In this study, a total of 112,913 children born after classical IVF or ICSI were compared with 4,471,368 naturally conceived children. Calculated from these data, the relative risk of having a child with a congenital defect after ART was 1.40, with a 95% confidence interval of 1.31-1.49. There was no difference between IVF children and ICSI children in this regard. The difference in risk between ART twins and “spontaneous” twins was smaller and not statistically significant. Apart from the reported increased risk of a child with a congenital defect after ART, results can vary locally. For instance, a relatively higher percentage was found in the USA after

ICSI due to a male factor (2021). What could all this mean biologically? Is there a genetic and/or epigenetic explanation?

First, let me focus on the relevant genetic knowledge on this (see the heading of this paragraph). In 2002, a study was published that was very comprehensive for its time, using classical **cytogenetics** to search for chromosomal aberrations (Ch3.4) in amniotic fluid cells and biopsies of chorionic villi (**foetal** placental tissue) from pregnancies after ICSI treatment. The work was conducted at the Free University of Brussels, Belgium, where ICSI was discovered (Ch16.1), and it was a collaboration between the IVF lab and the medical genetics laboratory at the same university. The injected sperm usually came from the ejaculate. Of the 1,586 embryos that were examined, 3% were found to have a chromosomal aberration, and in approximately half of the cases (1.6%), the abnormality did not occur in the parents (compare with table 4). Of these, 40% concerned the **sex chromosomes** (aneuploidy). 60% represented autosomal chromosomal aberrations, half of them being aneuploidies and the other half structural chromosomal aberrations (Ch3.4). Usually, the majority of these are reciprocal translocations (Ch3.4.1, fig 12,) which occur remarkably frequently in humans (table 4) and may occasionally be associated with a **phenotype**, for example, male infertility and subfertility (Ch12.12.1, fig 37). Usually, however, there is no phenotypic indication. The substantial mortality during the peri-implantation stage associated with unbalanced chromosomal aberrations (Ch3.3), which are caused by reciprocal translocations during **meiosis**, largely goes unnoticed in humans (Ch3.4.1). This type of chromosomal aberration is, therefore, rarely identified in women experiencing recurrent miscarriage. The reason why reciprocal translocations are of such interest to a geneticist lies in the fact that they result from double-strand DNA breaks. They can be viewed as the result of a brave attempt by the zygote (or early blastomere) to repair that break (Ch13.5.2). Another peculiarity is that reciprocal translocations mainly occur between the paternal chromosomes. We have known this since 1988 based on classical cytogenetics, and this picture is confirmed by further examination using modern NGS techniques. According to classical data from human cytogenetics (table 4), the *de novo* occurrence of structural chromosome aberrations (i.e. reciprocal translocations) is approximately 10 times higher in this series from Brussels.

It is, therefore, possible that ICSI increases the chance that double-strand breaks in sperm DNA end up in the egg or alternatively, that they are produced there. Less successful repair attempts by the DNA repair machinery of the egg cell can then result in unnaturally high numbers of reciprocal translocations (fig 12) in zygotes and the embryos that develop from them. Unfortunately, the majority of currently available platforms for whole genome research (WGS based on

short read sequencing) are not good at detecting exchanges of entire segments between chromosomes. It is not impossible, but it is laborious. The further development of DNA sequencing techniques that make it possible to determine the base sequence of longer DNA fragments at once (long read sequencing, Ch3.3) is fortunately on the rise and will greatly expand the possibilities in this area.

Something that current standard DNA analysis techniques are good at is detecting changes that cause a deviation from the diploid dose, such as deletions, duplications, and *de novo* **copy number variation** (CNVs, Ch3.3). Base pair changes, the *de novo* SNPs (Ch3.2), can now also be accurately detected. However, this latter approach, which allows for the identification of DNA **replication** inaccuracies (fig 5) and the workload of the BER DNA repair system (fig 33), has so far hardly been used to make a reliable comparison between mutation frequencies in spontaneous reproduction and when using ART. A recent pilot study published from Radboudumc in Nijmegen, the Netherlands, which uses trio sequencing, shows no difference between children born using TESE and the control group.

A study that primarily focused on the incidence of *de novo* chromosomal instability (the dosage abnormalities falling under CIN) and that compared natural conception (62 children/families) with ART (49 children and their parents) was published in *Nature Medicine* in November 2019. Following this, it also swiftly made its way into the science news of quality newspapers. Based on the occurrence of CIN in at least half to the majority of IVF embryos *in vitro* (Ch16.3.1), the researchers expected more CNV in the ART children. Remarkably, this was not found. Since there are many places in the genome where CIN can occur, for example as CNV (Ch3.3, table 3), a reliable statement can be made based on a smaller number of families. The fact that no differences were found is encouraging and instils a justified hope in the effectiveness of selective advantage, which gives embryonic cells with a “normal” genome structure the greatest chance of survival. Relatively small aberrations are, therefore, less likely to become permanently established in the embryo.

However, our understanding of this is not yet complete, as a genome-wide comparison of **mutation frequencies** between ART and non-ART children is still lacking to date (see table 3 for the usual classification of mutation types). In short, we currently do not know whether the *de novo* mutation frequency, measured across the entire spectrum, may still be increased after different forms of ART. Researching this is difficult because the patient population is so heterogeneous. First of all, there is, of course, interest in the mutation incidence in children of older women and older men with extremely low sperm counts and reduced sperm motility in the epididymis (PESA)

or in the ejaculate. It would also be interesting to know if the congenital abnormalities occurring with ART fall into the same spectrum of syndromes as those occurring with natural reproduction. In the previously mentioned Dutch study comparing culture media (Ch16.2), which included 300 babies, 3% had a major malformation and 4% had a minor malformation; a statement that no one will take lightheartedly. DNA (and **chromatin**) targeted follow-up research on these types of abnormalities could possibly provide answers to the questions that still exist about the genetic risks that are associated with the use of IVF/ART on such a large scale.

Very recently (2024), a publication has introduced a new avenue of genetic research, which relates to chapter 14. It provides a first insight into the prevalence of mtDNA mutations as part of research into the biological background of a lower birth weight associated with ART. A total of 270 ART probands were compared with 181 spontaneous conceptions. A higher rate of spontaneous mutations, particularly concerning non-synonymous base changes and mitochondrion-specific rRNAs (Ch14), is observed in the ART group (i.e. mitochondria have their own ribosomes and DNA repair characteristics, Ch14). The researchers relate their findings to female factor subfertility (a higher maternal age) and a relatively high superovulation response occurring at that age. The findings are interesting in relation to the protection of the primary oocyte against oxidative stress in the baseline reserve of primordial follicles (Ch11.1). Moreover, the association with the superovulation response (i.e. the management of the dynamic reserve, Ch11.1) sparks the speculation that mitochondrial genetic health may be a factor that follicle selection acts on during the development of the Graafian follicle. Overall, this research uncovers new aspects of the practice of ART, thereby inviting further investigation into the management of the baseline and dynamic oocyte reserves over the reproductive lifespan.

16.6 Does ART involve epigenetic changes?

The first indications of epigenetic consequences that may be associated with ART emerged in the 90s in the form of case reports on so-called **imprinting** disorders (Ch10). From a select set of genes, either the maternal or paternal copy is expressed in the embryo and/or in later stages. The epigenetic code for the regulation of the expression of each parent's **allele** is anchored in the **germline** of either the father or mother and it can be read from the CpG methylation status of the ICR (Ch10, fig 27). While the chromatin is reset during the period from the zygote stage to the blastocyst stage, the ICRs should maintain their methylation status (Ch9, Ch10). Most of the observations that are relevant here have been made in a region with the complex

abbreviation KCNQ1OT1, the ICR involved in Beckwith-Wiedemann syndrome. The imprint is maternal, meaning that the maternal copy of this ICR is methylated. In the non-ART population, the frequency of this disease lies around 1:14,000, so it is rare. With a birth rate of approximately 170,000 births per year in the Netherlands, around 12 children will be affected. A conservative estimate of the relative increase due to ART is a factor of four. If one in 30 children is born using ART, an increase of one to two children with this syndrome is expected per year in the Netherlands. While searching for the molecular cause of Beckwith-Wiedemann as a result of ART, researchers discovered that it is usually rooted in an inability to maintain the methylation status of the involved ICR: loss of methylation occurs. The cause of this should probably be sought in the period of *in vitro* culture (fig 23). However, the quality of the oocyte also enters the picture because the maintenance of imprinting is linked to the maternal to zygotic transition (Ch13.5.3, fig 24).

From the year 2000, publications emerged on the influence that the use of reproductive technology can have on the maintenance of imprinting in mice. The methylation status of the ICR of the paternally imprinted Igf2-H19 cluster (fig 27) was the most studied area in this context. The overall impression gained from these studies is that the efficiency of maintaining the methylation status of the ICR is reduced in IVF-related interventions, including superovulation. This effect is stronger for the embryonic part of the placenta than for the embryo itself (its extraembryonic part). This pattern effectively repeats itself in humans.

When researchers working with mice would visit their colleagues who conduct research in humans, they were sometimes surprised by the little impact their results made in that field. The most vivid demonstration of epigenetic influence by superovulation and/or *in vitro* culture (from zygote to blastocyst) can be found in the A^{vy} system (Ch15.2.2, fig 45) of the mouse. The coat colour of IVF-conceived A^{vy}/A offspring of A^{vy}/a fathers and wildtype *A/A* mothers in which superovulation was induced is noticeably lighter than that of naturally conceived A^{vy}/A offspring. Normally, the paternal **promoter** of the A^{vy} allele has no epigenetic memory, but now it responds to the unnatural conditions of superovulation and residence in the culture medium.

As genetic research has gained momentum thanks to technological developments (Ch3), the same is true for research into the significance of **epigenetics**. To determine the **methylome**, methylated cytosines (fig 11) can be distinguished from unmethylated ones after a chemical modification, which can be read out on a chip (Ch15.5). Recently, two publications appeared in which such an analysis was performed in the context of NGS in an ART setting. The first one compares genome methylation in newborns and adults aged between 22-35 years. The other

study focused on genomic DNA in the placenta. The genome contains around 29 x 10^6 CpGs. An 850k chip was used to measure methylation levels in babies and adults, and a 450k chip was used for the placenta. Both types of chips are enriched for the CpG-rich promoter regions of genes (Ch2.3.3). The larger 850k chip also picks up (m)CpGs located further away, around genes involved in their regulation. This chip is also suitable for determining the methylation status of the imprint control regions (ICRs, Ch10). The researchers who examined the genomic DNA of the same individuals at birth and in adulthood observed that in the ART group, there was a little more variation in the level of CpG methylation at birth, but that this had largely disappeared by adulthood. The cohort of 193 ART births and 86 control births they used consisted of healthy individuals. The 450k chip, on which DNA from 44 ART births and 44 controls was analysed, revealed greater variation in the placentas of the ART group (which resembles data in mice, see above). Moreover, abnormal patterns, which the researchers labelled as outliers, were also more frequently observed in this group. These were related to the known ICRs of imprinting gene clusters (fig 27), but not exclusively. Remarkably, the methylation pattern of genomic DNA after ovulation induction and IUI could be distinguished from the pattern after ICSI in these placentas within the ART group. The researchers also found methylation footprints associated with the father's age and impaired **spermatogenesis**. Whether this has lasting biological significance for the child is unknown. At the current state of research and understanding, it remains difficult to determine whether the different technological versions of ART each induce their own distinctive epigenetic signature. However, there are some initial indications that this is indeed the case.

Earlier, in chapter 16.2, the effect of the culture medium on birth weight was discussed. In the first decade of this century, research conducted at the Free University Amsterdam on the development of test tube babies into adolescents reported subtle effects of culture medium on blood pressure and glucose tolerance/insulin sensitivity. This was also found in ART trials in mice. Also worth mentioning is the difference in sex ratio. When conventional IVF is used, more boys are born than with ICSI, which seems to be the case even after the transfer of slightly more developed blastocysts. In 1993, the British researcher Paul Burgoyne reported that in mice, male embryos divide faster before implantation. This could lead to selection in the lab at the time of transfer. If this phenomenon also occurs in human embryos, it does not explain the difference between IVF and ICSI.

For large parts of the methylome (for example for positions not directly related to the regulation of **transcription**), the cell biological significance of differences in DNA methylation is still not easy to interpret, as was actually already indicated in chapter 15.5. Our insight on this will

grow in the near future as a result of the enormous increase in data. Therefore, the expectation is justified that we will soon know more about the mechanistic backgrounds of the phenomena outlined above, for both health purposes and for explaining phenotypic variation.

16.7 Is there a need for more ART?

Progressive liberals in the Netherlands tend to be welcoming to new developments in the area of assisted reproduction. Anyone and any type of couple should have access to options. That means there is work to be done by ethicists. Experiments in mice using *in vitro* **gametogenesis** with laboratory-induced **pluripotent stem cell**s (iPS cells, **somatic** cells that are reprogrammed, box 2) are currently a great source of inspiration for those who believe that everything should be realisable. After all, this could finally solve the problem of a shortage of eggs in human reproduction (for research and treatment). It could also open up possibilities for reproduction within same-sex couples (see Ch8 for male-to-female genetic sex reversal in mice). Sex reversal of a **germ cell** from female to male would require the loss of one X chromosome and the introduction of the Y chromosome, which certainly is more difficult.

As was observed earlier in this chapter, the world of artificial reproduction does have a strong drive towards enabling fertility for everyone, preferably with offspring that are genetically "their own". Here are some examples. In men whose sperm cells do not develop an **acrosome** during **spermiogenesis**, the sperm have a round nucleus (globozoospermia, table 6, often linked to a genetic diagnosis). These sperm do not activate the egg (Ch13.5.1). However, the Ca^{2+} spikes required for this can also be artificially induced. In the paragraph on this in chapter 13, some experimental data from mouse studies are provided. The treatment based on this knowledge has been internationally accepted for some time (for example, in Belgium), and, as of recently, it is also offered in the Netherlands. The following example is biologically more challenging. Some men have an arrest at or around the stage during which the metamorphosis from a round spermatid to a sperm cell with a condensed protamine-DNA structure takes place (Ch12.6). It is highly likely that a genetic defect plays a role in this (Ch12.12). Around the late 90s, articles appeared on the reliable microscopic identification of the round spermatids of these men in a wet preparation of a testicular biopsy, a process that is not easy. Of course, the aim was to inject the isolated spermatids into an egg (ROSI, round spermatid injection). This whole procedure has been further developed in Japan, and since 2016, it offers a new option to treat infertility. A publication from 2019 describes how 90 babies have been conceived with this method. It is not a

very efficient procedure, with a success rate considerably lower than that of ICSI with testicular sperm in non-obstructive azoospermia (TESE, Ch16.1, Ch16.3.2). Epigenetic research in mice looking at the very beginning of zygote formation after ROSI indicates such major deviations in the processing of male chromatin that it is a miracle that baby mice are born at all after the use of ROSI (as in humans, where a lower developmental potential is also seen). Also, the fact that nothing abnormal can initially be seen in humans or mice born after injection with round spermatids does not necessarily prove that all bodily functions are normal. Moreover, the question remains whether problems may occur over the long term, for example during growth into adulthood or during ageing.

Therefore, thorough follow-up research into the safety of each new development in the field of artificial reproductive techniques is essential; unfortunately, the use of many developments concerning ART in humans cannot be properly evaluated based on animal studies.

The attention for *in vitro* gametogenesis has been boosted by stem cell technology (iPSC, box 2) and is likely to stay. Another development in this field is the generation of primordial germ cell-like cells (PGCLCs) *in vitro*. These cells are derived from **embryonic stem cell**s that are equivalent to the **inner cell mass**, epiblast of the blastocyst (Ch7, fig 43), and from the aforementioned iPSCs. As can be seen in a recent review by leading researcher Katsuhoki Hayashi from Japan, *in vitro* gametogenesis was already a subject of interest a long time ago. He refers to experiments from 1920 in which attempts were made to let pieces of testicular tissue function outside the body. So it was not until much later that this type of research got a real boost, after the discovery of the induced pluripotent stem cell (iPSC, box 2) and the development of various methods to let them differentiate and specialise *in vitro*.

Further cell biological insight into this is largely beyond the scope of this book. However, in chapter 7 it was explained how this biological tool has been used, especially in humans, to reconstruct the early germline from the embryonic epiblast (fig 43).

Starting from the PGCLCs, a long *in vitro* process has to be undertaken to obtain mature gametes (eggs and sperm). The supporting cells required for this, such as Sertoli, granulosa and other follicle cells (fig 31 and fig 28), must be harvested from natural tissue material or ideally also generated *in vitro* via a stem cell model. In 2016, Katsuhiko Hayashi's group described in *Nature* how iPGCLCs were obtained from iPSCs, in this case derived from embryonic connective tissue cells or fibroblast cells from a tail tip. These iPGCLCs were then used to enable *in vitro* **oogenesis**. Using support cells isolated from the embryonic ovary, they developed a

system that produces oocytes *in vitro*. After completing meiosis and undergoing fertilisation, these oocytes could develop into adult mice. In the literature, a result like this is called a proof of principle, a demonstration that something is possible. However, when you look at the details in the publication, you will notice that the authors list several very serious indications of biological issues. These are problems that frequently appear during prenatal development from fertilisation onwards, and they lead to a particularly low efficiency of the procedure. The mice that were born were also not screened for genetic or epigenetic deviations. The Japanese researchers reported on the progress and current state of affairs, including the clinical perspective, in a comprehensive review that was published in October 2021. Clearly, the issues with the *in vitro* oogenesis technique have not yet been resolved. However, efforts are ongoing, especially in Japan. In September 2021, the group led by Mitinori Saitou reported on the birth of normally functioning mice generated from embryonic stem cells using *in vitro* spermatogenesis. Previously, the *in vivo* situation still played a more or less significant role in the complicated experimental set-up. Now, spermatogenesis was completed *in vitro* up to the round spermatid stage. After the round spermatids were injected into eggs (ROSI), viable mice were born which showed no apparent abnormalities at first glance (but see above). In this small experiment, proof of principle was obtained once more. A question that always remains in trials like this is whether the germline is adequately protected *in vitro*. After all, cell biological mechanisms, including the regulation of epigenetic characteristics (Ch2.3.2), can behave very differently *in vitro* under culture conditions compared to the *in vivo* situation. This aspect has recently (2023) been analysed with state-of-the-art methods. An important reason for the low efficiency of *in vitro* oogenesis in mice could be identified this way. During the development from **primordial germ cell** to mature primary oocyte, the epigenetic processes deviate from the expected course, resulting in systematic aberrations in gene expression. The Polycomb Repressive Complexes 1 and 2 are involved in this (see also Ch11.2). A consequence of this phenomenon is that gene activation is not well regulated from the zygote stage onwards. At the same time, the demethylation of maternal DNA (fig 23) does not proceed normally either. The fact that mice can be born from *in vitro* gametogenesis is currently more indicative of stochastic variation in gene expression than proof of principle. An additional problem is that the higher spontaneous mutation frequency of the iPSC, which is a somatic cell population that already harboured additional newly formed mutations from before the reprogramming event, is now directly introduced into the germline. The expected higher mutation rate of the *in vitro* gametogenesis is added to this.

Male gametogenesis is still almost impossible to achieve in a 100% *in vitro* system. Therefore, researcher Katsuhiko Hayashi concludes that there is still a long way to go. That said, what would be the purpose of this approach? After all, when it comes to saving endangered species, an argument often mentioned in articles on the topic of male and female *in vitro* gametogenesis, environmental conservation is a first and much more important prerequisite. Therefore, the question remains whether this technological leap is at all applicable to many species and whether it will come in time. When it comes to treatment methods that utilise CRISPR technology for human infertility with a genetic basis, it is far too early to transition to an ART approach that relies entirely on *in vitro* gametogenesis. Moreover, one can also justifiably question why we should pursue this, now or in the future.

However, the momentum of scientific research is great, as demonstrated by a very recent publication in *Nature* reporting the large scale *in vitro* genesis of embryonic oogonia and gonocytes/prespermatogonia. Science journalists have quickly heralded this as a next step in the battle against infertility and in the decoupling of biological sex and fertility. Amid this enthusiasm, the primary focus here, the resetting of chromatin in the germline (Ch9) and the variations of this, may receive less attention. Moreover, the challenge of meiosis looms ahead.

CHAPTER 17

Environmental influences

17.1 Introduction

When it comes to fertility problems, does the environment not play a role at all? Wasn't the formula P = G + E (Ch6.1)? For convenience, let's split the environment into the natural environment and the chemical environment that is added to it. This makes up the environment in which we currently live: it is a mix of natural conditions and chemical compounds that were introduced by humans and that may carry risks. A discipline related to the theme of this book is reproductive toxicology, which encompasses genetic toxicology. This subdiscipline deals with the **mutagenic** aspect of the environment. In this book, that side of reproductive genetics has only been touched upon for didactic and anecdotal purposes, with **ionising radiation** (Ch11, 12 and 13) and the substance BPA (Ch11) as examples. The reason for this is that the complexity of the field of **chromatin**-disrupting substances is so great that it requires more prior knowledge than is provided here.

The environmental effects of chemicals regularly make headlines, and their (potential) mutagenicity is discussed. In most cases, this concerns a single substance or a group of related substances. The term combination toxicology refers to the knowledge needed to understand what complex mixtures of compounds can do to cells and tissues. For example, a small-scale research project in the Netherlands in 2019 on dead nestlings of great tits from habitats that were labelled as "nature" revealed the presence of eight different pesticides at 10 locations, albeit in very low concentrations. This number is still not much compared to body-foreign chemicals that can be found in, for example, house dust. These cocktails of substances make matters so tricky; there is a high chance that interactions between substances will be involved in determining the

final effect. Already in the early days of chemical mutagenesis studies, in the 1970s-1980s, Bruce Cattanach (Ch12.7, from the Harwell MRC Radiobiology Unit) started combining the effect of chemical agents in the **germline** with the administration of ionising radiation. Both the experimental design and the interpretation of outcomes became considerably more complicated when just two agents were used at a fairly high dose. The interactions between many different chemicals are even more difficult to analyse, if at all possible, even when each of them has a low dose. However, the combined potential effect on the **genome** and on the **methylome** could be determined.

In chapter 12.10, we saw that a solid scientific foundation is beginning to emerge for the impression that human sperm production is declining. Indeed, articles about environmental effects on sperm production are easy to find in the publicly accessible database for biomedical scientific articles, PubMed (from the National Centre of Biotechnology Information in Bethesda, MD, part of the National Institutes of Health). The most valuable papers are the so-called meta-analyses, in which attempts are made to gather all publications on one topic, for example, the effect of antidepressants on **spermatogenesis**. A first screening is conducted to identify studies with a solid experimental design. Subsequently, all the selected research is combined in a model for statistical analysis. In many meta-analysis publications, observations in humans and in animal models complement each other. In 2018 and 2019, several meta-analyses were published regarding influences on the quantity and quality of human sperm.

Usually, there is no unambiguous, clear-cut conclusion but rather a situation of evidence accumulation (see introduction), which indicates advancing insight. Below (Ch17.2-17.4) are some examples of this. To be clear, this is not about tight underwear, sedentary occupations, whether or not combined with obesity, nor about underlying conditions, such as diabetes.

17.2 Social drugs: alcohol, smoking and cannabis

Normal consumption of alcohol is not associated with any abnormalities in male fertility. This is the conclusion based on publications covering a total of 16,395 men. The situation is more complex for smoking. The authoritative USA fertility journal *Fertility and Sterility* published a committee opinion about this in 2018, stating that smoking impairs fertility. This is most evident for women and for couples referred to an IVF clinic. For men, the situation is a little more subtle. There is a reasonable degree of certainty that smoking also has a negative effect on male

fertility, but it is not large enough to easily demonstrate an effect on conception. What is also mentioned is that maternal smoking could possibly affect the sperm production of her son.

Back in the 1970s, studies already emerged about a negative effect of cannabis on spermatogenesis in laboratory animals. In 2019, the entire field was scrutinised once again by the medical community. For humans, this has not yet led to a definitive disapproval of cannabis use, but it did lead to a call for doctors to mention the possibility of a negative effect when prescribing it as medicine. It was also recommended to inquire about marijuana use when men come in for fertility screening.

The harmful effects of smoking (and cannabis use), as well as chemical pollution in general, are now also being linked to changes in the methylome of sperm cells. In a recent review article, this phenomenon is seen as an important area of interest for further evidence-based research. It has already been known for some time that children of men who smoke are statistically more likely to develop leukaemia. An explanation for this that has been accepted among professionals is that smoking may actually have (epi)mutagenic effects in the male germline.

17.3 Chemical and physical influences

As mentioned in the introduction, reproductive toxicology, which includes the study of chemical and physical influences on the transmission of information to the next generation, is an extensive and complex scientific field. The reason to address it here is rooted in the observations already referred to in this book: the trend towards reduced sperm cell production in the Western world (Ch12.10). **Endocrine disruptor**s quickly come to mind as factors that contribute to this. However, many heavy metals, going by the name metalloestrogens, can also be disruptive via the same pathway. Furthermore, much attention is focused on the negative effects of pesticides. Chemical compounds that affect fertility are often classified into two classes: persistent and non-persistent (or easily degradable). Much of the literature on these topics is epidemiological in nature and exposure during the reproductive lifespan is studied. Far fewer articles study the effect of these compounds during the **embryo**nic and foetal stages. Obviously, compounds that are suspected to be harmful during the latter stages are also often harmful during the reproductive years, and vice versa. It is beyond the scope of this book to give an overview of the many chemicals with (potential) toxic effects on fertility. The reference list includes several review articles that list many candidate compounds. I am not aware of a meta-analysis with a systematic classification of these compounds. For most, there is cumulative evidence for their involvement

in fertility. As such, the indications do not look positive, which is consistent with the argument made by Niels Skakkebaek (the researcher who proposed the testicular dysgenesis syndrome hypothesis, Ch12.10) and colleagues. Their review from 2022 presents the hypothesis that the (chemical) effects of fossil fuel use are a denominator for the “infertility epidemic” (cited). In recent years, reviews have appeared on the effects of PFAS (an umbrella term for a group of many chemical formulations linked to endocrine disruptors) on spermatogenesis. Additionally, some papers have explored the potential effects of glyphosate on sperm production (see the list of references). Very recently, several review articles have been published on the effects of microplastics on reproductive health, specifically mentioning their effect in the blood-testis or Sertoli cell barrier (Ch12.5). The findings of these studies indicate that there is reason for caution, and they can also intuitively be associated with the increase in the rate of decline of human sperm production (Ch12.10).

Endocrine disruptors have received an increasing amount of attention from the scientific community when it concerns female reproductive health as well. The effect of substances like BPA (Ch11.4.3, Ch15.2.1,2), the pesticide methoxychlor (MXC), the dioxin TCDD, phthalates (plasticisers), and the soy compound genistein (Ch15.2.2) on the ovary have been well documented (see the list of references).

Research into the effects of medical drugs on **gametogenesis** is difficult. The effect of many drugs on, for example, spermatogenesis has not been thoroughly investigated yet. The extent of knowledge that is available is held by medical specialists, who need to be aware of the effects of cytostatic drugs. A drug like ibuprofen, a well-known anti-inflammatory and analgesic agent, seems suspicious, as shown by an analysis of semen samples from Jordan in 2019. Antidepressants are also mentioned; they have negative but reversible effects.

The negative effect of ionising radiation has been known for a long time and has been adequately covered earlier in this book. In animal experiments, the sensitivity of the spermatogonial stem cells to radiation has received extensive attention (Ch12.7). Discussions about these experiments often revolved around the balance between **DNA** repair and cell death. In fact, it was already known early on that ionising radiation can primarily cause complex damage to DNA, which can be visualised as a structural chromosomal aberration (such as reciprocal translocations, fig 12) with the use of **cytogenetic** techniques (Ch3.4.1, Ch12.7). Nowadays, we can complement this type of research with NGS and/or fancy new approaches such as Bionano (Ch3.4.1).

In terms of health, nothing remarkable has been found in descendants of survivors of the atomic bombings of Hiroshima and Nagasaki. However, large-scale genetic research utilising all currently available techniques has not (yet) been done. This type of research, conducted as trio sequencing (Ch3.2), is currently being performed on the DNA of children of the rescuers of the 1986 Chernobyl nuclear power plant explosion (the liquidators). At the single base level, this does not yield a significant effect. However, there are strong indications that a man who has been exposed to radiation can transmit more complex DNA damage (including reciprocal translocations, fig 12) to his children (see also Ch12.7). In humans, it is still unclear how DNA damage (especially double-strand DNA breaks) is processed by spermatogonial stem cells, which are, on average, not very mitotically active. Experiments that can shed light on this are only feasible *in vitro* (Ch16.7). And how much room does this leave for the selection of stem cells with less DNA damage and/or better DNA repair, which then become more dominant when the time between exposure and conception increases (Ch12.7)? Incidentally, the background levels of ionising radiation are so low that they do not play a role here.

For some years now, there has been more attention for the effects of non-ionising radiation that reaches us via mobile phones, laptops, microwave ovens, Wi-Fi connections, and so on. This type of radiation is commonly referred to as electrosmog, a popular umbrella term for electromagnetic pollution. The more official scientific abbreviation is RF-EMF, which stands for Radiofrequency Electromagnetic Fields. While the energy that tissues absorb from this electrosmog is not high enough to induce a break in a DNA strand, there are indications of a negative effect on spermatogenesis. Data miners from Kuala Lumpur sifted through 526 published papers on the subject (focusing on Wi-Fi connections), which they eventually narrowed down to 23 articles. Five of these involved human sperm samples, and the others focused on sperm cells and observations of testicular tissue from rats (most articles) and mice (a few articles). The variation in exposure conditions (all below the WHO standard) in this material was large. The processed publications frequently mentioned a temperature effect, and signals of **oxidative stress** were also regularly observed. As with ionising radiation, it would be helpful to know which cell types of spermatogenesis are most sensitive to RF-EMF, and also how spermatogenetic cells experience this type of radiation. If DNA lesions occur infrequently or not at all, would the methylome be a good readout parameter to use in answering this fundamental question? Studying effects on small non-coding (snc)RNAs (Ch12.5.2, Ch15.2.3) could be another potential starting point for this type of research (Ch15.2.3).

17.4 Oxidative stress and nutrition

Can the composition of our diet affect fertility? Possibly, if it induces a prolonged imbalance in the cellular reduction/oxidation (redox) balance. In fact, excessive oxidative stress can lead to lipid peroxidation which damages cell membranes in particular, and sperm cells are very sensitive to this (Ch12.7, Ch12.9.1, Ch12.9.4). “Lipid peroxidation” is a search term on its own. Meanwhile, several renowned researchers have attached their names to dietary supplements intended to help combat oxidative stress in the male reproductive system. For example, researcher John Aitken has been involved in the development of a dietary supplement with the lovely name “Fertilix”, the composition of which has been published. Supplements are also listed on a website for fertility medicine based on dietary supplements, run by Professor Frank Comhaire who was formerly based in Ghent, Belgium (Ch12.8). Besides selenium, which is directly involved in regulating the redox balance, zinc is also often present in these supplements.

Zinc aids in DNA packaging (with the help of protamines, Ch12.6 and Ch12.9.1), and it thereby contributes to maintaining DNA stability during ejaculation. Reliable scientific evidence for the effectiveness of food supplements on fertility is unfortunately lacking to date. It is difficult to obtain such evidence due to the heterogeneity and anonymity of the user group. The many anecdotal stories on this subject still lack scientific support.

17.5 Final remarks

Which of the effects discussed here, in particular those on spermatogenesis, fall entirely within the field of reproductive toxicology? Which ones can be categorised more specifically within the domain of genetic toxicology, and how should we deal with **epigenetic** effects that might be caused by aspects of nutrition (Ch15.2.1, Ch15.2.3, Ch15.5)?

Answers to these questions are generally not fully available because our knowledge has often not reached a sufficient level yet. The scientific value of a finding in the research field where it primarily belongs only becomes clear(er) when the molecular-cellular mechanism behind the effect is better understood. The concepts of safe (for chemicals) and minimal (for food supplements) dosage are relevant here. The unfamiliarity with potential interactions with other foreign substances, based on their presence as low dose multiples, resurfaces again. For ionising radiation, things are still rather straightforward: theoretically, there is no threshold dose for achieving an effect: any dose, no matter how small, causes proportional DNA damage.

In this chapter, I have mainly focused on compounds with a negative connotation in the field of spermatogenesis. Of course, there are also compounds like this that affect the oocyte. Unfortunately, much less is known about this. The role of nutrition in the molecular aspects of **oogenesis** is also a poorly developed topic (but see Ch15.2.5). Therefore, those who are interested in nutritional aspects affecting the female germline are referred to literature on farm animals.

In summary, and coming back to the opening lines of this chapter, one could say that, on a global scale, we are constantly artificially modifying our surrounding environment and diet without considering the high sensitivity of **gametes** and early embryos to environmental factors. The group led by Ann Van Soom from the veterinary faculty in Ghent wrote a beautiful book chapter about that sensitivity in 2017. This prompts the notion that humans themselves are the main cause of the systematic decline in at least male fertility. Could this trend also affect the load of new mutations that accumulates with every generation (table 3)?

An interesting observation that supports this suspicion is the significant reduction of 7% in the appearance of *de novo* base changes (table 3) among the Amish. The authors link this finding from 2020 to their lifestyle, which is inspired by the preindustrial period. Trio sequencing in isolated human communities that still live in an almost natural way could provide an even more interesting picture.

CHAPTER 18

From individual to population and vice versa

18.1 The concept of genetic load

In principle, genetic variation increases with each generation. This is because **DNA replication** in the **germline** is not flawless and because there are always cellular threats that can also alter the **genetic code**. Essentially, **mutation** is the result of that (table 3). Ancient variation can be observed in various ways, and it is nicely demonstrated by **SNP**s with an **allele frequency** of at least 1% in the human population (Ch3.2). Those are the SNPs that companies like 23andMe use to determine your genetic ancestry, and that can be used to determine genetic associations with traits (Ch3.2, Ch6.2). However, our total genetic variation, which is of course also located in the coding regions of **gene**s (box 1), is much larger. By aggregating the analyses of the coding DNA (box 1, fig 9) of the genes of 60,706 people, researchers concluded in 2016 that there can be a variant on every 8 bases. On top of this, there is also variation in significant sections beyond the coding DNA, such as **promoter**s and other base sequences involved in the regulation of **gene** activity and chromosome behaviour. A similar picture is obtained when the whole **genome** is analysed (table 3). The majority of genetic variation produces very small effects, and it is extremely difficult to attribute these effects to a single altered base position. A large part of the variation is neutral, meaning that there is no (or very little) selection pressure that acts on it. A minority has larger, adverse effects, such as the mutations in genes that appear in the OMIM (Ch4.2). Finally, there is a small minority of mutations with a positive effect. These are the mutations that enable Darwinian selection. Together, the mutations with a smaller or larger negative contribution to the individual are referred to as the genetic or mutational load, and logically, something is added to this in each generation (table 3). In the context of ART (Ch16.5),

the question has arisen whether an artificial method of reproduction promotes the introduction of new mutations. The technical procedures used in ART could potentially have an additional **mutagenic** aspect. Fortunately, natural selection at the **embryo** level ensures that a barrier is raised against this. However, we do not yet know whether this barrier is high enough to protect against all types of mutations (table 3).

The renowned quantitative and population geneticist Michael Lynch has summarised the situation under the title “Mutation and human exceptionalism; our future genetic load”. Michael Lynch is known, among other things, for the excellent textbook on quantitative genetics with the broad title *Evolution and selection of quantitative traits* (Ch6.1). Here, “human exceptionalism” means that due to the influence we have on our environment, and partly due to the social and knowledge systems, for instance those used in healthcare, we have curbed the role of natural selection processes. In the theory of population genetics, the generation of mutations with a negative effect and their elimination due to reproductive disadvantage are in balance. Natural selection thus has important functional significance. When it no longer operates to its full extent, humanity is confronted with the negative aspect of mutation due to DNA changes that have occurred over generations. For our species, this could result in a decline in overall health and deteriorating fertility. Such a scenario is also known as genetic erosion.

That was, in a nutshell, the theme of Rolf Hoekstra’s valedictory lecture, which he gave at the end of his career as professor of genetics at Wageningen University in 2010. After all, we have already experienced a few generations of greatly reduced natural selection in the Western world. If we extrapolate this into the future, how many more generations would it take before the resulting problems are really starting to be noticeable around you and possibly within yourself? Rolf, known from the university textbook *Evolution, an introduction*, stood before a gradually more bewildered audience. From his lecture, I remember a number of 10 generations before a declining **fitness** effect becomes visible (an estimate based on work with model organisms). A similar estimate is also mentioned by Michael Lynch. With a decrease of 1% per generation, he calculates that we will already be experiencing it after a few generations (over a period of about 100 years) in a society with a high level of medical care. That 1% is a conservative assumption. Is there really no more selection at the individual level? This is still the case, but as more early lives and lives up to reproductive age are saved, the selection becomes less and less. At the individual level and at the group level, saving lives is something that is deeply embedded in our culture. This argument is still separate from the growth of ART; the extent of population genetic effects of this on overall health and fertility cannot be determined yet (see, for example, Ch12.11).

And of course, there are also smaller trends that can increase fitness, such as the increase in marriages between members of different global populations with a longer genetic history.

The practice of IVF also contributes to the reproduction of fitness-reducing mutations. This aspect was already encountered in the description of AZFc deletions (Ch12.12.2). The first sperm counts of men conceived via ICSI were published in 2016. Their sperm concentration was a little less than half the level of their time- and age-matched peers. This is just one example of genetic erosion, but one that will also apply to female subfertility and infertility. Fertility simply has an extensive genetic basis, with many genes involved in it.

One way in which the genetic load manifests itself in a very intrusive way is through monogenic, single gene-based **congenital disorders**. More than 4,000 of these are now described in the OMIM (Ch4.2, 4,526). Via derived statistics, it has been determined that 55.0% are inherited **recessive**ly, 37.0% **dominant**ly, 2.7% both recessively and dominantly, and 5.3% are localised on the X chromosome (and are thus subject to X chromosome inactivation, box 4). The vast majority of these are very rare, but together they constitute quite a number! Measured in generations and interconnectedness (Ch6.2), our evolutionary history is old enough to make us all carriers of at least two recessively inherited mutations, and that still seems to be an underestimation. This, in turn, has the consequence that one in 100 couples will genetically experience a **heterozygous** x heterozygous mating which, when it leads to a **homozygous** (-/-) severely affected child, impacts the family like a tragic accident that seemingly emerged out of nowhere. One must learn to live with a situation like this using all the medical and social help that is available.

Monogenic disorders manifest at different stages of life. An estimate of their impact up to the age of 25 years, based on the literature, is that 1.7 in 1,000 are affected through recessive inheritance and 1.4 in 1,000 through a dominant inheritance pattern. Based on the prevalence of a recessive pattern in one in 100 couples, one would expect 2.5 in 1,000 to be affected by this category over the entire lifespan. All of this applies to a population in which the degree of genetic relatedness between partners is low. The concept of consanguinity applies to 0.06% of couples with Dutch ancestry. Of the approximately 2,500 currently known genes with “negative” **allele**s that are recessive, 115, those with the highest allele frequencies, are involved in the conditions of around 90% of the children affected in this way.

For the population, the experienced disease burden of monogenic disorders is part of a much larger group of mostly rare diseases that have been of interest to physicians and geneticists for some time due to suspected genetic causes. The definition of this category of conditions is that they occur in less than one in 2,000 individuals. It is estimated that there are approximately

10,000 such conditions in total, meaning that the disease incidence in the population could be as high as 6%. When considering the use of a **polygenic** model (Ch6) versus a monogenic model for these diseases, there are several arguments to start with the latter. Even in the polygenic model, a single gene can be a game changer. And secondly, we are much more aware nowadays that many genes are involved in various bodily processes, with brain functioning as an extreme example of this. It is, therefore, not surprising that in 2014, the Genetics Department at Radboudumc in Nijmegen reported a breakthrough for the monogenic scenario. The coding DNA (the **exome**) of offspring and parents was compared in a small sample of 50 patients with an IQ score lower than 50. In this way, researchers identified 20 dominant *de novo* mutations that could pass the criterion of a reliable diagnosis. Trio sequencing (Ch3.2) thus proved to be a very promising strategy to investigate the genetic basis of our bodily processes, in this case brain functions, via mutations that have spontaneously occurred. This strategy also draws attention to mutations that have not yet been accepted as markers for medical diagnosis, while the genes in which they occur may also be involved in the affected process. About 80-90% of new mutations fall into the base change and **indel** categories (table 3). When the *de novo* **mutation frequency** given in table 3 is corrected for the coding part of the genome (box 1), a number of one to two mutations per generation is obtained. Around 80% of these are expected to affect the protein, and a slightly lower percentage will have an effect on its functioning (see also box 1). These events are surely rare, yet extremely effective for the detection of the genetic basis of a **phenotype**.

Trio sequencing is now being used on an increasingly large scale. An example of this is the international consortium “Deciphering Developmental Disorder” (DDD). The genomic DNAs of 31,058 trios were analysed in a report from 2020, involving Christian Gilissen and colleagues from the Genetics Department of Radboudumc. In total, 24,348 *de novo* base changes and 4,229 indels were found in the coding regions (table 3). From this harvest, the researchers distilled 257 genes that can diagnostically reliably be associated with developmental disorders based on the literature and databases. In addition, 28 candidate genes emerged, some of which had already been reported via the OMIM as well. The use of **WGS** results in an increase in mutations of about 10%; logically, these are more often classified as coarse DNA changes (the structural variants of table 3, including **copy number variation** (**CNV**)). Since very few mutations that are directly linked to the exome are added per generation per person, a large sample size is necessary to study a trait that involves a lot of genes in order to reach some kind of saturation of the mutation target (read: the number of genes involved in the trait). In other words,

all genes in which a dominant mutation can lead to the clinical picture (usually due to **haploinsufficiency**) have now had a chance to mutate. This study mentions that the number of patients that would need to be examined in relation to their parents' DNA to reach this criterion is more than 300,000. This is still an (overly) simple representation of this complicated branch of genetics, but it does indicate the direction of research developments for the near future. All in all, the research conducted by this consortium led to the discovery of a *de novo* mutation in about 25% of patients. This percentage does depend on the disease under investigation. For characteristics such as deafness, visual impairment, or intellectual disability, the genetically resolved fraction lies around 40% when all genetic information is used (DNA testing in relatives is often involved). This statistic was also found in the most recent DDD study from the UK and Ireland, which was based on family SNP analysis and **WES**, and which included approximately 13,500 children. In another 22% of these children, mutations were identified in genes that are highly likely to have a cell biological relationship with a gene in which a mutation is already known to have a negative effect on development. However, the mutation found in a known gene or candidate gene may not be associated with a distinct phenotype in some individuals or families. This phenomenon is known as incomplete penetrance, a very old term.

This is a direct consequence of the fact that, as mentioned earlier, gene products do not operate in isolation but often have their role in large to highly complex cellular complexes, or "protein machines". These complexes are often regulated by **PTM**s and the genetics behind them, which means that the functions of individual subunits in such a complex are not always expressed. In that case, it is even more surprising that a dominant loss-of-function mutation has a negative effect on the phenotype so frequently. The gene dosage, expressed by the term haploinsufficiency, is thus important for many genes (Ch4.2). In line with this, we have known for a long time that a deletion of a section of the genome is more severe than a duplication, also in the case of **diploid**y (Ch3.4.2, Ch11.4, Ch16.3.1). *De novo* mutations in the germline mainly occur during **gametogenesis**, but early embryonic development is also sensitive to these mutations: a small fraction, less than 10%, emerge during this stage of life. The results of population studies on the origin of an abnormal phenotype thus arise from collaboration between experts in the clinical field, **genomics**, and bioinformatics, all performing high-quality work.

From the above, it is apparent that in the analysis of the role of mutations in explaining the disorder, *de novo* mutagenesis is quantitatively more important than the Mendelian inheritance of mutations that are present in the population. Recessive Mendelian inheritance explains 3.6% of developmental disorders in a population with a European background. In couples with a high

chance of genetic relatedness, such as people with a Pakistani background, where marriages between relatives are frequent, this is 31%, which makes the picture quite different.

18.2 How do we deal with the burden of genetic load?

Genetic erosion is a sensitive issue. When we breed crops for the purpose of food production, the goal is to eliminate genetic erosion, or those minor detrimental alleles, also known as slight detrimentals (Ch12.12.3), and selection is the proven method to achieve this. Some genetic variation should still be preserved, since there is no selection response without it. Agreements are made about the increase in inbreeding that is inevitable with a small effective population size (Ne, Ch6.2). In farm animals, breeding organisations (these are companies, often multinationals) try to eliminate newly occurring congenital abnormalities as quickly as possible.

While the international Holstein breeding industry of dairy cows raised for milk production operates with a Ne value of around 100, the estimate of this number for the original European human population is 10,000 (Ch6). This is an extremely large number for a species. With such a high reproductive capacity as we witness today, every type of mutation that can occur in DNA is bound to occur at some point. However, when such mutations are recessive, they will initially be invisible for very many generations, which is how they get a chance to spread among the population.

On the other side of the broad spectrum caused by the social manifestation of the genetic load are the honourable medical professionals who are confronted with the suffering that this generates on an almost daily basis. It is, therefore, not surprising that scientists sometimes wonder whether DNA screening and genetic engineering could be used to stop this. For technical and logistical reasons, this is not feasible for new spontaneously emerging *de novo* genetic abnormalities, and as noted earlier, that includes a considerably large portion of congenital abnormalities. But what could and should we do about recessively inherited alleles that already have a long history in the population?

After all, one in 100 couples is comprised of two individuals who are heterozygous for a negative allele of the same gene, which often leads to a severe phenotype when it is homozygously present. The search for the heterozygous presence of such an allele in couples who wish to conceive is called preconception carrier screening. There are currently no uniform agreements on this procedure in the Netherlands. The services that are offered around this depend on initiatives in university medical centres. For instance, Amsterdam UMC offers screening for specific

target groups, such as people from Volendam (a town in the Netherlands) and people with an Ashkenazi Jewish background. In both groups, there is an increased frequency of carriers of a harmfully mutated allele. Other groups that are mentioned in this context are people with Mediterranean and Surinamese/Antillean backgrounds. The number of genes that is examined during this screening is 50. In the Groningen region (the Netherlands), people can get tested for heterozygosity for 90 genes involved in 70 monogenic recessively inherited conditions after referral by a local general practitioner. A much larger and newer initiative comes from the Clinical Genetics Departments of the university medical centres in Maastricht and Nijmegen, directed at couples who share ancestors in more recent generations: there is consanguinity. The risk of having a child with a congenital abnormality is about 2.5x higher in these situations, with full-cousin marriages carrying the greatest risk. The increase in pathologies with a monogenic recessive inheritance pattern very significantly contributes to this. In Maastricht, couples with a family history of consanguinity can be screened for carrier status of negative mutations in 2,337 genes (as of 31-08-2023) based on the OMIM, which will result in the recessive phenotype when inherited from both parents. Based on theory, it was already clear that, especially in cases of consanguinity, all genes that can result in a homozygous recessive phenotype should be included in the screening using exome sequencing. The mutation of a single gene that is present in both mother and father is not necessarily the same at the DNA level. The results of the first 100 couples were published in 2021. In 28 couples, the partners shared a mutation of the same gene. Statistically expressed, the chance of a child with a recessive phenotype is then 7% for this group. Consanguinity in partner choice is culturally determined, has an element of social and economic advantage, and happens in large parts of the world. It occurs in as many as 20-50% of marriages, depending on the society. The more closely related the parents are, the larger the portion of the genome that is shared according to a theoretically supported estimate (identity by descent, with a confidence interval). The current state of affairs in the Netherlands is that the government has been given the advice to implement preconceptual screening for the entire population using WES. This would be conducted in a study setting that includes periconceptual care. The prospect is that the birth of several hundred severely handicapped children can be prevented this way, and at the same time, this strategy increases the reproductive autonomy of prospective parents.

Technologically, the most spectacular way to prevent the birth of monogenic genetic pathology is PGT, preimplantation genetic testing of the embryo (see list of abbreviations for variants). This always involves an IVF procedure with the aim of enabling embryo selection. Since the

genetic diagnosis of a single **blastomere** (8-cell embryos) or single cells (from the trophoblast of **blastocysts**, fig 43) takes time, the embryos need to be frozen, as embryo transfer can obviously only be done after the results are known. In the Netherlands, PGT has been operational at the University Medical Centre of Maastricht since 1995, covering a steadily expanding number of monogenic disorders. This includes dominantly inherited genes (the majority of requests/analyses, 62% as reported in 2022) and recessively inherited genes (15%), or X chromosome-linked genes (13%). Families in which a structural chromosomal abnormality (reciprocal and Robertsonian translocations, PGT-SR, Ch3.4.1, fig 12) is present and where there is a realistic chance of affected offspring are also eligible for testing (10%).

In the case of a dominantly inherited condition, such as Huntington's disease, the fraction of embryos that cannot be transferred is the highest, 50%. In recessive conditions, it is the lowest (25%), because carriers do not express the phenotype. In cases of X-linked conditions, half of the male embryos are affected (25%), but the biological and social process is more complicated (box 4).

The following example serves as an explanation. A fairly common (affecting about one in 4,000 boys who are born) and dramatic example is muscular dystrophy (Duchenne and Becker). The mutated dystrophin gene that is responsible for this, which is one of the largest genes in the human genome, has a high *de novo* mutation frequency of 8×10^{-5} per generation. Via women who are carriers, the mutated gene can become fully expressed in sons who have inherited the "wrong" X chromosome, and the mutation can persist in the population through the 50% of daughters of these women who will also be carriers. In principle, selection against the affected X chromosome at the embryo level has an efficiency (the extent to which the mutant allele is removed) of 100%. This means that, on average, 50% of the embryos will be left, namely those carrying the "right" X chromosome. However, selection to this extent is not permitted under the current embryo law in the Netherlands because female embryos carrying the mutation could lead a "normal" postnatal life. The reality around the decision-making in these types of X-linked hereditary disorders is, therefore, inevitably nuanced and depends, among other things, on the mutated gene. Carriers are not always completely symptom-free. Also, knowing that a daughter is also a carrier and that she may face the same dilemmas later on is psychologically stressful for families. This can also be an argument in favour of using embryo selection on female embryos. A decision like that statistically further reduces the number of embryos that can be returned by one-third.

Does PGT, more specifically PGT-M, change the frequencies of the affected alleles in the population? For dominant conditions, yes. The balance between the appearance and disappearance of the mutation will shift to a lower allele frequency. For recessive conditions, one would also need to select against the heterozygotes, because that is where the majority of mutant alleles are found. This is obviously not possible because we are all carriers (Ch18.1). For X-linked conditions, the balance between creation and elimination shifts in a more favourable direction when selecting against male embryos, but not as rapidly as it does for dominantly inherited diseases. In summary, PGT is thus mainly aimed at preventing the continuation of genetic suffering in families. A population genetic effect is most evident for the dominantly inherited monogenic disorders, which currently constitute the vast majority of cases. The actual bottleneck of PGT lies in the number of viable embryos after superovulation in normally fertile couples at a reproductive age that is around the average (Ch16.1, Ch16.3). After embryo transfers in the Netherlands in 2021, 183 children were born following PGT, a low number compared to a total of nearly 180,000 births. Embryo selection as described here does not result in significantly lower chances of a successful pregnancy than those commonly seen in IVF practice (Ch16.3).

In the context of reducing the genetic load of monogenic disorders, attention is also directed towards the possibilities that the CRISPR-Cas (box 3) methodology can offer in terms of mutation correction. First, let me mention something about the technical side of this. After initial euphoria about this method, the attitude has become much more cautious; the technique must be 100% reliable when used in the **zygote**, even better than nature itself, so to speak. In addition, the **genotype** of the zygote is unknown at the time of application. Also, the technique should not mutate the unaffected positive allele. As introduced in box 3, the molecular toolset around the CRISP-Cas system is constantly evolving. The double-strand break can be replaced by a single-strand break, and the accuracy with which mutations can be repaired down to the level of a single base or several bases can be controlled through so-called base editing and prime editing. However, the desired result is not yet guaranteed and needs to be controlled. Results on the efficiency of prime editing in the human zygote were not yet known in 2022, and research into the applicability of the CRISPR-Cas system in the human germline does not seem to be of high priority everywhere.

Early on, experiments focused on gene editing (box 3) in human zygotes were already carried out in China. In that sense, dr Jiankui He's experiment that resulted in the birth of gene-edited babies, for which he was sentenced to 3 years in prison in late 2019, was not an isolated incident. The experiments were possible due to the lack of some form of ethical and societal

control over this work; once again, application outpaced insight. In Europe and in the USA, it is still legally prohibited to introduce genetic changes at the zygote level that will be integrated into the germline. Reproductive geneticists connect the application of CRISPR-Cas with PGT. They consider the possibility of genetically repairing all embryos, including those that are heterozygous for a recessive monogenic disorder (from +/- to +/+). This would reduce the load of recessive negative alleles in the population, but the rate of reduction will be very slow. Also, every mutation requires an appropriate strategy to be worked out to achieve DNA repair. That is related to a phenomenon called **mutation spectrum**, which refers to the full range of DNA base sequence modifications that can disable the gene and that are found in the population. The extensiveness of that spectrum varies between genes. The advantage is clearer at the family level, as preconception screening for the affected gene would then no longer be necessary for future generations.

A new branch in PGT has received the addition P, for polygenic. Due to the extremely low amount of DNA that can be extracted from cellular biopsies of 8-cell embryos or blastocysts, the technological challenge of this form of testing is high; the score is based on an SNP profile of the preimplantation embryo biopsy. A *Nature* Comment from 2022 lists three companies that are active in this field in the USA, where, contrary to Europe, PGT-P is permitted. Merely for the reason that there are few embryos to choose from (good embryos are relatively scarce, Ch16.3.1, Ch16.3.2), selection within a family is not very effective. However, the first babies from this type of embryo selection have been born by now. Several approaches are undertaken to tackle the notoriously large confidence intervals of personal **polygenic risk score**s (Ch6.2 and appendix). This form of testing covers diseases in which the number of genes with a large impact is not very high, such as type 1 diabetes, certain heart conditions, and susceptibility to cancer: the major part of the **genetic architecture** (Ch6.2) of the trait must be known. In order to increase predictive power, one approach is to focus on one end of the score distribution from a population of embryos. An additional, alternative approach is to include parental DNA (using WGS) and to reconstruct the genome of the embryo based on its genome-wide SNP profile. This method gives accurate predictions when it is tested on monofactorial conditions such as *BRCA1*. Some people dream of the use of PGT-P for genetic enhancements, which obviously sparks public debate, also in the USA. Another general consideration around PGT-P is whether it should be offered as a public good or as a for-profit service, the latter being much more likely.

De novo genetic chromosomal aberrations (mainly number aberrations) can be detected prenatally with the NIPT test (non-invasive prenatal test). In the Netherlands, this test is now

offered to anyone who is pregnant. The test relies on the presence of small quantities of DNA from the **foetal** part of the placenta in the mother's blood. Using NGS methodology, this DNA can be detected from as early as 11 weeks of pregnancy. The analysis will show the number of copies of an entire chromosome or a part of it. The resolving power for chromosome segments is 10-15 mb of DNA, comparable to the quantity of DNA in a chromosomal band (see fig 12). To evaluate this technology for groups at risk of having a child with a chromosomal aberration and for the entire population, some great and very robust research has been conducted in the Netherlands, in which all academic hospitals in the country are involved. The **sex chromosomes** are not analysed. Cases of **trisomies** (Ch3.4.2) involve the well-known chromosomes 13, 18 and 21, as these have the greatest potential to develop. Other trisomies are occasionally found as well. The NIPT result must be confirmed using an existing invasive procedure such as the culturing of amniotic fluid cells (from the embryo) or a chorionic villus sampling (from the foetal placenta). With the NIPT test and subsequent validation, an abnormal chromosome pattern is found in about 0.5% of pregnancies, the majority of which are trisomies 13, 18 and 21. Fortunately, the incidence of cytogenetic abnormalities other than the known trisomies, referred to as incidental findings, is very low (0.064%). In summary, the researchers report that the process of preventing the birth of a child with a congenital chromosomal abnormality is burdensome. The debate on population-wide screening for this will, therefore, continue, which is also reflected in the current participation rate of 60%. Since 2007, the triple test has been offered in the Netherlands to detect, among other things, the presence of a **foetus** with Down syndrome. The NIPT test has been offered since 2014, and the triple test stopped being offered in October 2021. Upon evaluation, it is now apparent that NIPT has not led to a further decline in the births of children with Down syndrome in the Netherlands. However, using table 4, it can be deduced that in Denmark, the proportion of children born with Down syndrome has decreased by about 70% due to prenatal testing.

For cattle breeders, the concept of genomic prediction can be considered as the water pump pliers of animal breeding through selection. The more accurately the genomic potential of a breeding animal can be determined, the faster the breeding goal is achieved (Ch6.2). In Holstein-Friesian breeding, part of the selection takes place at the level of preimplantation embryos and is then based on an SNP profile (H6.2). A chip with 10,000 SNP positions is truly sufficient to obtain a reliable estimate of the breeding value (Ch6.2) because the structure of cattle breeding is unique (the combination of a small effective population size (Ne) with large families of half-sisters (a bull used for artificial insemination produces many daughters). In this

way, embryos derived from superovulation and/or *in vitro* embryo production (as in horses, see Ch16.3.2) can be ranked for traits such as the genetic predisposition for milk production. This can shorten the **generation interval** and increase selection efficiency, but how useful is that (fig 18)?

What else can be done to reduce the genetic load when genetic selection at the reproductive age is not at play? A discussion about this question arose during a university meeting of colleagues in the field, at which the speaker gave a presentation on congenital hereditary disorders in the context of daily practice in a hospital. At the end of the impressive talk, the following question was posed to the attendees: should we introduce prenatal genetic testing for everyone? Then, follow-up questions arise: as a technological advancement of NIPT, what would such a test be able to detect? All disease-associated alleles as they appear in the OMIM (Ch4.2), or even more? Will genes that are susceptible to *de novo* mutation and that are more or less absolutely linked to a severe dominant phenotype also be in the spotlight? While the biological and technical aspects of these questions were being contemplated, issues of a very different nature came up. One that quickly arose is the following: as the amount of DNA information that can be given biological significance increases, and technology to perform prenatal DNA analysis continues to develop, where are the limits? What is considered abnormal, and what is not? The sliding scale in the answers that can be given to this will naturally emerge. Also, what effect does it have on the peace of mind during an early pregnancy(?), and what other ethical and moral considerations are there? During the meeting, no one was well prepared for these questions and the discussions about them. The main question lingered; it was all too much for that time. Nevertheless, the available technology combined with the severity of congenital disorders will influence how we deal with these cases in our society in the future. “As we search, we will find”, in the words of an adept. However, genetic variation is simply a continuum promoted by what Michael Lynch refers to as our human exceptionalism (Ch18.1).

Genetic variation is thus a key aspect of the **nature-nurture** paradigm. Genetic change occurs slowly in humans. It is driven by *de novo* mutagenesis and also by partner choice (which also includes the concept of social monogamy).

As is obvious, the environment changes much faster, and the drivers of these changes are not unknown. The autonomy of any person does not end at reproduction. Rather, it is a behavioural paradigm that can be the norm or that is pursued, which manifests itself in the unlimited availability of goods and services to anyone who can afford them. Since the group of people seeking and able to pursue autonomy is growing rapidly, the environmental implications are

enormous. The countermovement, a less materialistic attitude, with a more natural lifestyle and pursuit of alternative values in art, spirituality, and the experience of nature, has always been much smaller and will only continue to shrink in relative size. The environment, including our actions in the social context, ultimately dominates the natural genetic changes that humans undergo. This also applies to the effects on our fertility and genetic health.

CHAPTER 19

Concluding remarks

A journey through the **germline**, it has been and still is a fascinating expedition. The technical possibilities within molecular and genetic research continue to expand, especially towards the characterisation of **DNA** and **RNA** in individual cells, thereby making variation between cells visible. The advancements also extend to the molecular architecture of cells in general. The global field of ART is the gateway to a better understanding of the cellular and genetic aspects of **embryo**nic development in humans, especially the initial stages. The results of DNA analysis using **SNP** and **NGS** platforms per single **blastomere** are already extraordinary, and the **methylome** and characterisation of the library of RNA molecules (the **transcriptome**) are added to this. Will that lead to the discovery of **biomarker**s that will accurately predict which embryo will make it and which will not? The characterisation of individual cells is also advancing in the field of **spermatogenesis** research, for example for the purpose of understanding the behaviour of stem cells.

As single cells and the populations they form are now also being subjected to big data analysis, curiosity arises about the answers this may bring. So far, these glimpses into cell biology have mainly come from individual genes in mouse **knockout** models with a protein as the missing product. However, the enhancement in understanding cell biological processes is becoming increasingly difficult, especially as we gain more system-level knowledge about the interconnectedness of molecular interactions within the cell. The unfolding spectrum of the functions of non-coding RNA is also a part of this. The fact that the entire molecular biology of the cell is comprised of flexible interactions is incomprehensible in its total complexity (Ch1, a simple example of this in fig 24).

19.1 Sexual reproduction, a miracle

A journey through the germline is equivalent to embracing sexual reproduction, with its fabulous ability to generate variation, and, therefore, also the ability to adapt when the environment changes. Even for professional enthusiasts, the literature is vast, maybe too vast.

Biology is sometimes described as the science of variation. So where does the obsession with cloning come from? It is an obsession that is perhaps not so noticeable in the Netherlands, but certainly in many other cultures, particularly in East Asia and South America.

My mind goes back to the summer of 1969, when I was visiting the ARC Reproductive Physiology Unit on Huntington Road in Cambridge. After my first microscopy experience with oocytes of cattle in Göttingen in 1969, I was eager to see how an expert would mount these miraculous cells onto a glass slide, place a coverslip over them, and how they would stain them in order to observe the meiotic stage. Chris Polge, the inventor of **cryopreservation** of sperm (in 1949), was happy to demonstrate this technique on a sunny summer afternoon; it was really not complicated. There was someone else who also watched him do this: now the late Ian Wilmut, the intellectual father of the first cloned mammal, Dolly, the sheep named after Dolly Parton (*Nature* 1996).

In 1998, I ran into Wilmut again at a Gordon conference in New Hampshire. He had a tough time when he gave a talk about the experiments in Edinburgh to an expert and critical audience. The low efficiency of the process and the many deformities during embryonic development were the cause of that. An experiment like the one around Dolly can be considered a proof of principle. Yes, an egg can reprogramme a **somatic** nucleus (Ch9), but it is a special egg that can do this and it will never be completely flawless. Although the process will probably be a bit better now, cloning still leads to a spectrum of incompletely reprogrammed **chromatin** landscapes with variable efficiency, resulting in a risk of embryonic developmental disorders. Insight into the resetting of the germline (Ch9) was already available back then, Dolly was not necessary for that. When you revisit the rationale that was behind cloning at the time, with genetic manipulation certainly being a driver of this, you will realise that the development of CRISPR technology has rendered it obsolete. Besides being a tool in research, cloning is too complicated and causes far too many welfare issues, thereby touching the boundaries of what is ethical. One thing cloning does well is reiterate, perhaps redundantly, that female **gametogenesis** and the female reproductive tract are the foundations of reproduction. It is all the more painful that this needs to be abused to

achieve the birth of a clone, due to the need for egg donation and the repeated use of surrogate mothers.

In Europe, cloning is actually prohibited due to considerations around the use of experimental animals, but this is not the case in South Korea, for example. Cloning is truly everything that sexual reproduction is not. The goal is to achieve something that is avoided in sexual reproduction (uniformity), and on top of that, cloning involves a second-hand aspect.

In contrast to what you might expect after reading about DNA repair in the germline in this book, the **genome** of **germ cell**s is certainly protected against **mutation**s, and better than the genome of somatic cells when it comes to DNA replication errors (see the legend of table 3). Somatic mutations acquired during life are transferred to the next generation through cloning. Molecular chromatin research shows that cloning causes additional variation in the methylome and **histone PTM**s (table 1). Problems with **genomic imprinting** (Ch10) are also encountered. Sometimes the effects are truly surprising: when tissue from a male dog is sent to South Korea, it can result in a dog with a female appearance (in 10 of 44 cases). Due to the spontaneous occurrence of methylation inactivation, the *SRY* **gene** (Ch8) is not expressed, resulting in the absence of testes, and due to the absence of sex hormones, internal and external male characteristics do not develop. In summary, to revisit the societal visions from the introduction, cloning is more like a mediocre project of the wizard than an invention of the prophet. Nevertheless, it continues to intrigue the scientific community, and the drive becomes even greater when primates are involved. This is especially true when they can contribute to research involving laboratory animals, in this case, the rhesus monkey. In China, a small "success" has now been achieved through the creation of a **chimera** with a placenta originating from an IVF embryo, while the "embryo proper" is derived from a somatic cell nucleus that has replaced the genetic material of a **zygote**. The fact that the **epigenetic** profile of the placenta was normal is key, realising that the placenta has the highest expression of imprinted genes (Ch10) and that the maintenance of imprinting (Ch9) is a problem associated with cloning. Is this result that has been reported for one animal a proof of principle? Stochastic processes will continue to play a role in resetting the chromatin of the somatic cell nucleus in the **cytoplasm** of the zygote and daughter cells. The production of cloned rhesus monkeys for the purpose of drug testing for safety seems far off. On top of that, one could easily argue that basing conclusions on one **genotype** is not really on par with the human population structure with its great genetic variation.

Fertility lies at the core of biology: it is one of the most, if not the most, important concepts (introduction).

Let it remain in the domain of the common people, of what affects us all. In other words, we should not sell fertility and genetics out to the corporate sector. The genetically perfect society really does not exist; it is a mirage. Instead, accommodate the genetic variation that is the defining characteristic of sexual reproduction. That also means that you must be able to recognise and critically follow social trends that lead to more uniformity, which are, for example, promoted by the digitalisation of society. Also, individualism (in this case, as a counterpart to collective interest) is actually a poor fit with our genetic exceptionalism (see Michael Lynch's explanation of its definition in Ch18). If individualism swings like a pendulum, then I hope it has now reached its zenith. I also hope that the balance between the individual, the population, and the environment is restored in all ethical and moral reflections on this subject.

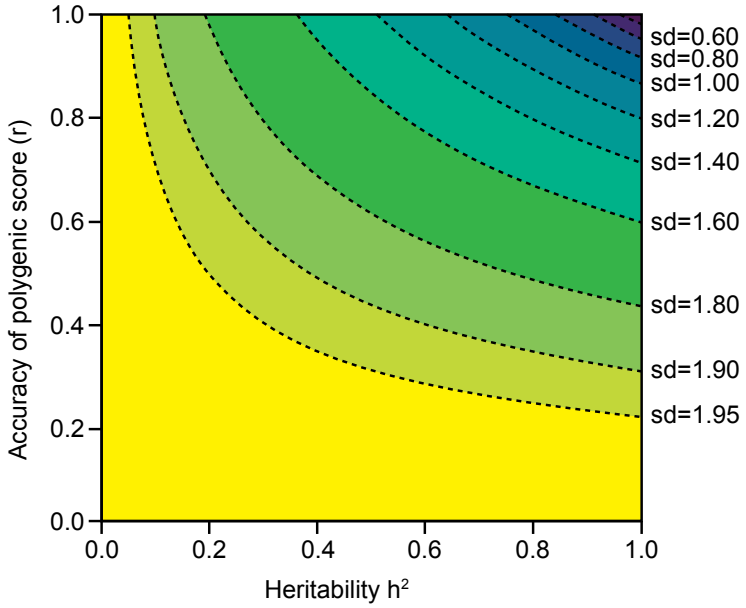
Back to the introduction, in which I already mentioned the wizard and the prophet from Mann's (2018) book. *In vitro* gametogenesis and CRISPR-Cas germline modification characteristically fit into the wizard's arsenal. A greater focus on the environment, including the sociocultural environment, aligns more with the prophet. In this book, the balance leans more towards the prophet, but the wizard is also needed. In the field of **somatic cell genetics**, this holds great promise. A better match between patient and drug, and between cancer cell and therapy, the start of the realisation of somatic gene therapy, or targeted RNA interventions; all thanks to the growth of **genomics** combined with molecular biology. As has been shown, the wizard's methods can also not be dismissed when trying to understand reproductive genetics. They have taught us a lot, and more is yet to come. However, in terms of the application of methods in the germline, my sympathy lies with the prophet.

The other thing I have tried to convey is the incredible complexity of life at the cellular level and within each individual; the balance between order and disorder and regulation at every level. Does all of that still fall under the concept of understanding? Beauty and wonder; the nearly religious feeling that is evoked here, could that be something more suited to the prophet? The idea in our culture, which is strongly dominated by the wizard, that any biological issue focused on reproduction can be solved with a little tinkering, exaggerates reality. There is a need for a broader perspective that is shared by a wider community. With this book, I hope to have made a small contribution to that.

Appendix

Through **DNA** research, it is possible to say something about an individual's genetic predisposition to certain diseases or traits. This appendix addresses genetic predisposition that is based on several to many **gene**s. In humans, the approximation of this genetic susceptibility is known as the **polygenic risk score** (the PRS, Ch6.2). The estimation of genetic predisposition based on DNA is subject to estimation error. Furthermore, genetic predisposition is not expressed in the **phenotype** in a one-to-one manner. Due to environmental influences, the expression level may be higher or lower. The total estimation error thus has two sources: the genetic predisposition is not known with certainty, and secondly, environmental influences play a role. The first estimation error depends on the accuracy (r) with which the genetic predisposition is estimated, and the second depends on the **heritability** (the h^2, Ch6.1, fig 17). Here, the accuracy r represents the statistical correlation between the "true" genetic predisposition and its estimate. The true genetic predisposition is a key concept in quantitative genetics. For example, if a father were to have a fairly large number of offspring, brought into the world by a variety of mothers in a range of different environments, one could exactly determine the father's true genetic predisposition: the average of the offspring would exactly reflect this. However, even with a smaller number of offspring, it is possible to obtain a reliable impression of the true genetic predisposition. The better the **genetic architecture** of a trait is understood (and this is especially true for traits for which a relatively small number of genes have a large influence) (Ch6.2), the higher the r and the smaller the estimation error of the genetic predisposition. And the higher the heritability, the less likely one is to be wrong due to environmental influences.

You may then ask the question: if the estimate of an individual's genetic predisposition is known, between which values can the expression of that genetic predisposition (the phenotype) then vary? To increase the probability that the actual expression falls within the limits of variation, you have to choose a wider range. It is quite common to use the estimate minus two times the estimation error and the estimate plus two times the estimation error as limits. When using such wide limits, there is a 2.5% chance that the actual expression will still be lower than the lower limit, and also a 2.5% chance that it will be higher than the upper limit. Expressed in statistical terms, the phenotypic range extends between the limits of plus two times the standard deviation and minus two times the standard deviation (fig 17).



What is shown in the figure? The X-axis represents the h^2 and the Y-axis represents r. The transitions between the colour zones show an indication of the phenotypic interval within which the **genotype** estimate falls. This is the estimation error mentioned above, expressed in the phenotypic standard deviation as a unit, to which we have assigned the value 1. The limits then extend from minus this number multiplied by the standard deviation, to plus this number multiplied by the standard deviation. In the theoretical case where h^2 and r both equal 1, the phenotype coincides with the genetic information, the genotype. With a larger h^2 and r, the interval between the lower and upper limits becomes smaller. However, as shown in the figure, this happens very slowly. Even in the situation where both are equal to 0.8 (which is much higher than a realistic situation), the limit is still 1.42. The reader might be surprised by how large the intervals are. In part, this has nothing to do with DNA, but rather with the fact that the focus lies on phenotypes. As an example: an h^2 of 0.6 is a pretty high value in practice. In that case, even if you are certain about the genotype, the estimation error will still be $\sqrt{1-0{,}6^2} = 0{,}8$ times the distribution of the phenotypes. That simple fact thus continues to haunt you, no matter how well you get to know the genotype. It is only at very high degrees of heritability, meaning that there are virtually no environmental influences on the expression of the genotype, that the estimation error truly decreases. In an article from 2019 published in the renowned American journal *Genetics*, Peter Visscher's group

provides a formula from which one can deduce that the size of the reference population on which the distribution of the PRS is based especially contributes to a more accurate estimate. The bottom line is that the complexity of the **genome**, in terms of the number of **loci** involved in a trait, needs to be tackled with the largest possible sample size for the statistical calculation of the PRS, i.e. one million people or more. Rather than using **SNP**s, it is even better to use the entire genotype (see the legend of table 3), and even then, large numbers of unrelated individuals are needed. The genome of such a group must have been sequenced (**WGS**, now transitioning from short read sequencing to long read sequencing; an example is the UK biobank), and the phenotype under study must be available. The sample size is so large because mutations that are rare but which often have a greater than average effect on the phenotype help to establish the influence of the genotype on the phenotype. Currently, the populations used for this are of European descent, but there are enough indications that the relationship between genotype and phenotype may vary between different populations on Earth (Ch6.2).

The formula used to calculate the accuracy of genomic information is as follows:

$$\sigma_{P/A} = \sqrt{(1-r_{AA}^2 h^2}\, \sigma_P$$

In this formula, $\sigma_{P/A}$ represents the estimation error of the estimated phenotype. It is a function of σ_P, the distribution of phenotypes (P), as well as the accuracy with which the genotype is estimated (r), and the heritability (h^2).

In practice, you can influence the level of h^2 to some extent by correcting phenotypes for systematic influences such as age and sex. To do this, you look at the average phenotype for men and for women that, for example, is seen at the age of 30. The systematic influences that play a role depend on the trait of interest. For diseases, one might consider correcting for risk factors such as smoking, alcohol consumption, and obesity. However, even after doing that, a heritability of 0.6 is still a pretty high value.

The value of r can thus be increased by gaining more knowledge of the situation at the DNA level, the concept of genetic architecture mentioned earlier. For example, this can be achieved by understanding the involvement of (many) more genes in the (often disease-related) phenotype, and by identifying the variation (the **mutation spectrum**) throughout the entire personal

genome (using WGS). Furthermore, it is also possible to include data from family members in the estimation, as well as all aspects of the phenotype, such as clinical research results, that provide any insight into the trait in question. In livestock breeding, all possible sources are used, both DNA information and phenotypes of the individual and family members. When many relatives are included (hundreds of close relatives), one can obtain an r value very close to 1, but in humans, 0.6 is probably quite high for this parameter as well. In a human situation, one could imagine that attempts can be made to standardise the influence of the environment, for example through diet. In that case, the h^2 increases because the influence of the environment on the phenotype is reduced, and in that way, the confidence interval also becomes smaller. In any case, it is clear that, according to a review article from 2024 (see below), PRS is moving towards the domain of clinical implementation. This will require a much greater understanding within the medical sector, as well as how to convey the uncertainty surrounding the estimation of genetic predisposition to patients.

NR Wray et al., Complex Trait Prediction From Genome Data: Contrasting EBV in Livestock to PRS in Humans: Genomic Prediction. Genetics 2019 Apr;211(4):1131-1141. PMID: 30967442, DOI: 10.1534/genetics.119.301859 [Review]

Afterword

During the more than five years spent on this book, many people contributed to it in one way or another. Whether by providing encouragement, supplying material, fostering a collegial environment, or by offering feedback on smaller to sometimes quite substantial parts of the book.

Without the willingness of Martien Groenen, head of the Department of Animal Breeding and Genomics at Wageningen University and Research, to provide accommodation within the department, none of this would have been possible.

Willy Baarends, Aniek Bouwman, Han Brunner, Mario Calus, Alwin Derijck, Godfried van der Heijden (title of the book), Kamlesh Madan (table 4), Johan van Ooijen, Manon Oud, Joke de Pater, Antoine Peters, Jeroen van Reeuwijk, Dick de Rooij, Jeroen Speksnijder and Henk Sulkers kindly provided me with material, as well as several anonymous reviewers. Pim Brascamp laid the groundwork for the appendix (with the figure by Mario Calus).

Feedback from and discussions with proofreaders are invaluable. Thank you Willy Baarends, Mario Calus, Edwin Cuppen, Christian Gilissen, Martien Groenen, Anton Grootegoed, Godfried van der Heijden, Ewart Kuijk, Aafke van Montfoort, Manon Oud, Antoine Peters, Liliana Ramos, Hein te Riele, Dick de Rooij, Klaas Swart, Elmar Tobi, and Marieke de Vries, thank you all.

A special word for Bé Wieringa. The first chapter, “Impression of a Cell,” is based on his valedictory lecture in Nijmegen (2018). Thank you, Bé, for bringing the cell biology in chapter 1 up to standard, and for your work on box 1; it makes a difference. An even greater difference was made when you were assigned a significant role as editor on behalf of Radboud University Press. No stone was left unturned, which greatly enhanced the solidity of the work. Your contribution to the English version of the book should also be acknowledged. The reviewers of the Dutch version are warmly thanked for their contributions.

As previously mentioned, Mario Calus and Pim Brascamp are largely responsible for the appendix on the accuracy of polygenic risk scores. However, conveying their genetic-statistical insights in logical language remains challenging; judge for yourself.

This book brings together a wealth of knowledge, covering the period from the 1950s onwards. Interpretation dilemmas are inevitable, and errors will also be found. The author remains accountable. A book like this will in parts be overtaken by time. This applies, for example, to the

population-level NGS results. However, the initial findings of this technology are mentioned in the book, despite the fact that they are still mainly limited to the Western world.

The further this project was delved into, the larger and more impressive the small world of the cell in general, and gametes in particular, became. One can hardly fathom it. Cells as life forms and what they build, alone and in conjunction with other cell types; a world that is hardly comprehensible, which therefore leaves space for the experience of religious sentiment, or deep wonder if one lacks that feeling. Weighing this extremely long evolutionary history against the very slow addition of new genetic variation, is looking back easier than looking forward? It feels like a stark contrast to the short-term nature of our culture; extreme specialisation drives rapid changes, both locally and globally, with the environment as a last consideration. After all, the phenotype depends on both genotype and environment (Ch6).

Approaching and hopefully organising the broader relationship between genetics and reproduction was my main driving force during the past few years. This also carried risks, such as the possibility of getting lost in the chaos of the overwhelming amount of information. Making high-quality figures is the most effective strategy to find order in this chaos. The collaboration with Marc was exceptional. This also applies to Judith Swart, who has created a beautiful translation of the book. The Genetics Department of Wageningen University and Research played a central role in meeting both of them. Without them, this project would have failed.

The idea of writing a book, and especially of producing an English edition, owes much to my mentors, Jaap Sybenga (Wageningen) and Tony Searle (Harwell, UK). Meeting Tony was pivotal in shaping my view of a human way of dealing with knowledge; he was my main inspiration during the early part of my career.

The working relationship with Radboud University Press started in the final phase of this project and proceeded smoothly, also for the English version. Thank you.

Dear Alie, without your presence, it would once again have been impossible to manage the tunnel vision that accompanies projects like this. Thank you very much!

Glossary

Acrosome Vesicle that covers the nucleus of the sperm cell and that is essential for fertilisation (fig 39, 42).

Acrosome reaction Fusion reaction between the membrane of the **acrosome** and the cell membrane of the sperm, which causes the release of the content of the acrosome as an initiation of fertilisation (fig 39, 42).

Allele One of the different variants of a **gene** at a particular **locus** (**chromosomal** position) (fig 13). Also applicable to **SNP**s, see the term **SNP**.

Allele frequency The measure of the presence of a specific **allele** in the population. It is calculated as the absolute number of this specific allele divided by two times the number of individuals in which it has been determined.

Amino acid The organic molecule that serves as the fundamental building block of proteins. There are 20 different types of amino acids in the **genetic code**.

Androgenote An embryo with genetic material solely derived from the father.

Androgens Group of male sex hormones.

Aneuploid Relatively small deviations from the species-specific number of **chromosomes**, at **haploid** and **diploid** levels.

Apoptosis The cellular mechanism responsible for programmed cell death.

Atresia The process of tissue going into regression as a result of **apoptosis**, in this case applicable to ovarian follicles (fig 28, 29).

Autoradiography Microscopic technique whereby radioactive radiation from an incorporated molecule becomes visible on the specimen after the processing of a photographic emulsion.

Autosomes All **chromosomes** that are not involved in sex determination (fig 12).

Azoospermia The absence of sperm cells in the ejaculate.

Biomarker Measurable indicator, for instance the quantity of an organic molecule, used to gain insight into the state of a cell, a clump of cells, or an individual.

Bivalents The paired **homologous chromosomes** during **meiosis** (fig 13, 14, 15) (for the **sex chromosomes** see **pseudoautosomal region**).

Blastocyst Preimplantation embryo in which a cavity has appeared within a clump of cells (see fig 43).

Blastomeres The individual cells in a preimplantation embryo up to the **blastocyst** stage.

Capacitation The process in which sperm cells acquire the ability to undergo the **acrosome reaction** (fig 39, 42).

Centromere A specialised region (in terms of **DNA** base sequence) on **eukaryotic chromosomes** where proteins organise themselves in a way that enables assembly of the **spindle apparatus** in order to facilitate cell division (fig 3).

Chiasma The X-shaped configuration of **homologous chromosomes** that forms at the site of **crossing over** when the **bivalents** condense, which is often observable with a light microscope (fig 14).

Chimera (from Greek: chimaera) Individual consisting of (usually) two different **genotypes**, resulting from two different **zygotes**.

Chip array (or DNA microarray) A systematic arrangement of **DNA** fragments of which the **genomic** position is known at the micro level. It is, for instance, used for automated detection of deviations from **diploidy** at a **chromosomal** and subchromosomal level. Another example for its use is the determination of a **SNP** profile.

Chromatids The two replicas of a **chromosome** after chromosome duplication during the S phase, also known as sister chromatids (see fig 3).

Chromatin Collective term for the molecular material of the **chromosome**.

Chromosome The organisational form of a single long **DNA** molecule with all its associated **chromatin** proteins and **RNA** types (fig 3, 7).

Chromosome mutation A change in the number of chromosomes and/or the **DNA** composition of a **chromosome** (or chromosomes), when the alteration involves an area larger than 50 bases. Other than a change in chromosome number these are now more often named **structural variants.**

CNV See **copy number variation**.

Cohesin complex The protein complex that holds sister **chromatids** together from the S phase onwards (fig 4, 6).

Complementation The production of a wildtype **phenotype** from two gametes that are not fully **haploid**. The absence of genetic material in one **gamete** is counterbalanced by its double presence in the other, which results in a **diploid genotype**.

Congenital disorders Disorders present from birth.

Copy number variation, CNV Variation between and within individuals regarding the number of repetitions of a variably sized **DNA** segment (in tandem).

Crossing over The molecular process of reciprocal exchange between **homologous chromosome** segments (fig 14).

Cryopreservation The controlled freezing and storage in liquid nitrogen at -196 degrees Celsius of, in this case, sperm cells, oocytes, and preimplantation **embryos**.

Cytogenetics The branch of genetics concerned with the shape and behaviour of **chromosomes** at a microscopic level, in **mitosis**, in **meiosis**, and as a result of **DNA** damage and repair.

Cytoplasm The entire contents of the cell excluding the nucleus.

DDR See **DNA damage response**.

Dimer A structure consisting of two equal subunits, such as two identical proteins. The term is also used for two consecutive **DNA** bases.

Diploid A cell or individual with two sets of **chromosomes**, each containing essentially the same information (with the exception of the X and Y chromosomes in males).

Disomy When two copies of one particular chromosome are present in a **haploid** situation.

DNA Chains of linked **nucleotides** characterised by the DNA bases A, C, G, T, with deoxyribose as a sugar molecule. One of the two types of nucleic acid, alongside **RNA**. In **chromosomes**, as in mitochondria, the two chains run in opposite orientations in the form of a double helix (double-stranded, fig 5, 7). Short stretches can also occur as single-strand DNA during chromosome functioning.

DNA damage response DDR The orchestrated response of a cell to damage, including chemical modifications of DNA.

DNA fingerprint A banding pattern (bar code) that is unique for each individual, created by cutting the **DNA** (with a restriction enzyme), electrophoresing the fragments, and then **hybridising** the DNA with a labelled piece of search DNA containing the sequence of a minisatellite (see Ch2.3.3).

Dominant The **allele** whose **phenotype** is visible in a **heterozygote**.

Dosage compensation The mechanism that broadly ensures that the amount of gene product is equal in both sexes (XX and XY), more specifically for the X **chromosome**.

Effective population size, Ne A way to calculate population size based on the contribution of each parent to subsequent generations.

Embryo The developmental trajectory of new life from fertilisation to the stage in which the organs are present, in the 8th week in humans (in principle), and after 14 days in mice (more completely formed) (see also **foetus**).

Embryonic stem cell Cells derived from the **inner cell mass** or **epiblast** (fig 43) that have retained the ability to differentiate into any type of body cell, in other words, they are **pluripotent**.

Endocrine disruptor Chemical substance/compound which, also prenatally, interferes with the endocrine system, including the sex hormones. Can be of natural origin, for example from plant sources.

Endonuclease A class of enzymes that cut the **DNA** double helix, often guided by a simple DNA code of at least four up to eight bases.

Epiblast The cells that develop from the **inner cell mass** around implantation and from which the three **germ layers**, endoderm, mesoderm, and ectoderm arise (fig 43).

Epigenetics The transfer of a pattern of gene expression to (a) subsequent generation(s) of cells and individuals, which does not involve a change in the **DNA** code.

Epistasis Describes the interaction between genes, whereby a mutation in one **gene** alters or prevents the expression of another gene (or other genes).

Epithelium Tissue layer that covers an underlying or overlying layer. Applies to both external surface (skin) and internal surface, such as cavities and blood vessels, the efferent duct of the testis, and the internal lining of the oviducts and uterus.

Euchromatin The majority of **chromatin** which either has an open structure, or can easily transform into this less condensed structure, compatible with **transcription**.

Eukaryotes Organisms consisting of a single cell or (usually) multiple cells with a distinct nucleus surrounded by **cytoplasm**.

Euploid Indicates the presence of the correct number of **chromosomes** in the correct form for the species.

Exome The complete set of expressed **DNA** in a cell or tissue, denoted as the production of mature **messenger RNA**.

Exon, Exons Those parts of the **gene** that are translated into protein according to the **genetic code**. To achieve this, the exons are cut from the primary transcript (pre-**messenger RNA, mRNA**) in a maturation process (fig 9).

F1, F2, F3 etc. F stands for “filial”, the number represents the generation. F1 is the product of the first cross. This notation is used when the subsequent matings occur within the first generation of offspring.

FISH Fluorescence in situ hybridisation Microscopic examination of **chromosomal DNA** (on a glass slide or in the cell nucleus) which involves **hybridisation** (formation of a double helix) of a labelled piece of search DNA (the probe) with the chromosomal DNA under investigation. The piece of search DNA that is, for example, specific to a gene, can be visualised after hybridisation using **fluorescence microscopy**.

Fitness Measure of the relative reproductive success of an individual, **genotype**, or (sub)population.

Fluorochrome, fluorescence microscope A form of microscopy in which a fluorochrome is utilised; a molecule that emits light of a longer wavelength (in the simplest form blue, green or red) when illuminated with light of a specific wavelength range. Such a fluorochrome acts as a detection molecule and is usually attached to a specific antibody, although there are also some that bind directly to a target, such as **DNA**.

Foetus The unborn offspring in the prenatal stages of development from the embryonic stage until birth. Growth and further maturation dominate during this stage (see also **embryo**).

Gametes The male and female (**haploid**) reproductive cells.

Gametogenesis The process of cell differentiation and division that produces the **haploid gametes**.

Gene The unit of heredity consisting of a regulatory domain (to enable **transcription**) and a domain for which **transcription** into **RNA** is essential.

Gene conversion One of the outcomes of the repair of double-strand breaks during the meiotic prophase, where an **allele** on one **homologue** changes to the allele on the other **homologue** (2:2 >> 3:1).

Gene dosage effect Proportionality between the degree of functionality of the **gene** and the number of gene copies.

Gene map Graphical representation of the location of genes on the **chromosome**, with the likelihood of **recombination** as a measure of the distances between them.

Generation interval The average age of the parents at the time of birth of the "average" child (general offspring).

Genetic architecture The genetic basis of a trait that relies on the expression of multiple (often many) genes, which can be coding as well as non-coding (box 1).

Genetic background Also known as residual genotype. This refers to the overall genotype in which, for example, a **mutant** is studied.

Genetic code The translation of the **DNA** base sequence (in **triplets**) into **amino acids**. The code is written in **RNA** language.

Genetic dissection The unravelling of the **genetic architecture** of essentially all **phenotypic** traits.

Genetic linkage The phenomenon that **allele**s located on the same **chromosome** can remain together during transmission to the next generation, because the chance of **crossing over** between genes with those different **allele**s is lower than one.

Genetic redundancy The phenomenon where the absence of genetic material is not noticeable in the **phenotype**. This can, for instance, be the case due to the presence of **genes** with overlapping functions.

Genome The total genetic information across all **chromosomes**. The term is used in both a **haploid** (one set) and **diploid** context (two sets). Within **genomics**, it refers to the diploid situation.

Genome-wide association study (GWAS) A method that uses many genetic polymorphisms (read: **SNP**s) scattered throughout the genome to discover statistical associations between a subset of those polymorphisms (read: **allele**s) and the trait under study in large populations. GWAS can serve as an initial step towards discovering the **genetic architecture** of that trait.

Genomic imprinting The **epigenetic** phenomenon that the activity of a **gene** in offspring can depend on the origin of the **gamete** (from the father or mother). Expression is thus inherited solely via the father or the mother. Less absolute forms also occur (high expression versus low expression).

Genomics The science concerned with the information content of the complete **DNA** of species.

Genotype Usually used to indicate the **allele** composition of a single **gene** or a number of genes, but it also refers to the entire **genome** in the **haploid** and **diploid** sense.

Germ cell Cell in the **germline**, starting from when it is first identifiable in early embryogenesis and extending to the stage of mature **gametes**.

Gonadotropins The messenger hormones originating from the pituitary gland which, through receptors on **somatic** cells in the gonad, are essential for the regulation of **gametogenesis**.

GWAS See **genome-wide association study**.

Gy (gray) Measure of the amount of radiation energy absorbed by tissue.

Gynogenote An embryo with genetic material solely from the mother.

h^2 See heritability.

Haploid A cell or organism with only one set of **chromosomes**.

Haploinsufficiency Situation where a single gene copy is insufficient to generate the normal **phenotype** (or for multiple genes: single copies).

Haplotype The **genotype** of a segment of a **chromosome** in singular form (so in the **haploid** state).

Heritability Also known as h^2. The genetic variance expressed as a fraction of the phenotypic variance. This measure is specific to the studied population and applies to the environment in which the **phenotype** measurements are taken (fig 17).

Heterochromatin, constitutive Chromatin that stains intensely, often located around the **centromeres**, sometimes elsewhere in the **chromosome** profile. It is always inactive in an interphase nucleus.

Heterochromatin, facultative Chromatin that stains intensely depending on its inactivity. The best-known example is the X **chromosome**.

Heterosis The phenomenon of the **F1** being more vigorous than the parental lines, also known as hybrid vigour.

Heterozygous The presence of two different **allele**s at a **locus**/of a **gene** in a **diploid** organism (fig 13). Applicable to any **DNA** motif, specially **SNPs** (fig 16).

Histology The branch of biology that studies the microscopic anatomy of cells and tissues.

Histones Family of conserved small basic proteins that **nucleosomes** are composed of (fig 7, 8).

Homeostasis The ability of multicellular organisms to maintain a balanced internal environment despite changes in the organism's external environment.

Homologues The same (or corresponding) **chromosomes** received from both parents, which can pair during **meiosis** (fig 13, 14).

Homologous recombination The process of exchange of homologous **chromosome** segments, facilitated through **meiotic** pairing (fig 13, 14).

Homozygous The presence of two identical **allele**s at a **locus/gene** in a **diploid** organism (see also **heterozygous**).

Human Genome Project The international research project that was completed in 2003 with the unravelling of the **DNA** code of the human **genome**.

Hybridisation The formation of a double **DNA** helix from two single-strand DNA molecules in a laboratory environment. **RNA** can also participate in helix formation.

Hybrids The offspring resulting from crosses between partners with a certain genetic distance, which is manifested in **allele** differences.

ICM See **inner cell mass**.

Immunofluorescence The microscopic detection of often minute molecular structures using **fluorescence** and antibodies.

Impact factor Measure used in the scientific world to indicate the importance of a journal. It is based on the number of citations of articles over the past two years.

Imprinting See **genomic imprinting**.

Indel Compound word of insertion and deletion, referring to a small-scale mutation event involving one to a few bases, in which a **DNA** base (or bases) are respectively added or removed.

Inner cell mass (ICM) The cluster of cells that appears on one side when a cavity is formed in the preimplantation **embryo** (at the transition from the morula to the blastocyst stage, fig 43).

Intron In contrast to the **exon**, these are the segments of the **gene** that are **transcribed**, but are spliced out in the nucleus when the mature **messenger (m)RNA** is formed (fig 9).

Ionising radiation Collective term for high-energy radiation that is capable of breaking the covalent bonds that lead to molecule formation. Causes damage to e.g. **DNA**.

Karyogram Systematic arrangement of **mitotic** chromosomes, often after application of **chromosome** banding to assemble the **homologous** pairs.

Karyotype The individual- or species-specific **karyogram**.

Germ layers The products of initial differentiation into the three primary **embryo**nic cell types – ectoderm, mesoderm, and endoderm – originating from the **epiblast**, which forms in the late blastocyst (fig 43).

Germline The consecutive cell stages connected by **mitotic** and **meiotic** divisions from the formation of the first identifiable **germ cells** up to and including **gametogenesis**, which links the generations.

Kinase A class of enzymes that can attach phosphate groups to proteins and thereby contribute to regulating the activity of those proteins.

Knock-in Altering a **gene** or adding a new gene via genetic manipulation.

Knockout Term used for disabling a **gene** via genetic manipulation.

Locus Usually the location on the **chromosome** where the protein-coding **gene** is located, but can refer to any distinguishable **DNA** motif that is transcribed and is non-coding (box 1) or that affects the phenotype in other ways.

Maternal to zygotic transition (MtZ) The phase during the early cleavage divisions of the **embryo** in which the influence of the maternal **cytoplasm** decreases and the influence of the early embryonic gene expression of the embryo's own **genome** increases.

Meiosis The two successive divisions in **gametogenesis** that lead to the **haploid gametes** (fig 13).

messenger RNA, mRNA Term for the **transcript** that forms in the nucleus after maturation. The mature form only contains the protein-coding information from the **exons** (fig 9). The primary transcript is also called pre-messenger RNA.

Metabolism The totality of chemical reactions that take place in a living cell, tissue, and individual, with **metabolites** as intermediate or final products.

Methylome The methylation status of cytosine, usually in the context of a CpG dimer, and ideally across the entire **genome** (fig 11).

Mismatch repair (MMR) One of the **DNA** base-directed main pathways for DNA repair, most active during DNA replication (fig 5). Mismatch refers to the presence of incorrect base pairs (such as no T opposite an A, and so forth).

Mitosis A nuclear division, usually accompanied by a cell division, that produces two genetically identical daughter cells.

MMR See under **mismatch repair**.

Monosomy Situation in which only one copy of a single **chromosome** is present in a **diploid** cell or organism.

Mosaic Organism composed of cells with differences in **genotype**, arising from a single **zygote** after fertilisation. Due to X **chromosome** inactivation, any individual with two or more X chromosomes is a mosaic, in this case caused by differences in gene expression.

MtZ See **Maternal-to-zygotic**.

Mutagen, mutagenic substance or physical treatment Substances (or treatments such as **ionising radiation**) that have the ability to damage **DNA**, which may result in a change in DNA after repair (the repair is not always flawless).

Mutation A **mutagen**-induced or spontaneously occurring change in **DNA**. The cell or individual with the mutation is a **mutant**.

Mutation frequency The chance that an **allele** copy changes when transmitted to the next generation. This concept is also applied to a generation of a **somatic** cell.

Mutation spectrum The range of genetic changes that have occurred throughout history for a single **gene/locus**. Especially relevant for disorders based on a single gene.

Mutator phenotype Greater genetic variation in offspring caused by an increase in spontaneous **mutation frequency** due to a genetic defect affecting the maintenance (repair) of **DNA**.

Nature versus nurture Misleading statement stemming from a disputable supremacy of either **genotype** or environment on the manifestation of the **phenotype**. For most traits, the genotype is in constant interaction with the environment, with sometimes a more dominant role for the genotype, sometimes for the environment, and usually something in between.

Ne See **effective population size**.

Next generation sequencing (NGS) Automated large-scale determination of **DNA** base sequences, made possible by microtechnology and data processing. Diagnostic applications often use(d) the **exome**, the genes that are expressed, but the whole **genome** is increasingly being analysed.

NGS See above.

Nondisjunction An incorrect distribution of **chromosomes** and **chromatids** among daughter cells in **mitosis** and **meiosis**. The literal translation, “not separating”, does not cover all mechanisms (fig 30).

Nucleolus Dense structure in the nucleus near the genes for ribosomal **RNA** (rRNA, present in many copies and processed in the nucleolus). rRNA is needed for protein synthesis (fig 2).

Nucleosome The basic unit of **chromosome** organisation, a “ball” of eight histone proteins around which **DNA** is wound (fig 7, 8).

Nucleotide The basic building block of **DNA** and **RNA**, consisting of the nitrogen-containing base indicated by A, C, G and T (in DNA) or U (in RNA), a sugar molecule (deoxyribose in DNA, ribose in RNA), and a phosphate group (box 1).

Oestrogens Class of female sex hormones.

Oestrus The period in the ovarian cycle of mammals during which mating activity with the male partner is accepted by the female.

Oligospermia Sperm count in the ejaculate falling below a threshold specified by professionals in the field (table 5).

Oogenesis The cellular differentiation process leading to the production of mature fertilisable egg cells.

Organelle A microscopically identifiable structure or subcompartment of a **eukaryotic** cell, with defined functions.

Orthologs Genes in different species with a still recognisable evolutionary origin in a common ancestor.

Oxidative stress An excessive density of reactive oxygen molecules in and around cells.

Oxygen free radicals Reactive forms of oxygen with a free electron.

PAR See **pseudoautosomal region**.

Paralogs Closely related genes within the genome, generated by the duplication of genetic material (term related to the concept of gene family).

Parthenogenesis, parthenogenote Organism formed by activation of the egg cell without the involvement of a sperm cell.

Phenotype The usually, but not necessarily outward appearance of an individual, which can refer to the overall appearance, but often focuses on a specific trait.

Pluripotency The ability to differentiate in any type of body cell.

Pluripotent stem cell An early embryonic or induced **somatic** cell that can differentiate into any type of body cell. The embryonic version is also called embryonic stem cell. Pluripotent stem cells can multiply “infinitely” under the right conditions. For the induced form, see box 2.

Polygenic Used for traits determined by multiple, often many **gene**s. Other **DNA** motifs, which are for example involved in gene regulation, may also play a role. The contrasting concept is monogenic.

Polyploidy The presence of more than two sets of **chromosomes** (a set containing the **haploid** number) per cell or individual.

Post-translational modification (PTM) A modification of a protein immediately after translation and thereafter, in which a variety of molecules, of which phosphate is the most well known, are attached to specific sites on the protein. There are certainly more than 50 of such modifying molecules.

Primordial germ cells The first specialised **germ cell** stage that can be identified in the embryo (fig 22).

Promoter That part of the **gene** that is strongly involved in the regulation of its activity and to which the **RNA** polymerase (fig 9) can attach.

Protein kinase See kinase.

Proteome The collection of proteins in a cell or tissue. “Proteomics” is the science focusing on this collection and its functions.

PRS, Polygenic risk score Statistical measure (estimate) per individual, indicating the genetic contribution to the variance of physiological traits.

Pseudoautosomal region, PAR A relatively small part of a **sex chromosome** that, after having evolved from an originally fully **homologous** pair, has retained homology, so that **crossing over** between the now largely different sex chromosomes remains possible (fig 15, 37, 38). This heteromorphic pair is also called sex **bivalent.**

PTM See **post translational modification**.

Recessive, recessive allele An allele whose **phenotypic** effect is not visible in a **heterozygote**.

Recombinase A collective term for enzymes with a key function in the processes that lead to **recombination**.

Recombination Key concept (alongside **mutation**) in genetics. The creation of new combinations of **allele**s for different **gene**s. When these are located on different **chromosomes**, recombination is described by Mendel's second law (Ch5.3). When the genes for these alleles are located on the same chromosome, **crossing over** leads to recombination.

Replication The copying of **DNA**, usually in preparation for the next cell cycle.

Retrotransposons A class of mobile genetic elements, which move via a) **transcription** and then b) the activity of a **reverse transcriptase**, followed by reintegration into the **genome**.

Reverse transcriptase An enzyme that can synthesise **DNA** using **RNA** as a template.

RNA One of the two nucleic acids. Chains of **nucleotides** that are formed via **transcription** of **DNA**. Two categories are distinguished: coding, i.e.; translatable into the amino acid sequence of a protein (the mature **messenger RNAs**), and non-coding. This second group roughly divides into two subgroups, long and small non-coding. Within the small non-coding group, several fractions are distinguished, including the tRNAs essential for protein synthesis (fig 9). In addition, there is non-coding ribosomal RNA, also necessary for protein synthesis (rRNA, in the ribosomes: (fig 2), see also under the term nucleolus). RNA usually functions in a single-stranded form, but it can also be double-stranded.

RNA polymerase Class of enzymes that synthesise **RNA** using **DNA** as a template.

Sex body The **chromatin** compartment in the male prophase of the first **meiotic** division that contains the **sex chromosomes** (fig 15, 37).

Sex chromosomes The **chromosomes** that are involved in sex determination in a causative manner and whose absence or presence corresponds to the sex.

Single nucleotide polymorphism (SNP) Variants that exist for a single base position in the **DNA**, usually consisting of two different options (bases, in this case these are **alleles**). In practical genetic analysis, the least common one still occurs in one in 50 individuals in that population (as an average, see also **allele frequency**). SNPs are common.

SNP See above.

Soma, somatic Everything related to body cells, in contrast to cells of the **germline**.

Somatic cell genetics The branch of genetics concerned with **mutation** and possibly **recombination** in body cells.

Spermatogenesis The cellular differentiation process that leads to the production of spermatozoa (sperm, sperm cells), which are in principle ready for fertilisation.

Spermiation The detachment of the elongated spermatids from the wall of the seminiferous tubules, which may now be called spermatozoa, sperm, or sperm cells.

Spermiogenesis The metamorphosis of the round spermatids produced in meiosis into elongated spermatids (fig 31), ending with **spermiation**.

Spermiogram The numerical representation of sperm quantity and quality, in terms of number, morphology and motility of cells (see table 5).

Spindle apparatus The complete structure of tubulin fibres that keeps the condensed **chromosomes** in balance, ready for **mitosis** (nuclear >> cell division, fig 3).

Structural variants, see **chromosome mutation.**

Synaptonemal complex The protein complex consisting of three parts that acts as a connector of the **homologous chromosomes** in the prophase of the first **meiotic** division, thereby facilitating **crossing over**. The three parts of the complex can be visualised using an electron microscope.

Telomerase An enzyme, a **reverse transcriptase**, that can elongate **chromosome** ends using a short **RNA** sequence as a template. This generates a large number of tandem repeats in the **DNA**, based on the instruction of the short RNA template.

Telomere The complex of short **DNA** repeats and proteins that cap the ends of **chromosomes** to prevent them from being recognised as double-strand DNA breaks.

Teratocarcinoma A tumour originating from a **germ cell** in which chaotic **embryo**nic structures can be recognised. Can arise in both the female and male **germline**, at an early time in males and later in females.

Totipotency The ability of typically an early embryonic cell to develop into any cell of the embryonic membranes (as components of the placenta) and any type of body cell.

Transcription The synthesis of **RNA** guided by a **DNA** template (fig 9).

Transcription factor A protein that binds to specific **DNA** motifs (usually to the **promoter** of a gene) to enable or regulate **transcription**. Several (transcription) factors work together in a complex.

Transcriptome The entirety of **RNA** molecules of a cell or tissue.

Transgenic Individual with extra genetic material, usually with the structure of a gene. This **gene DNA** can be from the same species or from a different species.

Translation Protein synthesis based on mature **messenger RNA**, utilising the **genetic code**.

Transposon, transposable elements Collective term for short **DNA** sequences that, whether or not through an **RNA** intermediate, appear at different locations in the **genome**. See also **retrotransposon**.

Triplet See **genetic code**.

Triploidy The presence of three sets of **chromosomes** (see also **haploidy** and **diploidy**).

Trisomy The presence of three copies of a **chromosome** in a **diploid** cell or organism instead of the usual two.

Univalent Unpaired **chromosome** during the first **meiotic** division.

WES See below.

WGS See below.

Whole exome sequencing (WES) Automated base sequence determination of all the mature **messenger RNA** in a cell or tissue.

Whole genome sequencing (WGS) Automated base sequence determination of the entire **genome** of a cell, tissue or organism.

Zona pellucida The gelatinous layer that is formed and maintained around the oocyte in growing ovarian follicles. Is lost prior to implantation of the preimplantation embryo (fig 39, 42).

Zygote All stages between fusion of the **gametes** and the first cleavage division of the preimplantation embryo.

List of abbreviations

The abbreviations of genes are always written in italics, using all capital letters for human genes, while only the first letter is capitalised for mouse genes. The gene product is written in regular, non-italicised font, in capital letters for both humans and mice.
* can also be found in the glossary.

AI	artificial insemination
AMC	Amsterdam Medical Centre, location of Amsterdam University Medical Centre
APC	anaphase promoting complex
ARC	Animal Research Council
ART	artificial/assisted reproductive techniques/technology
Atm	the gene Ataxia-telangiectasia mutated
ATP	adenosine triphosphate
Atr	the gene Ataxia-telangiectasia and Rad3-related
BER	base excision repair
Blimp-1	the gene B-lymphocyte-induced maturation protein-1
BMI	body mass index
bp	base pair
BPA	the chemical substance bisphenol A
BRCA	the genes breast cancer 1/2, early onset
CeRA	Centrum für Reproduktionsmedizin und Andrologie Münster Centre for Reproductive Medicine and Andrology Münster
CIN	chromosomal instability

CNV*	copy number variation
CRISPR	clustered regularly interspaced short palindromic repeats
DDR*	DNA damage response
DMR	differentially methylated region
DNA*	deoxyribonucleic acid
ESC	embryonic stem cell
ESHRE	European Society of Human Reproduction and Embryology
FISH*	fluorescence in situ hybridisation
GWAS*	genome-wide association study
Gy*	gray
HCG	human chorionic gonadotropin
HR*	homologous recombination
HRR	homologous recombination repair
ICM*	inner cell mass
ICR	imprinting control region
ICSI	intracytoplasmic sperm injection
IMPC	International Mouse Phenotyping Consortium
Indel*	compound word of insertion and deletion
INRA	Institut National de la Recherche Agronomique National Institute of Agricultural Research
iPS, iPSC	induced pluripotent stem cell
IUI	intrauterine insemination
IVF	*in vitro* fertilisation
LH	luteinising hormone
LINE	long interspersed element
MMR*	mismatch repair
MRC	Medical Research Council
mRNA*	messenger RNA
mt	mitochondrial
MtZ*	maternal to zygotic transition
ncRNA*	non-coding RNA, see RNA*
Ne*	effective population size
NER	nucleotide excision repair
NGS*	next generation sequencing

NHEJ	non-homologous end joining
NIH	National Institutes of Health
NIPT	noninvasive prenatal testing
NKI	Netherlands Cancer Institute
NVOG	Dutch Society for Obstetrics and Gynaecology
OAT	oligoasthenoteratozoospermia
OMIM	Online Mendelian Inheritance in Man
PAH	polycyclic aromatic hydrocarbons
PAR*	pseudoautosomal region
PESA	percutaneous epididymal sperm aspiration
PGC*	primordial germ cell
PGD	preimplantation genetic diagnostics
PGT-A	preimplantation genetic testing-aneuploidy
PGT-M	preimplantation genetic testing-monogenic
PGT-P	preimplantation genetic testing-polygenic disease
PGT-SR	preimplantation genetic testing-structural rearrangements
PHA	phytohemagglutinin
PMSG	pregnant mare serum gonadotropin
Prdm9	PR/SET domain containing protein 9
PRS*	polygenic risk score
PTM*	post-translational modification
Rad51	the gene Radiation 51 (yeast nomenclature)
RNA*	ribonucleic acid
lncRNA*	long non-coding RNA
mRNA*	messenger RNA
miRNA	micro RNA
ncRNA*	non-coding RNA
piRNA	with protein Piwi interacting RNA
rRNA*	ribosomal RNA
siRNA	small interfering RNA
sncRNA*	small non-coding RNA
tRNA*	transfer RNA
tsRNA	transfer RNA-derived small RNA

ROSI	round spermatid injection
SAC	spindle assembly checkpoint
SINE	short interspersed nuclear element
SNP*	single nucleotide polymorphism
SOS	severe/strong oxidative stress
SOS	Sertoli cell-only syndrome
STR	short tandem repeat
TAD	topologically associating domain
TDS	testicular dysgenesis syndrome
TESE	testicular sperm extraction
UMC	University Medical Centre
WES*	whole exome sequencing
WGS*	whole genome sequencing
WHO	World Health Organisation
WUR	Wageningen University and Research

bp base pair

kb kilobase 1000

mb megabase 1,000,000

µm 1/1000 mm

µl 1/1000 ml

References

References of greater importance are marked with the symbol *. The character # has been added to indicate which publications have contributed to which figure, and the character ^ serves the same purpose for tables. These references stand out in the layout.

Introduction

*B Alberts et al., *Essential Cell Biology*, 4th edition. Garland Science, 2022

*AJF Griffiths, SR Wessler, SB Carroll, RC Lewontin, *Introduction to genetic analysis*, 12th edition. WH Freeman and Co, 2020

*CC Mann, *The wizard and the prophet*. Picador New York, 2019. ISBN 9781509884186

Chapter 1

B Henderson, MA Fares, ACR Martin, *Protein Moonlighting in Biology and Medicine*. Wiley-Blackwell, 2016

DHEW Huberts, IJ van der Klei, Moonlighting proteins: An intriguing mode of multitasking. Biochim Biophys Acta - Mol Cell Res. 2010;1803(4):520-525. PMID: 20144902, DOI:10.1016/j.bbamcr.2010.01.022

C.Jefferey, An introduction to protein moonlighting. Biochem Soc Trans. 2014;42(6):1679-1683. PMID: 25399589, DOI: 10.1042/ BST20140226

Y Liu et al., Impact of Alternative Splicing on the Human Proteome. Cell Rep. 2017 Aug 1;20(5):1229-1241. PMID: 28768205, DOI: 10.1016/j.celrep.2017.07.025

*#2 R Milo, R Phillips, N Orme, *Cell biology by the numbers*. Garland Science, 2015. ISBN 9780815345374

S Plaza, G Menschaert, F Payre, In Search of Lost Small Peptides. Annu Rev Cell Dev Biol. 2017 Oct 6;33:391-416. PMID: 28759257, DOI: 10.1146/annurev-cellbio-100616-060516 [Review]

* B. Wieringa: ".....is 't grote niet weerd!?", (abbreviated version of) the speech delivered upon his retirement as professor of Cell Biology at Radboud University/Radboudumc on Friday, June 1, 2018.© Prof Dr B. Wieringa, Nijmegen, 2018

M Wu Orr, Y Mao, G Storz, S-B Qian, Alternative ORFs and small ORFs: shedding light on the dark proteome. Nucleic Acids Res. 2020 Feb 20;48(3):1029-1042. PMID: 31504789, DOI: 10.1093/nar/gkz734 [Review]

Chapter 2 tabel 2

C Ade, AM Roy-Engel, PL Deininger, Alu Elements: An Intrinsic Source of Human Genome Instability. Curr Opin Virol. 2013 Dec;3(6):639-645. PMID: 24080407, DOI: 10.1016/j.coviro.2013.09.002 [Review]

* B Alberts, D Bray, K Hopkins, A Johnson, J Lewis, M Raff, K Roberts, P Walter (Eds). *Essential Cell Biology*, 4th edition. Garland Science, 2014. ISBN 9780815344551

*^2 ME Aldrup-Macdonald, BA Sullivan, The past, present, and future of human centromere genomics. Genes (Basel) 2014 Jan 24;5(1):33-50. PMID: 24683489, DOI: 10.3390/genes5010033

N Altemose et al., Complete genomic and epigenetic maps of human centromeres. Science 2022 Apr;376(6588):eabl4178. PMID: 35357911, DOI: 10.1126/science.abl4178

* SKr Anand,A Sharma, N Singh, P Kakkar, Entrenching Role of Cell Cycle Checkpoints and Autophagy for Maintenance of Genomic Integrity. DNA Repair (Amst) 2020 Feb;86:102748. PMID: 31790874, DOI: 10.1016/j.dnarep.2019.102748 [Review]

FH Andrews, BD Strahl, TG Kutateladze, Insights Into Newly Discovered Marks and Readers of Epigenetic Information. Nat Chem Biol. 2016 Aug 18;12(9):662-668. PMID: 27538025, DOI: 10.1038/nchembio.2149 [Review]

GA Armeev et al., Linking Chromatin Composition and Structural Dynamics at the Nucleosome Level. Curr Opin Struct Biol. 2019 Jun;56:46-55. PMID: 30529788, DOI: 10.1016/j.sbi.2018.11.006 [Review]

TD Babatz, KH Burns, Functional Impact of the Human Mobilome. Curr Opin Genet Dev. 2013 Jun;23(3):264-270. PMID: 23523050, DOI: 10.1016/j.gde.2013.02.007 [Review]

* G Bernardi, Chromosome Architecture and Genome Organization. PLoS One 2015 Nov 30;10(11):e0143739. PMID: 26619076, DOI: 10.1371/journal.pone.0143739. eCollection 2015

PRJ Bois, Hypermutable Minisatellites, a Human Affair? Genomics 2003 Apr;81(4):349-355. PMID: 12676558, DOI: 10.1016/s0888-7543(03)00021-1 [Review]

T Cremer et al., The 4D Nucleome: Evidence for a Dynamic Nuclear Landscape Based on Co-Aligned Active and Inactive Nuclear Compartments. FEBS Lett. 2015 Oct 7;589(20 Pt A):2931-2943. PMID: 26028501, DOI: 10.1016/j.febslet.2015.05.037

* Z Dou et al., Recent Progress on the Localization of the Spindle Assembly Checkpoint Machinery to Kinetochores. Cells 2019 Mar 23;8(3):278. PMID: 30909555, DOI: 10.3390/cells8030278 [Review]

M Escalera-Zamudio, AD Greenwood, On the classification and evolution of endogenous retrovirus: human endogenous retroviruses may not be 'human' after all. APMIS. 2016 Jan-Feb;124(1-2):44-51. PMID: 26818261, DOI: 10.1111/apm.12489 [Review]

*#7 TM Escobar et al., Active and Repressed Chromatin Domains Exhibit Distinct Nucleosome Segregation during DNA Replication. Cell 2019 Oct 31;179(4):953-963.e11. PMID: 31675501, DOI: 10.1016/j.cell.2019.10.009

*^2 M Friedli, D Trono, The developmental control of transposable elements and the evolution of higher species. Annu Rev Cell Dev Biol. 2015;31:429-451. PMID: 26393776, DOI: 10.1146/annurev-cellbio-100814-125514 [Review]

*^2 JM Goldmann, *Characterization of de novo mutations in the human germline*. Proefschrift afdeling human genetica Radboud umc december 2018 https://commons.wikimedia.org/w/index.php?curid=50582882

*^4 AJF Griffiths, J Doebley, C Peichel, DA Wassarman, *Achieve for Introduction to Genetic Analysis*, 12th edition. WH Freeman and Co, 2020. ISBN 9781319401375

*^2 DC Hancks, HH Kazazian Jr., Roles for Retrotransposon Insertions in Human Disease. Mob DNA 2016 May 6;7:9. PMID: 27158268, DOI: 10.1186/s13100-016-0065-9. eCollection 2016 [Review]

* https://nl.wikipedia.org/wiki/Centraal_dogma_van_de_moleculaire_biologie: F Crick 1958 en 1970

*^2 https://commons.wikimedia.org/w/index.php?curid=50582882

https://www.mirbase.org/cgi-bin/browse.pl

http://www.noncode.org/analysis.php

*#6 K-I Ishiguro, The cohesin complex in mammalian meiosis. Genes Cells 2019 Jan;24(1):6-30. PMID: 30479058, DOI: 10.1111/gtc.12652 [Review]

* AJ Jeffreys, V Wilson, SL Thein, Hypervariable 'minisatellite' regions in human DNA. Nature 1985 Mar 7-13;314(6006):67-73. PMID: 3856104, DOI: 10.1038/314067a0

AJ Jeffreys et al., Complex gene conversion events in germline mutation at human minisatellites. Nat Genet. 1994 Feb;6(2):136-145. PMID: 8162067, DOI: 10.1038/ng0294-136

* T Jenuwein, CD Allis, Translating the Histone Code. Science 2001 Aug 10;293(5532):1074-1080. PMID: 11498575, DOI: 10.1126/science.1063127 [Review]

* WL Johnson, AF Straight, RNA-mediated Regulation of Heterochromatin. Curr Opin Cell Biol. 2017 Jun;46:102-109. PMID: 28614747, DOI: 10.1016/j.ceb.2017.05.004 [Review]

M Kataria, H Yamano, Interplay Between Phosphatases and the Anaphase-Promoting Complex/Cyclosome in Mitosis. Cells 2019 Aug 2;8(8):814. PMID: 31382469, DOI: 10.3390/cells8080814 [Review]

H Kimura, Histone modifications for human epigenome analysis. Journal of Human Genetics, 06 Jun 2013, 58(7):439-445. PMID: 23739122, DOI: 10.1038/jhg.2013.66 [Review]

C Maison et al., Higher-order Structure in Pericentric Heterochromatin Involves a Distinct Pattern of Histone Modification and an RNA Component. Nat Genet. 2002 Mar;30(3):329-334. PMID: 11850619, DOI: 10.1038/ng843

SM McNulty ,LL Sullivan, BA Sullivan, Human Centromeres Produce Chromosome-Specific and Array-Specific Alpha Satellite Transcripts That Are Complexed With CENP-A and CENP-C. Dev Cell 2017 Aug 7;42(3):226-240.e6. PMID: 28787590, DOI: 10.1016/j.devcel.2017.07.001

* A Mendez-Bermudez, M-J Giraud-Panis, J Ye, E Gilson, Heterochromatin replication goes hand in hand with telomere protection. Nat Struct Mol Biol. 2020 Apr;27(4):313-318. PMID: 32231287, DOI: 10.1038/s41594-020-0400-1 [Review]

*#6 M Moronta-Gines, TRH van Staveren, KS Wendt, One ring to bind them - Cohesin's interaction with chromatin fibers. Essays Biochem. 2019 Apr 23;63(1):167-176. PMID: 31015387, DOI: 10.1042/EBC20180064 [Review]

AL Olins, DE Olins, Spheroid chromatin units (v bodies). Science 1974 Jan 25;183(4122):330-332. PMID: 4128918, DOI: 10.1126/science.183.4122.330

* N Palmer, P Kaldis, Regulation of the Embryonic Cell Cycle During Mammalian reimplantation Development. Current Topics in Developmental Biology 2016;120:1-52. PMID 27475848, DOI: 10.1016/bs.ctdb.2016.05.001 [Review]

* El Prieto, K Maeshima, Dynamic chromatin organization in the cell. Essays in Biochemistry 2019;63:133-145. PMID: 30967477, DOI: 10.1042/EBC20180054

*^2 SR Richardson et al., Influence of LINE-1 and SINE Retrotransposons on Mammalian Genomes. Microbiol Spectr. 2015 Apr;3(2):MDNA3-0061-2014. PMID: 26104698, DOI: 10.1128/microbiolspec.MDNA3-0061-2014 [Review]

T Sexton, G Cavalli, The Role of Chromosome Domains in Shaping the Functional Genome. Cell 2015 Mar 12;160(6):1049-1059. PMID: 25768903, DOI: 10.1016/j.cell.2015.02.040

* JW Shay, WE Wright, Telomeres and Telomerase: Three Decades of Progress. Nat Rev Genet. 2019 May;20(5):299-309. PMID: 30760854, DOI: 10.1038/s41576-019-0099-1 [Review]

S Sivakumar, GJ Gorbsky, Spatiotemporal regulation of the anaphase-promoting complex in mitosis. Nat Rev Mol Cell Biol. 2015 Feb;16(2):82-94. PMID: 25604195, DOI: 10.1038/nrm3934 [Review]

* Z Su1, JM Denu, Reading the Combinatorial Histone Language. ACS Chem Biol. 2016 Mar 18;11(3):564-574. PMID: 26675328, DOI: 10.1021/acschembio.5b00864 [Review]

* Q Szabo, F Bantignies, G Cavalli, Principles of genome folding into topologically associating domains. Sci Adv. 2019 Apr 10;5(4):eaaw1668. PMID: 30989119, DOI: 10.1126/sciadv.aaw1668 [Review]

*#10 S Viollet, C Monot, G Cristofari, L1 retrotransposition, The snap-velcro model and its consequences. Mob Genet Elements 2014 Jan 1;4(1):e28907. PMID: 24818067, DOI: 10.4161/mge.28907

P Weller, AJ Jeffreys, V Wilson, A Blanchetot, Organization of the Human Myoglobin Gene. EMBO J. 1984 Feb;3(2):439-446. PMID: 6571704, DOI: 10.1002/j.1460-2075.1984.tb01825.x

* H Yamano, APC/C: current understanding and future perspectives. F1000Res. 2019 May 23;8:F1000 Faculty Rev-725. PMID: 31164978, DOI: 10.12688/f1000research.18582.1 [Review]

Box 1

F Abascal et al., Loose ends: almost one in five human genes still have unresolved coding status. Nucleic Acids Res. 2018 Aug 21;46(14):7070-7084. PMID: 29982784, DOI: 10.1093/nar/gky587

M Pertea et al., CHESS: a new human gene catalog curated from thousands of large-scale RNA sequencing experiments reveals extensive transcriptional noise. Genome Biol. 2018 Nov 28;19(1):208. PMID: 30486838, DOI: 10.1186/s13059-018-1590-2

DR Zerbino, A Frankish, P Flicek, Progress, Challenges, and Surprises in Annotating the Human Genome. Annu Rev Genomics Hum Genet. 2020 Aug 31;21:55-79. PMID: 32421357, DOI: 10.1146/annurev-genom-121119-083418 [Review]

Chapter 3, tabel 3 en 4

*^3 S Aganezov et al., A complete reference genome improves analysis of human genetic variation. Science 2022 Apr;376(6588):eabl3533. PMID: 35357935, DOI: 10.1126/science.abl3533

PA Audano et al., Characterizing the Major Structural Variant Alleles of the Human Genome. Cell 2019 Jan 24;176(3):663-675.e19. PMID: 30661756, DOI: 10.1016/j.cell.2018.12.019

*^4 R Bandyopadhyay et al., Parental origin and timing of de novo Robertsonian translocation formation. Am J Hum Genet. 2002 Dec;71(6):1456-1462. PMID: 12424707, DOI: 10.1086/344662 [Review]

*#11 TH Bestor, JR Edwards, M Boulard, Notes on the role of dynamic DNA methylation in mammalian development. Proc Natl Acad Sci U S A. 2015 Jun 2;112(22):6796-6799. PMID: 25368180, DOI: 10.1073/pnas.1415301111 [Review]

J Boué, A Boué, P Lazar, Retrospective and Prospective Epidemiological Studies of 1500 Karyotyped Spontaneous Human Abortions. Birth Defects Res A Clin Mol Teratol. 2013 Jul;97(7):471-486. PMID: 23873813, DOI: 10.1002/bdra.23174 [Classical Article]

J Byrne et al., Morphology of early fetal deaths and their chromosomal characteristics. Teratology 1985 Oct;32(2):297-315. PMID: 4049288, DOI: 10.1002/tera.1420320218

BP Coe et al., Neurodevelopmental disease genes implicated by de novo mutation and copy number variation morbidity. Nat Genet. 2019 Jan;51(1):106-116. PMID: 30559488, DOI: 10.1038/s41588-018-0288-4

L Coelewij, D Curtis, Mini-review: Update on the genetics of schizophrenia. Ann Hum Genet. 2018 Sep;82(5):239-243. PMID: 29923609, DOI: 10.1111/ahg.12259 [Review]

*^3 RL Collins et al., A structural variation reference for medical and population genetics. Nature 2020 May;581(7809):444-451. PMID: 32461652, DOI: 10.1038/s41586-020-2287-8

* R Essers et al., Prevalence of chromosomal alterations in first-trimester spontaneous pregnancy loss. Nat Med. 2023 Dec;29(12):3233-3242. PMID: 37996709, DOI: 10.1038/s41591-023-02645-5

T Ghi et al., Prenatal diagnosis versus first-trimester screening of trisomy 21 among pregnant women aged 35 or more. J Matern Fetal Neonatal Med. 2015 Apr;28(6):674-678. PMID: 24866349, DOI: 10.3109/14767058.2014.928852

* S Gudmundsson et al., Variant interpretation using population databases: Lessons from gnomAD. Hum Mutat. 2022 Aug;43(8):1012-1030. PMID: 34859531, DOI: 10.1002/humu.24309 [Review]

*^3 DC Hancks, HH Kazazian Jr., Roles for Retrotransposon Insertions in Human Disease. Mob DNA 2016 May 6;7:9. PMID: 27158268, DOI: 10.1186/s13100-016-0065-9. eCollection 2016 [Review]

PS Harper, The discovery of the human chromosome number in Lund, 1955-1956. Hum Genet. 2006 Mar;119(1-2):226-232. PMID: 16463025, DOI: 10.1007/s00439-005-0121-x

^4 T Hassold, P Hunt, To err (meiotically) is human: the genesis of human aneuploidy. Nat Rev Genet . 2001 Apr;2(4):280-291. PMID: 11283700, DOI: 10.1038/35066065 [Review]

*^3 SS Ho, AE Urban, RE Mills, Structural variation in the sequencing era. Nat Rev Genet 2020 Mar;21(3):171-189. PMID: 31729472, DOI: 10.1038/s41576-019-0180-9 [Review]

EJ Hollox, LW Zuccherato, S Tucci, Genome structural variation in human evolution. Trends Genet. 2022 Jan;38(1):45-58. PMID: 34284881, DOI: 10.1016/j.tig.2021.06.015 [Review]

https://bionano.com

* https://gnomad.broadinstitute.org/

*#12 https://www.sciencedirect.com/topics/medicine-and-dentistry/chromosome-identification

https://www.ed.ac.uk/mrc-human-genetics-unit/about/mrc-hgu-history/john-evans. Creation of population-level chromosome portraits of >10,000 babies

L Hu et al., Clinical significance of germline copy number variation in susceptibility of human diseases. J Genet Genomics 2018 Jan 20;45(1):3-12. PMID: 29396143, DOI: 10.1016/j.jgg.2018.01.001 [Review]

ISCN 2020 An International System for Human Cytogenetic Nomenclature. J McGowan-Jordon, RJ Hastings, S Moore (eds). Cytogenet Genome Res 2020; 160:341-503. DOI 10.1159/000510090 and DOI: 10.1159/isbn.978-3-318-06867-2. Erratum. Cytogenet Genome Res 2021;161:476-477. DOI: 10.1159/000520838

*^4 PA Jacobs et al., Estimates of the frequency of chromosome abnormalities detectable in unselected newborns using moderate levels of banding. J Med Genet. 1992 Feb;29(2):103-108. PMID: 1613759, DOI: 10.1136/jmg.29.2.103 [Comparative Study]

ED Jarvis et al., Semi-automated assembly of high-quality diploid human reference genomes. Nature. 2022 Nov;611(7936):519-531. PMID: 36261518, DOI: 10.1038/s41586-022-05325-5

*^3 KJ Karczewski et al., The mutational constraint spectrum quantified from variation in 141,456 humans. Nature 2020 May;581(7809):434-443. PMID: 32461654, DOI: 10.1038/s41586-020-2308-7

*^3 MD Kessler et al., De novo mutations across 1,465 diverse genomes reveal mutational insights and reductions in the Amish founder population. Proc Natl Acad Sci U S A. 2020 Feb 4;117(5):2560-2569. PMID: 31964835, DOI: 10.1073/pnas.1902766117

^3 M Lek et al., Analysis of protein-coding genetic variation in 60,706 humans. Nature 2016 Aug 18;536(7616):285-291. PMID: 27535533, DOI: 10.1038/nature19057

*^3 G Logsdon, MR Vollger, EE Eichler, Long-read human genome sequencing and its applications. Nat Rev Genet. 2020 Oct;21(10):597-614. PMID: 32504078, DOI: 10.1038/s41576-020-0236-x [Review]

CR Marshall et al., Contribution of copy number variants to schizophrenia from a genome-wide study of 41,321 subjects. Nat Genet. 2017 Jan;49(1):27-35. PMID: 27869829, DOI: 10.1038/ng.3725 [Comparative Study]

^3 L Moore et al., The mutational landscape of human somatic and germline cells. Nature 2021 Sep;597(7876):381-386. PMID: 34433962, DOI: 10.1038/s41586-021-03822-7

SI Nagaoka, TJ Hassold, PA Hunt, Human aneuploidy: mechanisms and new insights into an age-old problem. Nat Rev Genet. 2012 Jun 18;13(7):493-504. PMID: 22705668, DOI: 10.1038/nrg3245 [Review]

*^4 J Nielsen, M Wohlert, Chromosome abnormalities found among 34,910 newborn children: results from a 13-year incidence study in Arhus, Denmark. Hum Genet. 1991 May;87(1):81-83. PMID: 2037286, DOI: 10.1007/BF01213097

* VS Pandey, A Bhardwaj, From the reference human genome to human pangenome: Premise, promise and challenge. Front Genet. 2022 Nov 10:13:1042550. PMID: 36437921, DOI: 10.3389/fgene.2022.1042550

O Pös et al., DNA copy number variation: Main characteristics, evolutionary significance, and pathological aspects. Biomed J. 2021 Oct;44(5):548-559. PMID: 34649833, DOI: 10.1016/j.bj.2021.02.003 [Review]

S Ratcliffe, Long-term Outcome in Children of Sex Chromosome Abnormalities. Arch Dis Child 1999 Feb;80(2):192-195. PMID: 10325742, DOI: 10.1136/adc.80.2.192

*^3 SR Richardson et al., The Influence of LINE-1 and SINE Retrotransposons on Mammalian Genomes. Microbiol Spectr. 2015 Apr;3(2):MDNA3-0061-2014. PMID: 261046982, DOI: 10.1128/microbiolspec.MDNA3-0061-2014 [Review]

J Sebat et al., Strong association of de novo copy number mutations with autism. Science 2007 Apr 20;316(5823):445-449. PMID: 17363630, DOI: 10.1126/science.1138659

*^3 D Taliun et al., Sequencing of 53,831 diverse genomes from the NHLBI TOPMed program. bioRxiv 563866. DOI: 10.1101/563866

*^4 NS Thomas et al., De novo apparently balanced translocations in man are predominantly paternal in origin and associated with a significant increase in paternal age. J Med Genet 2010 Feb;47(2):112-115. PMID: 19638350, DOI: 10.1136/jmg.2009.069716

C Vinson, R Chatterjee, CG methylation. Epigenomics . 2012 Dec;4(6):655-63. PMID: 23244310, DOI: 10.2217/epi.12.55 [Review]

T Wang et al., The Human Pangenome Project: a global resource to map genomic diversity. Nature. 2022 Apr;604(7906):437-446. PMID: 35444317, DOI: 10.1038/s41586-022-04601-8 [Review]

*^4 D Warburton, De novo balanced chromosome rearrangements and extra marker chromosomes identified at prenatal diagnosis: clinical significance and distribution of breakpoints. Am J Hum Genet. 1991 Nov;49(5):995-1013. PMID: 1928105, PMCID: PMC1683246

F Zhang, W Gu, ME Hurles, JR Lupski, Copy number variation in human health, disease, and evolution. Annu Rev Genomics Hum Genet . 2009;10:451-481. PMID: 19715442, DOI: 10.1146/annurev.genom.9.081307.164217 [Review]

Chapter 4

S Aganezov et al., A complete reference genome improves analysis of human genetic variation. Science 2022 Apr;376(6588):eabl3533. PMID: 35357935, DOI: 10.1126/science.abl3533

* I Bartha, J di Iulio, J Craig Venter, A Telenti, Human Gene Essentiality, Nat Rev Genet. 2018 Jan;19(1):51-62. PMID: 29082913, DOI: 10.1038/nrg.2017.75 [Review]

* P Cacheiro, MA Haendel, D Smedley, International Mouse Phenotyping Consortium and the Monarch Initiative, New Models for Human Disease From the International Mouse Phenotyping Consortiu. Mamm Genome 2019 Jun;30(5-6):143-150. PMID: 31127358, DOI: 10.1007/s00335-019-09804-5 [Review]

M Eriksson et al., Recurrent De Novo Point Mutations in Lamin A Cause Hutchinson-Gilford Progeria Syndrome. Nature 2003 May 15;423(6937):293-8. PMID: 12714972, DOI: 10.1038/nature01629

Z Gao et al., An estimate of the average number of recessive lethal mutations carried by humans. Genetics 2015 Apr;199(4):1243-1254. PMID: 25697177, DOI: 10.1534/genetics.114.173351

S Gonzalo, R Kreienkamp, P Askjaer, Hutchinson-Gilford Progeria Syndrome: A premature aging disease caused by LMNA gene mutations. Ageing Res Rev. 2017 Jan;33:18-2. PMID: 27374873, DOI: 10.1016/j.arr.2016.06.007 [Review]

* S Gudmundsson et al., Variant interpretation using population databases: Lessons from gnomAD. Hum Mutat. 2022 Aug;43(8):1012-1030. PMID: 34859531, DOI: 10.1002/humu.24309 [Review]

https://en.wikipedia.org/wiki/Birth_rate#/media/File:Countries_by_Birth_Rate_in_2017.svg

https://en.wikipedia.org/wiki/Human_Genome_Project

https://gnomad.broadinstitute.org/ The Genome Aggregation Database (gnomAD) is a resource developed by an international coalition of investigators, with the goal of aggregating and harmonizing both exome and genome sequencing data from a wide variety of large-scale sequencing projects, and making summary data available for the wider scientific community.

https://www.jax.org/search?q=Knockout+Mouse+Project+%28KOMP%29

Knockout mouse project: https://www.komp.org Knockout Mouse Project (KOMP)

Scientists around the world are working to generate a targeted knockout mutation for every gene in the mouse genome through the Knockout Mouse Project (KOMP) which is providing critical tools for understanding gene function and the genetic causes of human diseases

https://www.mousephenotype.org/. International mouse phenotype consortium

https://www.ncbi.nlm.nih.gov/omim, #176670 - Hutchinson-Gilford progeria syndrome; HGPS

Progeria syndrome, childhood-onset, included. Cytogenetic locations: 1q22. OMIM

http://www.progeria.nu/index3.html

https://www.proteinatlas.org/humanproteome/tissue/

KJ Karczewski et al., The mutational constraint spectrum quantified from variation in 141,456 humans. Nature 2020 May;581(7809):434-443. PMID: 32461654, DOI: 10.1038/s41586-020-2308-7

LW Koblan et al., In vivo base editing rescues Hutchinson-Gilford progeria syndrome in mice. Nature 2021 Jan;589(7843):608-614. PMID: 33408413, DOI: 10.1038/s41586-020-03086-7

DG MacArthur et al., A systematic survey of loss-of-function variants in human protein-coding genes. Science 2012 Feb 17;335(6070):823-828. PMID: 22344438, DOI: 10.1126/science.1215040

* OMIM: https://www.ncbi.nlm.nih.gov/omim

S Nurk et al., The complete sequence of a human genome. Science 2022 Apr;376(6588):44-53. PMID: 35357919, DOI: 10.1126/science.abj6987

EGM Reynolds et al., Non-additive association analysis using proxy phenotypes identifies novel cattle syndromes. Nat Genet. 2021 Jul;53(7):949-954. PMID: 34045765, DOI: 10.1038/s41588-021-00872-5

* J Rozman et al., Identification of Genetic Elements in Metabolism by High-Throughput Mouse Phenotyping. Nat Commun. 2018 Jan 18;9(1):288. PMID: 29348434, DOI: 10.1038/s41467-017-01995-2

A De Sandre-Giovannoli et al., Lamin a Truncation in Hutchinson-Gilford Progeria. Science 2003 Jun 27;300(5628):2055. PMID: 12702809, DOI: 10.1126/science.1084125

D Taliun et al., Sequencing of 53,831 diverse genomes from the NHLBI TOPMed program. bioRxiv 563866. DOI: 10.1101/563866

* Welcome to ClinGen, https://clinicalgenome.org

DR Zerbino, A Frankish, P Flicek, Progress, Challenges, and Surprises in Annotating the Human Genome. Annu Rev Genomics Hum Genet. 2020 Aug 31;21:55-79. PMID: 32421357, DOI: 10.1146/annurev-genom-121119-083418 [Review]

Chapter 5

*#14,15 B Arbeithuber, AJ Betancourt, T Ebner, I Tiemann-Boeg, Crossovers are associated with mutation and biased gene conversion at recombination hotspots. Proc Natl Acad Sci U S A. 2015 Feb 17;112(7):2109-2114. PMID: 25646453, DOI: 10.1073/pnas.1416622112

SM Baker et al., Male Mice Defective in the DNA Mismatch Repair Gene PMS2 Exhibit Abnormal Chromosome Synapsis in Meiosis. Cell 1995 Jul 28;82(2):309-319. PMID: 7628019, DOI: 10.1016/0092-8674(95)90318-6

SM Baker et al., Involvement of Mouse Mlh1 in DNA Mismatch Repair and Meiotic Crossing Over. Nat Genet. 1996 Jul;13(3):336-342. PMID: 8673133, DOI: 10.1038/ng0796-336

AN Blackburn et al., Genotype phasing in pedigrees using whole-genome sequence data. Eur J Hum Genet. 2020 Jun;28(6):790-803. PMID: 31996801, DOI: 10.1038/s41431-020-0574-3

KA Boateng et al., Homologous pairing preceding SPO11-mediated double-strand breaks in mice. Dev Cell 2013 Jan 28;24(2):196-205. PMID: 23318132, DOI: 10.1016/j.devcel.2012.12.002

*#14,15 J Brooks Crickard, EC Greene, Biochemical attributes of mitotic and meiotic presynaptic complexes . DNA Repair (Amst) 2018 Nov;71:148-157. PMID: 30195641, DOI: 10.1016/j.dnarep.2018.08.018 [Review]

CL Campbell et al., Escape from crossover interference increases with maternal age. Nat Commun. 2015 Feb 19;6:6260. PMID: 25695863, DOI: 10.1038/ncomms7260

S Claerhout et al., CSYseq: The first Y-chromosome sequencing tool typing a large number of Y-SNPs and Y-STRs to unravel worldwide human population genetics. PLoS Genet. 2021 Sep 7;17(9):e1009758 . PMID: 34491993, DOI: 10.1371/journal.pgen.1009758

F Cole, S Keeney, M Jasin, Preaching about the converted: how meiotic gene conversion influences genomic diversity. Ann N Y Acad Sci. 2012 Sep;1267:95-102. PMID: 22954222, DOI: 10.1111/j.1749-6632.2012.06595.x [Review]

W Edelmann et al., Mammalian MutS homologue 5 is required for chromosome pairing in meiosis. Nat Genet. 1999 Jan;21(1):123-127. PMID: 9916805, DOI: 10.1038/5075

* Y Erlich, T Shor, I Pe'er, S Carmi, Identity Inference of Genomic Data Using Long-Range Familial Searches. Science 2018 Nov 9;362(6415):690-694. PMID: 30309907, DOI: 10.1126/science.aau4832

S Gilad et al., Ataxia-telangiectasia: founder effect among north African Jews. Hum Mol Genet. 1996 Dec;5(12):2033-2037. PMID: 8968760, DOI: 10.1093/hmg/5.12.2033

RE Green et al., A Draft Sequence of the Neandertal Genome. Science 2010 May 7;328(5979):710-722. PMID: 20448178, DOI: 10.1126/science.1188021

* P Gunz et al., Neandertal Introgression Sheds Light on Modern Human Endocranial Globularity. Curr Biol. 2019 Jan 7;29(1):120-127.e5. PMID: 30554901, DOI: 10.1016/j.cub.2018.10.065

BV Halldorsson et al., Characterizing mutagenic effects of recombination through a sequence-level genetic map. Science 2019 Jan 25;363(6425):eaau1043. PMID: 30679340, DOI: 10.1126/science.aau1043

K Hayashi, K Yoshida, Y Matsui, A Histone H3 Methyltransferase Controls Epigenetic Events Required for Meiotic Prophase. Nature 2005 Nov 17;438(7066):374-378. PMID: 16292313, DOI: 10.1038/nature04112

WG Hill, BS Weir, Variation in actual relationship as a consequence of Mendelian sampling and linkage. Genet Res (Camb) 2011 Feb;93(1):47-64. PMID: 21226974, DOI: 10.1017/S0016672310000480

* R Hinch, P Donnelly, A Gupta Hinch, Meiotic DNA breaks drive multifaceted mutagenesis in the human germ line. Science 2023 Dec;382(6674):eadh2531. PMID: 38033082, DOI: 10.1126/science.adh2531

https://allofus.nih.gov/

K-I Ishiguro et al., Meiosis-specific cohesin mediates homolog recognition in mouse spermatocytes. Genes Dev. 2014 Mar 15;28(6):594-607. PMID: 24589552, DOI: 10.1101/gad.237313.113

K-I Ishiguro, The cohesin complex in mammalian meiosis. Genes Cells 2019 Jan;24(1):6-30. PMID: 30479058, DOI: 10.1111/gtc.12652 [Review]

AJ Jeffreys, CA May, Intense and highly localized gene conversion activity in human meiotic crossover hot spots. Nat Genet. 2004 Feb;36(2):151-156. PMID: 14704667, DOI: 10.1038/ng1287

KS Keegan et al., The Atr and Atm Protein Kinases Associate With Different Sites Along Meiotically Pairing Chromosomes. Genes Dev. 1996 Oct 1;10(19):2423-2437. PMID: 8843195, DOI: 10.1101/gad.10.19.2423

K Korunes, MAF Noor, Gene conversion and linkage: effects on genome evolution and speciation. Mol Ecol. 2017 Jan;26(1):351-364. PMID: 27337640, DOI: 10.1111/mec.13736

* J Lange et al., The landscape of mouse meiotic double-strand break formation, processing and repair. Cell 2016 October 20; 167(3): 695-708.e16. PMID: 27745971, DOI:10.1016/j.cell.2016.09.035

EYi Liu, High-resolution sex-specific linkage maps of the mouse reveal polarized distribution of crossovers in male germline. Genetics 2014 May;197(1):91-106. PMID: 24578350, DOI: 10.1534/genetics.114.161653

*#16 S Lu et al., Probing meiotic recombination and aneuploidy of single sperm cells by whole-genome sequencing. Science 2012 Dec 21;338(6114):1627-1630. PMID: 23258895, DOI: 10.1126/science.1229112

MF Lyon, S Rastan, SDM Brown, Genetic variants and strains of the laboratory mouse, third edition, vol 2. Oxford University Press, 1996. ISBN 0198548699

CM Manhart, E Alani, Roles for mismatch repair family proteins in promoting meiotic crossing over. DNA Repair (Amst) 2016 Feb;38:84-93. PMID: 26686657, DOI: 10.1016/j.dnarep.2015.11.024 [Review]

AJ McNairn, VD Rinaldi, JC Schimenti, Repair of Meiotic DNA Breaks and Homolog Pairing in Mouse Meiosis Requires a Minichromosome Maintenance (MCM) Paralog. Genetics 2017 Feb;205(2):529-537. PMID: 27986806, DOI: 10.1534/genetics.116.196808

A Mytlis et al., Control of meiotic chromosomal bouquet and germ cell morphogenesis by the zygotene cilium. Science 2022 Jun 17;376(6599):eabh3104. PMID: 35549308, DOI: 10.1126/science.abh3104

* VM Narasimhan et al., Health and Population Effects of Rare Gene Knockouts in Adult Humans With Related Parents. Science 2016 Apr 22;352(6284):474-477. PMID: 26940866, DOI: 10.1126/science.aac8624

* D Reich, *Who we are and how we got there*. Oxford University Press, 2019. ISBN 9780198821267

EP Rogakou, C Boon, C Redon, WM Bonner, Megabase Chromatin Domains Involved in DNA Double-Strand Breaks in Vivo. J Cell Biol. 1999 Sep 6;146(5):905-916. PMID: 10477747, DOI: 10.1083/jcb.146.5.905

AR Rogers, NS Harris, AA Achenbach, Neanderthal-Denisovan ancestors interbred with a distantly related hominin. Sci Adv. 2020 Feb 21;6(8):eaay5483. PMID: 32128408, DOI: 10.1126/sciadv.aay5483

C Rothblum-Oviatt et al., Ataxia telangiectasia: a review. Orphanet J Rare Dis. 2016 Nov 25;11(1):159. PMID: 27884168, DOI: 10.1186/s13023-016-0543-7 [Review]

R Rowsey et al., Examining variation in recombination levels in the human female: a test of the production-line hypothesis. Am J Hum Genet. 2014 Jul 3;95(1):108-112. PMID: 24995869, DOI: 10.1016/j.ajhg.2014.06.008

S Rubinacci, D M Ribeiro, RJ Hofmeister, O Delaneau, Efficient phasing and imputation of low-coverage sequencing data using large reference panels. Nat Genet. 2021 Jan;53(1):120-126. PMID: 33414550, DOI: 10.1038/s41588-020-00756-0

*#14,15 A Sanchez, G Reginato, P Cejka, Crossover or non-crossover outcomes: tailored processing of homologous recombination intermediates. Curr Opin Genet Dev. 2021 Jul 19;71:39-47. PMID: 34293660, DOI: 10.1016/j.gde.2021.06.012 [Review]

F Sa´nchez-Quinto, C Lalueza-Fox, Almost 20 years of Neanderthal palaeogenetics: adaptation,admixture, diversity, demography and extinction. Phil. Trans. R. Soc. B 2015 Jan 19;370(1660):20130374. PMID: 25487326, DOI: 10.1098/rstb.2013.0374 [Review]

S Sankararaman et al., The date of interbreeding between Neandertals and modern humans. PLoS Genet. 2012;8(10):e1002947. PMID: 23055938, DOI: 10.1371/journal.pgen.1002947

S Sankararaman et al., The genomic landscape of Neanderthal ancestry in present-day humans. Nature 2014 Mar 20;507(7492):354-357. PMID: 24476815, DOI: 10.1038/nature12961

H Scherthan et al., Centromere and Telomere Movements During Early Meiotic Prophase of Mouse and Man Are Associated With the Onset of Chromosome Pairing. J Cell Biol. 1996 Sep;134(5):1109-1125. PMID: 8794855, DOI: 10.1083/jcb.134.5.1109

K Sonehara, Y Okada, Obelisc: an identical-by-descent mapping tool based on SNP streak. Bioinformatics 2021 Apr 5;36(24):5567-5570. PMID: 33135050, DOI: 10.1093/bioinformatics/btaa940

*#15 F Sun et al., The relationship between meiotic recombination in human spermatocytes and aneuploidy in sperm. Hum Reprod. 2008 Aug;23(8):1691-7. PMID: 18482994, DOI: 10.1093/humrep/den027

FAT de Vries et al., Mouse Sycp1 Functions in Synaptonemal Complex Assembly, Meiotic Recombination, and XY Body Formation. Genes Dev. 2005 Jun 1;19(11):1376-1389. PMID: 15937223, DOI: 10.1101/gad.329705

SS de Vries et al., Mouse MutS-like Protein Msh5 Is Required for Proper Chromosome Synapsis in Male and Female Meiosis. Genes Dev. 1999 Mar 1;13(5):523-531. PMID: 10072381, DOI: 10.1101/gad.13.5.523

* B Walsh, M Lynch, *Evolution and Selection of Quantitative Traits*. Oxford University Press, 2018: table 4.2, p. 106. ISBN 9780198830870

Y Xu et al., Targeted disruption of ATM leads to growth retardation, chromosomal fragmentation during meiosis, immune defects, and thymic lymphoma. Genes Dev. 1996 Oct 1;10(19):2411-2422. PMID: 8843194, DOI: 10.1101/gad.10.19.2411

Chapter 6

* A Abdellaoui et al., 15 years of GWAS discovery: Realizing the promise. Am J Hum Genet. 2023 Feb 2;110(2):179-194. PMID: 36634672, DOI: 10.1016/j.ajhg.2022.12.011 [Review]

I Azarova, A Polonikov, E Klyosova, Molecular Genetics of Abnormal Redox Homeostasis in Type 2 Diabetes Mellitus. Int J Mol Sci. 2023 Mar 1;24(5):4738. PMID: 36902173, DOI: 10.3390/ijms24054738 [Review]

AC Bouwman et al., Meta-analysis of genome-wide association studies for cattle stature identifies common genes that regulate body size in mammals. Nat Genet. 2018 Mar;50(3):362-367. PMID: 29459679, DOI: 10.1038/s41588-018-0056-5 [Meta-Analysis]

S Chen et al., A genomic mutational constraint map using variation in 76,156 human genomes. Nature 2024 Jan;625(7993):92-100. PMID: 38057664, DOI: 10.1038/s41586-023-06045-0

M Claussnitzer et al., A brief history of human disease genetics. Nature 2020 Jan;577(7789):179-189. PMID: 31915397, DOI: 10.1038/s41586-019-1879-7

* L Geddes, News in focus: Height's missing heritablity found. Nature 2019;568:444-445. DOI: 10.1038/d41586-019-01157-y

* https://allofus.nih.gov/

https://sourcebooks.fordham.edu/source/pop-in-eur.asp

* https://www.ukbiobank.ac.uk/

* https://www.uscdcb.com

*#18 Council on Dairy Cattel breeding: https;//queries.uscdcb.com/eval/summary/trend.cfm

* https://www.nhlbiwgs.org/. The Trans-Omics for Precision Medicine (TOPMed) program, sponsored by the National Institutes of Health (NIH) National Heart, Lung and Blood Institute (NHLBI), is part of a broader Precision Medicine Initiative, which aims to provide disease treatments tailored to an individual's unique genes and environment. TOPMed contributes to this Initiative through the integration of whole-genome sequencing (WGS) and other omics (e.g., metabolic profiles, epigenomics, protein and RNA expression patterns) data with molecular, behavioral, imaging, environmental, and clinical data.

https://en.wikipedia.org/wiki/Medieval_demography

V Trubetskoy et al., Mapping genomic loci implicates genes and synaptic biology in schizophrenia. Nature . 2022 Apr;604(7906):502-508. PMID: 35396580, DOI: 10.1038/s41586-022-04434-5

* PM Visscher, NR Wray, CS Haley, “Arte Et Labore”-A Blackburn Rovers Fan’s Legacy in Human Complex Trait Genetics. J Anim Breed Genet. 2019 Jul;136(4):273-278. PMID: 31247684, DOI: 10.1111/jbg.12384 [Review]

* PM Visscher, L Yengo, NJ Cox, NR Wray, Discovery and implications of polygenicity of common diseases. Science 2021 Sep 24;373(6562):1468-1473. PMID: 34554790, DOI: 10.1126/science.abi8206 [Review]

* P Wainschtein et al., Assessing the contribution of rare variants to complex trait heritability from whole-genome sequence data. Nat Genet. 2022 Mar;54(3):263-273. PMID: 35256806, DOI: 10.1038/s41588-021-00997-7

* B Walsh, Michael Lynch, *Evolution and Selection of Quantitative Traits*. Oxford University Press, 2018. ISBN: 9780198830870

NR Wray et al., Common Disease Is More Complex Than Implied by the Core Gene Omnigenic Model. Cell 2018 Jun 14;173(7):1573-1580. PMID: 29906445, DOI: 10.1016/j.cell.2018.05.051 [Review]

* NR Wray et al., Complex Trait Prediction From Genome Data: Contrasting EBV in Livestock to PRS in Humans: Genomic Prediction. Genetics 2019 Apr;211(4):1131-1141. PMID: 30967442, DOI: 10.1534/genetics.119.301859

A Zajec et al., Pathogenesis of Type 1 Diabetes: Established Facts and New Insights. Genes (Basel) 2022 Apr 16;13(4):706. PMID: 35456512, DOI: 10.3390/genes13040706 [Review]

Chapter 7

* T Boroviak, J Nichols, Primate Embryogenesis Predicts the Hallmarks of Human naïve Pluripotency. Development 2017 Jan 15;144(2):175-186. PMID: 28096211, DOI: 10.1242/dev.145177 [Review]

SH Geyer et al., The venous system of E14.5 mouse embryos-reference data and examples for diagnosing malformations in embryos with gene deletions. J Anat. 2022 Jan;240(1):11-22. PMID: 34435363, DOI: 10.1111/joa.13536

* https://en.wikipedia.org/wiki/Chimera

https://www.nuffieldfoundation.org

https://www.nuffieldbioethics.org/publications/time-limits-on-maintaining-human-embryos-in-research (2017)

*#43 N Irie, A Sybirna, MA Surani, What Can Stem Cell Models Tell Us About Human Germ Cell Biology? Curr Top Dev Biol. 2018;129:25-65. PMID: 29801530, DOI: 10.1016/bs.ctdb.2018.02.010 [Review]

T Kobayashi et al., Principles of early human development and germ cell program from conserved model systems. Nature 2017 Jun 15;546(7658):416-420. PMID: 28607482, DOI: 10.1038/nature22812

*#43 T Kobayashi, MA Surani, On the origin of the human germline. Development 2018 Jul 23;145(16):dev150433. PMID: 30037844, DOI: 10.1242/dev.150433 [Review]

K Madan, Natural human chimeras: A review. Eur J Med Genet. 2020 Sep;63(9):103971. PMID: 32565253, DOI: 10.1016/j.ejmg.2020.103971 [Review]

J-A Mäkelä, JJ Koskenniemi, HE Virtanen, J Toppari, Testis Development. Endocr Rev. 2019 Aug 1;40(4):857-905. PMID: 30590466, DOI: 10.1210/er.2018-00140 [Review]

* A McLaren, Primordial germ cells in the mouse. Dev Biol. 2003 Oct 1;262(1):1-15. PMID: 14512014, DOI: 10.1016/s0012-1606(03)00214-8 [Review]

* Y Ohinata et al., Blimp1 is a critical determinant of the germ cell lineage in mice. Nature. 2005 Jul 14;436(7048):207-13. PMID: 15937476, DOI: 10.1038/nature03813

MN Shahbazi et al., Self-organization of the human embryo in the absence of maternal tissues. Nat Cell Biol. 2016 Jun;18(6):700-708. PMID: 27144686, DOI: 10.1038/ncb3347

*#43 A Sybirna, FCK Wong, MA Surani, Genetic Basis for Primordial Germ Cells Specification in Mouse and Human: Conserved and Divergent Roles of PRDM and SOX Transcription Factors. Curr Top Dev Biol. 2019;135:35-89. PMID: 31155363, DOI: 10.1016/bs.ctdb.2019.04.004

Box 2

* R Abu-Dawud et al., Pluripotent Stem Cells: Induction and Self-Renewal. Philos Trans R Soc Lond B Biol Sci. 2018 Jul 5;373(1750):20170213. PMID: 29786549, DOI: 10.1098/rstb.2017.0213 [Review]

H Bolton et al., Mouse model of chromosome mosaicism reveals lineage-specific depletion of aneuploid cells and normal developmental potential. Nat Commun. 2016 Mar 29;7:11165. PMID: 27021558, DOI: 10.1038/ncomms11165

MJ Evans, MH Kaufman, Establishment in Culture of Pluripotential Cells From Mouse Embryos. Nature 1981 Jul 9;292(5819):154-156. PMID: 7242681, DOI: 10.1038/292154a0

M De Felici, S Dolci, From Testis to Teratomas: A Brief History of Male Germ Cells in Mammals. Int J Dev Biol. 2013;57(2-4):115-121. PMID: 23784821, DOI: 10.1387/ijdb.130069md [Review]

CE Ford, EP Evans, RL Gardner, Marker chromosome analysis of two mouse chimaeras. J Embryol Exp Morphol. 1975 Apr;33(2):447-457. PMID: 1236926

R L Gardner, Contributions of blastocyst micromanipulation to the study of mammalian development. Bioessays 1998 Feb;20(2):168-180. PMID: 9631662, DOI: 10.1002/(SICI)1521-1878(199802)20:2<168::AID-BIES9>3.0.CO;2-P

* HG Leitch, A Smith, The Mammalian Germline as a Pluripotency Cycle. Development 2013 Jun;140(12):2495-2501. PMID: 23715543, DOI: 10.1242/dev.091603

* J-A Mäkelä, JJ Koskenniemi, HE Virtanen, J Toppari, Testis Development. Endocr Rev. 2019 Aug 1;40(4):857-905. PMID: 30590466, DOI: 10.1210/er.2018-00140 [Review]

GR Martin, Isolation of a Pluripotent Cell Line From Early Mouse Embryos Cultured in Medium Conditioned by Teratocarcinoma Stem Cells. Proc Natl Acad Sci U S A. 1981 Dec;78(12):7634-7638. PMID: 6950406, DOI: 10.1073/pnas.78.12.7634

NE Skakkebaek, Possible carcinoma-in-situ of the testis. Lancet 1972 Sep 9;2(7776):516-517. PMID: 4115573, DOI: 10.1016/s0140-6736(72)91909-5

LC Stevens, CC Little, Spontaneous Testicular Teratomas in an Inbred Strain of Mice. Proc Natl Acad Sci U S A. 1954 Nov;40(11):1080-1087. PMID: 16578442, DOI: 10.1073/pnas.40.11.1080

Box 3

AV Anzalone, LW Koblan, DR Liu, Genome editing with CRISPR-Cas nucleases, base editors, transposases and prime editors. Nat Biotechnol. 2020 Jul;38(7):824-844. PMID: 32572269, DOI: 10.1038/s41587-020-0561-9 [Review]

* B Davies, The technical risks of human gene editing. Hum Reprod. 2019 Nov 1;34(11):2104-2111. PMID: 31696232, DOI: 10.1093/humrep/dez162

*#21 JA Doudna, E Charpentier, Genome editing. The new frontier of genome engineering with CRISPR-Cas9. Science 2014 Nov 28;346(6213):1258096. PMID: 25430774, DOI: 10.1126/science.1258096 [Review]

* J Doudna, S Sternberg, *A crack in creation, the new power to control evolution*. London: Vintage, 2018. ISBN 9781784702762

* JA Doudna, The promise and challenge of therapeutic genome editing. Nature 2020 Feb;578(7794):229-236. PMID: 32051598, DOI: 10.1038/s41586-020-1978-5 [Review]

D Egli et al., Inter-homologue repair in fertilized human eggs'? Nature 2018;560:E5-E7. Arising from H. Ma et al. Nature 2017;548 :413-419. DOI: 10.1038/nature23305

H Gu et al., Deletion of a DNA polymerase beta gene segment in T cells using cell type-specific gene targeting. Science 1994 Jul 1;265(5168):103-106. PMID: 8016642, DOI: 10.1126/science.8016642

IJ Huijbers, Generating Genetically Modified Mice: A Decision Guide. Methods Mol Biol. 2017;1642:1-19. PMID: 28815490, DOI: 10.1007/978-1-4939-7169-5_1 [Review]

* https://www.nap.edu/catalog/25665/heritable-human-genome-editing

*#20 https://en.wikipedia.org/wiki/Cre-Lox_recombination

* H Ledford, CRISPR gene editing in human embryos wreaks chromosomal mayhem. Three studies showing large DNA deletions and reshuffling heighten safety concerns about heritable genome editing. Nature 2020;583:17-18. DOI: 10.1038/d41586-020-01906-4 [News]

* S Mastenbroek, S Repping, Crispr-baby's komen eraan. NRC 4 december 2018

RD Palmiter, RL Brinster, Transgenic mice. Cell 1985 Jun;41(2):343-345. PMID: 2985274, DOI: 10.1016/s0092-8674(85)80004-0 [Review]

K Prelle, N Zink, E Wolf, Pluripotent stem cells--model of embryonic development, tool for gene targeting, and basis of cell therapy. Anat Histol Embryol. 2002 Jun;31(3):169-186. PMID: 12479360, DOI: 10.1046/j.1439-0264.2002.00388.x [Review]

AJ Song, RD Palmiter, Detecting and Avoiding Problems When Using the Cre-lox System. Trends Genet. 2018 May;34(5):333-340. PMID: 29336844, DOI: 10.1016/j.tig.2017.12.008 [Review]

KR Thomas, C Deng, MR Capecchi, High-fidelity gene targeting in embryonic stem cells by using sequence replacement vectors. Mol Cell Biol. 1992 Jul;12(7):2919-2923. PMID: 1620105, DOI: 10.1128/mcb.12.7.2919

H Yang, H Wang, R Jaenisch, Generating genetically modified mice using CRISPR/Cas-mediated genome engineering. Nat Protoc. 2014 Aug;9(8):1956-1968. PMID: 25058643, DOI: 10.1038/nprot.2014.134

Chapter 8

A Berglund et al., Incidence, Prevalence, Diagnostic Delay, and Clinical Presentation of Female 46,XY Disorders of Sex Development. J Clin Endocrinol Metab. 2016 Dec;101(12):4532-4540. PMID: 27603905, DOI: 10.1210/jc.2016-2248

* PS Burgoyne et al., Cell-autonomous Action of the Testis-Determining Gene: Sertoli Cells Are Exclusively XY in XX----XY Chimaeric Mouse Testes. Development 1988 Feb;102(2):443-450. PMID: 3166423, DOI: 10.1242/dev.102.2.443

* B Capel, Women in reproductive science: To Be or Not to Be a Testis. Reproduction 2019 Dec;158(6):F101-F111. PMID: 31265995, DOI: 10.1530/REP-19-0151

EM Eicher, LL Washburn, JB Whitney 3rd, KE Morrow, Mus poschiavinus Y chromosome in the C57BL/6J murine genome causes sex reversal. Science 1982 Aug 6;217(4559):535-537. PMID: 7089579, DOI: 10.1126/science.7089579

T Endo et al., Retinoic Acid and Germ Cell Development in the Ovary and Testis. Biomolecules 2019 Nov 24;9(12):775. PMID: 31771306, DOI: 10.3390/biom9120775 [Review]

EP Evans, CE Ford, MF Lyon, Direct evidence of the capacity of the XY germ cell in the mouse to become an oocyte. Nature 1977 Jun 2;267(5610):430-431. PMID: 559943, DOI: 10.1038/267430a0

M De Felici, FG Klinger, Programmed cell death in mouse primordial germ cells. Int J Dev Biol. 2015;59(1-3):41-49. PMID: 26374524, DOI: 10.1387/ijdb.150064md [Review]

JK Findlay, KJ Hutt, M Hickey, RA Anderson, How Is the Number of Primordial Follicles in the Ovarian Reserve Established? Biol Reprod. 2015 Nov;93(5):111. PMID: 26423124, DOI: 10.1095/biolreprod.115.133652

N Frydman et al., Human foetal ovary shares meiotic preventing factors with the developing testis. Hum Reprod. 2017 Mar 1;32(3):631-642. PMID: 28073973, DOI: 10.1093/humrep/dew343

N Gonen, R Lovell-Badge, The regulation of Sox9 expression in the gonad. Curr Top Dev Biol. 2019;134:223-252. PMID: 30999977, DOI: 10.1016/bs.ctdb.2019.01.004 [Review]

https://en.wikipedia.org/wiki/Chimera

N Irie, A Sybirna, MA Surani, What Can Stem Cell Models Tell Us About Human Germ Cell Biology? Curr Top Dev Biol. 2018;129:25-65. PMID: 29801530, DOI: 10.1016/bs.ctdb.2018.02.010 [Review]

PA Jost, Le controle hormonale de la différentiation du sexe. Biol Rev. 1948;23:201-236. DOI: 10.1111/j.1469-185X.1948.tb00462.x

K Kashimada, P Koopman, Sry: the master switch in mammalian sex determination. Development 2010 Dec;137(23):3921-3930. PMID: 21062860, DOI: 10.1242/dev.048983 [Review]

* P Koopman et al., Male Development of Chromosomally Female Mice Transgenic for Sry. Nature 1991 May 9;351(6322):117-121. PMID: 2030730, DOI: 10.1038/351117a0

* J-A Mäkelä, JJ Koskenniemi, HE Virtanen, J Toppari, Testis Development. Endocr Rev. 2019 Aug 1;40(4):857-905. PMID: 30590466, DOI: 10.1210/er.2018-00140 [Review]

K Murakami et al., Generation of functional oocytes from male mice in vitro. Nature 2023 Mar;615(7954):900-906. PMID: 36922585, DOI: 10.1038/s41586-023-05834-x

*#22 S Nef, I Stévant, A Greenfield, Characterizing the bipotential mammalian gonad. Curr Top Dev Biol. 2019;134:167-194. PMID: 30999975, DOI: 10.1016/bs.ctdb.2019.01.002 [Review]

* I Stévant, MD Papaioannou, S Nef, A brief history of sex determination. Mol Cell Endocrinol. 2018 Jun 15;468:3-10. PMID: 29635012, DOI: 10.1016/j.mce.2018.04.004 [Review]

N Vernet et al., The expression of Y-linked Zfy2 in XY mouse oocytes leads to frequent meiosis 2 defects, a high incidence of subsequent early cleavage stage arrest and infertility. Development 2014 Feb;141(4):855-866. PMID: 24496622, DOI: 10.1242/dev.091165

M Villemure et al., The presence of X- and Y-chromosomes in oocytes leads to impairment in the progression of the second meiotic division. Dev Biol. 2007 Jan 1;301(1):1-13. PMID: 17123505, DOI: 10.1016/j.ydbio.2006.10.034

W Yamazaki et al., Effects of the Sex Chromosome Complement, XX, XO, or XY, on the Transcriptome and Development of Mouse Oocytes During Follicular Growth. Front Genet. 2021 Dec 20;12:792604. PMID: 34987552, DOI: 10.3389/fgene.2021.792604

* J-Q Zhu, SL Tan, T Taketo, A Lack of Coordination Between Sister-Chromatids Segregation and Cytokinesis in the Oocytes of B6.Y ^{TIR} (XY) Sex-Reversed Female Mice. Sci Rep. 2017 Apr 19;7(1):960. PMID: 28424461, DOI: 10.1038/s41598-017-00922-1

Chapter 9

T Dahlet et al., Genome-wide analysis in the mouse embryo reveals the importance of DNA methylation for transcription integrity. Nat Commun. 2020 Jun 19;11(1):3153. PMID: 32561758, DOI: 10.1038/s41467-020-16919-w

S Gkountela et al., DNA Demethylation Dynamics in the Human Prenatal Germline

Cell 2015 Jun 4;161(6):1425-1436. PMID: 26004067, DOI: 10.1016/j.cell.2015.05.012

H Guo et al., The DNA methylation landscape of human early embryos. Nature 2014 Jul 31;511(7511):606-610. PMID: 25079557, DOI: 10.1038/nature13544

*#23 CW Hanna, H Demond, G Kelsey, Epigenetic regulation in development: is the mouse a good model for the human? Hum Reprod Update 2018 Sep 1;24(5):556-576. PMID: 29992283, DOI: 10.1093/humupd/dmy021 [Review]

*#24 M Leseva, BB Knowles, DM Messerschmidt, Davor Solter, Erase-Maintain-Establish: Natural Reprogramming of the Mammalian Epigenome. Cold Spring Harb Symp Quant Biol. 2015;80:155-163. PMID: 26763985, DOI: 10.1101/sqb.2015.80.027441 [Review]

M Monk, M Boubelik, S Lehnert, Temporal and Regional Changes in DNA Methylation in the Embryonic, Extraembryonic and Germ Cell Lineages During Mouse Embryo Development. Development 1987 Mar;99(3):371-382. PMID: 3653008, DOI: 10.1242/dev.99.3.371

D Schübeler, Function and Information Content of DNA Methylation. Nature 2015 Jan 15;517(7534):321-326. PMID: 25592537, DOI: 10.1038/nature14192 [Review]

* MKY Seah, DM Messerschmidt, From Germline to Soma: Epigenetic Dynamics in the Mouse Preimplantation Embryo. Curr Top Dev Biol. 2018;128:203-235. PMID: 29477164, DOI: 10.1016/bs.ctdb.2017.10.011

* A Stäubli, AHFM Peters, Mechanisms of maternal intergenerational epigenetic inheritance. Curr Opin Genet Dev. 2021 Apr;67:151-162. PMID: 33607605, DOI: 10.1016/j.gde.2021.01.008 [Review]

*#23 WWC Tang et al., A Unique Gene Regulatory Network Resets the Human Germline Epigenome for Development. Cell 2015 Jun 4;161(6):1453-1467. PMID: 26046444, DOI: 10.1016/j.cell.2015.04.053

* WWC Tang et al., Specification and Epigenetic Programming of the Human Germ Line. Nat Rev Genet. 2016 Oct;17(10):585-600. PMID: 27573372, DOI: 10.1038/nrg.2016.88 [Review]

G Wu, L Lei, HR Schöler, Totipotency in the mouse. J Mol Med (Berl). 2017 Jul;95(7):687-694. PMID: 28102431, DOI: 10.1007/s00109-017-1509-5 [Review]

P Zhu et al., Single-cell DNA methylome sequencing of human preimplantation embryos. Nat Genet. 2018 Jan;50(1):12-19. PMID: 29255258, DOI: 10.1038/s41588-017-0007-6

Chapter 10

* SC Barton, MA Surani, ML Norris, Role of Paternal and Maternal Genomes in Mouse Development. Nature 1984 Sep 27-Oct 3;311(5984):374-376. PMID: 6482961, DOI: 10.1038/311374a0

AC Bouwman, et al., Meta-analysis of Genome-Wide Association Studies for Cattle Stature Identifies Common Genes That Regulate Body Size in Mammals. Nat Genet. 2018 Mar;50(3):362-367. PMID: 29459679, DOI: 10.1038/s41588-018-0056-5 [Meta-Analysis]

FC Cassidy, M Charalambous, Genomic imprinting, growth and maternal-fetal interactions. J Exp Biol. 2018 Mar 7;221(Pt Suppl 1):jeb164517. PMID: 29514882, DOI: 10.1242/jeb.164517 [Review]

BM Cattanach, M Kirk, Differential activity of maternally and paternally derived chromosome regions in mice. Nature 1985 Jun;315(6019):496-8. PMID: **4000278,** DOI: 10.1038/315496a0

*#26 H Chen et al., Heads, Shoulders, Elbows, Knees, and Toes: Modular Gdf5 Enhancers Control Different Joints in the Vertebrate Skeleton. PLoS Genet. 2016 Nov 30;12(11):e1006454. PMID: 27902701, DOI: 10.1371/journal.pgen.1006454

CW Hanna, G Kelsey, The Specification of Imprints in Mammals. Heredity (Edinb) 2014 Aug;113(2):176-183. PMID: 24939713, DOI: 10.1038/hdy.2014.54 [Review]

* CW Hanna, H Demond, G Kelsey, Epigenetic Regulation in Development: Is the Mouse a Good Model for the Human? Hum Reprod Update 2018 Sep 1;24(5):556-576. PMID: 29992283,DOI: 10.1093/humupd/dmy021 [Review]

https://en.wikipedia.org/wiki/Molar_pregnancy

http://igc.otago.ac.nz

https://www.geneimprint.com/

* https://www.mousebook.org/introduction-data-imprinting-genes

* J McGrath, D Solter, Completion of Mouse Embryogenesis Requires Both the Maternal and Paternal Genomes. Cell 1984 May;37(1):179-183. PMID: 6722870, DOI: 10.1016/0092-8674(84)90313-1

G E Moore et al., The role and interaction of imprinted genes in human fetal growth. Philos Trans R Soc Lond B Biol Sci. 2015 Mar 5;370(1663):20140074. PMID: 25602077, DOI: 10.1098/rstb.2014.0074 [Review]

GT O'Neill, MH Kaufman, Cytogenetic analysis of ethanol-induced parthenogenesis. J Exp Zool. 1989 Feb;249(2):182-192. PMID: 2723605, DOI: 10.1002/jez.1402490211

MM Patten et al, The Evolution of Genomic Imprinting: Theories, Predictions and Empirical Tests. Heredity (Edinb) 2014 Aug;113(2):119-128. PMID: 24755983, DOI: 10.1038/hdy.2014.29 [Review]

* JD Perez, ND Rubinstein, C Dulac, New Perspectives on Genomic Imprinting, an Essential and Multifaceted Mode of Epigenetic Control in the Developing and Adult Brain. Annu Rev Neurosci. 2016 Jul 8;39:347-384. PMID: 27145912, DOI: 10.1146/annurev-neuro-061010-113708 [Review]

*#27 J Peters, The Role of Genomic Imprinting in Biology and Disease: An Expanding View. Nat Rev Genet. 2014 Aug;15(8):517-530. PMID: 24958438, DOI: 10.1038/nrg3766 [Review]

AG Searle, CV Beechey, Complementation Studies With Mouse Translocations. Cytogenet Cell Genet. 1978;20(1-6):282-303. PMID: 648183, DOI: 10.1159/000130859

* V Tucci et al., Genomic Imprinting and Physiological Processes in Mammals. Cell 2019 Feb 21;176(5):952-965. PMID: 30794780, DOI: 10.1016/j.cell.2019.01.043 [Review]

F Zink et al., Insights into imprinting from parent-of-origin phased methylomes and transcriptomes. Nat Genet. 2018 Nov;50(11):1542-1552. PMID: 30349119, DOI: 10.1038/s41588-018-0232-7

Box 4

AP Arnold et al., The Importance of Having Two X Chromosomes. Review Philos Trans R Soc Lond B Biol Sci. 2016 Feb 19;371(1688):20150113. PMID: 26833834, DOI: 10.1098/rstb.2015.0113

* BP Balaton, T Dixon-McDougall, SB Peeters, CJ Brown, The eXceptional Nature of the X Chromosome. Hum Mol Genet. 2018 Aug 1;27(R2):R242-R249. PMID: 29701779, DOI: 10.1093/hmg/ddy148 [Review]

* AC Christensen, Cats as an Aid to Teaching Genetics. Genetics 2000 Jul;155(3):999-1004. PMID: 10880464

* T Green, S Flash, AL Reiss, Sex Differences in Psychiatric Disorders: What We Can Learn From Sex Chromosome Aneuploidies. Neuropsychopharmacology 2019 Jan;44(1):9-21. PMID: 30127341, DOI: 10.1038/s41386-018-0153-2 [Review]

JBS Haldane, The mutation rate of the gene for haemophilia, and its segregation ratios in males and females. Ann Eugen. 1947 Jun;13(4):262-271. PMID: 20249869, DOI: 10.1111/j.1469-1809.1946.tb02367.x

N He et al., At Term, XmO and XpO Mouse Placentas Show Differences in Glucose Metabolism in the Trophectoderm-Derived Outer Zone. Front Cell Dev Biol. 2017 Jun 21;5:63. PMID: 28680878, DOI: 10.3389/fcell.2017.00063

https://en.wikipedia.org/wiki/Hardy%E2%80%93Weinberg_principle

PA Hunt, Survival of XO mouse fetuses: effect of parental origin of the X chromosome or uterine environment? Development 1991 Apr;111(4):1137-1141. PMID: 1879355, DOI: 10.1242/dev.111.4.1137

SD Koukoui, A Chaudhuri, Neuroanatomical, Molecular Genetic, and Behavioral Correlates of Fragile X Syndrome. Review Brain Res Rev. 2007 Jan;53(1):27-38. PMID: 16844227, DOI: 10.1016/j.brainresrev.2006.06.001

MF Lyon, SG Hawker, Reproductive Lifespan in Irradiated and Unirradiated Chromosomally XO Mice. Genet Res. 1973 Apr;21(2):185-194. PMID: 4735820, DOI: 10.1017/s0016672300013355

CI van der Made et al., Presence of Genetic Variants Among Young Men With Severe COVID-19. JAMA. 2020 Jul 24;324(7):1-11. PMID: 32706371, DOI: 10.1001/jama.2020.13719

LS Mamsen et al., Characterization of Follicles in Girls and Young Women With Turner Syndrome Who Underwent Ovarian Tissue Cryopreservation. Fertil Steril. 2019 Jun;111(6):1217-1225.e3. PMID: 30922638, DOI: 10.1016/j.fertnstert.2019.02.003

H Mira-Bontenbal, J Gribnau, New Xist-Interacting Proteins in X-Chromosome Inactivation. Curr Biol. 2016 Apr 25;26(8):R338-342. PMID: 27115694, DOI: 10.1016/j.cub.2016.03.022 [Review]

JL Mueller et al., Independent specialization of the human and mouse X chromosomes for the male germ line. Nat Genet. 2013 Sep;45(9):1083-1087. PMID: 23872635, DOI: 10.1038/ng.2705

* AK San Roman et al., The human inactive X chromosome modulates expression of the active X chromosome. Cell Genom. 2023 Feb 8;3(2):100259. PMID: 36819663, DOI: 10.1016/j.xgen.2023.100259

AG Searle, Gene Frequencies in London's Cats. J Genet. 1949 Dec;49(3):214-220. PMID: 24536670, DOI: 10.1007/BF02986074

AG Searle, *Comparative Genetics of Coat Colour in Mammals*. Academic Press, 1968. ISBN 978-0126334500

* M Souyris, JE Mejía, J Chaumeil, J Guéry, Female predisposition to TLR7-driven autoimmunity: gene dosage and the escape from X chromosome inactivation. Semin Immunopathol. 2019 Mar;41(2):153-164. PMID: 30276444, DOI: 10.1007/s00281-018-0712-y [Review]

* E Viggiano, M Ergoli, E Picillo, L Politano, Determining the Role of Skewed X-chromosome Inactivation in Developing Muscle Symptoms in Carriers of Duchenne Muscular Dystrophy. Hum Genet. 2016 Jul;135(7):685-698. PMID: 27098336, DOI: 10.1007/s00439-016-1666-6 [Review]

Chapter 11

RR Angell, Predivision in Human Oocytes at Meiosis I: A Mechanism for Trisomy Formation in Man. Hum Genet. 1991 Feb;86(4):383-387. PMID: 1999340, DOI: 10.1007/BF00201839

P de Boer, FA van der Hoeven, The Use of Translocation-Derived “Marker-Bivalents” for Studying the Origin of Meiotic Instability in Female Mice. Cytogenet Cell Genet. 1980;26(1):49-58. PMID: 6929247, DOI: 10.1159/000131421

P de Boer, FA van der Hoeven, Chromosome Damage and Non-Disjunction Measured at the First Cleavage Division in Normal and Chromosomally Mutant Female Mice Irradiated at the Diakinesis Stage of Female Meiosis. Mutat Res. 1991 May;248(1):155-162. PMID: 2030704, DOI: 10.1016/0027-5107(91)90097-8

JG Brewen, HS Payne, RJ Preston, X-ray-induced Chromosome Aberrations in Mouse Dictyate Oocytes. I. Time and Dose Relationships. Mutat Res. 1976 Apr;35(1):111-120. PMID: 58377, DOI: 10.1016/0027-5107(76)90173-1

J P de Bruin et al., The role of genetic factors in age at natural menopause. Hum Reprod. 2001 Sep;16(9):2014-2018. PMID: 11527915 , DOI: 10.1093/humrep/16.9.2014

A Caine, MF Lyon, The Induction of Chromosome Aberrations in Mouse Dictyate Oocytes by X-rays and Chemical Mutagens. Mutat Res. 1977 Dec;45(3):325-331. PMID: 201837, DOI: 10.1016/0027-5107(77)90142-7

* C Charalambous, A Webster, M Schuh, Aneuploidy in mammalian oocytes and the impact of maternal ageing. Nat Rev Mol Cell Biol. 2022 Sep 6. PMID: 36068367, DOI: 10.1038/s41580-022-00517-3 [Review]

S Cheng et al., Mammalian oocytes store mRNAs in a mitochondria-associated membraneless compartment. Science 2022 Oct 21;378(6617). PMID: 36264786, DOI: 10.1126/science.abq4835

EY Cheng, et al., Meiotic Recombination in Human Oocytes. PLoS Genet. 2009 Sep;5(9):e1000661. PMID: 19763179, DOI: 10.1371/journal.pgen.1000661

* G Coticchio, DF Albertini, L De Santis (eds), *Oogenesis*. London: Springer, 2013. ISBN 9780857298256, DOI: 10.1007/978-0-85729-826-3

RG Edwards, AG Searle, Genetic radiosensitivity of specific post-dictyate stages in mouse oocytes. Genet. Res. 1963;4:389-398. DOI: 10.1017/S0016672300003785

FE Duncan et al., Unscrambling the oocyte and the egg: clarifying terminology of the female gamete in mammals. Mol Hum Reprod. 2020 Nov 1;26(11):797-800. PMID: 33022047, DOI: 10.1093/molehr/gaaa066 [Editorial]

* C Emori, Z Boucher, E Bolcun-Filas, CHEK2 signaling is the key regulator of oocyte survival after chemotherapy. Sci Adv. 2023 Oct 20;9(42):eadg0898. PMID: 37862420, DOI: 10.1126/sciadv.adg0898

T Endo et al., Retinoic Acid and Germ Cell Development in the Ovary and Testis. Biomolecules 2019 Nov 24;9(12):775. PMID: 31771306, DOI: 10.3390/biom9120775 [Review]

JJ Eppig, Reproduction: Oocytes Call, Granulosa Cells Connect. Curr Biol. 2018 Apr 23;28(8):R354-R356. PMID: 29689210, DOI: 10.1016/j.cub.2018.03.005 [Comment]

X Fan et al., Single-cell reconstruction of follicular remodeling in the human adult ovary. Nat Commun. 2019 Jul 18;10(1):3164. PMID: 31320652, DOI: 10.1038/s41467-019-11036-9

* JK Findlay, KJ Hutt, M Hickey, RA Anderson, How Is the Number of Primordial Follicles in the Ovarian Reserve Established? Biol Reprod. 2015 Nov;93(5):111. PMID: 26423124, DOI: 10.1095/biolreprod.115.133652 [Review]

J Fontana et al., Metabolic cooperation in the ovarian follicle. Physiol Res. 2020 Feb 19;69(1):33-48. PMID: 31854191, DOI: 10.33549/physiolres.934233 [Review]

* JM Franasiak et al., The nature of aneuploidy with increasing age of the female partner: a review of 15,169 consecutive trophectoderm biopsies evaluated with comprehensive chromosomal screening. Fertil Steril. 2014 Mar;101(3):656-663.e1. PMID: 24355045, DOI: 10.1016/j.fertnstert.2013.11.004 [Review]

L Gahurova et al., Transcription and chromatin determinants of de novo DNA methylation timing in oocytes. Epigenetics Chromatin . 2017 May 12:10:25. PMID: 28507606, DOI: 10.1186/s13072-017-0133-5

S Ganesan, AF Keating, Bisphenol A-Induced Ovotoxicity Involves DNA Damage Induction to Which the Ovary Mounts a Protective Response Indicated by Increased Expression of Proteins Involved in DNA Repair and Xenobiotic Biotransformation. Toxicol Sci. 2016 Jul;152(1):169-180. PMID: 27208089, DOI: 10.1093/toxsci/kfw076

Z Gao et al., Overlooked roles of DNA damage and maternal age in generating human germline mutations. Proc Natl Acad Sci U S A. 2019 May 7;116(19):9491-9500. PMID: 31019089, DOI: 10.1073/pnas.1901259116

* JM Goldmann et al., Germline De Novo Mutation Clusters Arise During Oocyte Aging in Genomic Regions With High Double-Strand-Break Incidence. Nat Genet. 2018 Apr;50(4):487-492. PMID: 29507425, DOI: 10.1038/s41588-018-0071-6

BV Halldorsson et al., Characterizing mutagenic effects of recombination through a sequence-level genetic map. Science 2019 Jan 25;363(6425):eaau1043. PMID: 30679340, DOI: 10.1126/science.aau1043

SA Henderson, RG Edwards, Chiasma Frequency and Maternal Age in Mammals. Nature 1968 Apr 6;218(5136):22-28. PMID: 4230650, DOI: 10.1038/218022a0

CA Hodges et al., Experimental Evidence That Changes in Oocyte Growth Influence Meiotic Chromosome Segregation. Hum Reprod. 2002 May;17(5):1171-1180. PMID: 11980735, DOI: 10.1093/humrep/17.5.1171

CA Hodges et al., SMC1beta-deficient Female Mice Provide Evidence That Cohesins Are a Missing Link in Age-Related Nondisjunction. Nat Genet. 2005 Dec;37(12):1351-1355. PMID: 16258540, DOI: 10.1038/ng1672

Z Holubcová, M Blayney, K Elder, M Schuh, Error-prone chromosome-mediated spindle assembly favors chromosome segregation defects in human oocytes. Science 2015 Jun 5;348(6239):1143-1147. PMID: 26045437, DOI: 10.1126/science.aaa9529

* https://en.wikipedia.org/wiki/Miscarriage

https://en.wikipedia.org/wiki/Lionel_Penrose

https://nl.wikipedia.org/wiki/Bisfenol A

PA Hunt, Ethanol-induced Aneuploidy in Male Germ Cells of the Mouse. Cytogenet Cell Genet. 1987;44(1):7-10. PMID: 3816306, DOI: 10.1159/000132333

PA Hunt et al., Bisphenol a Exposure Causes Meiotic Aneuploidy in the Female Mouse. Curr Biol. 2003 Apr 1;13(7):546-553. PMID: 12676084, DOI: 10.1016/s0960-9822(03)00189-1

* IMA Jentoft et al., Mammalian oocytes store proteins for the early embryo on cytoplasmic lattices. Cell . 2023 Nov 22;186(24):5308-5327.e25. PMID: 37922900, DOI: 10.1016/j.cell.2023.10.003

MH Kaufman, The Chromosome Complement of Single-Pronuclear Haploid Mouse Embryos Following Activation by Ethanol Treatment. J Embryol Exp Morphol. 1982 Oct;71:139-154. PMID: 7153693

MH Kaufman, Ethanol-induced chromosomal abnormalities at conception. Nature 1983 Mar 17-23;302(5905):258-260. PMID: 6682181, DOI: 10.1038/302258a0

YK Kawamura et al., Preventing CpG island hypermethylation in oocytes safeguards mouse development. bioRxiv 2024.05.29.595726, DOI: 10.1101/2024.05.29.595726

* D-A Kim, E-K Suh, Defying DNA Double-Strand Break-Induced Death During Prophase I Meiosis by Temporal TAp63α Phosphorylation Regulation in Developing Mouse Oocytes. Mol Cell Biol. 2014 Apr;34(8):1460-1473. PMID: 24515437, DOI: 10.1128/MCB.01223-13

SIR Lane et al., DNA Damage Induces a Kinetochore-Based ATM/ATR-independent SAC Arrest Unique to the First Meiotic Division in Mouse Oocytes. Development 2017 Oct 1;144(19):3475-3486. PMID: 28851706, DOI: 10.1242/dev.153965

SD Limback, DF Albertini, Imaging Strategies for Studying Mammalian Oogenesis. In: Coticchio G., Albertini D., De Santis L. (eds), *Oogenesis*. London: Springer, London, 2013. First Online 12 October 2012. DOI: 10.1007/978-0-85729-826-3_1

* W Lin et al., Ovarian Aging in Women With BRCA Germline Mutations. J Clin Endocrinol Metab. 2017 Oct 1;102(10):3839-3847. PMID: 28938488, DOI: 10.1210/jc.2017-00765

* L Macurek, DNA damage in the oocytes SACs. Cell Cycle 2016;15(4):491-492. PMID: 26709886, DOI: 10.1080/15384101.2015.1136522 [Editorial]

S Malki et al., A Role for Retrotransposon LINE-1 in Fetal Oocyte Attrition in Mice. Dev Cell. 2014 Jun 9;29(5):521-533. PMID: 24882376, DOI: 10.1016/j.devcel.2014.04.027

S Malki, ME Tharp, A Bortvin, A Whole-Mount Approach for Accurate Quantitative and Spatial Assessment of Fetal Oocyte Dynamics in Mice. Biol Reprod. 2015 Nov;93(5):113. PMID: 26423126, DOI: 10.1095/biolreprod.115.132118

P Marangos, J Carroll, Oocytes progress beyond prophase in the presence of DNA damage. Curr Biol. 2012 Jun 5;22(11):989-994. PMID: 22578416, DOI: 10.1016/j.cub.2012.03.063

P Marangos et al., DNA damage-induced metaphase I arrest is mediated by the spindle assembly checkpoint and maternal age. Nat Commun. 2015 Nov 2;6:8706. PMID: 26522734, DOI: 10.1038/ncomms9706

* JH Martin, RJ Aitken, EG Bromfield, B Nixon, DNA Damage and Repair in the Female Germline: Contributions to ART. Hum Reprod Update 2019 Mar 1;25(2):180-201. PMID: 30541031, DOI: 10.1093/humupd/dmy040 [Review]

A Mayer et al., DNA Damage Response During Mouse Oocyte Maturation. Cell Cycle 2016;15(4):546-558. PMID: 26745237, DOI: 10.1080/15384101.2015.1128592

* BP Mihalas, KA Redgrove, EMcLaughlin, B Nixon, Molecular Mechanisms Responsible for Increased Vulnerability of the Ageing Oocyte to Oxidative Damage. Oxid Med Cell Longev. 2017;2017:4015874. PMID: 29312475, DOI: 10.1155/2017/4015874 [Review]

*#28,29 D Monniaux et al., The Ovarian Reserve of Primordial Follicles and the Dynamic Reserve of Antral Growing Follicles: What Is the Link? Biol Reprod. 2014 Apr 25;90(4):85. PMID: 24599291, DOI: 10.1095/biolreprod.113.117077 [Review]

D Monniaux, Driving Folliculogenesis by the Oocyte-Somatic Cell Dialog: Lessons From Genetic Models. Theriogenology 2016 Jul 1;86(1):41-53. PMID: 27155734, DOI: 10.1016/j.theriogenology.2016.04.017 [Review]

SI Nagaoka, TJ Hassold, PA Hunt, Human Aneuploidy: Mechanisms and New Insights Into an Age-Old Problem. Nat Rev Genet. 2012 Jun 18;13(7):493-504. PMID: 22705668, DOI: 10.1038/nrg3245 [Review]

*#29 SS Nussey, S Whitehead, *Endocrinology, An Integrated Approach*. Oxford: BIOS Scientific Publishers, 2001. ISBN-10: 1859962521, DOI: 10.1201/b15306

GT O'Neill, MH Kaufman, Cytogenetic Analysis of Ethanol-Induced Parthenogenesis. J Exp Zool. 1989 Feb;249(2):182-192. PMID: 2723605, DOI: 10.1002/jez.1402490211

LS Penrose, Parental Age and Mutation. Lancet 1955 Aug 13;269(6885):312-313. PMID: 13243724, DOI: 10.1016/s0140-6736(55)92305-9

JS Perry, The Incidence of Embryonic Mortality as a Characteristic of the Individual Sow. J Reprod Fertil. 1960 Feb;1:71-83. PMID: 14432033, DOI: 10.1530/jrf.0.0010071

S Quenby et al., Miscarriage matters: the epidemiological, physical, psychological, and economic costs of early pregnancy loss. Lancet 2021 May 1;397(10285):1658-1667. PMID: 33915094, DOI: 10.1016/S0140-6736(21)00682-6 [Review]

* A Rodríguez-Nuevo et al., Oocytes maintain ROS-free mitochondrial metabolism by suppressing complex I. Nature 2022 Jul;607(7920):756-761. PMID: 35859172, DOI: 10.1038/s41586-022-04979-5

* R Rowsey et al., Examining Variation in Recombination Levels in the Human Female: A Test of the Production-Line Hypothesis. Am J Hum Genet. 2014 Jul 3;95(1):108-112. PMID: 24995869, DOI: 10.1016/j.ajhg.2014.06.008

KS Ruth et al., Genetic insights into biological mechanisms governing human ovarian ageing. Nature 2021 Aug;596(7872):393-397. PMID: 34349265, DOI: 10.1038/s41586-021-03779-7

*#34 Y Sakakibara et al., Bivalent Separation Into Univalents Precedes Age-Related Meiosis I Errors in Oocytes. Nat Commun. 2015 Jul 1;6:7550. PMID: 26130582, DOI: 10.1038/ncomms8550

*#28,29 RM Schultz, P Stein, P Svoboda, The Oocyte-To-Embryo Transition in Mouse: Past, Present, and Future. Biol Reprod. 2018 Jul 1;99(1):160-174. PMID: 29462259, DOI: 10.1093/biolre/ioy013

AL Severance, U Midic, KE Latham, Genotypic divergence in mouse oocyte transcriptomes: possible pathways to hybrid vigor impacting fertility and embryogenesis. Physiol Genomics 2020 Feb 1;52(2):96-109. PMID: 31869285, DOI: 10.1152/physiolgenomics.00078.2019

VB Seplyarskiy et al., TOPMed Population Genetics Working Group, Population sequencing data reveal a compendium of mutational processes in the human germ line. Science 2021 Aug 27;373(6558):1030-1035. PMID: 34385354, DOI: 10.1126/science.aba7408

* P Smith, D Wilhelm, R Rodgers, Development of mammalian ovary. J Endocrinol. 2014 Jun;221(3):R145-161. PMID: 24741072, DOI: 10.1530/JOE-14-0062 [Review]

H Tateno, K Mikamo, Effects of neonatal ovarian X-irradiation in the Chinese hamster. II. Absence of chromosomal and developmental damages in surviving oocytes irradiated at the pachytene and resting dictyate stages. J Radiat Res. 1989 Sep;30(3):209-217. PMID: 2681722, DOI: 10.1269/jrr.30.209

C Tease, G Fisher, X-ray-induced Chromosome Aberrations in Immediately Preovulatory Oocytes. Mutat Res. 1986 Mar;173(3):211-215. PMID: 3951471, DOI: 10.1016/0165-7992(86)90038-2

ME Tharp, S Malki, A Bortvin, Maximizing the Ovarian Reserve in Mice by Evading LINE-1 Genotoxicity. Nat Commun. 2020 Jan 16;11(1):330. PMID: 31949138, DOI: 10.1038/s41467-019-14055-8

* S Titus et al., Impairment of BRCA1-related DNA Double-Strand Break Repair Leads to Ovarian Aging in Mice and Humans. Sci Transl Med. 2013 Feb 13;5(172):172ra21. PMID: 23408054, DOI: 10.1126/scitranslmed.3004925

S Wang et al., Inefficient Crossover Maturation Underlies Elevated Aneuploidy in Human Female Meiosis. Cell 2017 Mar 9;168(6):977-989.e17. PMID: 28262352, DOI: 10.1016/j.cell.2017.02.002

*#34 Al Webster, M Schuh, Mechanisms of Aneuploidy in Human Eggs. Trends Cell Biol. 2017 Jan;27(1):55-68. PMID: 27773484, DOI: 10.1016/j.tcb.2016.09.002 [Review]

*#28,29 AL Winship, JM Stringer, SH Liew, KJ Hutt, The importance of DNA repair for maintaining oocyte quality in response to anti-cancer treatments, environmental toxins and maternal ageing. Hum Reprod Update 2018 Mar 1;24(2):119-134. PMID: 29377997, DOI: 10.1093/humupd/dmy002

AP Zielinska, M Schuh, A Microscopy-Based Approach for Studying Meiosis in Live and Fixed Human Oocytes. Methods Cell Biol. 2018;145:315-333. PMID: 29957212, DOI: 10.1016/bs.mcb.2018.03.039

M Zuccotti, Meiotic and Developmental Competence of Mouse Antral Oocytes. Biol Reprod. 1998 Mar;58(3):700-704. PMID: 9510956, DOI: 10.1095/biolreprod58.3.700

Chapter 12, tabel 5 en 6

I Ahlmann Olesen et al., Clinical, Genetic, Biochemical, and Testicular Biopsy Findings Among 1,213 Men Evaluated for Infertility. Fertil Steril. 2017 Jan;107(1):74-82.e7. PMID: 27793385, DOI: 10.1016/j.fertnstert.2016.09.015

EA Ahmed et al., Genetic Probing of Homologous Recombination and Non-Homologous End Joining During Meiotic Prophase in Irradiated Mouse Spermatocytes. Mutat Res. 2010 Jun 1;688(1-2):12-18. PMID: 20167225, DOI: 10.1016/j.mrfmmm.2010.02.004

*#35 RP Amann, Considerations in Evaluating Human Spermatogenesis on the Basis of Total Sperm Per Ejaculate. J Androl. Nov-Dec 2009;30(6):626-641. PMID: 19443909, DOI: 10.2164/jandrol.108.006817 [Review]

LB Alexandrov et al., Signatures of mutational processes in human cancer. Nature 2013 Aug 22;500(7463):415-421. PMID: 23945592, DOI: 10.1038/nature12477

MJ Anderson et al., Functional evidence for differences in sperm competition in humans and chimpanzees. Am J Phys Anthropol. 2007 Oct;134(2):274-280. PMID: 17632799, DOI: 10.1002/ajpa.20674 [Comparative Study]

F Åsenius, AF Danson, SJ Marzi, DNA methylation in human sperm: a systematic review. Hum Reprod Update 2020 Nov 1;26(6):841-873. PMID: 32790874, DOI: 10.1093/humupd/dmaa025

*^6 E Van Assche et al., Cytogenetics of infertile men. Hum Reprod. 1996 Dec;11 Suppl 4:1-24; discussion 25-6. PMID: 9147109, DOI: 10.1093/humrep/11.suppl_4.1 [Review]

KI Aston et al., Aberrant Sperm DNA Methylation Predicts Male Fertility Status and Embryo Quality. Fertil Steril. 2015 Dec;104(6):1388-97.e1-5. PMID: 26361204, DOI: 10.1016/j.fertnstert.2015.08.019

D Bachtrog, The Y Chromosome as a Battleground for Intragenomic Conflict Trends Genet. 2020 Jul;36(7):510-522. PMID: 32448494, DOI: 10.1016/j.tig.2020.04.008 [Review]

R Baskar, KA Lee, R Yeo, K-W Yeoh, Cancer and radiation therapy: current advances and future directions. Int J Med Sci. 2012;9(3):193-199. PMID: 22408567, DOI: 10.7150/ijms.3635 [Review]

P Bearelly, R Oates, Recent advances in managing and understanding Klinefelter syndrome. F1000Res. 2019 Jan 28;8:F1000 Faculty Rev-112. PMID: 30755791, DOI: 10.12688/f1000research.16747.1 [Review]

RA Beatty et al, An Experiment With Heterospermic Insemination in Cattle. J Reprod Fertil. 1969 Aug;19(3):491-502. PMID: 5809467, DOI: 10.1530/jrf.0.0190491

RA Beatty, MC Lim, VJ Coulter, A quantitative study of the second meiotic metaphase in male mice (Mus musculus). Cytogenet Cell Genet. 1975;15(4):256-275. PMID: 1240796, DOI: 10.1159/000130523

K Bhutani et al., Widespread haploid-biased gene expression enables sperm-level natural selection. Science 2021 Mar 5;371(6533):eabb1723. PMID: 33446482, DOI: 10.1126/science.abb1723

P de Boer, JH Nijhoff, Incomplete Sex Chromosome Pairing of Oligospermic Male Hybrids of Mus Musculus and M. Musculus Molossinus in Relation to the Source of the Y Chromosome and the Presence or Absence of a Reciprocal Translocation. J Reprod Fertil. 1981 May;62(1):235-243. PMID: 7230129, DOI: 10.1530/jrf.0.0620235

P de Boer et al., Kinetics of Meiosis in Azoospermic Males: A Joint Histological and Cytological Approach. Cytogenet Genome Res. 2004;105(1):36-46. PMID: 15218256, DOI: 10.1159/000078007 [Clinical Trial]

F Boitrelle et al., The Sixth Edition of the WHO Manual for Human Semen Analysis: A Critical Review and SWOT Analysis. Life (Basel) 2021 Dec 9;11(12):1368. PMID: 34947899, DOI: 10.3390/life11121368 [Review]

* JS Brandt, MA Cruz Ithier, T Rosen, E Ashkinadze, Advanced Paternal Age, Infertility, and Reproductive Risks: A Review of the Literature. Prenat Diagn. 2019 Jan;39(2):81-87. PMID: 30520056, DOI: 10.1002/pd.5402

JG Brewen, RJ Preston, N Gengozian, Analysis of X-Ray-Induced Chromosomal Translocations in Human and Marmoset Spermatogonial Stem Cells. Nature 1975 Feb 6;253(5491):468-470. PMID: 803303, DOI: 10.1038/253468a0

JE Buizer-Voskamp et al., Paternal Age and Psychiatric Disorders: Findings From a Dutch Population Registry. Schizophr Res. 2011 Jul;129(2-3):128-132. PMID: 21489755, DOI: 10.1016/j.schres.2011.03.021

BM Cattanach, Spermatogonial Stem Cell Killing in the Mouse Following Single and Fractionated X-Ray Doses, as Assessed by Length of Sterile Period. Mutat Res. 1974 Oct;25(1):53-62. PMID: 4427607, DOI: 10.1016/0027-5107(74)90218-8

F Chalmel, AD Rolland, Linking transcriptomics and proteomics in spermatogenesis. Reproduction 2015 Nov;150(5):R149-R157. PMID: 26416010, DOI: 10.1530/REP-15-0073 [Review]

AC Chandley, Genetic Contribution to Male Infertility. Hum Reprod. 1998 Jun;13 Suppl 3:76-83; discussion 84-88. PMID: 9755416, DOI: 10.1093/humrep/13.suppl_3.76

Z Chang et al., Triptonide is a reversible non-hormonal male contraceptive agent in mice and non-human primates. Nat Commun. 2021 Feb 23;12(1):1253. PMID: 33623031, DOI: 10.1038/s41467-021-21517-5

* L Cheng et al., Testicular cancer. Nat Rev Dis Primers 2018 Oct 5;4(1):29. PMID: 30291251, DOI: 10.1038/s41572-018-0029-0 [Review]

Y Clermont, Renewal of Spermatogonia in Man. Am J Anat. 1966 Mar;118(2):509-524. PMID: 5917196, DOI: 10.1002/aja.1001180211

J Cocquet et al., A genetic basis for a postmeiotic X versus Y chromosome intragenomic conflict in the mouse. PLoS Genet 2012 Sep;8(9):e1002900. PMID: 23028340, DOI: 10.1371/journal.pgen.1002900

* M Cocuzza, C Alvarenga, R Pagani, The epidemiology and etiology of azoospermia. Clinics (Sao Paulo). 2013;68 Suppl 1(Suppl 1):15-26. PMID: 23503951, DOI: 10.6061/clinics/2013(sup01)03

*^6 S Colaco, D Modi, Genetics of the human Y chromosome and its association with male infertility. Reprod Biol Endocrinol. 2018 Feb 17;16(1):14. PMID: 29454353, DOI: 10.1186/s12958-018-0330-5 [Review]

*#34 ^5 TG Cooper et al., World Health Organization Reference Values for Human Semen Characteristics. Hum Reprod Update May-Jun 2010;16(3):231-245. PMID: 19934213, DOI: 10.1093/humupd/dmp048

G Corona et al., Sperm recovery and ICSI outcomes in Klinefelter syndrome: a systematic review and meta-analysis. Hum Reprod Update 2017 May 1;23(3):265-275. PMID: 28379559, DOI: 10.1093/humupd/dmx008 [Review]

* A Davis Bell et al., Insights into variation in meiosis from 31,228 human sperm genomes. Nature 2020 Jul;583(7815):259-264. PMID: 32494014, DOI: 10.1038/s41586-020-2347-0 [Review]

* JR Drevet, RJ Aitken, Oxidative Damage to Sperm DNA: Attack and Defense. Adv Exp Med Biol. 2019;1166:107-117. PMID: 31301049, DOI: 10.1007/978-3-030-21664-1_7

S van den Driesche et al., Experimentally induced testicular dysgenesis syndrome originates in the masculinization programming window. JCI Insight 2017 Mar 23;2(6):e91204. PMID: 28352662, DOI: 10.1172/jci.insight.91204

ML Eisenberg et al., The Relationship Between Male BMI and Waist Circumference on Semen Quality: Data From the LIFE Study. Hum Reprod. 2014 Feb;29(2):193-200. PMID: 24306102, DOI: 10.1093/humrep/det428

* ML Eisenberg et al., Semen Quality, Infertility and Mortality in the USA. Hum Reprod. 2014 Jul;29(7):1567-1574. PMID: 24838701, DOI: 10.1093/humrep/deu106 [Multicenter Study]

ML Eisenberg et al., Increased Risk of Cancer in Infertile Men: Analysis of U.S. Claims Data. J Urol. 2015 May;193(5):1596-1601. PMID: 25463997, DOI: 10.1016/j.juro.2014.11.080

ML Eisenberg, S Li, MR Cullen, LC Baker, Increased Risk of Incident Chronic Medical Conditions in Infertile Men: Analysis of United States Claims Data. Fertil Steril. 2016 Mar;105(3):629-636. PMID: 26674559, DOI: 10.1016/j.fertnstert.2015.11.011

T Endo et al., Retinoic Acid and Germ Cell Development in the Ovary and Testis
Biomolecules 2019 Nov 24;9(12):775. PMID: 31771306, DOI: 10.3390/biom9120775 [Review]

A Enguita-Marruedo et al., Meiotic arrest occurs most frequently at metaphase and is often incomplete in azoospermic men. Fertil Steril. 2019 Dec;112(6):1059-1070.e3. PMID: 31767154, DOI: 10.1016/j.fertnstert.2019.08.004 [Observational Study]

B Eskenazi et al., The association of age and semen quality in healthy men. Hum Reprod. 2003 Feb;18(2):447-454. PMID: 12571189, DOI: 10.1093/humrep/deg107

I Faisal, L Kauppi, Reduced MAD2 levels dampen the apoptotic response to non-exchange sex chromosomes and lead to sperm aneuploidy. Development 2017 Jun 1;144(11):1988-1996. PMID: 28506992, DOI: 10.1242/dev.149492

A Forand et al., Similarities and Differences in the in Vivo Response of Mouse Neonatal Gonocytes and Spermatogonia to Genotoxic Stress. Biol Reprod. 2009 May;80(5):860-873. PMID: 19144961, DOI: 10.1095/biolreprod.108.072884

CE Ford, AG Searle, EP Evans, BJ West, Differential Transmission of Translocations Induced in Spermatogonia of Mice by Irradiation. Cytogenetics 1969;8(6):447-470. PMID: 5365250, DOI: 10.1159/000130056

LR Fraser, I Maudlin, Analysis of aneuploidy in first-cleavage mouse embryos fertilized in vitro and in vivo. Environ Health Perspect. 1979 Aug;31:141-149. PMID: 387395, DOI: 10.1289/ehp.7931141

Z Gao et al., Overlooked roles of DNA damage and maternal age in generating human germline mutations. Proc Natl Acad Sci U S A. 2019 May 7;116(19):9491-9500. PMID: 31019089, DOI: 10.1073/pnas.1901259116

P Gharagozloo et al., A Novel Antioxidant Formulation Designed to Treat Male Infertility Associated With Oxidative Stress: Promising Preclinical Evidence From Animal Models. Hum Reprod. 2016 Feb;31(2):252-262. PMID: 26732620, DOI: 10.1093/humrep/dev302

ME Gill et al., De novo transcriptome assembly of mouse male germ cells reveals novel genes, stage-specific bidirectional promoter activity, and noncoding RNA expression. Genome Res. 2023 Dec 21;33(12):2060-2078. PMID: 38129075, DOI: 10.1101/gr.278060.123

*#33 C González-Marín, J Gosálvez, R Roy, Types, causes, detection and repair of DNA fragmentation in animal and human sperm cells. Int J Mol Sci. 2012 Oct 31;13(11):14026-14052. PMID: 23203048, DOI: 10.3390/ijms131114026 [Review]

*#32 M-C Grégoire et al., The DNA double-strand “breakome” of mouse spermatids. Cell Mol Life Sci. 2018 Aug;75(15):2859-2872. PMID: 29417179, DOI: 10.1007/s00018-018-2769-0

J Guo et al., The adult human testis transcriptional cell atlas. Cell Res. 2018 Dec;28(12):1141-1157. PMID: 30315278, DOI: 10.1038/s41422-018-0099-2

* P Hallast et al., Assembly of 43 human Y chromosomes reveals extensive complexity and variation. Nature 2023 Sep;621(7978):355-364. PMID: 37612510, DOI: 10.1038/s41586-023-06425-6

G Hamer et al., DNA Double-Strand Breaks and gamma-H2AX Signaling in the Testis. Biol Reprod. 2003 Feb;68(2):628-634. PMID: 12533428, DOI: 10.1095/biolreprod.102.008672

G Hamer, DG de Rooij, Mutations Causing Specific Arrests in the Development of Mouse Primordial Germ Cells and Gonocytes. Biol Reprod. 2018 Jul 1;99(1):75-86. PMID: 29590307, DOI: 10.1093/biolre/ioy075 [Review]

K Harris, R Nielsen, The Genetic Cost of Neanderthal Introgression. Genetics 2016 Jun;203(2):881-891. PMID: 27038113, DOI: 10.1534/genetics.116.186890

GW van der Heijden et al., Chromosome-wide Nucleosome Replacement and H3.3 Incorporation During Mammalian Meiotic Sex Chromosome Inactivation. Nat Genet. 2007 Feb;39(2):251-258. PMID: 17237782, DOI: 10.1038/ng1949

CG Heller, Y Clermont, Spermatogenesis in man: an estimate of its duration. Science 1963 Apr 12;140(3563):184-186. PMID: 13953583, DOI: 10.1126/science.140.3563.184

S Hilz, AJ Modzelewski, PE Cohen, A Grimson, The roles of microRNAs and siRNAs in mammalian spermatogenesis. Development 2016 Sep 1;143(17):3061-3073. PMID: 27578177, DOI: 10.1242/dev.136721 [Review]

*#33 JH Hoeijmakers, Genome maintenance mechanisms for preventing cancer

Nature 2001 May 17;411(6835):366-374. PMID: 11357144, DOI: 10.1038/35077232 [Review]

G van der Horst, L Maree, Sperm form and function in the absence of sperm competition. Mol Reprod Dev. 2014 Mar;81(3):204-216. PMID: 24273026, DOI: 10.1002/mrd.22277 [Review]

https://en.wikipedia.org/wiki/Percoll

* http://infertilegenome.org

* https://en.wikipedia.org/wiki/Percoll

https://www.proteinatlas.org/humanproteome/tissue/

SZ Jan et al., Distinct Prophase Arrest Mechanisms in Human Male Meiosis. Development 2018 Apr 16;145(16):dev160614. PMID: 29540502, DOI: 10.1242/dev.160614

B Jégou et al., Meiotic Genes Are Enriched in Regions of Reduced Archaic Ancestry. Mol Biol Evol. 2017 Aug 1;34(8):1974-1980. PMID: 28444387, DOI: 10.1093/molbev/msx141

TG Jenkins et al., Decreased fecundity and sperm DNA methylation patterns. Fertil Steril. 2016 Jan;105(1):51-7.e1-3. PMID: 26453269, DOI: 10.1016/j.fertnstert.2015.09.013

E Johanson, A Study of Schizophrenia in the Male: A Psychiatric and Social Study Based on 138 Cases With Follow Up. Acta Psychiatr Neurol Scand Suppl. 1958;125:1-132. PMID: 13594594

SG Johnsen, Testicular Biopsy Score Count--A Method for Registration of Spermatogenesis in Human Testes: Normal Values and Results in 335 Hypogonadal Males. Hormones 1970;1(1):2-25. PMID: 5527187, DOI: 10.1159/000178170

L Johnson, DD Varner, Effect of Daily Spermatozoan Production but Not Age on Transit Time of Spermatozoa Through the Human Epididymis. Biol Reprod. 1988 Nov;39(4):812-817. PMID: 3207807, DOI: 10.1095/biolreprod39.4.812

J Kaplanis et al., Genetic and chemotherapeutic influences on germline hypermutation. Nature 2022 May;605(7910):503-508. PMID: 35545669, DOI: 10.1038/s41586-022-04712-2

*^6 L Kasak, M Laan, Monogenic causes of non-obstructive azoospermia: challenges, established knowledge, limitations and perspectives. Hum Genet. 2021 Jan;140(1):135-154. PMID: 31955275, DOI: 10.1007/s00439-020-02112-y [Review]

* DL Keefe, Telomeres and Genomic Instability During Early Development. Eur J Med Genet. 2020 Feb;63(2):103638. PMID: 30862510, DOI: 10.1016/j.ejmg.2019.03.002 [Review]

T Kido, Y-F C Lau, Roles of the Y chromosome genes in human cancers. Asian J Androl. May-Jun 2015;17(3):373-380. PMID: 25814157, DOI: 10.4103/1008-682X.150842 [Review]

*#38 C Krausz, E Casamonti, Spermatogenic Failure and the Y Chromosome. Hum Genet. 2017 May;136(5):637-655. PMID: 28456834, DOI: 10.1007/s00439-017-1793-8 [Review]

* C Krausz, A Riera-Escamilla, Genetics of Male Infertility. Nat Rev Urol. 2018 Jun;15(6):369-384. PMID: 29622783, DOI: 10.1038/s41585-018-0003-3 [Review]

*^6 C Krausz et al., Genetic dissection of spermatogenic arrest through exome analysis: clinical implications for the management of azoospermic men. Genet Med. 2020 Dec;22(12):1956-1966. PMID: 32741963, DOI: 10.1038/s41436-020-0907-1

CL Labrecque, B Fuglestad, Electrostatic Drivers of GPx4 Interactions with Membrane, Lipids, and DNA. Biochemistry 2021 Sep 21;60(37):2761-2772. PMID: 34492183, DOI: 10.1021/acs.biochem.1c00492

D Langenstroth-Röwer et al., De novo methylation in male germ cells of the common marmoset monkey occurs during postnatal development and is maintained in vitro. Epigenetics 2017 Jul 3;12(7):527-539. PMID: 27786608, DOI: 10.1080/15592294.2016.1248007

EL Larson et al., The Composite Regulatory Basis of the Large X-Effect in Mouse Speciation. Mol Biol Evol. 2017 Feb 1;34(2):282-295. PMID: 27999113, DOI: 10.1093/molbev/msw243

T Latif et al., Semen Quality as a Predictor of Subsequent Morbidity: A Danish Cohort Study of 4,712 Men With Long-Term Follow-up. Am J Epidemiol. 2017 Oct 15;186(8):910-917. PMID: 28498890, DOI: 10.1093/aje/kwx067

S Laurentino, J Borgmann, J Gromoll, On the Origin of Sperm Epigenetic Heterogeneity. Reproduction 2016 May;151(5):R71-R78. PMID: 26884419, DOI: 10.1530/REP-15-0436 [Review]

HS Levin, Testicular Biopsy in the Study of Male Infertility: Its Current Usefulness, Histologic Techniques, and Prospects for the Future. Hum Pathol. 1979 Sep;10(5):569-584. PMID: 43278, DOI: 10.1016/s0046-8177(79)80100-8

H Levine et al., Temporal Trends in Sperm Count: A Systematic Review and Meta-Regression Analysis. Hum Reprod Update 2017 Nov 1;23(6):646-659. PMID: 28981654, DOI: 10.1093/humupd/dmx022 [Review]

* H Levine et al., Temporal trends in sperm count: a systematic review and meta-regression analysis of samples collected globally in the 20th and 21st centuries. Hum Reprod Update . 2023 Mar 1;29(2):157-176. PMID: 36377604, DOI: 10.1093/humupd/dmac035 [Meta-Analysis]

R Lilford et al., Case-control Study of Whether Subfertility in Men Is Familial. BMJ. 1994 Sep 3;309(6954):570-573. PMID: 8086942, DOI: 10.1136/bmj.309.6954.570

EA Lucotte et al., Dynamic Copy Number Evolution of X- And Y-Linked Ampliconic Genes in Human Populations. Genetics 2018 Jul;209(3):907-920. PMID: 29769284, DOI: 10.1534/genetics.118.300826

* J-A Mäkelä, JJ Koskenniemi, HE Virtanen, J Toppari, Testis Development. Endocr Rev. 2019 Aug 1;40(4):857-905. PMID: 30590466, DOI: 10.1210/er.2018-00140 [Review]

M Marcet-Ortega et al., p53 and TAp63 participate in the recombination-dependent pachytene arrest in mouse spermatocytes. PLoS Genet. 2017 Jun 15;13(6):e1006845. PMID: 28617799, DOI: 10.1371/journal.pgen.1006845

F Marchetti, J Bishop, J Gingerich, AJ Wyrobek, Meiotic interstrand DNA damage escapes paternal repair and causes chromosomal aberrations in the zygote by maternal misrepair. Sci Rep. 2015 Jan 8;5:7689. PMID: 25567288, DOI: 10.1038/srep07689

I Masliukaite et al., Establishing Reference Values for Age-Related Spermatogonial Quantity in Prepubertal Human Testes: A Systematic Review and Meta-Analysis. Fertil Steril. 2016 Dec;106(7):1652-1657.e2. PMID: 27717555, DOI: 10.1016/j.fertnstert.2016.09.002 [Review]

FL Mendez, The Divergence of Neandertal and Modern Human Y Chromosomes. Am J Hum Genet. 2016 Apr 7;98(4):728-734. PMID: 27058445, DOI: 10.1016/j.ajhg.2016.02.023

V Monesi, R Geremia, A D'Agostino, C Boitani, Biochemistry of Male Germ Cell Differentiation in Mammals: RNA Synthesis in Meiotic and Postmeiotic Cells. Curr Top Dev Biol. 1978;12:11-36. PMID: 352625, DOI: 10.1016/s0070-2153(08)60592-x [Review]

D Mortimer, The functional anatomy of the human spermatozoon: relating ultrastructure and function. Mol Hum Reprod. 2018 Dec 1;24(12):567-592. PMID: 30215807, DOI: 10.1093/molehr/gay040

*#31 B Muciaccia et al., Novel Stage Classification of Human Spermatogenesis Based on Acrosome Development. Biol Reprod. 2013 Sep 19;89(3):60. PMID: 23946533, DOI: 10.1095/biolreprod.113.111682

JL Mueller et al., The Mouse X Chromosome Is Enriched for Multicopy Testis Genes Showing Postmeiotic Expression. Nat Genet. 2008 Jun;40(6):794-799. PMID: 18454149, DOI: 10.1038/ng.126

K Mulryan et al., Reduced Vas Deferens Contraction and Male Infertility in Mice Lacking P2X1 Receptors. Nature 2000 Jan 6;403(6765):86-89. PMID: 10638758, DOI: 10.1038/47495

E Mulugeta Achame, WM Baarends, J Gribnau, JA Grootegoed, Evaluating the Relationship Between Spermatogenic Silencing of the X Chromosome and Evolution of the Y Chromosome in Chimpanzee and Human. PLoS One 2010 Dec 14;5(12):e15598. PMID: 21179482, DOI: 10.1371/journal.pone.0015598

A Noblanc et al., Epididymis Response Partly Compensates for Spermatozoa Oxidative Defects in snGPx4 and GPx5 Double Mutant Mice. PLoS One 2012;7(6):e38565. PMID: 22719900, DOI: 10.1371/journal.pone.0038565

A Noblanc et al., DNA oxidative damage in mammalian spermatozoa: where and why is the male nucleus affected? Free Radic Biol Med. 2013 Dec;65:719-723. PMID: 23954469, DOI: 10.1016/j.freeradbiomed.2013.07.044

MT Oetjens et al., Y-Chromosome Structural Diversity in the Bonobo and Chimpanzee Lineages. Genome Biol Evol. 2016 Aug 3;8(7):2231-2240. PMID: 27358426, DOI: 10.1093/gbe/evw150

A-K Olsen et al., How Do Male Germ Cells Handle DNA Damage? Toxicol Appl Pharmacol. 2005 Sep 1;207(2 Suppl):521-531. PMID: 16051290, DOI: 10.1016/j.taap.2005.01.060 [Review]

J Onofre, Y Baert, K Faes, E Goossens, Cryopreservation of testicular tissue or testicular cell suspensions: a pivotal step in fertility preservation. Hum Reprod Update 2016 Nov;22(6):744-761. PMID: 27566839, DOI: 10.1093/humupd/dmw029 [Review]

*^6 MS Oud et al., Validation and application of a novel integrated genetic screening method to a cohort of 1,112 men with idiopathic azoospermia or severe oligozoospermia. Hum Mutat. 2017 Nov;38(11):1592-1605. PMID: 28801929, DOI: 10.1002/humu.23312 [Meta-Analysis]

*^6 MS Oud et al., A Systematic Review and Standardized Clinical Validity Assessment of Male Infertility Genes. Hum Reprod. 2019 May 1;34(5):932-941. PMID: 30865283, DOI: 10.1093/humrep/dez022

*^6 MS Oud, *Genetic causes of male infertility*. PhD Thesis Radboud Universiteit Nijmegen, 26 november 2020. ISBN 9789464160215

* M S Oud et al., A de novo paradigm for male infertility. Nat Commun. 2022 Jan 10;13(1):154. PMID: 35013161, DOI: 10.1038/s41467-021-27132-8

DM Ozata et al., PIWI-interacting RNAs: small RNAs with big functions. Nat Rev Genet. 2019 Feb;20(2):89-108. PMID: 30446728, DOI: 10.1038/s41576-018-0073-3 [Review]

S Ozturk, Telomerase Activity and Telomere Length in Male Germ Cells. Biol Reprod. 2015 Feb;92(2):53. PMID: 25568305, DOI: 10.1095/biolreprod.114.124008 [Review]

S Pacheco et al., The ATM signaling cascade promotes recombination-dependent pachytene arrest in mouse spermatocytes. PLoS Genet. 2015 Mar 13;11(3):e1005017. PMID: 25768017, DOI: 10.1371/journal.pgen.1005017

* S Di Persio et al., Spermatogonial Kinetics in Humans. Development 2017 Oct 1;144(19):3430-3439. PMID: 28827392, DOI: 10.1242/dev.150284

A Rhie et al., The complete sequence of a human Y chromosome. Nature 2023 Sep;621(7978):344-354. PMID: 37612512, DOI: 10.1038/s41586-023-06457-y

A Riera-Escamilla et al., Sequencing of a 'Mouse Azoospermia' Gene Panel in Azoospermic Men: Identification of RNF212 and STAG3 Mutations as Novel Genetic Causes of Meiotic Arrest. Hum Reprod. 2019 Jun 4;34(6):978-988. PMID: 31125047, DOI: 10.1093/humrep/dez042

DG de Rooij, P de Boer, Specific arrests of spermatogenesis in genetically modified and mutant mice. Cytogenet Genome Res. 2003;103(3-4):267-276. PMID: 15051947, DOI: 10.1159/000076812 [Review]

K Rothkamm, M Löbrich, Evidence for a Lack of DNA Double-Strand Break Repair in Human Cells Exposed to Very Low X-Ray Doses. Proc Natl Acad Sci U S A. 2003 Apr 29;100(9):5057-5062. PMID: 12679524, DOI: 10.1073/pnas.0830918100

V Roukos, T Misteli, The Biogenesis of Chromosome Translocations. Nat Cell Biol. 2014 Apr;16(4):293-300. PMID: 24691255, DOI: 10.1038/ncb2941 [Review]

H Royo et al., Evidence that meiotic sex chromosome inactivation is essential for male fertility. Curr Biol. 2010 Dec 7;20(23):2117-2123. PMID: 21093264, DOI: 10.1016/j.cub.2010.11.010

CE Rübe et al., Protecting the Heritable Genome: DNA Damage Response Mechanisms in Spermatogonial Stem Cells. DNA Repair (Amst) 2011 Feb 7;10(2):159-168. PMID: 21123119, DOI: 10.1016/j.dnarep.2010.10.007

A Salas-Huetos et al., The role of miRNAs in male human reproduction: a systematic review. Andrology 2020 Jan;8(1):7-26. PMID: 31578810, DOI: 10.1111/andr.12714

D Santi, S De Vincentis, E Magnani, G Spaggiari, Impairment of sperm DNA methylation in male infertility: a meta-analytic study. Andrology 2017 Jul;5(4):695-703. PMID: 28718528, DOI: 10.1111/andr.12379 [Review]

*#37 RB Sciurano et al., The role of asynapsis in human spermatocyte failure. Int J Androl. 2012 Aug;35(4):541-549. PMID: 21977946, DOI: 10.1111/j.1365-2605.2011.01221.x [Case Reports]

AG Searle, CV Beechey, Sperm-count, Egg-Fertilization and Dominant Lethality After X-irradiation of Mice. Mutat Res. 1974 Jan;22(1):63-72. PMID: 4842084, DOI: 10.1016/0027-5107(74)90009-8

PB Selby, X-ray-induced Specific-Locus Mutation Rate in Newborn Male Mice. Mutat Res. 1973 Apr;18(1):63-75. PMID: 4572528, DOI: 10.1016/0027-5107(73)90021-3

VB Seplyarskiy et al., TOPMed Population Genetics Working Group. Population sequencing data reveal a compendium of mutational processes in the human germ line. Science 2021 Aug 27;373(6558):1030-1035. PMID: 34385354, DOI: 10.1126/science.aba7408

* R Sharma et al., Effects of Increased Paternal Age on Sperm Quality, Reproductive Outcome and Associated Epigenetic Risks to Offspring. Reprod Biol Endocrinol. 2015 Apr 19;13:35. PMID: 25928123, DOI: 10.1186/s12958-015-0028-x [Review]

S Sharma et al., Spermatogonial Stem Cells: Updates From Specification to Clinical Relevance. Hum Reprod Update 2019 May 1;25(3):275-297. PMID: 30810745, DOI: 10.1093/humupd/dmz006 [Review]

M Simard, C Laprise, SL Girard, Impact of Paternal Age at Conception on Human Health. Clin Chem. 2019 Jan;65(1):146-152. PMID: 30602478, DOI: 10.1373/clinchem.2018.294421 [Review]

L Simon et al., Paternal Influence of Sperm DNA Integrity on Early Embryonic Development. Hum Reprod. 2014 Nov;29(11):2402-2412. PMID: 25205757, DOI: 10.1093/humrep/deu228

* L Simon et al., A Systematic Review and Meta-Analysis to Determine the Effect of Sperm DNA Damage on *in vitro* Fertilization and Intracytoplasmic Sperm Injection Outcome. Asian J Androl. Jan-Feb 2017;19(1):80-90. PMID: 2734500, DOI: 10.4103/1008-682X.182822 [Review]

* NE Skakkebaek et al., Male Reproductive Disorders and Fertility Trends: Influences of Environment and Genetic Susceptibility. Physiol Rev. 2016 Jan;96(1):55-97. PMID: 26582516, DOI: 10.1152/physrev.00017.2015 [Review]

* NE Skakkebæk et al., Environmental factors in declining human fertility. Nat Rev Endocrinol. 2022 Mar;18(3):139-157. PMID: 34912078, DOI: 10.1038/s41574-021-00598-8 [Review]

DL Stewart et al., A Second Experiment With Heterospermic Insemination in Cattle. J Reprod Fertil. 1974 Jan;36(1):107-116. PMID: 4809693, DOI: 10.1530/jrf.0.0360107

RA Tegelenbosch, DG de Rooij, A Quantitative Study of Spermatogonial Multiplication and Stem Cell Renewal in the C3H/101 F1 Hybrid Mouse. Mutat Res. 1993 Dec;290(2):193-200. PMID: 7694110, DOI: 10.1016/0027-5107(93)90159-d

C Templado, L Uroz, A Estop, New Insights on the Origin and Relevance of Aneuploidy in Human Spermatozoa. Mol Hum Reprod. 2013 Oct;19(10):634-643. PMID: 23720770, DOI: 10.1093/molehr/gat039 [Review]

L Tiepolo, O Zuffardi, Localization of factors controlling spermatogenesis in the nonfluorescent portion of the human Y chromosome long arm. Hum Genet. 1976 Oct 28;34(2):119-124. PMID: 1002136, DOI: 10.1007/BF00278879

*^5 H Tournaye, C Krausz, RD Oates, Novel Concepts in the Aetiology of Male Reproductive Impairment. Lancet Diabetes Endocrinol. 2017 Jul;5(7):544-553. PMID: 27395771, DOI: 10.1016/S2213-8587(16)30040-7 [Review]

B Trombetta, F Cruciani, chromosome palindromes and gene conversion. Hum Genet. 2017 May;136(5):605-619. PMID: 28303348, DOI: 10.1007/s00439-017-1777-8 [Review]

*#34,36 ^5,6 F Tüttelmann, C Ruckert, A Röpke, Disorders of Spermatogenesis: Perspectives for Novel Genetic Diagnostics After 20 Years of Unchanged Routine. Med Genet. 2018;30(1):12-20. PMID: 29527098, DOI: 10.1007/s11825-018-0181-7 [Review]

M Uhlén et al., Proteomics. Tissue-based map of the human proteome. Science 2015 Jan 23;347(6220):1260419. PMID: 25613900, DOI: 10.1126/science.1260419

R Vegesna, M Tomaszkiewicz, P Medvedev, KD Makova, Dosage regulation, and variation in gene expression and copy number of human Y chromosome ampliconic genes. PLoS Genet. 2019 Sep 16;15(9):e1008369. PMID: 31525193, DOI: 10.1371/journal.pgen.1008369

ER te Velde, JP Bonde, Misconceptions About Falling Sperm Counts and Fertility in Europe. Asian J Androl. 2013 Mar;15(2):195-198. PMID: 23202698, DOI: 10.1038/aja.2012.122 [Review]

M de Vries et al., Human Male Meiotic Sex Chromosome Inactivation. PLoS One 2012;7(2):e31485. PMID: 22355370, DOI: 10.1371/journal.pone.0031485 [Comparative Study]

M De Vries, L Ramos, Z Housein, P De Boer, Chromatin Remodelling Initiation During Human Spermiogenesis. Biol Open 2012 May 15;1(5):446-457. PMID: 23213436, DOI: 10.1242/bio.2012844

H Wang et al., The Microbiome, an Important Factor That Is Easily Overlooked in Male Infertility. Front Microbiol. 2022 Mar 2;13:831272. PMID: 35308385, DOI: 10.3389/fmicb.2022.831272 [Review]

PJ Wang, JR McCarrey, F Yang, DC Page, An Abundance of X-linked Genes Expressed in Spermatogonia. Nat Genet. 2001 Apr;27(4):422-426. PMID: 11279525, DOI: 10.1038/86927

*#32 WS Ward, Organization of sperm DNA by the nuclear matrix. Am J Clin Exp Urol. 2018 Apr 1;6(2):87-92. eCollection 2018. PMID: 29666836, PMCID: PMC5902726 [Review]

MC Wessels-Kaalen, R Bakker, P de Boer, Transmission of X-ray-induced Reciprocal Translocations in Normal Male Mice and in Male Mice With a Reduced Sperm Count Due to Translocation Homozygosity. Mutat Res. 1986 Nov;163(2):181-191. PMID: 3762559, DOI: 10.1016/0027-5107(86)90047-3

*^5 *WHO laboratory manual for the examination and processing of human semen*. 6th edition. World Health Organization, 2021. ISBN 9789240030787

M Wohl, P Gorwood, Paternal Ages Below or Above 35 Years Old Are Associated With a Different Risk of Schizophrenia in the Offspring. Eur Psychiatry 2007 Jan;22(1):22-26. PMID: 17142012, DOI: 10.1016/j.eurpsy.2006.08.007 [Meta-Analysis]

MS Wosnitzer, M Goldstein, Obstructive azoospermia. Urol Clin North Am. 2014 Feb;41(1):83-95. PMID: 24286769, DOI: 10.1016/j.ucl.2013.08.013 [Review]

MJ Xavier et al., Paternal impacts on development: identification of genomic regions vulnerable to oxidative DNA damage in human spermatozoa. Hum Reprod. 2019 Oct 2;34(10):1876-1890. PMID: 31593593, DOI: 10.1093/humrep/dez153

Q Yin et al., Revisiting chromatin packaging in mouse sperm. Genome Res. 2023 Dec 21;33(12):2079-2093. PMID: 38129076, DOI: 10.1101/gr.277845.123

EV Zubkova, M Wade, B Robaire, Changes in Spermatozoal Chromatin Packaging and Susceptibility to Oxidative Challenge During Aging. Fertil Steril. 2005 Oct;84 Suppl 2:1191-1198. PMID: 16210011, DOI: 10.1016/j.fertnstert.2005.04.044 [Comparative Study]

Chapter 13

A Ahmadi, SC Ng, Fertilizing Ability of DNA-damaged Spermatozoa. J Exp Zool. 1999 Nov 1;284(6):696-704. PMID: 10531556, DOI: 10.1002/(sici)1097-010x(19991101)284:6<696::aid-jez11>3.0.co;2-e

* MA Ali et al., Odorant and Taste Receptors in Sperm Chemotaxis and Cryopreservation: Roles and Implications in Sperm Capacitation, Motility and Fertility. Genes (Basel) 2021 Mar 27;12(4):488. PMID: 33801624, DOI: 10.3390/genes12040488 [Review]

MJ Ashwood-Smith, RG Edwards, DNA Repair by Oocytes. Mol Hum Reprod. 1996 Jan;2(1):46-51. PMID: 9238657, DOI: 10.1093/molehr/2.1.46 [Review]

T Avidor-Reiss, M Mazur, EL Fishman, P Sindhwani, The Role of Sperm Centrioles in Human Reproduction - The Known and the Unknown. Front Cell Dev Biol. 2019 Oct 1;7:188. PMID: 31632960, DOI: 10.3389/fcell.2019.00188 [Review]

H Bakker, JH Wallinga, RD Politiek, Reproduction and body weight of mice after long-term selection for large litter size. Journal of Animal Science 1978;46(6):1572-1580. DOI: 10.2527/jas1978.4661572x

P de Boer, FA van der Hoeven, JA Chardon, The Production, Morphology, Karyotypes and Transport of Spermatozoa From Tertiary Trisomic Mice and the Consequences for Egg Fertilization. J Reprod Fertil. 1976 Nov;48(2):249-256. PMID: 994096, DOI: 10.1530/jrf.0.0480249

ML Boerjan, P de Boer, First cell cycle of zygotes of the mouse derived from oocytes aged postovulation in vivo and fertilized in vivo. Mol Reprod Dev. 1990 Feb;25(2):155-163. PMID: 2178641, DOI: 10.1002/mrd.1080250208

* H Bolton et al., Mouse Model of Chromosome Mosaicism Reveals Lineage-Specific Depletion of Aneuploid Cells and Normal Developmental Potential. Nat Commun. 2016 Mar 29;7:11165. PMID: 27021558, DOI: 10.1038/ncomms11165

S Bontekoe et al., Low Oxygen Concentrations for Embryo Culture in Assisted Reproductive Technologies. Cochrane Database Syst Rev. 2012 Jul 11;(7):CD008950. PMID: 22786519, DOI: 10.1002/14651858.CD008950.pub2 [Review]

E Casser et al., Totipotency Segregates Between the Sister Blastomeres of Two-Cell Stage Mouse Embryos. Sci Rep. 2017 Aug 15;7(1):8299. PMID: 28811525, DOI: 10.1038/s41598-017-08266-6

E Casser, S Israel, M Boiani, Multiplying Embryos: Experimental Monozygotic Polyembryony in Mammals and Its Uses. Int J Dev Biol. 2019;63(3-4-5):143-155. PMID: 31058293, DOI: 10.1387/ijdb.190016mb [Review]

* AW Chan et al., Clonal Propagation of Primate Offspring by Embryo Splitting. Science 2000 Jan 14;287(5451):317-319. PMID: 10634789, DOI: 10.1126/science.287.5451.317

J Cohen, DC McNaughton, Spermatozoa: The Probable Selection of a Small Population by the Genital Tract of the Female Rabbit. J Reprod Fertil. 1974 Aug;39(2):297-310. PMID: 4415860, DOI: 10.1530/jrf.0.0390297

JF Crow, The Origins, Patterns and Implications of Human Spontaneous Mutation. Nat Rev Genet. 2000 Oct;1(1):40-47. PMID: 11262873, DOI: 10.1038/35049558 [Review]

A Derijck et al., Motile Human Normozoospermic and Oligozoospermic Semen Samples Show a Difference in Double-Strand DNA Break Incidence. Hum Reprod. 2007 Sep;22(9):2368-2376. PMID: 17580297, DOI: 10.1093/humrep/dem166

* A Derijck et al., DNA double-strand break repair in parental chromatin of mouse zygotes, the first cell cycle as an origin of de novo mutation. Hum Mol Genet. 2008 Jul 1;17(13):1922-1937. PMID: 18353795, DOI: 10.1093/hmg/ddn090

A Destouni et al., Zygotes Segregate Entire Parental Genomes in Distinct Blastomere Lineages Causing Cleavage-Stage Chimerism and Mixoploidy. Genome Res. 2016 May;26(5):567-578. PMID: 27197242, DOI: 10.1101/gr.200527.115

* AF Dixson, Copulatory and Postcopulatory Sexual Selection in Primates. Folia Primatol (Basel) 2018;89(3-4):258-286. PMID: 29804108, DOI: 10.1159/000488105 [Review]

R Dumollard et al., Mitochondrial function and redox state in mammalian embryos. Semin Cell Dev Biol. 2009 May;20(3):346-353. PMID: 19530278, DOI: 10.1016/j.semcdb.2008.12.013 [Review]

DB Dunson, DD Baird, AJ Wilcox, CR Weinberg, Day-specific Probabilities of Clinical Pregnancy Based on Two Studies With Imperfect Measures of Ovulation. Hum Reprod. 1999 Jul;14(7):1835-1839. PMID: 10402400, DOI: 10.1093/humrep/14.7.1835 [Comparative Study]

*#41 DB Dunson, B Colombo, DD Baird, Changes With Age in the Level and Duration of Fertility in the Menstrual Cycle. Hum Reprod. 2002 May;17(5):1399-1403. PMID: 11980771, DOI: 10.1093/humrep/17.5.1399 [Comparative Study]

RG Edwards, Maturation in Vitro of Mouse, Sheep, Cow, Pig, Rhesus Monkey and Human Ovarian Oocytes. Nature 1965 Oct 23;208(5008):349-351. PMID: 4957259, DOI: 10.1038/208349a0

* M Van Eetvelde et al., The Importance of the Periconception Period: Immediate Effects in Cattle Breeding and in Assisted Reproduction Such as Artificial Insemination and Embryo Transfer. Adv Exp Med Biol. 2017;1014:41-68, PMID: 28864984, DOI: 10.1007/978-3-319-62414-3_3 [Review]

MAMM Ferraz et al., An Oviduct-On-A-Chip Provides an Enhanced in Vitro Environment for Zygote Genome Reprogramming. Nat Commun. 2018 Nov 22;9(1):4934. PMID: 30467383, DOI: 10.1038/s41467-018-07119-8

NL First, RE Short, JB Peters, FW Stratman, Transport of spermatozoa in estrual and luteal sows. Journal of Animal Science 1965;24:917 [Abstract]

B Fischer, W Künzel, J Kleinstein, H Gips, Oxygen tension in follicular fluid falls with follicle maturation. Eur J Obstet Gynecol Reprod Biol. 1992 Jan 9;43(1):39-43. PMID: 1737607, DOI: 10.1016/0028-2243(92)90241-p

*#39 EL Fishman et al., A novel atypical sperm centriole is functional during human fertilization. Nat Commun. 2018 Jun 7;9(1):2210. PMID: 29880810, DOI: 10.1038/s41467-018-04678-8

* R De La Fuente, C Baumann, MM Viveiros, ATRX Contributes to Epigenetic Asymmetry and Silencing of Major Satellite Transcripts in the Maternal Genome of the Mouse Embryo. Development 2015 May 15;142(10):1806-1817. PMID: 25926359, DOI: 10.1242/dev.118927

* L Gaspa-Toneu, AHFM Peters, Nucleosomes in mammalian sperm: conveying paternal epigenetic inheritance or subject to reprogramming between generations? Curr Opin Genet Dev. 2023 Apr:79:102034. PMID: 36893482, DOI: 10.1016/j.gde.2023.102034 [Review]

LV Gatica et al., Picomolar Gradients of Progesterone Select Functional Human Sperm Even in Subfertile Samples. Mol Hum Reprod. 2013 Sep;19(9):559-569. PMID: 23729411, DOI: 10.1093/molehr/gat037

WM Generoso, KT Cain, M Krishna, SW Huff, Genetic Lesions Induced by Chemicals in Spermatozoa and Spermatids of Mice Are Repaired in the Egg. Proc Natl Acad Sci U S A. 1979 Jan;76(1):435-437. PMID: 284360, DOI: 10.1073/pnas.76.1.435

* JG Hall, Twinning. Lancet 2003 Aug 30;362(9385):735-743. PMID: 12957099, DOI: 10.1016/S0140-6736(03)14237-7 [Review]

GW van der Heijden et al., Asymmetry in histone H3 variants and lysine methylation between paternal and maternal chromatin of the early mouse zygote. Mech Dev. 2005 Sep;122(9):1008-1022. PMID: 15922569, DOI: 10.1016/j.mod.2005.04.009

GW van der Heijden et al., Parental Origin of Chromatin in Human Monopronuclear Zygotes Revealed by Asymmetric Histone Methylation Patterns, Differs Between IVF and ICSI. Mol Reprod Dev. 2009 Jan;76(1):101-108. PMID: 18481364, DOI: 10.1002/mrd.20933

*#39 G Hernández-Silva, M Chirinos, Proteins From Male and Female Reproductive Tracts Involved in Sperm Function Regulation. Zygote 2019 Feb;27(1):5-16. PMID: 30621800, DOI: 10.1017/S096719941800062X [Review]

RH Hunter, The Effects of Delayed Insemination on Fertilization and Early Cleavage in the Pig. J Reprod Fertil. 1967 Feb;13(1):133-147. PMID: 6066768, DOI: 10.1530/jrf.0.0130133

P Hyttel, T Greve, H Callesen, Ultrastructural aspects of oocyte maturation and fertilization in cattle. J Reprod Fertil Suppl. 1989;38:35-47. PMID: 2677348 [Review]

* K Jacobs et al., Mitotic Spindle Disruption in Human Preimplantation Embryos Activates the Spindle Assembly Checkpoint but Not Apoptosis Until Day 5 of Development. Mol Hum Reprod. 2017 May 1;23(5):321-329. PMID: 28159965, DOI: 10.1093/molehr/gax007

* DL Keefe, Telomeres and genomic instability during early development. Eur J Med Genet. 2020 Feb;63(2):103638. PMID: 30862510, DOI: 10.1016/j.ejmg.2019.03.002 [Review]

KE Koehler et al., Sex-specific Differences in Meiotic Chromosome Segregation Revealed by Dicentric Bridge Resolution in Mice. Genetics 2002 Nov;162(3):1367-1379. PMID: 12454080, DOI: 10.1093/genetics/162.3.1367

* S Kölle et al., Ciliary Transport, Gamete Interaction, and Effects of the Early Embryo in the Oviduct: Ex Vivo Analyses Using a New Digital Videomicroscopic System in the Cow. Biol Reprod. 2009 Aug;81(2):267-274. PMID: 19299315, DOI: 10.1095/biolreprod.108.073874

* S Ladstätter, K Tachibana-Konwalski, A Surveillance Mechanism Ensures Repair of DNA Lesions During Zygotic Reprogramming. Cell 2016 Dec 15;167(7):1774-1787.e13. PMID: 27916276, DOI: 10.1016/j.cell.2016.11.009

JD Lee, Y Kamiguchi, R Yanagimachi, Analysis of Chromosome Constitution of Human Spermatozoa With Normal and Aberrant Head Morphologies After Injection Into Mouse Oocytes. Hum Reprod. 1996 Sep;11(9):1942-1946. PMID: 8921068, DOI: 10.1093/oxfordjournals.humrep.a019521

* HJ Leese, Metabolism of the Preimplantation Embryo: 40 Years On. Reproduction 2012 Apr;143(4):417-427. PMID: 22408180, DOI: 10.1530/REP-11-0484 [Review]

* T Lord, RJ Aitken, Fertilization Stimulates 8-hydroxy-2'-deoxyguanosine Repair and Antioxidant Activity to Prevent Mutagenesis in the Embryo. Dev Biol. 2015 Oct 1;406(1):1-13. PMID: 26234752, DOI: 10.1016/j.ydbio.2015.07.024

* X Lu, Z Gao, D Qin, L Li, A Maternal Functional Module in the Mammalian Oocyte-To-Embryo Transition. Trends Mol Med. 2017 Nov;23(11):1014-1023. PMID: 28993030, DOI: 10.1016/j.molmed.2017.09.004 [Review]

K Ly et al, Telomere length in early childhood is associated with sex and ethnicity. Sci Rep. 2019 Jul 17;9(1):10359. PMID: 31316106, DOI: 10.1038/s41598-019-46338-x [Comparative Study]

K Madan, Natural human chimeras: A review. Eur J Med Genet. 2020 Sep;63(9):103971. PMID: 32565253, DOI: 10.1016/j.ejmg.2020.103971 [Review]

* V Maillo et al., Oviductal Response to Gametes and Early Embryos in Mammals. Reproduction 2016 Oct;152(4):R127-R141. PMID: 27512123, DOI: 10.1530/REP-16-0120 [Review]

* LA McGraw, SS Suarez, MF Wolfner, On a matter of seminal importance

Bioessays 2015 Feb;37(2):142-147. PMID: 25379987, DOI: 10.1002/bies.201400117

*#43 MA Molè, A Weberling, M Zernicka-Goetz, Comparative analysis of human and mouse development: From zygote to pre-gastrulation. Curr Top Dev Biol. 2020;136:113-138. PMID: 31959285, DOI: 10.1016/bs.ctdb.2019.10.002 [Review]

Y Muro, M Okabe, Mechanisms of fertilization--a view from the study of gene-manipulated mice. J Androl. May-Jun 2011;32(3):218-225. PMID: 20966425, DOI: 10.2164/jandrol.110.010900 [Review]

R Musson, Ł Gąsior, S Bisogno, GE Ptak, DNA damage in preimplantation embryos and gametes: specification, clinical relevance and repair strategies. Hum Reprod Update 2022 May 2;28(3):376-399. PMID: 35021196, DOI: 10.1093/humupd/dmab046 [Review]

T Nakatani et al., DNA replication fork speed underlies cell fate changes and promotes reprogramming. Nat Genet. 2022 Mar;54(3):318-327. PMID: 35256805, DOI: 10.1038/s41588-022-01023-0

* KYB Ng et al., In vivo oxygen, temperature and pH dynamics in the female reproductive tract and their importance in human conception: a systematic review. Hum Reprod Update 2018 Jan 1;24(1):15-34. PMID: 29077897, DOI: 10.1093/humupd/dmx028 [Review]

M Okabe, Sperm-egg interaction and fertilization: past, present, and future . Biol Reprod. 2018 Jul 1;99(1):134-146. PMID: 29462236, DOI: 10.1093/biolre/ioy028 [Review]

SB Olsen, RE Magenis, Preferential paternal origin of de novo structural chromosome rearrangements. In: The cytogenetics of mammalian autosomal rearrangements 583-599. Alan LIss Inc, 1988

JP Ozil, D Huneau, Activation of Rabbit Oocytes: The Impact of the Ca2+ Signal Regime on Development. Development 2001 Mar;128(6):917-928. PMID: 11222146, DOI: 10.1242/dev.128.6.917

* J-P Ozil, T Sainte-Beuve, B Banrezes, [Mg^{2+}]/[Ca^{2+}] Determines Ca^{2+} Response at Fertilization: Tuning of Adult Phenotype? Reproduction 2017 Nov;154(5):675-693. PMID: 28851827, DOI: 10.1530/REP-16-0057

GD Palermo, LT Colombero, Z Rosenwaks, The human sperm centrosome is responsible for normal syngamy and early embryonic development. Rev Reprod. 1997 Jan;2(1):19-27. PMID: 9414462, DOI: 10.1530/ror.0.0020019 [Review]

* KL Palmerola et al., Replication stress impairs chromosome segregation and preimplantation development in human embryos. Cell 2022 Aug 4;185(16):2988-3007.e20. PMID: 35858625, DOI: 10.1016/j.cell.2022.06.028

*#42 L Papale, A Fiorentino, M Montag, G Tomas, The zygote. Hum Reprod. 2012 Aug;27 Suppl 1:i22-49. PMID: 22811310, DOI: 10.1093/humrep/des205

* HV Ramírez-Gómez, I Tuval, A Guerrero, A Darszon, Analysis of sperm chemotaxis. Methods Cell Biol. 2019;151:473-486. PMID: 30948027, DOI: 10.1016/bs.mcb.2018.12.002 [Review]

JP Rigby, The persistance of spermatozoa at the uterotubal junction (varken). J Reprod Fert. 1966;11:153

JB Roelofs, FJCM van Eerdenburg, NM Soede, B Kemp, Various behavioral signs of estrous and their relationship with time of ovulation in dairy cattle. Theriogenology 2005 Mar 15;63(5):1366-1377. PMID: 15725444, DOI: 10.1016/j.theriogenology.2004.07.009

K Rothkamm, M Löbrich, Evidence for a Lack of DNA Double-Strand Break Repair in Human Cells Exposed to Very Low X-Ray Doses. Proc Natl Acad Sci U S A. 2003 Apr 29;100(9):5057-5062. PMID: 12679524, DOI: 10.1073/pnas.0830918100

M Santa Rocca, C Foresta, A Ferlin, Telomere length: lights and shadows on their role in human reproduction. Biol Reprod. 2019 Feb 1;100(2):305-317. PMID: 30277496, DOI: 10.1093/biolre/ioy208

*#39 AH Sathananthan et al., The sperm centriole: its inheritance, replication and perpetuation in early human embryos. Hum Reprod. 1996 Feb;11(2):345-356. PMID: 8671223, DOI: 10.1093/humrep/11.2.345

AG Searle et al., The Induction of Translocations in Mouse Spermatozoa. I. Kinetics of Dose Response With Acute X-Irradiation. Mutat Res. 1974 Feb;22(2):157-174. PMID: 4842823, DOI: 10.1016/0027-5107(74)90096-7 [Comparative Study]

Q-Q Sha et al., Dynamics and clinical relevance of maternal mRNA clearance during the oocyte-to-embryo transition in humans. Nat Commun. 2020 Oct 1;11(1):4917. PMID: 33004802, DOI: 10.1038/s41467-020-18680-6

SS Suarez, AA Pacey, Sperm transport in the female reproductive tract. Hum Reprod Update Jan-Feb 2006;12(1):23-37. PMID: 16272225, DOI: 10.1093/humupd/dmi047 [Review]

* SS Suarez, Regulation of sperm storage and movement in the mammalian oviduct. Int J Dev Biol. 2008;52(5-6):455-462. PMID: 18649258, DOI: 10.1387/ijdb.072527ss [Review]

* SS Suarez, Mammalian sperm interactions with the female reproductive tract. Cell Tissue Res. 2016 Jan;363(1):185-194. PMID: 26183721, DOI: 10.1007/s00441-015-2244-2 [Review]

ZB Tong et al., Mater, a Maternal Effect Gene Required for Early Embryonic Development in Mice. Nat Genet. 2000 Nov;26(3):267-268. PMID: 11062459, DOI: 10.1038/81547

T Trapphoff et al., Postovulatory Aging Affects Dynamics of mRNA, Expression and Localization of Maternal Effect Proteins, Spindle Integrity and Pericentromeric Proteins in Mouse Oocytes. Hum Reprod. 2016 Jan;31(1):133-149. PMID: 26577303, DOI: 10.1093/humrep/dev279

H Van de Velde et al., The Four Blastomeres of a 4-cell Stage Human Embryo Are Able to Develop Individually Into Blastocysts With Inner Cell Mass and Trophectoderm. Hum Reprod. 2008 Aug;23(8):1742-1747. PMID: 18503052, DOI: 10.1093/humrep/den190

*#40 AJ Wilcox, D Dunson, DD Baird, The timing of the “fertile window” in the menstrual cycle: day specific estimates from a prospective study. BMJ. 2000 Nov 18;321(7271):1259-1262. PMID: 11082086, DOI: 10.1136/bmj.321.7271.1259

*#43 CC Wong et al, Non-invasive imaging of human embryos before embryonic genome activation predicts development to the blastocyst stage. Nat Biotechnol. 2010 Oct;28(10):1115-1121. PMID: 20890283, DOI: 10.1038/nbt.1686

WE Wright et al., Telomerase activity in human germline and embryonic tissues and cells. Dev Genet. 1996;18(2):173-179. PMID: 8934879, DOI: 10.1002/(SICI)1520-6408(1996)18:2<173::AID-DVG10>3.0.CO;2-3

YQ Yuan et al., Influence of Oxygen Tension on Apoptosis and Hatching in Bovine Embryos Cultured in Vitro. Theriogenology 2003 Apr 1;59(7):1585-1596. PMID: 12559463, DOI: 10.1016/s0093-691x(02)01204-9

MV Zaragoza et al., Parental origin and phenotype of triploidy in spontaneous abortions: predominance of diandry and association with the partial hydatidiform mole. Am J Hum Genet. 2000 Jun;66(6):1807-1820. PMID: 10801385, DOI: 10.1086/302951

*#39 MT Zenzes, R Bielecki, RF Casper, SP Leibo, Effects of chilling to 0 degrees C on the morphology of meiotic spindles in human metaphase II oocytes. Fertil Steril. 2001 Apr;75(4):769-777. PMID: 11287033, DOI: 10.1016/s0015-0282(00)01800-8 [Comparative Study]

K Zhang, GW Smith, Maternal Control of Early Embryogenesis in Mammals. Reprod Fertil Dev. 2015 Jul;27(6):880-896. PMID: 25695370, DOI: 10.1071/RD14441 [Review]

Chapter 14

R Allio, S Donega, N Galtier, B Nabholz, Large Variation in the Ratio of Mitochondrial to Nuclear Mutation Rate across Animals: Implications for Genetic Diversity and the Use of Mitochondrial DNA as a Molecular Marker. Mol Biol Evol. 2017 Nov 1;34(11):2762-2772. PMID: 28981721, DOI: 10.1093/molbev/msx197

* LM Cree et al., A Reduction of Mitochondrial DNA Molecules During Embryogenesis Explains the Rapid Segregation of Genotypes. Nat Genet. 2008 Feb;40(2):249-254. PMID: 18223651, DOI: https://doi.org/10.1038/ng.2007.63

* VI Floros et al., Segregation of Mitochondrial DNA Heteroplasmy Through a Developmental Genetic Bottleneck in Human Embryos. Nat Cell Biol. 2018 Feb;20(2):144-151. PMID: 29335530, DOI: 10.1038/s41556-017-0017-8

C Freyer et al., Variation in Germline mtDNA Heteroplasmy Is Determined Prenatally but Modified During Subsequent Transmission. Nat Genet. 2012 Nov;44(11):1282-1285. PMID: 23042113, DOI: 10.1038/ng.2427

MW Gray, Lynn Margulis and the endosymbiont hypothesis: 50 years later. Mol Biol Cell. 2017 May 15;28(10):1285-1287. PMID: 28495966, DOI: 10.1091/mbc.E16-07-0509 [Review]

* W Lee et al., Molecular basis for maternal inheritance of human mitochondrial DNA. Nat Genet. 2023 Oct;55(10):1632-1639. PMID: 37723262, DOI: 10.1038/s41588-023-01505-9

M Li et al., Transmission of human mtDNA heteroplasmy in the Genome of the Netherlands families: support for a variable-size bottleneck. Genome Res. 2016 Apr;26(4):417-426. PMID: 26916109, DOI: 10.1101/gr.203216.115

J van der Reest, G Nardini Cecchino, MC Haigis, P Kordowitzki, Mitochondria: Their relevance during oocyte ageing. Ageing Res Rev. 2021 Sep;70:101378. PMID: 34091076, DOI: 10.1016/j.arr.2021.101378 [Review]

W-H Song, J William, O Ballard, Y-J Yi, P Sutovsky, Regulation of Mitochondrial Genome Inheritance by Autophagy and Ubiquitin-Proteasome System: Implications for Health, Fitness, and Fertility. Biomed Res Int. 2014;2014:981867. PMID: 25028670, DOI: 10.1155/2014/981867 [Review]

* JB Stewart, PF Chinnery, The Dynamics of Mitochondrial DNA Heteroplasmy: Implications for Human Health and Disease. Nat Rev Genet. 2015 Sep;16(9):530-542. PMID: 26281784, DOI: 10.1038/nrg3966 [Review]

* H Zhang, SP Burr, PF Chinnery, The Mitochondrial DNA Genetic Bottleneck: Inheritance and Beyond. Essays Biochem. 2018 Jul 20;62(3):225-234. PMID: 29880721, DOI: 10.1042/EBC20170096 [Review]

Chapter 15

MD Anway, AS Cupp, M Uzumcu, MK Skinner, Epigenetic transgenerational actions of endocrine disruptors and male fertility. Science 2005 Jun 3;308(5727):1466-1469. PMID: 15933200, DOI: 10.1126/science.1108190

D J Barker, Fetal Origins of Coronary Heart Disease. BMJ. 1995 Jul 15;311(6998):171-174. PMID: 7613432, DOI: 10.1136/bmj.311.6998.171 [Review]

JT Bell, T Spector, DNA methylation studies using twins: what are they telling us? D.Genome Biol. 2012 Oct 18;13(10):172. PMID: 23078798, DOI: 10.1186/gb-2012-13-10-172

M Ben Maamar, I Sadler-Riggleman, D Beck, MK Skinner, Epigenetic Transgenerational Inheritance of Altered Sperm Histone Retention Sites. Sci Rep. 2018 Mar 28;8(1):5308. PMID: 29593303, DOI: 10.1038/s41598-018-23612-y

ME Blewitt et al., Dynamic reprogramming of DNA methylation at an epigenetically sensitive allele in mice. PLoS Genet. 2006 Apr;2(4):e49. PMID: 16604157, DOI: 10.1371/journal.pgen.0020049

B Chen et al., Maternal inheritance of glucose intolerance via oocyte TET3 insufficiency . Nature 2022 May;605(7911):761-766. PMID: 35585240, DOI: 10.1038/s41586-022-04756-4

* Q Chen, W Yan, E Duan, Epigenetic Inheritance of Acquired Traits Through Sperm RNAs and Sperm RNA Modifications. Nat Rev Genet. 2016 Dec;17(12):733-743. PMID: 27694809, DOI: 10.1038/nrg.2016.106 [Review]

* HJ Clarke, K-F Vieux, Epigenetic Inheritance Through the Female Germ-Line: The Known, the Unknown, and the Possible. Semin Cell Dev Biol. 2015 Jul;43:106-116. PMID: 26183189, DOI: 10.1016/j.semcdb.2015.07.003 [Review]

L Daxinger, E Whitelaw, Understanding Transgenerational Epigenetic Inheritance via the Gametes in Mammals. Nat Rev Genet. 2012 Jan 31;13(3):153-162. PMID: 22290458, DOI: 10.1038/nrg3188 [Review]

BG Dias, KJ Ressler, Parental Olfactory Experience Influences Behavior and Neural Structure in Subsequent Generations. Nat Neurosci. 2014 Jan;17(1):89-96. PMID: 24292232, DOI: 10.1038/nn.3594

J van Dongen et al., Epigenome-Wide Association Study of Aggressive Behavior. Twin Res Hum Genet. 2015 Dec;18(6):686-698 . PMID: 26508088, DOI: 10.1017/thg.2015.74

YE Dubrova et al., Transgenerational Mutation by Radiation. Nature 2000 May 4;405(6782):37. PMID: 10811208, DOI: 10.1038/35011135

S Erkek et al., Molecular determinants of nucleosome retention at CpG-rich sequences in mouse spermatozoa . Nat Struct Mol Biol. 2013 Jul;20(7):868-875 . PMID: 23770822, DOI: 10.1038/nsmb.2599

K Gapp et al., Implication of Sperm RNAs in Transgenerational Inheritance of the Effects of Early Trauma in Mice. Nat Neurosci. 2014 May;17(5):667-669. PMID: 24728267, DOI: 10.1038/nn.3695

L Gaspa-Toneu, AHFM Peters, Nucleosomes in mammalian sperm: conveying paternal epigenetic inheritance or subject to reprogramming between generations? Review. Curr Opin Genet Dev. 2023 Apr:79:102034. PMID: 36893482, DOI: 10.1016/j.gde.2023.102034

P A Ginno et al., A genome-scale map of DNA methylation turnover identifies site-specific dependencies of DNMT and TET activity. Nat Commun. 2020 May 29;11(1):2680. PMID: 32471981, DOI: 10.1038/s41467-020-16354-x

R Goyal, R Reinhardt, A Jeltsch, Accuracy of DNA methylation pattern preservation by the Dnmt1 methyltransferase. Nucleic Acids Res. 2006 Feb 25;34(4):1182-1188. PMID: 16500889, DOI: 10.1093/nar/gkl002

E Heard, RA Martienssen, Transgenerational epigenetic inheritance: myths and mechanisms. Cell 2014 Mar 27;157(1):95-109. PMID: 24679529, DOI: 10.1016/j.cell.2014.02.045 [Review]

BT Heijmans et al., Persistent epigenetic differences associated with prenatal exposure to famine in humans. Proc Natl Acad Sci U S A. 2008 Nov 4;105(44):17046-17049. PMID: 18955703, DOI: 10.1073/pnas.0806560105

I Hernando-Herraez, R Garcia-Perez, AJ Sharp, T Marques-Bonet, DNA Methylation: Insights Into Human Evolution. PLoS Genet. 2015 Dec 10;11(12):e1005661. PMID: 26658498, DOI: 10.1371/journal.pgen.1005661 [Review]

* https://en.wikipedia.org/wiki/Jean-Baptiste_Lamarck

https://en.wikipedia.org/wiki/Modern_synthesis_(20th_century)

* https://en.wikipedia.org/wiki/Trofim_Lysenko

RF Hillary et al., Genome and epigenome wide studies of neurological protein biomarkers in the Lothian Birth Cohort 1936. Nat Commun. 2019 Jul 18;10(1):3160. PMID: 31320639, DOI: 10.1038/s41467-019-11177-x

ML Holland et al., Early-life Nutrition Modulates the Epigenetic State of Specific rDNA Genetic Variants in Mice. Science 2016 Jul 29;353(6298):495-498. PMID: 27386920, DOI: 10.1126/science.aaf7040

M Hua et al., Identification of small non-coding RNAs as sperm quality biomarkers for in vitro fertilization. Cell Discov. 2019 Apr 9;5:20. PMID: 30992999, DOI: 10.1038/s41421-019-0087-9

* E Jablonka, The Evolutionary Implications of Epigenetic Inheritance Interface Focus 2017 Oct 6;7(5):20160135. PMID: 28839916, DOI: 10.1098/rsfs.2016.0135 [Review]

*#45]RL Jirtle, MK Skinner, Environmental Epigenomics and Disease Susceptibility. Nat Rev Genet. 2007 Apr;8(4):253-262. PMID: 17363974, DOI: 10.1038/nrg2045 [Review]

N Katzmarski et al., Transmission of trained immunity and heterologous resistance to infections across generations. Nat Immunol. 2021;22:1382-1390. PMID: 34663978, DOI: 10.1038/s41590-021-01052-7

A Kazachenka et al., Identification, Characterization, and Heritability of Murine Metastable Epialleles: Implications for Non-genetic Inheritance. Cell 2018 Nov 15;175(5):1259-1271.e13. PMID: 30454646, DOI: 10.1016/j.cell.2018.09.043

NJ Kessler, RA Waterland, AM Prentice, MJ Silver, Establishment of environmentally sensitive DNA methylation states in the very early human embryo. Sci Adv. 2018 Jul 11;4(7):eaat2624. PMID: 30009262, DOI: 10.1126/sciadv.aat2624

SE King, MK Skinner, Epigenetic Transgenerational Inheritance of Obesity Susceptibility. Trends Endocrinol Metab. 2020 Jul;31(7):478-494. PMID: 32521235, DOI: 10.1016/j.tem.2020.02.009 [Review]

D Kubsad et al., Assessment of Glyphosate Induced Epigenetic Transgenerational Inheritance of Pathologies and Sperm Epimutations: Generational Toxicology. Sci Rep. 2019 Apr 23;9(1):6372. PMID: 31011160, DOI: 10.1038/s41598-019-42860-0

K Kvaløy, CM Page, T L Holmen, Epigenome-wide methylation differences in a group of lean and obese women - A HUNT Study. Sci Rep. 2018 Nov 5;8(1):16330. PMID: 30397228, DOI: 10.1038/s41598-018-34003-8

* L Ly, D Chan, JM Trasler, Developmental Windows of Susceptibility for Epigenetic Inheritance Through the Male Germline. Semin Cell Dev Biol. 2015 Jul;43:96-105. PMID:26265561, DOI: 10.1016/j.semcdb.2015.07.006 [Review]

KB Michels et al., Recommendations for the design and analysis of epigenome-wide association studies. Nat Methods 2013 Oct;10(10):949-955. PMID: 24076989, DOI: 10.1038/nmeth.2632 [Review]

* EA Miska, AC Ferguson-Smith, Transgenerational Inheritance: Models and Mechanisms of non-DNA Sequence-Based Inheritance. Science 2016 Oct 7;354(6308):59-63. PMID: 27846492, DOI: 10.1126/science.aaf4945 [Review]

* D Nätt et al., Human sperm displays rapid responses to diet. PLoS Biol. 2019 Dec 26;17(12):e3000559. PMID: 31877125, DOI: 10.1371/journal.pbio.3000559

RC Painter et al., Transgenerational effects of prenatal exposure to the Dutch famine on neonatal adiposity and health in later life. BJOG 2008 Sep;115(10):1243-1249. PMID: 18715409, DOI: 10.1111/j.1471-0528.2008.01822.x

ME Pembrey et al., Sex-specific, Male-Line Transgenerational Responses in Humans. Eur J Hum Genet. 2006 Feb;14(2):159-166. PMID: 16391557, DOI: 10.1038/sj.ejhg.5201538 [Comparative Study]

* M Pembrey et al., Human Transgenerational Responses to Early-Life Experience: Potential Impact on Development, Health and Biomedical Research. J Med Genet. 2014 Sep;51(9):563-572. PMID: 25062846, DOI: 10.1136/jmedgenet-2014-102577 [Review]

OJ Rando, KJ Verstrepen, Timescales of Genetic and Epigenetic Inheritance. Cell 2007 Feb 23;128(4):655-668. PMID: 17320504, DOI: 10.1016/j.cell.2007.01.023 [Review]

M Rassoulzadegan et al., RNA-mediated Non-Mendelian Inheritance of an Epigenetic Change in the Mouse. Nature 2006 May 25;441(7092):469-474. PMID: 16724059, DOI: 10.1038/nature04674

*#44 VM Sales, AC Ferguson-Smith, M-E Patti, Epigenetic Mechanisms of Transmission of Metabolic Disease across Generations. Cell Metab. 2017 Mar 7;25(3):559-571. PMID: 28273478, DOI: 10.1016/j.cmet.2017.02.016 [Review]

A Salas-Huetos et al., The role of miRNAs in male human reproduction: a systematic review. Andrology 2020 Jan;8(1):7-26. PMID: 31578810, DOI: 10.1111/andr.12714

* D Schübeler, Function and Information Content of DNA Methylation. Nature 2015 Jan 15;517(7534):321-326. PMID: 25592537, DOI: 10.1038/nature14192 [Review]

U Sharma et al., Biogenesis and Function of tRNA Fragments During Sperm Maturation and Fertilization in Mammals. Science 2016 Jan 22;351(6271):391-396. PMID: 26721685, DOI: 10.1126/science.aad6780

K Siklenka et al., Disruption of Histone Methylation in Developing Sperm Impairs Offspring Health Transgenerationally. Science 2015 Nov 6;350(6261):aab2006. PMID: 26449473, DOI: 10.1126/science.aab2006

* AL Teh et al, The Effect of Genotype and in Utero Environment on Interindividual Variation in Neonate DNA Methylomes. Genome Res. 2014 Jul;24(7):1064-1074. PMID: 24709820, DOI: 10.1101/gr.171439.113

* EW Tobi et al., DNA Methylation Signatures Link Prenatal Famine Exposure to Growth and Metabolism. Nat Commun. 2014 Nov 26;5:5592. PMID: 25424739, DOI: 10.1038/ncomms6592

S Tyebji, AJ Hannan, CJ Tonkin, Pathogenic Infection in Male Mice Changes Sperm Small RNA Profiles and Transgenerationally Alters Offspring Behavior. Cell Rep. 2020 Apr 28;31(4):107573. PMID: 32348768, DOI: 10.1016/j.celrep.2020.107573

D Vågerö, PR Pinger, V Aronsson, GJ van den Berg, Paternal grandfather's access to food predicts all-cause and cancer mortality in grandsons. Nat Commun. 2018 Dec 11;9(1):5124. PMID: 30538239, DOI: 10.1038/s41467-018-07617-9

MVE Veenendaal et al.,Transgenerational effects of prenatal exposure to the 1944-45 Dutch famine. BJOG 2013 Apr;120(5):548-553. PMID: 23346894, DOI: 10.1111/1471-0528.12136

Y Wang et al., Sperm microRNAs confer depression susceptibility to offspring. Sci Adv. 2021 Feb 10;7(7):eabd7605. PMID: 33568480, DOI: 10.1126/sciadv.abd7605

E Whitelaw, DI Martin, Retrotransposons as Epigenetic Mediators of Phenotypic Variation in Mammals. Nat Genet. 2001 Apr;27(4):361-365. PMID: 11279513, DOI: 10.1038/86850

Z Xiong et al., EWAS Data Hub: a resource of DNA methylation array data and metadata. Nucleic Acids Res. 2020 Jan 8;48(D1):D890-D895. PMID: 31584095, DOI: 10.1093/nar/gkz840

Y Zhang et al., Dnmt2 mediates intergenerational transmission of paternally acquired metabolic disorders through sperm small non-coding RNAs. Nat Cell Biol. 2018 May;20(5):535-540. PMID: 29695786, DOI: 10.1038/s41556-018-0087-2

* Y Zhang et al., Sperm RNA Code Programmes the Metabolic Health of Offspring. Nat Rev Endocrinol. 2019 Aug;15(8):489-498. PMID: 31235802, DOI: 10.1038/s41574-019-0226-2 [Review]

Chapter 16

* E Aizawa et al., Epigenetic regulation limits competence of pluripotent stem cell-derived oocytes. EMBO J. 2023 Dec 1;42(23):e113955. PMID: 37850882, DOI: 10.15252/embj.2023113955

KI Aston et al., Aberrant Sperm DNA Methylation Predicts Male Fertility Status and Embryo Quality. Fertil Steril. 2015 Dec;104(6):1388-97.e1-5. PMID: 26361204, DOI: 10.1016/j.fertnstert.2015.08.019

EB Baart et al., Preimplantation Genetic Screening Reveals a High Incidence of Aneuploidy and Mosaicism in Embryos From Young Women Undergoing IVF. Hum Reprod. 2006 Jan;21(1):223-233. PMID: 16155075, DOI: 10.1093/humrep/dei291

H Bakker, JH Wallinga, RD Politiek, Reproduction and body weight of mice after long term selection for large litter size. Journal of Animal Science 1978; 46(6):1572-1580. DOI: https://doi.org/10.2527/jas1978.4661572x

J Barberet et al., Can Novel Early Non-Invasive Biomarkers of Embryo Quality Be Identified With Time-Lapse Imaging to Predict Live Birth? Hum Reprod. 2019 Aug 1;34(8):1439-1449. PMID: 31287145, DOI: 10.1093/humrep/dez085

MA Bellis, K Hughes, S Hughes, JR Ashton, Measuring paternal discrepancy and its public health consequences. J Epidemiol Community Health 2005 Sep;59(9):749-754. PMID: 16100312, DOI: 10.1136/jech.2005.036517 [Review]

P de Boer, M de Vries, Are There Benefits From Having Two Genetic Fathers? Biol Reprod. 2011 Mar;84(3):409-411. PMID: 21178174, DOI: 10.1095/biolreprod.110.090472 [Comment]

* H Bolton et al., Mouse Model of Chromosome Mosaicism Reveals Lineage-Specific Depletion of Aneuploid Cells and Normal Developmental Potential. Nat Commun. 2016 Mar 29;7:11165. PMID: 27021558, DOI: 10.1038/ncomms11165

* M Bonduelle et al., Prenatal testing in ICSI pregnancies: incidence of chromosomal anomalies in 1586 karyotypes and relation to sperm parameters. Hum Reprod. 2002 Oct;17(10):2600-2614. PMID: 12351536, DOI: 10.1093/humrep/17.10.2600

S Bontekoe et al., Low Oxygen Concentrations for Embryo Culture in Assisted Reproductive Technologies. Cochrane Database Syst Rev. 2012 Jul 11;(7):CD008950. PMID: 22786519, DOI: 10.1002/14651858.CD008950.pub2 [Review]

PS Burgoyne, A Y-chromosomal effect on blastocyst cell number in mice. Development 1993 Jan;117(1):341-345. PMID: 8223257, DOI: 10.1242/dev.117.1.341

A Capalbo et al., On the reproductive capabilities of aneuploid human preimplantation embryos. Am J Hum Genet. 2022 Sep 1;109(9):1572-1581. PMID: 36055209, DOI: 10.1016/j.ajhg.2022.07.009 [Review]

SR Catford, RI McLachlan, MK O'Bryan, JL Halliday, Long-term follow-up of intra-cytoplasmic sperm injection-conceived offspring compared with in vitro fertilization-conceived offspring: a systematic review of health outcomes beyond the neonatal period. Andrology 2017 Jul;5(4):610-621. PMID: 28632930, DOI: 10.1111/andr.12369 [Review]

SR Catford, RI McLachlan, MK O'Bryan, JL Halliday, Long-term follow-up of ICSI-conceived offspring compared with spontaneously conceived offspring: a systematic review of health outcomes beyond the neonatal period. Andrology 2018 Sep;6(5):635-653. PMID: 30296010, DOI: 10.1111/andr.12526

M Ceelen et al., Cardiometabolic Differences in Children Born After in Vitro Fertilization: Follow-Up Study. J Clin Endocrinol Metab. 2008 May;93(5):1682-1688. PMID: 18285409, DOI: 10.1210/jc.2007-2432

M Ceelen et al., Growth During Infancy and Early Childhood in Relation to Blood Pressure and Body Fat Measures at Age 8-18 Years of IVF Children and Spontaneously Conceived Controls Born to Subfertile Parents. Hum Reprod. 2009 Nov;24(11):2788-2795. PMID: 19648588, DOI: 10.1093/humrep/dep273

* S Choufani et al., Impact of Assisted Reproduction, Infertility, Sex and Paternal Factors on the Placental DNA Methylome. Hum Mol Genet. 2019 Feb 1;28(3):372-385. PMID: 30239726, DOI: 10.1093/hmg/ddy321

A Claes, TAE Stout, Success rate in a clinical equine in vitro embryo production program. Theriogenology 2022 Jul 15:187:215-218. PMID: 35623226, DOI: 10.1016/j.theriogenology.2022.04.019

* S Cornelisse, Preimplantation genetic testing for aneuploidies (abnormal number of chromosomes) in in vitro fertilisation. Cochrane Database Syst Rev. 2020 Sep 8;9:CD005291. PMID: 32898291, DOI: 10.1002/14651858.CD005291.pub3 [Meta-Analysis]

G Corona et al., Sperm recovery and ICSI outcomes in men with non-obstructive azoospermia: a systematic review and meta-analysis. Hum Reprod Update 2019 Nov 5;25(6):733-757. PMID: 31665451, DOI: 10.1093/humupd/dmz028 [Meta-Analysis]

G Coticchio et al., Focused time-lapse analysis reveals novel aspects of human fertilization and suggests new parameters of embryo viability. Hum Reprod. 2018 Jan 1;33(1):23-31. PMID: 29149327, DOI: 10.1093/humrep/dex344

DR Cox, SA Smith, LB Epstein, CJ Epstein, Mouse Trisomy 16 as an Animal Model of Human Trisomy 21 (Down Syndrome): Production of Viable Trisomy 16 Diploid Mouse Chimeras. Dev Biol. 1984 Feb;101(2):416-424. PMID: 6229437, DOI: 10.1016/0012-1606(84)90156-8

J Deng et al., The impact of culture conditions on blastocyst formation and aneuploidy rates: a comparison between single-step and sequential media in a large academic practice. J Assist Reprod Genet. 2020 Jan;37(1):161-169. PMID: 31950455, DOI: 10.1007/s10815-019-01621-8 [Comparative Study]

J Donckers, JLH Evers, JA Land, The long-term outcome of 946 consecutive couples visiting a fertility clinic in 2001-2003. Fertil Steril. 2011 Jul;96(1):160-164. PMID: 21550038, DOI: 10.1016/j.fertnstert.2011.04.019 [Comparative Study]

AJCM van Dongen et al., Reasons for dropping out from a waiting list for in vitro fertilization. Fertil Steril. 2010 Oct;94(5):1713-1716 . PMID: 19896662, DOI: 10.1016/j.fertnstert.2009.08.066

J van Echten-Arends et al., Chromosomal Mosaicism in Human Preimplantation Embryos: A Systematic Review. Hum Reprod Update Sep-Oct 2011;17(5):620-627. PMID: 21531753, DOI: 10.1093/humupd/dmr014 [Review]

RG Edwards, BD Bavister, PC Steptoe, Early Stages of Fertilization in Vitro of Human Oocytes Matured in Vitro. Nature 1969 Feb 15;221(5181):632-635. PMID: 4886881, DOI: 10.1038/221632a0

RG Edwards, PC Steptoe, J M Purdy, Fertilization and Cleavage in Vitro of Preovulator Human Oocytes. Nature 1970 Sep 26;227(5265):1307-1309. PMID: 4916973, DOI: 10.1038/2271307a0

CJ Epstein et al., Production of Viable Adult Trisomy 17 Reversible Diploid Mouse Chimeras. Proc Natl Acad Sci U S A. 1982 Jul;79(14):4376-4380. PMID: 6956868, DOI: 10.1073/pnas.79.14.4376

F Federici et al., Round Spermatid Injection Rescues Female Lethality of a Paternally Inherited Xist Deletion in Mouse. PLoS Genet. 2016 Oct 7;12(10):e1006358. PMID: 27716834, DOI: 10.1371/journal.pgen.1006358

* JM Franasiak et al., The nature of aneuploidy with increasing age of the female partner: a review of 15,169 consecutive trophectoderm biopsies evaluated with comprehensive chromosomal screening. Fertil Steril. 2014 Mar;101(3):656-663.e1. PMID: 24355045, DOI: 10.1016/j.fertnstert.2013.11.004 [Review]

L Gaspa-Toneu, AHFM Peters, Nucleosomes in mammalian sperm: conveying paternal epigenetic inheritance or subject to reprogramming between generations?. Curr Opin Genet Dev. 2023 Apr:79:102034. PMID: 36893482, DOI: 10.1016/j.gde.2023.102034 [Review]

D Glujovsky, D Blake, C Farquhar, A Bardach, Cleavage Stage Versus Blastocyst Stage Embryo Transfer in Assisted Reproductive Technology. Cochrane Database Syst Rev. 2012 Jul 11;(7):CD002118. PMID: 22786480, DOI: 10.1002/14651858.CD002118.pub4 [Review]

* D Glujovsky et al., Cleavage Stage Versus Blastocyst Stage Embryo Transfer in Assisted Reproductive Technology. Cochrane Database Syst Rev. 2016 Jun 30;(6):CD002118. PMID: 27357126, DOI: 10.1002/14651858.CD002118.pub5 [Review]

* K Hayashi, In Vitro Reconstitution of Germ Cell development. Biol Reprod. 2019 Sep 1;101(3):567-578. PMID: 31295346, DOI: 10.1093/biolre/ioz111

AT Hertig, J Rock, EC Adams, A Description of 34 Human Ova Within the First 17 Days of Development. Am J Anat. 1956 May;98(3):435-493. PMID: 13362122, DOI: 10.1002/aja.1000980306

* O Hikabe et al., Reconstitution in Vitro of the Entire Cycle of the Mouse Female Germ Line. Nature 2016 Nov 10;539(7628):299-303. PMID: 27750280, DOI: 10.1038/nature20104

https://www.proteinatlas.org/humanproteome/tissue/

https://psycnet.apa.org/record/2014-10209-013

https://strathprints.strath.ac.uk/72556/1/Jones_etal_TCS_2018_Ovulation_sex_hormones_and_womens_mating_psychology.pdf

Y Ishikura et al., In vitro reconstitution of the whole male germ-cell development from mouse pluripotent stem cells. Cell Stem Cell 2021 Sep 1;S1934-5909(21)00342-8. PMID: 34496297, DOI: 10.1016/j.stem.2021.08.005

* K Jacobs et al., Mitotic Spindle Disruption in Human Preimplantation Embryos Activates the Spindle Assembly Checkpoint but Not Apoptosis Until Day 5 of Development. Mol Hum Reprod. 2017 May 1;23(5):321-329. PMID: 28159965, DOI: 10.1093/molehr/gax007

TG Jenkins et al., Age-associated Sperm DNA Methylation Alterations: Possible Implications in Offspring Disease Susceptibility. PLoS Genet. 2014 Jul 10;10(7):e1004458. PMID: 25010591, DOI: 10.1371/journal.pgen.1004458

* DL Keefe, Telomeres and Genomic Instability During Early Development. Eur J Med Genet. 2020 Feb;63(2):103638. PMID: 30862510, DOI: 10.1016/j.ejmg.2019.03.002 [Review]

DC Kieslinger et al., Clinical outcomes of uninterrupted embryo culture with or without time-lapse-based embryo selection versus interrupted standard culture (SelecTIMO): a three-armed, multicentre, double-blind, randomised controlled trial. Lancet 2023 Apr 29;401(10386):1438-1446. PMID: 37004670, DOI: 10.1016/S0140-6736(23)00168-X [Randomized Controlled Trial]

* SHM Kleijkers et al., Influence of Embryo Culture Medium (G5 and HTF) on Pregnancy and Perinatal Outcome After IVF: A Multicenter RCT. Hum Reprod. 2016 Oct;31(10):2219-2230. PMID: 27554441, DOI: 10.1093/humrep/dew156 [Randomized Controlled Trial]

* CB Lambalk, M Van Wely, K Kirkegaard, C De Geyter, Ethics Beyond Ethics. Hum Reprod. 2020 Jan 1;35(1):1-2. PMID: 31886856, DOI: 10.1093/humrep/dez250

T Lappalainen, JM Greally, Associating Cellular Epigenetic Models With Human Phenotypes. Nat Rev Genet. 2017 Jul;18(7):441-451. PMID: 28555657, DOI: 10.1038/nrg.2017.32 [Review]

S Laurentino, J Borgmann, J Gromoll, On the Origin of Sperm Epigenetic Heterogeneity. Reproduction 2016 May;151(5):R71-8. PMID: 26884419, DOI: 10.1530/REP-15-0436 [Review]

M Lek et al., Analysis of Protein-Coding Genetic Variation in 60,706 Humans. Nature 2016 Aug 18;536(7616):285-291. PMID: 27535533, DOI: 10.1038/nature19057

B Luke et al., The risk of birth defects with conception by ART. Hum Reprod. 2021 Jan 1;36(1):116-129. PMID: 33251542, DOI: 10.1093/humrep/deaa272

J Mertens et al., Children born after assisted reproduction more commonly carry a mitochondrial genotype associating with low birthweight. Nat Commun. 2024 Feb 9;15(1):1232. PMID: 38336715, DOI: 10.1038/s41467-024-45446-1

APA van Montfoort et al., Assisted Reproduction Treatment and Epigenetic Inheritance. Hum Reprod Update Mar-Apr 2012;18(2):171-197. PMID: 22267841, DOI: 10.1093/humupd/dmr047 [Review]

* APA Van Montfoort et al., Reduced oxygen concentration during human IVF culture improves embryo utilization and cumulative pregnancy rates per cycle. Hum Reprod Open 2020 Jan 22;2020(1):hoz036. PMID: 31989044, DOI: 10.1093/hropen/hoz036

HD Morgan et al., The Culture of Zygotes to the Blastocyst Stage Changes the Postnatal Expression of an Epigentically Labile Allele, Agouti Viable Yellow, in Mice. Biol Reprod. 2008 Oct;79(4):618-623. PMID: 18562706, DOI: 10.1095/biolreprod.108.068213

CL Mulder et al., Comparison of DNA Methylation Patterns of Parentally Imprinted Genes in Placenta Derived From IVF Conceptions in Two Different Culture Media. Hum Reprod. 2020 Mar 27;35(3):516-528. PMID: 32222762, DOI: 10.1093/humrep/deaa004

S Munné et al., Detailed Investigation Into the Cytogenetic Constitution and Pregnancy Outcome of Replacing Mosaic Blastocysts Detected With the Use of High-Resolution Next-Generation Sequencing. Fertil Steril. 2017 Jul;108(1):62-71.e8. PMID: 28579407, DOI: 10.1016/j.fertnstert.2017.05.002

* S Munné et al., Clinical outcomes after the transfer of blastocysts characterized as mosaic by high resolution Next Generation Sequencing- further insights. Eur J Med Genet. 2020 Feb;63(2):103741. PMID: 31445143, DOI: 10.1016/j.ejmg.2019.103741

* S Munné et al., First PGT-A Using Human in Vivo Blastocysts Recovered by Uterine Lavage: Comparison With Matched IVF Embryo controls. Hum Reprod. 2020 Jan 1;35(1):70-80. PMID: 31886877, DOI: 10.1093/humrep/dez242

Y Murase et al., *In vitro* reconstitution of epigenetic reprogramming in the human germ line. Nature (2024). PMID 38768632, DOI: 10.1038/s41586-024-07526-6

* KK Niakan et al., Human Pre-Implantation Embryo Development. Development 2012 Mar;139(5):829-841. PMID: 22318624, DOI: 10.1242/dev.060426 [Review]

A Noblanc et al., Epididymis Response Partly Compensates for Spermatozoa Oxidative Defects in snGPx4 and GPx5 Double Mutant Mice. PLoS One 2012;7(6):e38565. PMID: 22719900, DOI: 10.1371/journal.pone.0038565

* B Novakovic et al., Assisted Reproductive Technologies Are Associated With Limited Epigenetic Variation at Birth That Largely Resolves by Adulthood. Nat Commun. 2019 Sep 2;10(1):3922. PMID: 31477727, DOI: 10.1038/s41467-019-11929-9

SB Olsen, RE Magenis, Preferential paternal origin of de novo structural chromosome arrangements. In: *The Cytogenetics of Mammalian Autosomal Rearrangements 583-599*. Alan Liss Inc, 1988

G Palermo, H Joris, P Devroey, AC Van Steirteghem, Pregnancies After Intracytoplasmic Injection of Single Spermatozoon Into an Oocyte. Lancet 1992 Jul 4;340(8810):17-18. PMID: 1351601, DOI: 10.1016/0140-6736(92)92425-f

M Popovic et al., Chromosomal Mosaicism in Human Blastocysts: The Ultimate Challenge of Preimplantation Genetic Testing? Hum Reprod. 2018 Jul 1;33(7):1342-1354. PMID: 29796631, DOI: 10.1093/humrep/dey106

W Pryll, Kohabitionstermin und kindergeschlecht. Muenchener Medizinische Wochenschrift 1916;45:1579-1582.

J Rock, AT Hertig, The Human Conceptus During the First Two Weeks of Gestation. Am J Obstet Gynecol. 1948 Jan;55(1):6-17. PMID: 18918949, DOI: 10.1016/0002-9378(48)90152-5

* M Saitou, K Hayashi, Mammalian in vitro gametogenesis. Science 2021 Oct;374(6563):eaaz6830. PMID: 34591639, DOI: 10.1126/science.aaz6830 [Review]

I Schneider et al., Dual spindles assemble in bovine zygotes despite the presence of paternal centrosomes. J Cell Biol. 2021 Nov 1;220(11):e202010106. PMID: 34550316, DOI: 10.1083/jcb.202010106

K Servick, “Unsettled Questions Trail IVF’s Success.” Science 2014;345(6198):744-746. https://www.jstor.org/stable/24917156

Asian J Androl. 2017 Jan-Feb;19(1):80-90. PMID: 27345006, DOI: 10.4103/1008-682X.182822

* Q-Q Sha et al., Dynamics and clinical relevance of maternal mRNA clearance during the oocyte-to-embryo transition in humans. Nat Commun. 2020 Oct 1;11(1):4917. PMID: 33004802, DOI: 10.1038/s41467-020-18680-6

RM Smits et al., De novo mutations in children born after medical assisted reproduction. Hum Reprod 2022 May 30;37(6):1360-1369. PMID: 35413117, DOI: 10.1093/humrep/deac068 [Observational Study]

J Speksnijder et al., Living in a neighbourhood with an average low socioeconomic status is associated with lower cumulative ongoing pregnancy rate after in vitro fertilization treatment. Reproductive BioMedicine Online, 2024, DOI: 10.1016/j.rbmo.2024.103908

A Van Steirteghem, M Bonduelle, P Devroey, I Liebaers, Follow-up of Children Born After ICSI. Hum Reprod Update Mar-Apr 2002;8(2):111-116. PMID: 12099626, DOI: 10.1093/humupd/8.2.111 [Review]

TAE Stout, Clinical Application of in Vitro Embryo Production in the Horse. J Equine Vet Sci. 2020 Jun;89:103011. PMID: 32563449, DOI: 10.1016/j.jevs.2020.103011

MC Summers, JD Biggers, Chemically Defined Media and the Culture of Mammalian Preimplantation Embryos: Historical Perspective and Current Issues. Hum Reprod Update Nov-Dec 2003;9(6):557-582. PMID: 14714592, DOI: 10.1093/humupd/dmg039 [Review]

PR Supramaniam et al., Secondary sex ratio in assisted reproduction: an analysis of 1 376 454 treatment cycles performed in the UK. Hum Reprod Open 2019 Oct 5;2019(4):hoz020. PMID: 31598568, DOI: 10.1093/hropen/hoz020

A Tanaka et al., Fourteen babies born after round spermatid injection into human oocytes. Proc Natl Acad Sci U S A. 2015 Nov 24;112(47):14629-14634. PMID: 26575628, DOI: 10.1073/pnas.1517466112

A Tanaka et al., Ninety babies born after round spermatid injection into oocytes: survey of their development from fertilization to 2 years of age. Fertil Steril. 2018 Aug;110(3):443-451. PMID: 30098696, DOI: 10.1016/j.fertnstert.2018.04.033 [Clinical Trial]

*#46 A Taylor, ABC of subfertility: extent of the problem. BMJ. 2003 Aug 23;327(7412):434-436. PMID: 12933733, DOI: 10.1136/bmj.327.7412.434 [Review]

O Tšuiko et al., Genome Stability of Bovine in Vivo-Conceived Cleavage-Stage Embryos Is Higher Compared to in Vitro-Produced Embryos. Hum Reprod. 2017 Nov 1;32(11):2348-2357. PMID: 29040498, DOI: 10.1093/humrep/dex286

* E Vanneste et al., Chromosome Instability Is Common in Human Cleavage-Stage Embryos. Nat Med. 2009 May;15(5):577-583. PMID: 19396175, DOI: 10.1038/nm.1924

G Verheyen, E Crabbé, H Joris, A Van Steirteghem, Simple and Reliable Identification of the Human Round Spermatid by Inverted Phase-Contrast Microscopy. Hum Reprod. 1998 Jun;13(6):1570-1577. PMID: 9688394, DOI: 10.1093/humrep/13.6.1570

V Vloeberghs et al., How successful is TESE-ICSI in couples with non-obstructive azoospermia? Hum Reprod. 2015 Aug;30(8):1790-1796. PMID: 26082482, DOI: 10.1093/humrep/dev139

LLM Welling, DA Puts, Female adaptations to ovulation. In: VA Weekes-Shackelford, TK Shackelford (Eds), *Evolutionary psychology. Evolutionary perspectives on human sexual psychology and behavior*. Springer, 2014: 243-260. DOI: 10.1007/978-1-4939-0314-6_13

* AJ Wilcox et al., On the Frequency of Intercourse Around Ovulation: Evidence for Biological Influences. Hum Reprod. 2004 Jul;19(7):1539-1543. PMID: 15190016, DOI: 10.1093/humrep/deh305

J Yan et al., Live Birth with or without Preimplantation Genetic Testing for Aneuploidy. N Engl J Med. 2021 Nov 25;385(22):2047-2058. PMID: 34818479, DOI: 10.1056/NEJMoa2103613 [Randomized Controlled Trial]

M Yang et al., Depletion of aneuploid cells in human embryos and gastruloids. Nat Cell Biol. 2021 Apr;23(4):314-321. PMID: 33837289, DOI: 10.1038/s41556-021-00660-7

* M Zamani Esteki et al., In Vitro Fertilization Does Not Increase the Incidence of De Novo Copy Number Alterations in Fetal and Placental Lineages. Nat Med. 2019 Nov;25(11):1699-1705. PMID: 31686035, DOI: 10.1038/s41591-019-0620-2

* J Zhao, Y Yan, X Huang, Y Li, Do the Children Born After Assisted Reproductive Technology Have an Increased Risk of Birth Defects? A Systematic Review and Meta-Analysis. J Matern Fetal Neonatal Med. 2020 Jan;33(2):322-333. PMID: 30189770, DOI: 10.1080/14767058.2018.1488168 [Meta-Analysis]

AB Adewoye, SJ Lindsay, YE Dubrova, ME Hurles, The genome-wide effects of ionizing radiation on mutation induction in the mammalian germline. Nat Commun. 2015 Mar 26;6:6684. PMID: 25809527, DOI: 10.1038/ncomms7684

F Åsenius, AF Danson, SJ Marzi, DNA methylation in human sperm: a systematic review. Hum Reprod Update 2020 Nov 1;26(6):841-873. PMID: 32790874, DOI: 10.1093/humupd/dmaa025

SA Banihani, Effect of Ibuprofen on Semen Quality. Andrologia 2019 May;51(4):e13228. PMID: 30623461, DOI: 10.1111/and.13228

* MA Beal, CL Yauk, F Marchetti, From sperm to offspring: Assessing the heritable genetic consequences of paternal smoking and potential public health impacts. Mutat Res. 2017 Jul;773:26-50. PMID: 28927533, DOI: 10.1016/j.mrrev.2017.04.001

* LA Beeder, MK Samplaski, Effect of Antidepressant Medications on Semen Parameters and Male Fertility. Int J Urol. 2020 Jan;27(1):39-46. PMID: 31542895, DOI: 10.1111/iju.14111 [Review]

L Björndahl, U Kvist, Human Sperm Chromatin Stabilization: A Proposed Model Including Zinc Bridges. Mol Hum Reprod. 2010 Jan;16(1):23-29. PMID: 19933313, DOI: 10.1093/molehr/gap099 [Review]

* L Björndahl, U Kvist, A Model for the Importance of Zinc in the Dynamics of Human Sperm Chromatin Stabilization After Ejaculation in Relation to Sperm DNA Vulnerability. Syst Biol Reprod Med. 2011 Feb;57(1-2):86-92. PMID: 21204594, DOI: 10.3109/19396368.2010.516306 [Review]

BM Cattanach, AJ Crocker, Modified Genetic Response to X-irradiation of Mouse Spermatogonial Stem Cells Surviving Treatment With TEM. Mutat Res. 1980 Apr;70(2):211-220. PMID: 7374659, DOI: 10.1016/0027-5107(80)90161-x

* JR Drevet, RJ Aitken, Oxidative Damage to Sperm DNA: Attack and Defense. Adv Exp Med Biol. 2019;1166:107-117. PMID: 31301049, DOI: 10.1007/978-3-030-21664-1_7

S van den Driesche et al., Experimentally induced testicular dysgenesis syndrome originates in the masculinization programming window. JCI Insight 2017 Mar 23;2(6):e91204. PMID: 28352662, DOI: 10.1172/jci.insight.91204

* M Van Eetvelde et al., The Importance of the Periconception Period: Immediate Effects in Cattle Breeding and in Assisted Reproduction Such as Artificial Insemination and Embryo Transfer. Adv Exp Med Biol. 2017;1014:41-68. PMID: 28864984, DOI: 10.1007/978-3-319-62414-3_3 [Review]

* Fertility and Sterility, Practice Committee of the American Society for Reproductive Medicine, Smoking and Infertility: A Committee Opinion. Fertil Steril. 2018 Sep;110(4):611-618. PMID: 30196946, DOI: 10.1016/j.fertnstert.2018.06.016 [Review]

EJ Fletcher et al., Effects of Phthalate Mixtures on Ovarian Folliculogenesis and Steroidogenesis. Toxics 2022 May 16;10(5):251. PMID: 35622664, DOI: 10.3390/toxics10050251 [Review]

* P Gharagozloo et al., A Novel Antioxidant Formulation Designed to Treat Male Infertility Associated With Oxidative Stress: Promising Preclinical Evidence From Animal Models. Hum Reprod. 2016 Feb;31(2):252-262. PMID: 26732620, DOI: 10.1093/humrep/dev302

A Gorga et al., In vitro effects of glyphosate and Roundup on Sertoli cell physiology. Toxicol In Vitro 2020 Feb:62:104682. PMID: 31626902, DOI: 10.1016/j.tiv.2019.104682

R Hauser et al., Male reproductive disorders, diseases, and costs of exposure to endocrine-disrupting chemicals in the European Union. J Clin Endocrinol Metab. 2015 Apr;100(4):1267-1277. PMID: 25742517, DOI: 10.1210/jc.2014-4325

M Holtgrewe et al., Multisite de novo mutations in human offspring after paternal exposure to ionizing radiation. Sci Rep. 2018 Oct 2;8(1):14611. PMID: 30279461, DOI: 10.1038/s41598-018-33066-x

* https://www.fertilix.com/men

https://www.figo.org Welcome to FIGO, The International Federation of Gynecology and Obstetrics. Comment PFAS

M Ješeta et al., Overview of the Mechanisms of Action of Selected Bisphenols and Perfluoroalkyl Chemicals on the Male Reproductive Axes. Front Genet 2021 Sep 27:12:692897. PMID: 34646297, DOI: 10.3389/fgene.2021.692897 [Review]

J Jiang, Z Shu, L Qiu, Adverse effects and potential mechanisms of polystyrene microplastics (PS-MPs) on the blood-testis barrier. Environ Geochem Health 2024 Jun 7;46(7):238. PMID: 38849627, DOI: 10.1007/s10653-024-02033-z [Review]

J Jurewicz, E Dziewirska, M Radwan, W Hanke, Air Pollution From Natural and Anthropic Sources and Male Fertility. Reprod Biol Endocrinol. 2018 Dec 23;16(1):109. PMID: 30579357, DOI: 10.1186/s12958-018-0430-2 [Review]

K Jurkowska, EM Kratz, E Sawicka, A Piwowar, The Impact of Metalloestrogens on the Physiology of Male Reproductive Health as a Current Problem of the XXI Century. J Physiol Pharmacol. 2019 Jun;70(3):337-355. PMID: 31539881, DOI: 10.26402/jpp.2019.3.02 [Review]

KK Kesari, A Agarwal, R Henkel, Radiations and Male Fertility. Reprod Biol Endocrinol. 2018 Dec 9;16(1):118. PMID: 30445985, DOI: 10.1186/s12958-018-0431-1 [Review]

* MD Kessler et al., De novo mutations across 1,465 diverse genomes reveal mutational insights and reductions in the Amish founder population. Proc Natl Acad Sci U S A. 2020 Feb 4;117(5):2560-2569. PMID: 31964835, DOI: 10.1073/pnas.1902766117

MP Little, DT Goodhead, BA Bridges, SD Bouffler, Evidence relevant to untargeted and transgenerational effects in the offspring of irradiated parents. Mutat Res. Jul-Sep 2013;753(1):50-67. PMID: 23648355, DOI: 10.1016/j.mrrev.2013.04.001 [Review]

* SE Martenies, MJ Perry, Environmental and occupational pesticide exposure and human sperm parameters: a systematic review. Toxicology 2013 May 10;307:66-73. PMID: 23438386, DOI: 10.1016/j.tox.2013.02.005 [Review]

AM Neff, JA Flaws, The effects of plasticizers on the ovary. Curr Opin Endocr Metab Res. 2021 Jun:18:35-47. PMID: 33997465, DOI: 10.1016/j.coemr.2021.01.004

V Otasevic et al., Reactive Oxygen, Nitrogen, and Sulfur Species in Human Male Fertility. A Crossroad of Cellular Signaling and Pathology. Biofactors 2020 Mar;46(2):206-219. PMID: 31185138, DOI: 10.1002/biof.1535 [Review]

S Patel, C Zhou, S Rattan, JA Flaws, Effects of Endocrine-Disrupting Chemicals on the Ovary. Biol Reprod. 2015 Jul;93(1):20. PMID: 26063868, DOI: 10.1095/biolreprod.115.130336 [Review]

* KS Payne, DJ Mazur, JM Hotaling, AW Pastuszak, Cannabis and Male Fertility: A Systematic Review. J Urol. 2019 Oct;202(4):674-681. PMID: 30916627, DOI: 10.1097/JU.0000000000000248 [Review]

* V Pergialiotis et al., Diabetes Mellitus and Functional Sperm Characteristics: A Meta-Analysis of Observational Studies. J Diabetes Complications 2016 Aug;30(6):1167-1176. PMID: 27107613, DOI: 10.1016/j.jdiacomp.2016.04.002 [Review]

TH Pham et al., Perinatal Exposure to Glyphosate and a Glyphosate-Based Herbicide Affect Spermatogenesis in Mice. Toxicol Sci. 2019 May 1;169(1):260-271. PMID: 30785197, DOI: 10.1093/toxsci/kfz039

CH Ramlau-Hansen et al., Is prenatal exposure to tobacco smoking a cause of poor semen quality? A follow-up study. Am J Epidemiol. 2007 Jun 15;165(12):1372-1379. PMID: 17369608, DOI: 10.1093/aje/kwm032

* E Ricci et al., Semen Quality and Alcohol Intake: A Systematic Review and Meta-Analysis. Reprod Biomed Online 2017 Jan;34(1):38-47. PMID: 28029592, DOI: 10.1016/j.rbmo.2016.09.012 [Review]

A Salas-Huetos et al., The Effect of Nutrients and Dietary Supplements on Sperm Quality Parameters: A Systematic Review and Meta-Analysis of Randomized Clinical Trials. Adv Nutr. 2018 Nov 1;9(6):833-848. PMID: 30462179, DOI: 10.1093/advances/nmy057 [Meta-Analysis]

N Sermondade et al., BMI in relation to sperm count: an updated systematic review and collaborative meta-analysis. Hum Reprod Update. May-Jun 2013;19(3):221-231. PMID: 23242914, DOI: 10.1093/humupd/dms050 [Review]

U Sharma et al., Biogenesis and function of tRNA fragments during sperm maturation and fertilization in mammals. Science 2016 Jan 22;351(6271):391-396. PMID: 26721685, DOI: 10.1126/science.aad6780

* NE Skakkebæk et al., Environmental factors in declining human fertility. Nat Rev Endocrinol. 2022;18:139-157. PMID: 34912078, DOI: 10.1038/s41574-021-00598-8 [Review]

B Sopko et al., Glyphosate Interaction with eEF1α1 Indicates Altered Protein Synthesis: Evidence for Reduced Spermatogenesis and Cytostatic Effect. ACS Omega 2021 Jun 2;6(23):14848-14857. PMID: 34151066, DOI: 10.1021/acsomega.1c00449

P Tarapore, B Ouyang, Perfluoroalkyl Chemicals and Male Reproductive Health: Do PFOA and PFOS Increase Risk for Male Infertility? Int J Environ Res Public Health 2021 Apr 5;18(7):3794. PMID: 33916482, DOI: 10.3390/ijerph18073794 [Review]

M Wang et al., The hidden threat: Unraveling the impact of microplastics on reproductive health. Sci Total Environ. 2024 Jul 20:935:173177. PMID: 38750730, DOI: 10.1016/j.scitotenv.2024.173177 Epub 2024 May 13 [Review]

M Yeager et al., Lack of transgenerational effects of ionizing radiation exposure from the Chernobyl accident. Science 2021 Apr 22;eabg2365. PMID: 33888597, DOI: 10.1126/science.abg2365

D Zamkowska, A Karwacka, J Jurewicz, M Radwan, Environmental exposure to non-persistent endocrine disrupting chemicals and semen quality: An overview of the current epidemiological evidence. Int J Occup Med Environ Health 2018 Jul 4;31(4):377-414. PMID: 30160090, DOI: 10.13075/ijomeh.1896.01195 [Review]

* Y Zhang et al., Sperm RNA Code Programmes the Metabolic Health of Offspring. Nat Rev Endocrinol. 2019 Aug;15(8):489-498. PMID: 31235802, DOI: 10.1038/s41574-019-0226-2 [Review]

Death of great tits, 2019 report of CLM Culemborg, the Netherlands

Chapter 18

R Acuna-Hidalgo, JA Veltman, A Hoischen, New insights into the generation and role of de novo mutations in health and disease. Genome Biol. 2016 Nov 28;17(1):241. PMID: 27894357, DOI: 10.1186/s13059-016-1110-1 [Review]

SE Antonarakis, Carrier screening for recessive disorders. Nat Rev Genet. 2019 Sep;20(9):549-561. PMID: 31142809, DOI: 10.1038/s41576-019-0134-2 [Review]

* B Bekaert et al., CRISPR/Cas gene editing in the human germline. Semin Cell Dev Biol. 2022 Nov:131:93-107. PMID: 35305903, DOI: 10.1016/j.semcdb.2022.03.012 [Review]

F Belva et al., Semen Quality of Young Adult ICSI Offspring: The First Results. Hum Reprod. 2016 Dec;31(12):2811-2820. PMID: 27707840, DOI: 10.1093/humrep/dew245

JR Belyeu et al., De novo structural mutation rates and gamete-of-origin biases revealed through genome sequencing of 2,396 families. Am J Hum Genet. 2021 Apr 1;108(4):597-607. PMID: 33675682, DOI: 10.1016/j.ajhg.2021.02.012

HG Brunner, Preconceptual screening for the entire population? Utrecht: De Snoo symposium, 2019

* M Claussnitzer et al., A brief history of human disease genetics. Nature 2020 Jan;577(7789):179-189. PMID: 31915397, DOI: 10.1038/s41586-019-1879-7 [Review]

Deciphering Developmental Disorders Study, Prevalence and architecture of de novo mutations in developmental disorders. Nature 2017 Feb 23;542(7642):433-438. PMID: 28135719, DOI: 10.1038/nature21062 [Meta-Analysis]

H Fridman et al., The landscape of autosomal-recessive pathogenic variants in European populations reveals phenotype-specific effects. Am J Hum Genet. 2021 Apr 1;108(4):608-619. PMID: 33740458, DOI: 10.1016/j.ajhg.2021.03.004

C Gilissen et al., Genome sequencing identifies major causes of severe intellectual disability. Nature 2014 Jul 17;511(7509):344-347. PMID: 24896178, DOI: 10.1038/nature13394

M de Groot-van der Mooren et al., Does non-invasive prenatal testing affect the livebirth prevalence of Down syndrome in the Netherlands? A population-based register study. Prenat Diagn. 2021 Sep;41(10):1351-1359. PMID: 34176145, DOI: 10.1002/pd.6003

Heritable human genome editing, a consensus study report. International Commission on the Clinical Use of Human Germline Genome Editing, National Academy of Medicine, National Academy of Sciences, the Royal Society (UK), sept 2020

* https://www.23andme.com/en-int/

* https://www.deciphergenomics.org/ Database of genomic variants and phenotype in humans using ensemble resources

https://www.economist.com/science-and-technology/2019/11/07/modern-genetics-will-improve-health-and-usher-in-designer-children

* https://www.nap.edu/catalog/25665/heritable-human-genome-editing

J Kaplanis et al., Evidence for 28 genetic disorders discovered by combining healthcare and research data. Nature 2020 Oct;586(7831):757-762. PMID: 33057194, DOI: 10.1038/s41586-020-2832-5

* E Karavani et al., Screening Human Embryos for Polygenic Traits Has Limited Utility. Cell 2019 Nov 27;179(6):1424-1435.e8. PMID: 31761530, DOI: 10.1016/j.cell.2019.10.033

M Kozlov, The controversial embryo tests that promise a better baby. Nature 2022 Sep;609(7928):668-671. PMID: 36131039, DOI: 10.1038/d41586-022-02961-9

A Kumar et al., Whole-genome risk prediction of common diseases in human preimplantation embryos. Nat Med. 2022 Mar;28(3):513-516. PMID: 35314819, DOI: 10.1038/s41591-022-01735-0

* H Ledford, CRISPR gene editing in human embryos wreaks chromosomal mayhem. Three studies showing large DNA deletions and reshuffling heighten safety concerns about heritable genome editing. Nature 2020;583:17-18. DOI: 10.1038/d41586-020-01906-4 [news]

*^3 M Lek et al., Analysis of protein-coding genetic variation in 60,706 humans. Nature 2016 Aug 18;536(7616):285-291. PMID: 27535533, DOI: 10.1038/nature19057

T Lencz et al., Utility of polygenic embryo screening for disease depends on the selection strategy. Elife 2021 Oct 12:10:e64716. PMID: 34635206, DOI: 10.7554/eLife.64716

F Leturcq, S Tuffery-Giraud, Genetics and molecular aspects of dystrophinopathies [Article in French]. Arch Pediatr. 2015 Dec;22(12 Suppl 1):12S3-11. PMID: 26773583, DOI: 10.1016/S0929-693X(16)30002-1 [Review]

* M Lynch, Mutation and Human Exceptionalism: Our Future Genetic Load. Genetics 2016 Mar;202(3):869-875. PMID: 26953265, DOI: 10.1534/genetics.115.180471

HC Martin et al., Quantifying the contribution of recessive coding variation to developmental disorders. Science 2018 Dec 7;362(6419):1161-1164. PMID: 30409806, DOI: 10.1126/science.aar6731 [Deciphering Developmental Disorders Study]

* KRM van der Meij et al., TRIDENT-2: National Implementation of Genome-wide Non-invasive Prenatal Testing as a First-Tier Screening Test in the Netherlands. Am J Hum Genet. 2019 Dec 5;105(6):1091-1101. PMID: 31708118, DOI: 10.1016/j.ajhg.2019.10.005

E Mullaart, D Wells, Embryo biopsies for genomic selection. In: Animal Biotechnology 2, H Niemann, C Wrenzycki (Eds). Springer, 2018:81-93. DOI: 10.1007/978-3-319-92348-2_5

* D van Opstal et al., Origin and Clinical Relevance of Chromosomal Aberrations Other Than the Common Trisomies Detected by Genome-Wide NIPS: Results of the TRIDENT Study. Genet Med. 2018 Apr;20(5):480-485. PMID: 29121006, DOI: 10.1038/gim.2017.132

DB Paul, HG Spencer, The hidden science of eugenics. Nature 1995 Mar 23;374(6520):302-304. PMID: 7885463, DOI: 10.1038/374302a0

FA Reed, CF Aquadro, Mutation, Selection and the Future of Human Evolution. Trends Genet. 2006 Sep;22(9):479-484. PMID: 16857288, DOI: 10.1016/j.tig.2006.07.005 [Review]

SCEH Sallevelt et al., Diagnostic exome-based preconception carrier testing in consanguineous couples: results from the first 100 couples in clinical practice. Genet Med. 2021 Jun;23(6):1125-1136. PMID: 33742171, DOI: 10.1038/s41436-021-01116-x

* E Sheridan et al., Risk Factors for Congenital Anomaly in a Multiethnic Birth Cohort: An Analysis of the Born in Bradford Study. Lancet 2013 Oct 19;382(9901):1350-1359. PMID: 23830354, DOI: 10.1016/S0140-6736(13)61132-0

D Smedley et al.,100,000 Genomes Project Pilot Investigators; 100,000 Genomes Pilot on Rare-Disease Diagnosis in Health Care - Preliminary Report. N Engl J Med. 2021 Nov 11;385(20):1868-1880. PMID: 34758253, DOI: 10.1056/NEJMoa2035790

SC Stearns, R Hoekstra, *Evolution: An Introduction*. Oxford University Press, 2000. ISBN 9780198549680

P Turley et al., Problems with Using Polygenic Scores to Select Embryos. N Engl J Med. 2021 Jul 1;385(1):78-86. PMID: 34192436, DOI: 10.1056/NEJMsr2105065

* CF Wright et al., Genomic Diagnosis of Rare Pediatric Disease in the United Kingdom and Ireland. N Engl J Med. 2023 Apr 27;388(17):1559-1571. PMID: 37043637, DOI: 10.1056/NEJMoa2209046 [Multicenter Study]

Chapter 19

* N Anaparthy et al., Single-Cell Applications of Next-Generation Sequencing. Cold Spring Harb Perspect Med. 2019 Oct 1;9(10):a026898. PMID: 30617056, DOI: 10.1101/cshperspect.a026898 [Review]

* KH Campbell, J McWhir, WA Ritchie, I Wilmut, Sheep Cloned by Nuclear Transfer From a Cultured Cell Line. Nature 1996 Mar 7;380(6569):64-66. PMID: 8598906, DOI: 10.1038/380064a0

https://en.wikipedia.org/wiki/Christopher_Polge

https://ec.europa.eu/food/animals/welfare/other_aspects/cloning/citizens_en

* https://en.wikipedia.org/wiki/Commons

Y-H Jeong et al., Stochastic Anomaly of Methylome but Persistent SRY Hypermethylation in Disorder of Sex Development in Canine Somatic Cell Nuclear Transfer. Sci Rep. 2016 Aug 9;6:31088. PMID: 27501986, DOI: 10.1038/srep31088

Z Liao et al., Reprogramming mechanism dissection and trophoblast replacement application in monkey somatic cell nuclear transfer. Nat Commun. 2024 Jan 16;15(1):5. PMID: 38228612, DOI: 10.1038/s41467-023-43985-7

* CC Mann, *The wizard and the prophet*. Picador New York, 2019. ISBN 9781509884186

S Matoba et al., Loss of H3K27me3 Imprinting in Somatic Cell Nuclear Transfer Embryos Disrupts Post-Implantation Development. Cell Stem Cell 2018 Sep 6;23(3):343-354.e5. PMID: 30033120, DOI: 10.1016/j.stem.2018.06.008

M Naddaf, Cloned rhesus monkey lives to adulthood for the first time. Nature News 16 January 2024. https://www.nature.com/articles/d41586-024-00136-2

* S Sharma et al., Spermatogonial Stem Cells: Updates From Specification to Clinical Relevance. Hum Reprod Update 2019 May 1;25(3):275-297. PMID: 30810745, DOI: 10.1093/humupd/dmz006 [Review]

H Yao et al., scm6A-seq reveals single-cell landscapes of the dynamic m6A during oocyte maturation and early embryonic development. Nat Commun. 2023 Jan 19;14(1):315. PMID: 36658155, DOI: 10.1038/s41467-023-35958-7

D Yu et al., Silencing of Retrotransposon-Derived Imprinted Gene RTL1 Is the Main Cause for Postimplantational Failures in Mammalian Cloning. Proc Natl Acad Sci U S A. 2018 Nov 20;115(47):E11071-E11080. PMID: 30381455, DOI: 10.1073/pnas.1814514115

P Zhu et al., Single-cell DNA methylome sequencing of human preimplantation embryos. Nat Genet. 2018 Jan;50(1):12-19. PMID: 29255258, DOI: 10.1038/s41588-017-0007-6

Index

Terms that also appear in the glossary are marked with a *

Reproduction and genetics are the central themes of biology. They relate to big questions of life such as: who are we, where do we come from, and where are we going? In the combined science of reproductive genetics, generations are interconnected and fertility plays a significant role.

The cell lineage responsible for this is known as the germline, with gametes as the functional products and heralds of information. The book follows the germline and attempts, through a more liberal and narrative use of language, to bridge the gap between an academic textbook and society. Personal experiences are not shunned but do not dominate.

The book is intended for those professionally involved in or entering this field, including those in education and journalism. However, it is also suitable for enthusiasts without a professional background, partly thanks to its beautiful colour illustrations and an up-to-date glossary.

Peter de Boer graduated with a master's degree in animal husbandry from the Agricultural University, Wageningen, the Netherlands (1971). He obtained his PhD in mouse cytogenetics in 1975 and then worked as a senior staff member in the Department of Genetics at what is now Wageningen University. From 2001 to 2012, Peter was a senior researcher in the Department of Obstetrics and Gynaecology at Radboudumc in Nijmegen, the Netherlands.

Marc Maas graduated with a master's degree (Biology) from Wageningen University, the Netherlands (2000). He then completed his PhD in genetics at the same institution in 2005. After a postdoctoral period, he developed into a science illustrator. In that role, he works for various scientific institutions.

Judith Swart graduated with a master's degree in Behavioural and Cognitive Neurosciences from the University of Groningen, the Netherlands (2018). She then moved to New Zealand for several years, where she completed her PhD in reproductive neuroendocrinology at the University of Otago (2022).

ISBN 978-94-9329-652-7

9 789493 296527 >

Radboud University

www.ru.nl/radbouduniversitypress