



Age-related changes in sperm DNA methylation and their forensic and clinical implications

Lucie Kotková¹  & Jiří Drábek^{*,1} 

¹Institute of Molecular & Translational Medicine, Faculty of Medicine & Dentistry, Palacky University Olomouc and University Hospital Olomouc, 77900, Czech Republic

*Author for correspondence: jiri.drabek@upol.cz

As a link between a stable genome and a dynamic environment, epigenetics is a promising tool for mapping age-related changes in human DNA. Methylated cytosine changes at specific loci are generally less studied in sperm DNA than in somatic cell DNA. Age-related methylation changes can be connected to various reproductive health problems and multiple disorders in offspring. In addition, they can be helpful in forensic fields, where testing of specific loci in semen samples found at sexual assault crime scenes can predict a perpetrator's age and narrow down the police investigation. This review focuses on age-related methylation changes in sperm. It covers the biological role of methylation, methylation testing techniques and the implications of methylation changes in forensics and clinical practice.

Plain language summary: DNA methylation is a biological process that can change the activity of a gene without changing its sequence. We do not know much about DNA methylation in sperm and what changes methylation undergoes during the lifespan. These changes can, however, be important both for health and solving crimes. Presperm cells renew themselves, which gives rise to new sperm cells, from youth to death, accumulating cell divisions prone to error. This is why sperm cells are affected by age more than nondividing eggs. Methylation is specific in different tissues of the body. The ratio between number of sperm cells, white blood cells, and other cell types is highly variable and hardly predicted, which may distort the results. Clinical studies have shown that older fathers have worse reproductive health. Their children can develop metabolic, neurological and behavioral disorders. This also applies to younger men whose DNA methylation pattern is similar to that of older men. Methylation changes allow us to build a model capable of predicting the age of an unknown person with a mean error of about 5 years. This can be helpful for police investigators in cases of sexual assault, when biological material is found but there is no match in the police database.

Tweetable abstract: A number of methyl groups on DNA changes with age. In sperm DNA, these changes can influence the successful conception and health of offspring or help police find an unknown criminal offender by predicting their age.

First draft submitted: 29 August 2023; Accepted for publication: 20 October 2023; Published online: 21 November 2023

Keywords: age prediction • epigenome • forensic phenotyping • methylation • semen • sperm

Previously, it was believed that sperm had little or no contribution to conception beyond the safe delivery of paternal nuclear DNA to the oocyte. Recently, evidence has arisen that the sperm epigenome plays an important role in embryonic development through gene imprinting and coding-specific RNA [1,2]. Also, it was found that an older age of the father negatively affects reproductive health outcomes that were previously attributed only to an older age of the mother [3]. Although age-related changes in DNA methylation in somatic cells have been extensively studied, especially since the publication of "epigenetic clocks" by Hannum *et al.* [4] and Horvath [5], the impact of aging on the methylation profile of germ cells is much less known, and existing studies of germ cell methylation are mostly focused on known imprinted regions [2,3].

Although methylation of oocytes has only reproductive consequences, methylation of sperm can also be used forensically since ejaculate is the primary biological material used for perpetrator identification in sexual assault cases. In this review, we describe the biological processes behind DNA methylation in sperm cells in comparison with somatic cells as well as some state-of-the-art forensic applications of age estimation from sperm methylation profiles.

The role of DNA methylation

A correct methylation signal is crucial for maintaining proper gene expression control and is one of the tools for cell type specification [6]. Methylation can silence gene expression in two ways: by physically preventing the binding of transcription factors to the promoter and by participating in shaping chromatin into a form such that the DNA is not open to transcription [7]. Proper methylation and subsequent silencing of specific genes in embryonic totipotent stem cells are two of the factors sending individual cell types on their specific developmental pathways and restricting their proliferative potential [8]. The dynamic character of DNA methylation helps facilitate cellular adaptation to a changing environment [6].

Defects in proper methylation signals can lead to developmental defects and carcinogenesis, with global hypomethylation and local hypermethylation being two of the hallmarks of cancer [7]. This can promote insertional mutagenesis by dysregulation of retrotransposons or tumorigenesis by the overexpression of oncogenes and silencing of tumor suppressors. In sperm cells, an aberrant DNA methylation signal can result in improper chromatin condensation in the sperm head and abnormal gene expression, leading to infertility [9].

Development of sperm DNA methylation signals

Although oocytes are in a state of meiotic arrest, spermatogonial cells have a self-renewing potential and divide continuously, giving rise to differentiating daughter cells. Until the beginning of puberty, they undergo approximately 35 divisions, and this number increases roughly every 2 or 3 weeks to more than 800 replication cycles at the age of 50 years [2,10]. This could represent an evolutionary benefit, producing a more diverse population of germ cells, but each replication cycle brings a higher risk of error, not only during replication of the DNA sequence but also during the replication of epigenetic markers. Because replication mechanisms have an error rate at least one order of magnitude higher for epigenetic information compared with genetic information, spermatozoa from older men are burdened with more epigenetic than sequence changes. Epigenetic markers are known to be heritable, but the methylation profiles of the sperm fertilizing an egg and the mature sperm of this new individual are separated by several waves of extensive methylation reprogramming (Figure 1), allowing the establishment of a specific male methylation profile [11].

The first global remodeling of the methylation profile takes place 10–12 h after fertilization, and although it is mostly understood as demethylation, single-cell sequencing shows that it is a dynamic process characterized by massive global methylation erasure and local remethylation (Figure 2) [2,12,13]. Demethylation goes from around 80% in the sperm to around 50% in the male pronucleus in the zygote, where enhancers and gene body regions, particularly, are demethylated [12]. Methylation erasure continues during the first cell division and the formation of morula, when introns and short interspersed nuclear elements, especially Alu elements, are significantly enriched.

Additionally, *de novo* methylation starts during preimplantation development and is more prevalent than demethylation in the development of the eight-cell stage embryo. Potentially active repeat elements (i.e., short interspersed nuclear elements, long interspersed nuclear elements and long terminal repeats) are methylated to protect genome stability [12].

Demethylation of the paternal genome is much faster and more profound than that of the maternal genome [12]. Despite starting with higher methylation levels (82% methylation compared with 55% methylation of the oocyte), the average methylation of the paternal genome drops lower than that of the maternal genome in the two-cell stage of the embryo (15% of the paternal genome is methylated compared with 23% of the maternal genome). This difference also persists in the postimplantation embryo, although it is lower.

Analysis of differentially methylated regions (DMRs) – namely, analysis of genomic regions with different methylation statuses among a set of samples, such as different tissues, different individuals, different time points for an individual and diseased and healthy samples, depending on the context of the study – shows that methylation of the paternal genome is enriched in somatic cell-specific enhancers, whereas methylation of the maternal genome is enriched in CpG islands and gene promoters. In both genomes, short interspersed nuclear elements, especially Alu repeats, are methylated.

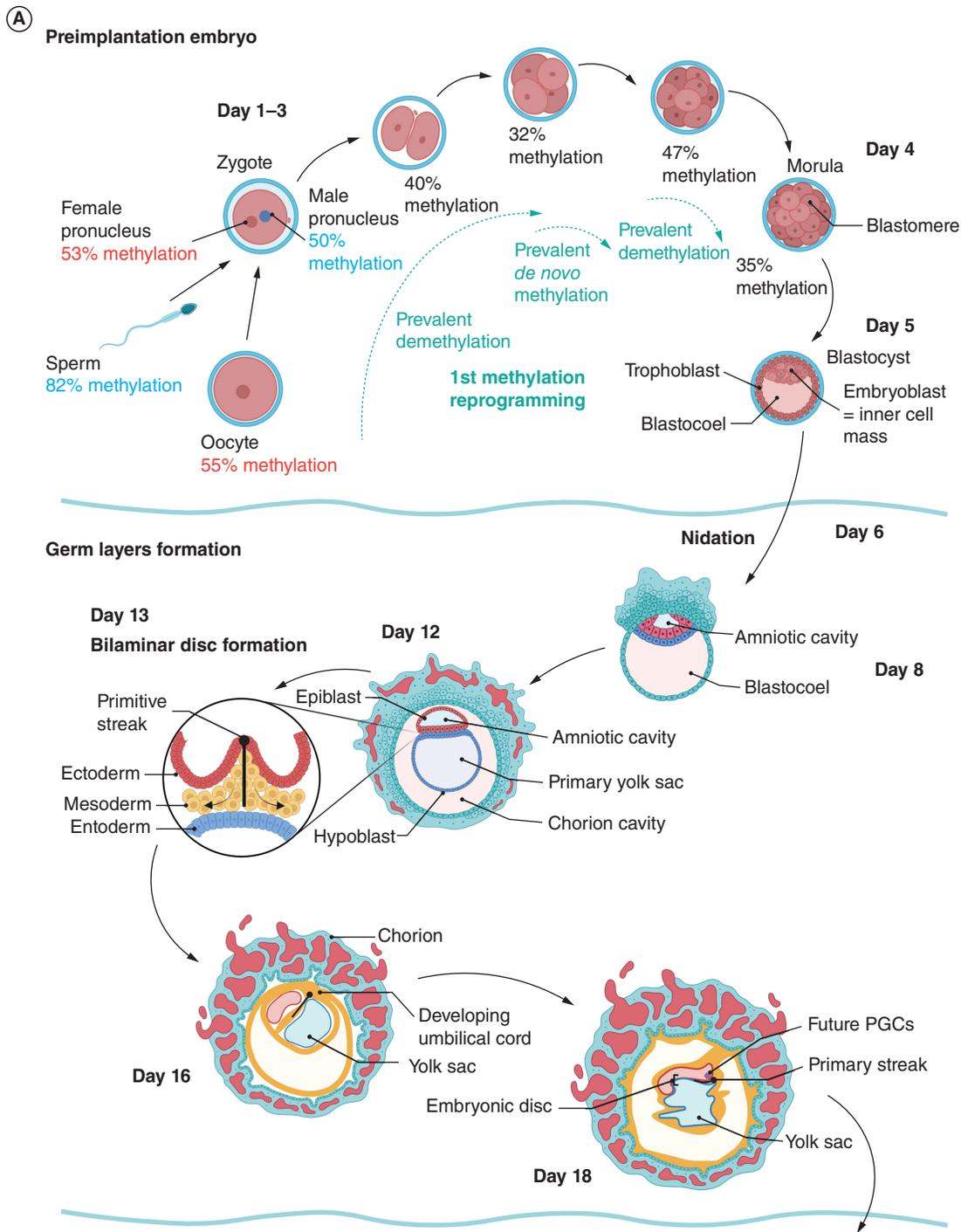


Figure 1. Development of male germ cells from conception to mature sperm. (A) Preimplantation and germ layer formation stage. In the Day 13, bilaminar disc formation is shown only in close-up for the sake of clarity. The rest of the structures are similar to the Day 12 stage. Ectodermal (red), mesodermal (orange), and entodermal (blue) structures are coded in the same color in each stage. **(B)** Second half of embryonic development. Embryo in the Days 22–28 stage is shown in the overview and cross-section to better demonstrate the migration route of PGCs as they pass the hindgut to the genital ridge. In the Week 7 stage, the digestive system is not shown for the sake of clarity. **(C)** Development of mature sperm cells in puberty. Prospermatogonia are called spermatogonia postnatally and have self-renewing function, providing a constant pool of precursor cells throughout the male lifetime. BMP: Bone morphogenetic protein; PGC: Primordial germ cell.

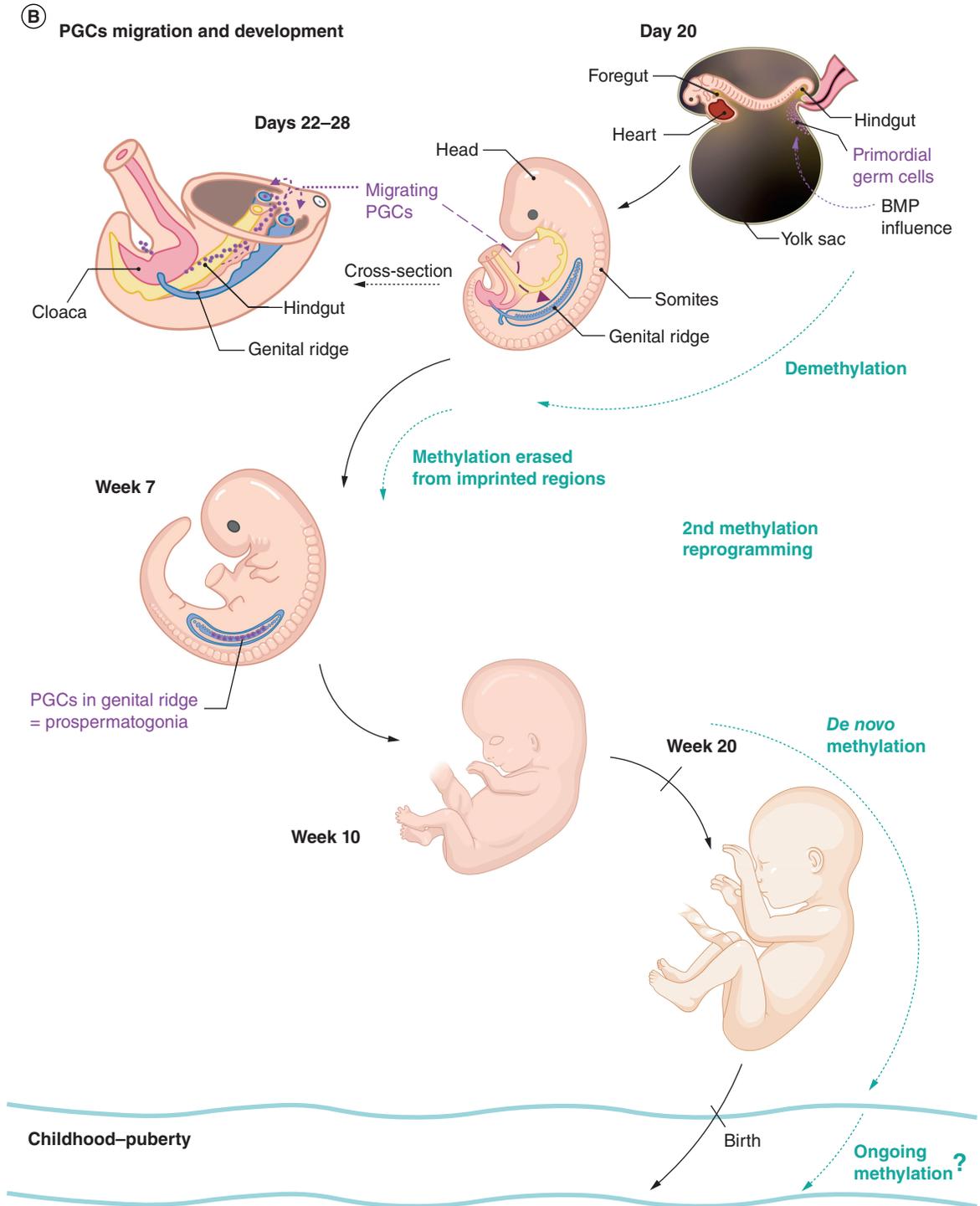


Figure 1. Development of male germ cells from conception to mature sperm (cont.). (A) Preimplantation and germ layer formation stage. In the Day 13, bilaminar disc formation is shown only in close-up for the sake of clarity. The rest of the structures are similar to the Day 12 stage. Ectodermal (red), mesodermal (orange), and entodermal (blue) structures are coded in the same color in each stage. **(B)** Second half of embryonic development. Embryo in the Days 22–28 stage is shown in the overview and cross-section to better demonstrate the migration route of PGCs as they pass the hindgut to the genital ridge. In the Week 7 stage, the digestive system is not shown for the sake of clarity. **(C)** Development of mature sperm cells in puberty. Prospermatogonia are called spermatogonia postnatally and have self-renewing function, providing a constant pool of precursor cells throughout the male lifetime. BMP: Bone morphogenetic protein; PGC: Primordial germ cell.

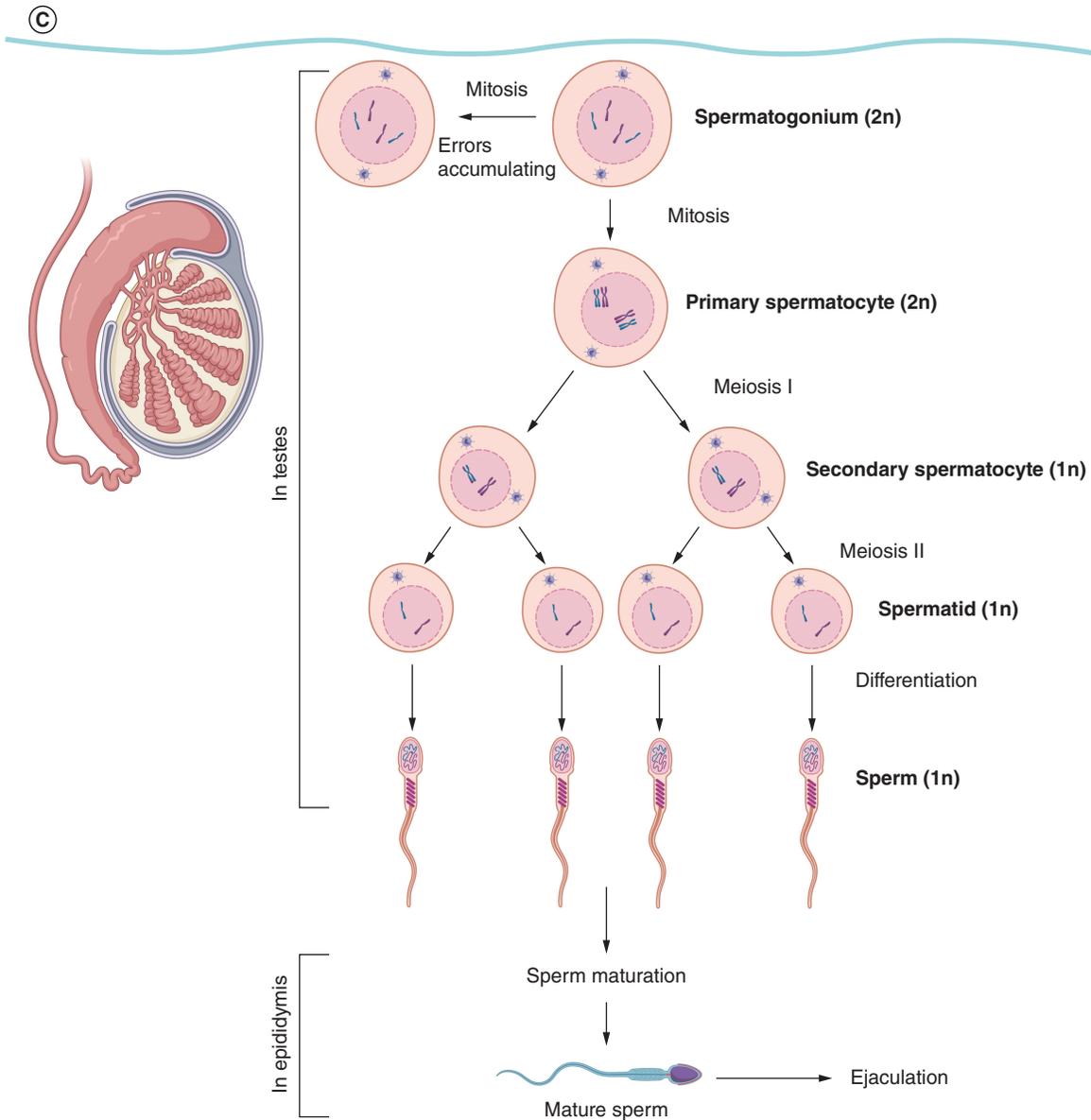


Figure 1. Development of male germ cells from conception to mature sperm (cont.). (A) Preimplantation and germ layer formation stage. In the Day 13, bilaminar disc formation is shown only in close-up for the sake of clarity. The rest of the structures are similar to the Day 12 stage. Ectodermal (red), mesodermal (orange), and endodermal (blue) structures are coded in the same color in each stage. **(B)** Second half of embryonic development. Embryo in the Days 22–28 stage is shown in the overview and cross-section to better demonstrate the migration route of PGCs as they pass the hindgut to the genital ridge. In the Week 7 stage, the digestive system is not shown for the sake of clarity. **(C)** Development of mature sperm cells in puberty. Prospermatogonia are called spermatogonia postnatally and have self-renewing function, providing a constant pool of precursor cells throughout the male lifetime. BMP: Bone morphogenetic protein; PGC: Primordial germ cell.

At the beginning of their development, primordial germ cells (PGCs) are epigenetically identical to other cells of the epiblast, showing stable values of DNA methylation and X-inactivation [2]. Under the influence of bone morphogenetic protein, PGCs undergo a second wave of DNA methylation remodeling, with almost complete erasure of DNA methylation. During this process, PGCs start to migrate past the hindgut to the genital ridge, where DNA methylation is also erased from imprinting control regions and meiotic genes [14]. High levels of DNA methylation are conserved in only several specific genomic regions (termed ‘escapees’) generally connected with genes expressed in the brain and associated with neurological and metabolic disorders, implying the possible transgenerational inheritance of epigenetic markers in these areas.

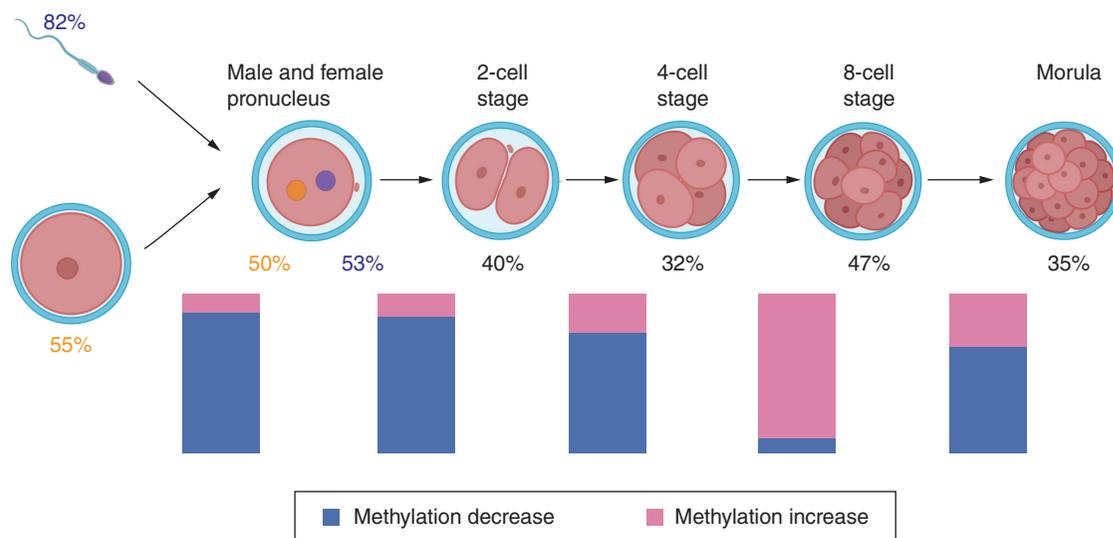


Figure 2. First wave of global epigenetic reprogramming. The ratio between methylation decrease (blue) and methylation increase (red) taking place during development from one stage to another is shown in the bar graph. Numerical values correspond to the percentage of methylated genome, and values highlighted in orange and purple correspond to the female and male pronucleus, respectively. Data taken from [12].

Under the influence of the SRY protein, somatic cells of the genital ridge undergo sex determination for the development of male gonads, differentiating into Sertoli cells. Subsequently, PGCs become developmentally restricted to the male sex, becoming gonocytes, also called prospermatogonia [15]. Sex-specific subpopulations of cells can be detected from week 9 [14]. By the seventh to tenth week of embryonic development, the global level of methylation in prospermatogonia is 5.0–8.0%, whereas somatic cells are mostly hypermethylated, with a mean methylation of 67.5–80.0% [2,11,14,16].

The window for extreme hypomethylation is quite long in human male germ cell precursors compared with, for example, those in mice. The process of *de novo* methylation starts in prospermatogonia in mitotic arrest around the 20th week, but because of developmental asynchrony, hypomethylated prospermatogonia may be found until the 26th week [11,14]. *De novo* methylation starts in the DNA repetition areas, especially retrotransposons and imprinted genes, creating a sex-specific methylation profile of germline cells [17]. DNA transposons and long interspersed nuclear elements are significantly overrepresented among DMRs in younger and older men [18]. It is unclear at which point we could consider the methylation signal established. Global methylation levels seem not to change significantly after birth, whereas spermatogonia remain in mitotic arrest [17], but some studies [11] suggest that this process continues after birth and is not completed until the start of puberty. More research is needed to investigate this postnatal phase of germ cell development.

Spermatogenesis itself starts in puberty, and prominent epigenetic changes include chromatin packaging and the exchange of the majority of histones for protamines. The number of nucleosomes bound to histones in humans is around 12.5% [2,19]. The areas conserved on histones are mostly CpG rich and contain promoters of developmentally important genes, imprinted areas and genes that code for miRNA [19]. Spermatogonia can mitotically either self-renew or give rise to a primary spermatocyte entering meiosis, producing four spermatids from one primary spermatocyte [20]. During the final stage of spermatogenesis (called spermiogenesis), the structure of the round spermatid polarizes, an acrosomal cap and a tail develop and an early sperm cell is oriented to the epididymis, where it undergoes a maturation process, finally acquiring the capacity to fertilize an egg.

Interesting question is whether there is a possibility of a third wave of methylation reprogramming during post-testicular sperm maturation, as the sperm enters *caput epididymis*. To further investigate this possibility, Galan *et al.* studied seven stages of developing sperm in mice and reported modest global hypomethylation in *caput epididymis* sperm, which was not observed in the preceding or succeeding developmental stages [13]. Cell-free DNA, presumably derived from the somatic cells of the epididymis, was identified as the source of this shift in methylation

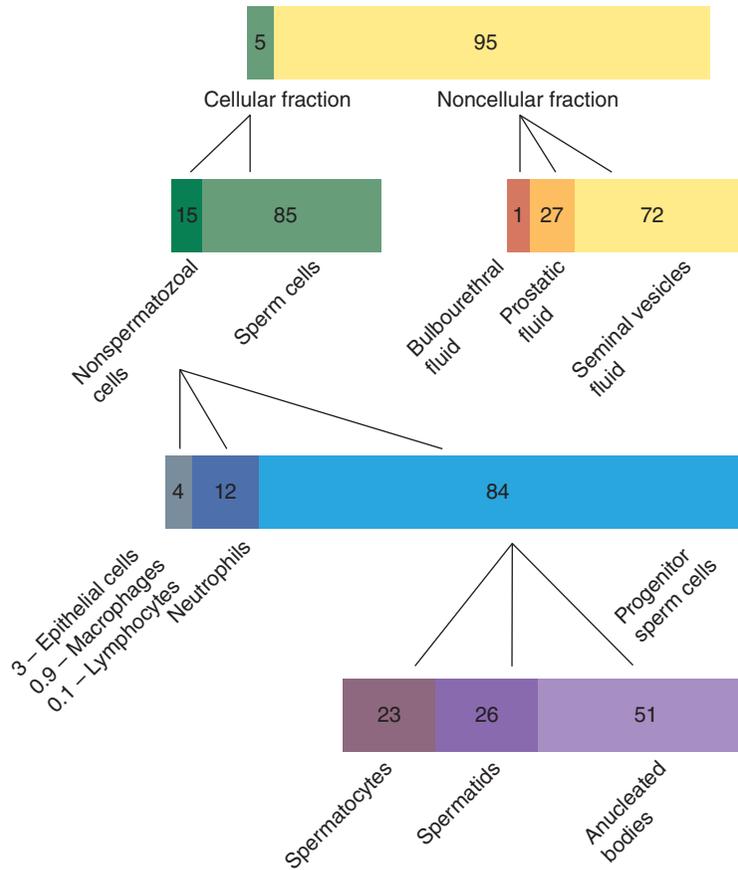


Figure 3. Composition of human semen. Values are listed as percentages. Data taken from [23–25].

profiles, emphasizing the differences between somatic and germline profiles. Interestingly, this cell-free material was found only in virgin mice and disappeared after mating.

Errors in the erasure or re-establishment of DNA methylation signals in individual PGCs may lead to the development of spermatogonia that produce, via clonal division, a subpopulation of sperm with defective methylation profiles [11]. Incomplete methylation of retrotransposons may also lead to their uncontrolled propagation throughout the genome, cause insertional mutagenesis and promote carcinogenesis [17]. Although the continuous division of spermatogonia creates and accumulates more epigenetic than sequencing errors, waves of reprogramming partly prevent their passing on to another generation.

A study by Jenkins *et al.* found no difference between patients with younger and older grandparents regarding levels of global methylation and only subtle changes in selected age-related regions, suggesting that age-related changes do not persist for two generations [21]. This is consistent with a study by Leitão *et al.*, who compared almost 100 patients with severely impaired spermatogenesis with healthy controls and detected no recurrent epigenetic mutations [11]. By contrast, it has already been shown that epigenetic changes caused by other factors (e.g., exposure to pollutants such as the pesticide dichlorodiphenyltrichloroethane) obtain imprinting-like features, evading methylation erasure and creating an abnormal epigenome in embryonic stem cells, affecting all forming somatic cells and tissues and subsequently producing altered germline cells [20].

Sperm cell sampling

Mature sperm cells capable of fertilization are stored in the epididymides. Their concentrated suspension is diluted by prostatic fluid, and subsequently seminal vesicle fluid in the urethra, to form the ejaculate during sexual activity [22]. Considering this anatomical succession, sequential ejaculate fractions are not equal, and their composition is influenced by circumstances such as abstinence time before ejaculation as well as the level and/or duration of sexual arousal, which affects the smooth muscle cell contractions emptying glands and actively transporting spermatozoa. Samples collected in a container in a laboratory setting can thus have different compositions compared with the individual fractions ejaculated during sexual intercourse. Figure 3 depicts the proportional composition of human

semen; however, the literature sources vary, so this information cannot be considered definitive.

Human semen consists of a 2–5% cellular fraction, whereas 95–98% is a seminal plasma rich in sugars, glycans, lipids, ions, metabolites, cell-free DNA, miRNAs, peptides and proteins secreted by the seminal vesicles, prostate, epididymides and bulbourethral glands [23]. Of semen's cellular fraction, Long *et al.* report that approximately 85% is allotted to sperm cells and 84% of the rest is progenitor sperm cells [24]. Although the authors also report the proportion of white blood cells in semen to be around 13%, there is no official consensus defining the physiological ratio of leukocytes to other cell types in semen [22]. During genitourinary tract or accessory gland infections such as prostatitis or other inflammatory processes, the proportion of leukocytes rises. A higher proportion of immature germ cells in semen might occur during various developmental disorders of the genitourinary tract, with a varicocele or as an adverse effect during cytotoxic treatments [24]. The remaining portion of nonsperm cells consists of squamous epithelial cells from the excretory ducts, epithelial cells from the seminal vesicles and exfoliated Sertoli cells [25]. In addition, the amount of these cells can rise during seminal vesiculitis.

Differences between somatic & germ cell methylation

Contamination of somatic DNA in sperm samples is described as one of the major confounding factors in sperm methylation studies [26]. The set of 217 DMRs, potentially useful as diagnostic markers of male idiopathic infertility published by Luján *et al.* [27], was unable to distinguish normozoospermic from oligozoospermic samples in independent testing by Leitão *et al.* [11] but pinpointed samples containing somatic DNA among clean sperm samples. In the study by Leitão *et al.*, 2462 genes showed differential expression between blood and sperm samples, with promoters of these genes being mostly methylated in blood and unmethylated in sperm.

Although the epigenetic clock by Horvath is described as a multitissue model [5], it cannot be used for age prediction using testicular tissue and sperm cells [1]. When tested on a small set of semen samples, the mean absolute error (MAE) of the prediction was 13.3 years compared with an MAE between 4 and 5 years for somatic tissues [28]. Testing of another smaller three-CpG blood-based prediction model yielded an MAE of 37.3 years [29]. This confirms that the loci used for age prediction in somatic cells have no predictive value in sperm [30].

Global hypomethylation with regional gene-associated hypermethylation is considered an important feature of aging in somatic cells. In sperm DNA, the opposite situation was identified, with significant global hypermethylation and specific hypomethylated loci [1,19,21]. Hypermethylation has been associated with CpG islands and hypomethylation with CpG shores in sperm DNA.

Frequently dividing cells are burdened with accumulating epigenetic mutations, or epigenetic drift [31]. However, models that successfully predict age from various tissues, such as the clocks of Hannum *et al.* [4] and Horvath [5], confirm that epigenetic age does not measure proliferation rate, as the included tissues have variable – or even zero (e.g., neuronal cells) – proliferation rates [32].

Although the average methylation change at a single CpG in somatic cells is 0.05–0.15% per year, Jenkins *et al.* reported a stronger effect of age on methylation in sperm cells, with an average change of >0.4% per year in 39 genomic regions [1]. In these especially altered regions, age-related changes can be within the range of 10–12% during the male reproductive lifespan. By contrast, according to Heidegger *et al.*, the predictive effect of an individual CpG in sperm cells is smaller than in somatic cells (R^2 0.34–0.35 in sperm vs 0.66–0.68 in blood) [33]. The diverse character of these studies (one worked with 1 kbp long genomic windows [3], whereas the other used single CpGs [30]) may contribute to these contradictory results, but further and more uniform research is needed.

The magnitude of DNA methylation changes in sperm DNA

In a study by Oluwayiose *et al.* [3], 1698 CpGs were associated with age in tested sperm DNA, 91% of which showed an increase in methylation of 0.2–1.7% for every 5 years of age [3]. Among this set of CpGs, none overlapped with the epigenetic clocks of Hannum *et al.* [4] and Horvath [5] based on somatic cells and only 30% overlapped with the model developed by Jenkins *et al.* [50] for chronological age prediction from sperm DNA methylation.

When wider DMRs were considered, 1146 were detected by Oluwayiose *et al.* [3]. In contrast to individual CpG methylation, only 57% DMRs showed hypermethylation, and 59% in another similarly designed study by Pisarek *et al.* [34]. A study by Bernhardt *et al.* detected a similar number of DMRs (1565), but the ratio between hypermethylated and hypomethylated was reversed, with only 26% of DMRs hypermethylated [10]. In that study, hypermethylated regions were mostly associated with biological processes connected to synapsis and abnormal behavior, whereas the majority of hypomethylated areas were enriched in male infertility phenotypes. This may be

Table 1. Number of CpGs and age-associated differentially methylated regions and percentage of loci hypermethylated with higher age in included studies.

Author and Year	Number of CpGs	Percentage of hypermethylated loci	Ref.
Lee, 2015	10710	43%	[28]
Cao, 2020	21855	62%	[49]
Oluwayiose, 2021	1698	91%	[3]
Pisarek, 2021	14916	59%	[34]

Author and Year	Number of DMRs	Percentage of hypermethylated loci	Ref.
Denomme, 2020	3405	24%	[19]
Laurentino, 2020	236	51%	[18]
Oluwayiose, 2021	1146	57%	[3]
Pilsner, 2021	2364	44%	[30]
Bernhardt, 2023	1565	26%	[10]

one of the reasons behind age-associated idiopathic male infertility. The results of these studies are summarized in Table 1.

Clinical impacts of age-related epigenetic sperm changes

Age-related changes in DNA methylation patterns are associated with three main reproductive health outcomes: lower probability of successful conception, both naturally and with assisted reproduction methods; higher risk of abnormal embryonic and fetal development; and development of various disorders in offspring [3]. As these are all complex multifactorial processes that include genetic and epigenetic influences, we could expect that the influence of paternal age comprises a higher number of small-scale changes, rather than just one or two highly penetrating epigenetic variants [10].

The study by Oluwayiose *et al.*, which looked at 47 infertile couples, identified characteristic CpG hypermethylation patterns in the oldest fathers compared with the youngest [3]. Results from two intermediate age groups suggested a dose-dependent impact of age-related methylation changes on successful conception. The linear relationship between age and methylation was confirmed in another study [2].

In the study by Oluwayiose *et al.*, successful live birth was reported in 57% of couples in the youngest fathers group, 35% of couples in the intermediate group and none of the couples in the oldest fathers group [3]. Among the younger patients with methylation patterns corresponding to the older group's pattern, there were also no successful live births. The probability of conception decreased by 4% each year. Pilsner *et al.* reported the same trend, with a 17% lower cumulative probability of achieving pregnancy for couples with a predicted older father, together with a longer time to achieving pregnancy and gestational age of the developing embryo [30].

Age-related changes in sperm DNA methylation are located mostly in the area of embryonic development and neurological development genes [3,19,35,36]. The most enriched ontology terms include flavonoid glucuronidation, transmission on chemical synapses, histone ubiquitination [30], opioid and neuronal nitric oxide synthase signaling pathways [19], and DNA binding [18]. Also, homeobox genes were significantly enriched, specifically between hypomethylated regions [18].

Age-associated methylation changes were reported in genes connected to diabetes mellitus [1], hypertension [1], autism spectrum disorder [1,19], bipolar disorder [1,19], schizophrenia [1,19] and other neurological and developmental disorders, especially those affecting the cerebellum (e.g., ataxia and general learning disabilities) [35]. The same effect was experimentally demonstrated in mice, where DNA methylation changes in the sperm of older males manifested at the level of gene expression in areas associated with schizophrenia and bipolar disorder as well as at the behavioral level [3]. Fewer exploratory activities, lower response rates to acoustic startle signals and decreased cognitive functions were also observed in the offspring of older males [35]. This effect was preserved even in the next generation, with hypomethylation persisting in 75% of the offspring of older grandfathers [1].

Table 2. Summary of methods used as a starting point in studies of age-related methylation changes in semen samples.

Whole-genome sequencing approaches		Array-based approaches		Ref.
WGBS	RRBS	450K Infinium	850K EPIC	
Leitão, 2020	Denomme, 2020	Jenkins, 2014	Jenkins, 2019	[1,11,19,21]
Denomme, 2020	Prell, 2022	Lee, 2015	Pilsner, 2022	[19,28,30,36]
Laurentino, 2020	Bernhardt, 2023	Jenkins, 2018	Pisarek, 2021	[10,18,34,50]
Cao, 2020		Oluwayiose, 2021		[3,49]
di Persio, 2021				[26]

RRBS: Reduced representation bisulfite sequencing; WGBS: Whole-genome bisulfite sequencing.

Most of differentially methylated genes reported in several comparable studies are associated with embryonic neural development [11]. However, there is a little overlap in tangible genes. Of the seven studies reviewed by Bernhardt *et al.*, only three genes – namely, *DLGAP2*, *PRDM16* and *SLC22A18AS* – were reported in at least five studies, with 90% of the reported genes included in only the original study [10].

Age-induced DNA methylation changes could become a useful diagnostic tool. Array-based DNA methylation profiles from blood samples can highlight differences between infertile men and fertile controls [17]. The previously mentioned studies have shown that the methylation of sperm DNA can be even more informative. Candidate genes, which can be considered epigenetic markers of lower fertility and higher risk of neurodevelopmental disorders in offspring, are starting to emerge.

Forensic applications of age-related epigenetic sperm changes

Age prediction from DNA is important in the field of forensics. The ability to predict an unknown person's different phenotype traits based on their DNA could be an important forensic tool. This can provide information in cases where biological traces have been found but comparing DNA profiles with police databases or suspect profiles yields no match. Various visible traits, such as age, pigmentation, occurrence of freckles and premature hair graying or balding, together with lifestyle traits, such as diet, exercise, smoking and alcohol or drug consumption, can help to significantly narrow down the pool of potential suspects. Age prediction is currently one of the most pursued tools of DNA phenotyping [37]. Forensic age prediction from blood samples has been deeply studied [38,39], but we lack a predictive model for other forensically relevant tissues, such as semen samples, which are often found after sexual assaults [28,34].

Age prediction models

The majority of studies researching age-related changes in DNA methylation of semen samples use whole genome sequencing or array-based approaches as the starting point (Table 2).

The consistency between chronological age and methylation in specific loci suggests the existence of regions more prone to epigenetic changes, but there is only a little overlap between the most informative and predictive loci in individual studies. Different predictive models with similar outcomes use different loci, making it impossible to identify a set of the most informative. Pilsner *et al.* built two independent age prediction models, one based on individual CpGs and the other based on the entire DMRs [30]. Prediction errors were very similar, but of the 318 CpGs included in the DMR-based model, only ten (less than 1%) overlapped with the 120 CpGs included in the CpG-based model. From a practical point of view, the authors decided to use the CpG-based model for further clinical testing, as it requires fewer loci and minimizes the cost of the array, but it is possible that both models have an independent association with reproductive health outcomes.

For clinically focused prediction models, it can be advantageous to use a large number of markers to make the resulting model as precise as possible even if the effect of individual loci on the predicted age is low. However, these models are generally inapplicable for forensic practices, where the generally low quality and quantity of available DNA are the biggest limiting factors [33,34,40].

Clinically and forensically focused models also selected their most informative CpGs based on different criteria. Although for medical purposes age acceleration or deceleration compared with the real chronological age is the key factor in predicting health-related outcomes [1,3], forensic predictors must be focused on chronological age itself, with a minimum influence of other lifestyle, environmental and health factors.

Table 3. Markers used for individual prediction models.

Author, year	Locus	Coordinates in hg38	Ref.
Lee, 2015	<i>TTC7B</i>	chr14:90817262	[28]
	<i>NOX4 (FOLH1B)</i>	chr11:89589683	
	cg12837463	chr7:35260617	
Li, 2020	<i>TTC7B</i>	chr14:90817262	[45]
	<i>NOX4 (FOLH1B)</i>	chr11:89589683	
Pisarek, 2021	<i>SH2B2</i>	chr7:102288454	[34]
	<i>NOX4 (FOLH1B)</i>	chr11:89589683	
	<i>EXOC3</i>	chr5:525617	
	<i>IFITM2</i>	chr11:312518	
	<i>GALR2_1</i>	chr17:76077680	
	<i>GALR2_2</i>	chr17:76077748	

The first forensically focused prediction model using sperm DNA methylation was developed by Lee *et al.* using the three most informative CpGs identified with a 450K Infinium array (Illumina, CA, USA) in 12 semen samples [28]. The model itself was constructed on a set of 68 samples (31 as the training set and 37 as the testing set, with an MAE of 4.2 and 5.4 years, respectively), and after retraining with the whole set, it reached an MAE of 4.7 years. These results were confirmed by an independent validation study, which found an MAE of 4.8 years for 12 volunteer samples and 5.2 years for 19 real-life casework samples, with various methods of sampling and a storage time from 3 months to 17 years [41].

Although the prediction model developed by Lee *et al.* was based on the single-base extension method [28] and a newer study by Pisarek *et al.* used next-generation sequencing [34], the resulting performance of both models was similar, with an MAE of 4.3 and 5.1 years in the training and testing data sets, respectively. Lee's prediction model applied to Pisarek's samples resulted in an MAE of 5.7 years. This difference can be explained by the use of different technology as well as interpopulation variability in DNA methylation, considering that these studies were conducted in South Korea and Poland, respectively.

Despite Pisarek *et al.* testing three CpGs [34] from the model developed by Lee *et al.* [28], together with their top hits from the 850K EPIC array (Illumina, CA, USA), only one was included in the final six-CpG prediction model (*NOX4*, also known as *FOLH1B*) (Table 3). This final model, along with a wider panel comprising 13 loci and 36 CpGs, was validated by five laboratories from the VISAGE Consortium, which assessed its reproducibility and sensitivity [33]. For two mock crime scene samples, the average difference between predicted and chronological age was 2.0 and 5.4 years.

A variety of blood-based age prediction studies describe the phenomenon of decreasing prediction sensitivity with increasing age [42–44]. A validation study by Lee *et al.* confirmed this for sperm DNA as well [41]. Although the general MAE was 4.8 years, the mean error increased from 2.9 years for individuals in their 20s to 7.2 years for individuals in their 50s. The results and parameters of the aforementioned studies are summarized in Table 4.

The decreasing precision of the prediction with increasing age must also be taken into account when comparing prediction results from the different studies shown in Table 4. The model developed by Li *et al.* offers the lowest MAE but is built on samples from individuals aged 21–54 years and validated in individuals younger than 46 years [45], whereas Pisarek *et al.* worked with individuals aged 26–57 years [34] and Lee *et al.* included individuals up to the age of 73 years [28]. Mutual comparability is also limited by the use of differential extraction by Li *et al.* [45] and the swim-up sperm fraction by Laurentino *et al.* [18] compared with other studies that worked with whole semen samples.

Limiting factors

Population sample size

The main limitation of age-related methylation studies in sperm is very small sample size. In most studies, the entire sample set is less than 100 samples, and even the largest use sets smaller than 400 samples (Table 4). Similar studies that have developed age prediction models for blood samples generally use at least 100 samples, with average sample sets consisting of 200–300 samples and the largest exceeding 1000 samples [29,46]. Apart from the fact that the population of potential sample donors is approximately halved by sex, semen is considered a much more

Table 4. Comparison of performance of various prediction models based on sperm DNA methylation changes.

Clinically focused studies				Ref.
Author, year	Number of samples	Number of loci	Prediction results	
Jenkins, 2018	329	51 DMRs	MAE = 2.0 years	[50]
Cao, 2020	94	5000 CpGs	MAE _{training} = 2.7 years	[49]
	12		MAE _{validation} = 4.4 years	
Pilsner, 2022	379	120 CpG	MAE _{training} = 1.6 years	[30]
			MAE _{testing} = 2.2 years	
		117 DMRs (318 CpGs)	MAE _{training} = 1.7 years	
			MAE _{testing} = 2.3 years	
Forensically focused studies				
	Number of samples	Number of loci	Prediction results	
Lee, 2015	68	3 CpGs	MAE = 4.7 years	[28]
Li, 2020	38	2 CpGs	MAE _{training} = 3.0 years	[45]
	17		MAE _{validation} = 4.3 years [†]	
Laurentino, 2020	42	6 DMRs	MAE _{training} = 7.8 years	[18]
	33		MAE _{validation} = 9.8 years	
Pisarek, 2021	125	6 CpGs	MAE _{training} = 4.3 years	[34]
	54		MAE _{validation} = 5.1 years	

[†] Mean value for MAEs calculated from prediction results for liquid semen samples and fresh and aged semen stains.
DMR: Differentially methylated region; MAE: Mean absolute error.

private biological material than, for example, blood, and the number of volunteers willing to provide a sample for research is generally lower. Cooperation with assisted reproduction clinics may help to increase the total number of samples but burdens the sample set with potential selection bias, as most donors are looking for help with fertility issues. With the decline in sperm quality and quantity among the general population over the last few decades and many reproductive health problems passing undetected, we must choose between a specific known bias in samples from individuals with fertility problems and a bias we cannot fully define but that is representative of the general population in samples from healthy volunteers [22]. Small sample sets like this tend to misrepresent biological variability in the population and lead to the overfitting of prediction models, which then perform significantly worse on independent samples.

Differentially methylated regions

Comparing the results of different studies can be complicated because of unclear definitions of the differentially methylated area in terms of length and difference thresholds. Leitão *et al.* defined the DMR as a region of at least ten CpGs with a methylation difference of at least 80% between two individuals of different ages [11]. However, Oluwayiose *et al.*, for example, used clustering algorithms starting with two CpGs in 1000 bp [3].

Typing platform

Pyrosequencing is generally considered the gold standard, but many phenotyping studies use the single-base extension method SNaPshot (Applied Biosystems, MA, USA) or the mass spectrometry-based EpiTYPER (Agena Bioscience, CA, USA), and massively parallel sequencing is becoming more and more common. The use of different typing platforms can also be a source of inconsistent results, as loss of accuracy of methylation data has been reported during technology transfer among different methods [46,47].

Sample composition

The composition of the sample itself can significantly influence the entire study. As already mentioned, ejaculate is not created until it leaves the body, and its composition changes during ejaculation, between instances of sexual activity and depending on many individual physical and psychological factors [22]. Because collecting the entire sample in a single container erases differences in individual fractions of the ejaculate, studies focused on reproduction lose specific information from the initial fraction of ejaculate, which is most effective for conception, whereas laboratory studies work with samples that do not correspond to forensic samples, such as individual semen stains and vaginal swabs. The ratio between individual cell types in the semen of an unknown person can hardly

be estimated. The number of sperm per milliliter of semen is highly variable and a low sperm count can result in relative enrichment of nonspermatozoal cells; the number of white blood cells increases during infection; and atypical cells with aberrant methylation profiles can be present during cancer [24,25].

Purification methods

Considering the significantly different methylation profiles of somatic cells, most studies use some kind of purification of the semen so as to analyze only the sperm DNA. However, sample purity can itself be a confounding factor. Pilsner *et al.* [30] detected a number of age-associated CpGs comparable to Oluwayiose *et al.* [3] (2364 and 1698, respectively), but fewer of these regions were hypermethylated (44 and 91%, respectively). Although Oluwayiose *et al.* worked with samples that had been checked using microscopy for any somatic contamination, Pilsner *et al.* used whole semen samples after one step of 40% gradient centrifugation. It is possible that significant differences in obtained values are the result of DNA methylation of somatic cells.

There are a wide variety of methods applicable to the purification of semen samples. Swim-up, chemotactic methods and microfluidic chips are based on sperm motility; density gradient centrifugation, with a variable number of steps to select a fraction of the sample, is based on sedimentation; and advanced methods use the specific binding of sperm to, for example, hyaluronan. The effectiveness of these methods is variable, and there is no consensus regarding their priority for DNA methylation studies. Most of the aforementioned studies used one- or two-step gradient centrifugation [3,30], in some cases with swim-up enrichment of the motile sperm fraction [10,19].

The use of whole semen samples can increase predictive error as a result of the unpredictable ratio of sperm and nonsperm cells but may result in epigenetic profiles more applicable to real-life samples, especially in forensic settings, where semen stains from clothing or bedsheets are often analyzed. A whole semen prediction model could be more robust for this kind of situation than a model built on purified sperm cells. However, purified sperm DNA models have their place in forensics as well because when analyzing vaginal swabs, differential lysis must be performed to eliminate the genetic material of the woman so only sperm DNA remains. The forensically focused studies by Lee *et al.* [28], Li *et al.* [45], and Pisarek *et al.* [34] refer to the same genetic loci and to each other despite the use of differential dithiothreitol-based extraction of sperm DNA by Li *et al.* and the use of whole semen samples by Lee *et al.* and Pisarek *et al.*

Biological factors

Lifestyle and environmental factors can be significant confounding factors in studies focusing on age-related changes in DNA methylation since they can influence epigenetic patterns as well [3]. Although age is considered to be the most profound and predictable modifier of sperm epigenetic profiles [21], smoking, alcohol and drug consumption, diet, physical activity, stress levels and chemical exposure can also cause significant changes in sperm DNA patterns and contribute to male infertility [3,17]. Both hypomethylation and hypermethylation can have positive and negative impacts on health, depending on the affected loci.

Although metabolic disorders in offspring, such as diabetes mellitus, have been linked to the hypomethylation pattern in sperm from older fathers, imprinted genes controlling growth and development (neuronatin, *PEG3* and *MEST*) were also hypomethylated in blood samples from children with obese fathers [17]. An unhealthy paternal diet also negatively influences the methylation of genes encoding some liver functions connected to cholesterol and lipid metabolism. By contrast, oral supplementation of micronutrients such as folate, vitamin B, zinc and cysteine showed a positive increase in DNA methylation levels [17].

A supervised yoga regimen promoted methylation changes, with positive health impacts in almost 400 genes, whereas tobacco smoking had a strong negative impact on health through methylation changes in many genes, not just those related to male infertility [17]. The effect of smoking on sperm DNA methylation is similar to the effect of aging, as smokers tend to be predicted as older than nonsmokers [1,30]. Together, these findings suggest a global link between a healthy lifestyle, proper methylation patterns and reproductive health [17].

Future perspective

This whole field of study is quite fragmented, with often contradictory results. Inconsistencies in study design limit the possibility for inferring any consensus. Unifying the terminology would definitely help (e.g., a clear definition of DMRs in terms of total length, minimum number of contained CpGs and minimum difference between compared sample sets). The easiest way might be to completely stop using DMRs in studies and instead use an unambiguous individual CpG approach. Array-based studies also do not cover all positions where significant differences and

associations can be found by the whole genome approach. The use of whole genome sequencing technology in the discovery phase of studies seems feasible, as it provides more data from less researched areas like noncoding and repetitive regions. Those regions can also be interesting targets for future aging-related epigenetic studies. The biggest limitation of sperm and semen DNA methylation studies is the small sample sets, so it is necessary to validate predictive models on larger sets of well-defined samples covering various characteristics of donors in terms of general and reproductive health, lifestyle factors, like obesity and smoking, and age from puberty to senescence.

The differences between methylation profiles from purified sperm cells and whole semen samples should also be researched further. Defining the magnitude of error that the variable ratio of somatic cells in the semen sample brings to age prediction is important for real-life use of this tool and would minimize inconsistencies caused by different sample preprocessing methods. Single-cell sequencing methods should be implemented more widely, as they hold the potential to overcome the problem of sample purity and epigenetic differences between somatic and germ cells. The development of techniques suitable for working with smaller samples would also be very beneficial, especially for forensic use, where the low quantity and quality of crime scene DNA is the biggest limiting factor. For the same reason, bisulfite conversion, which is considered the gold standard for methylation studies, should be replaced with some less destructive method, such as enzymatic conversion or TET-assisted pyridine borane sequencing. Recently, a new quantitative PCR method capable of distinguishing between methylated and unmethylated cytosine in untreated DNA has been created [48]. With this approach, the use of intercalating pseudonucleotides increases the melting temperature of sequences, including methylated cytosine, allowing their preferential amplification.

Author contributions

L Kotková contributed to conceptualization, preparing of the original draft, graphics and writing. J Drábek contributed to conceptualization, reviewing, editing and administration. Both authors have read and agreed to the submitted version of the manuscript.

Acknowledgments

Graphics were created with BioRender.com.

Financial disclosure

This work was supported by European Infrastructure for Translational Medicine and funded by LM2023033, CZ.02.1.01/0.0/0.0/16_026/0008448, EF16_013/0001674, LX22NPO5102 and IGA LF UP 2023_006. The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

Competing interests disclosure

The authors have no competing interests or relevant affiliations with any organization or entity with an interest in or conflict with the subject matter or materials discussed in the manuscript. This includes employment, consultancies, honoraria, stock ownership or options, expert testimony, grants or patents received or pending or royalties.

Writing disclosure

The authors thank B Stenglová for help with graphics and J Parrott for proofreading.

Open access

This work is licensed under the Attribution-NonCommercial-NoDerivatives 4.0 Unported License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>

Executive summary

Background

- Compared with somatic cells, age-related methylation changes in sperm are understudied, although they have both clinical and forensic implications.

The role of DNA methylation

- The correct methylation signal is important for proper gene expression control, maintaining genome stability, proper cell type differentiation and cellular adaptation to the changing environment.

Development of sperm DNA methylation signals

- Because of continuous cell division and higher epigenetic information reproduction error rate, sperm DNA is more affected by epigenetic changes with increasing age than somatic cells or ovum DNA.
- During embryonic development, embryonic cells and future sperm cell precursors undergo almost complete remodeling of the methylation signal.

Sperm cell sampling

- Methylation is tissue-specific; thus, the highly variable ratio of individual cell types in semen might significantly affect age prediction results.

Differences between somatic & germ cell methylation

- Although aging in sperm cells is characterized by global hypermethylation with specific hypomethylated loci, the situation in somatic cells is exactly the opposite.
- Multitissue age prediction models report a high mean absolute error for semen samples compared with somatic tissues.

The magnitude of DNA methylation changes in sperm DNA

- Results of studies comparing methylation in semen samples of young and old individuals are inconclusive in terms of the ratio between hypomethylated and hypermethylated loci during aging.

Clinical impacts of age-related epigenetic sperm changes

- Clinical studies show the connection between older paternal age and longer time to conception, lower percentage of live births and various disorders (mainly metabolic, neurological and behavioral) in offspring. The same applies to younger men with a higher epigenetic age.

Forensic applications of age-related epigenetic sperm changes

- Forensic models allow the prediction of an unknown sperm trace donor's age, which can help to narrow down police investigations, primarily in cases of sexual assault.

Age prediction models

- Clinically relevant age prediction models can use a larger number of markers with a smaller individual impact and are more focused on detecting age acceleration or deceleration compared with chronological age. The mean absolute error of prediction is approximately 2.5 years.
- Forensically relevant age prediction models must use a lower number of predictors because of the limited quantity of DNA samples and are focused on chronological age itself. The mean absolute error of prediction is approximately 5.0 years.

Limiting factors

- Existing studies are mostly burdened with insufficient sample sizes and significant inconsistencies regarding types of samples, sample preprocessing, typing methods and terminology used to describe results.

Future perspective

- The entire field of study is fragmented, with contradictory results, and inconsistencies in study designs and terminology limit the possibilities for reaching consensus.
- Larger sample sets with various donor characteristics, use of whole genome sequencing in the discovery phase, development of techniques suitable for less DNA input, and alternatives to bisulfite sequencing could lead to the development of better prediction models.

References

Papers of special note have been highlighted as: • of interest; •• of considerable interest

1. Jenkins TG, Aston KI, Pflueger C, Cairns BR, Carrell DT. Age-associated sperm DNA methylation alterations: possible implications in offspring disease susceptibility. *PLoS Genet.* 10(7), e1004458 (2014).
 2. Ashapkin V, Suvorov A, Pilsner JR, Krawetz SA, Sergeyev O. Age-associated epigenetic changes in mammalian sperm: implications for offspring health and development. *Hum. Reprod. Update* 29(1), 24–44 (2023).
 3. Oluwayiose OA, Wu H, Saddiki H *et al.* Sperm DNA methylation mediates the association of male age on reproductive outcomes among couples undergoing infertility treatment. *Sci. Rep.* 11(1), 3216 (2021).
- **Maps the dose-dependent relationship between male age and sperm methylation in four father groups stratified by age and finds that paternal methylation patterns are associated with offspring health.**

4. Hannum G, Guinney J, Zhao L *et al.* Genome-wide methylation profiles reveal quantitative views of human aging rates. *Mol. Cell* 49(2), 359–367 (2013).
5. Horvath S. DNA methylation age of human tissues and cell types. *Genome Biol.* 14(10), R115 (2013).
6. Bormann F, Rodríguez-Paredes M, Hagemann S *et al.* Reduced DNA methylation patterning and transcriptional connectivity define human skin aging. *Aging Cell* 15(3), 563–571 (2016).
7. Jin B, Li Y, Robertson KD. DNA methylation: superior or subordinate in the epigenetic hierarchy? *Genes Cancer* 2(6), 607–617 (2011).
8. Ghazimoradi MH, Hasegawa K, Zolghadr E, Montazeri S, Farivar S. Reprogramming of fibroblast cells to totipotent state by DNA demethylation. *Sci. Rep.* 13(1), 1154 (2023).
9. Aston KI, Uren PJ, Jenkins TG *et al.* Aberrant sperm DNA methylation predicts male fertility status and embryo quality. *Fertil. Steril.* 104(6), 1388–1397.e1–5 (2015).
10. Bernhardt L, Dittrich M, Prell A *et al.* Age-related methylation changes in the human sperm epigenome. *Aging (Albany NY)* 15(5), 1257–1278 (2023).
- **Maps loci with significantly changed methylation status associated with age and compares obtained results with seven different corresponding studies.**
11. Leitão E, Di Persio S, Laurentino S *et al.* The sperm epigenome does not display recurrent epimutations in patients with severely impaired spermatogenesis. *Clin. Epigenetics* 12(1), 61 (2020).
12. Zhu P, Guo H, Ren Y *et al.* Single-cell DNA methylome sequencing of human preimplantation embryos. *Nat. Genet.* 50(1), 12–19 (2018).
- **Uniquely maps changes in methylation levels in individual stages of the preimplantation embryo.**
13. Galan C, Serra RW, Sun F, Rinaldi VD, Conine CC, Rando OJ. Stability of the cytosine methylome during post-testicular sperm maturation in mouse. *PLoS Genet.* 17(3), e1009416 (2021).
14. Ramakrishna NB, Murison K, Miska EA, Leitch HG. Epigenetic regulation during primordial germ cell development and differentiation. *Sex. Dev.* 15(5–6), 411–431 (2021).
15. Ben Maamar M, Beck D, Nilsson E, McCarrey JR, Skinner MK. Developmental alterations in DNA methylation during gametogenesis from primordial germ cells to sperm. *iScience* 25(2), 103786 (2022).
16. Li L, Li L, Li Q *et al.* Dissecting the epigenomic dynamics of human fetal germ cell development at single-cell resolution. *Cell Res.* 31(4), 463–477 (2021).
17. Rotondo JC, Lanzillotti C, Mazziotta C, Tognon M, Martini F. Epigenetics of male infertility: the role of DNA methylation. *Front. Cell Dev. Biol.* 9, 689624 (2021).
18. Laurentino S, Cremers JF, Horsthemke B *et al.* A germ cell-specific ageing pattern in otherwise healthy men. *Aging Cell* 19(10), e13242 (2020).
19. Denomme MM, Haywood ME, Parks JC, Schoolcraft WB, Katz-Jaffe MG. The inherited methylome landscape is directly altered with paternal aging and associated with offspring neurodevelopmental disorders. *Aging Cell* 19(8), e13178 (2020).
20. Ben Maamar M, King SE, Nilsson E, Beck D, Skinner MK. Epigenetic transgenerational inheritance of parent-of-origin allelic transmission of outcross pathology and sperm epimutations. *Dev. Biol.* 458(1), 106–119 (2020).
21. Jenkins TG, James ER, Aston KI *et al.* Age-associated sperm DNA methylation patterns do not directly persist trans-generationally. *Epigenetics Chromatin* 12(1), 74 (2019).
22. WHO. *WHO Laboratory Manual for the Examination and Processing of Human Semen*. WHO, Geneva, Switzerland (2021).
23. Anamthakmakula P, Winuthayanon W. Mechanism of semen liquefaction and its potential for a novel non-hormonal contraception. *Biol. Reprod.* 103(2), 411–426 (2020).
24. Long S, Kenworthy S. Round cells in diagnostic semen analysis: a guide for laboratories and clinicians. *Br. J. Biomed. Sci.* 79, 10129 (2022).
25. Fedder J. Nonsperm cells in human semen: with special reference to seminal leukocytes and their possible influence on fertility. *Arch. Androl.* 36(1), 41–65 (1996).
26. Di Persio S, Leitão E, Wöste M *et al.* Whole-genome methylation analysis of testicular germ cells from cryptozoospermic men points to recurrent and functionally relevant DNA methylation changes. *Clin. Epigenetics* 13(1), 160 (2021).
27. Luján S, Caroppo E, Niederberger C *et al.* Sperm DNA methylation epimutation biomarkers for male infertility and FSH therapeutic responsiveness. *Sci. Rep.* 9(1), 16786 (2019).
28. Lee HY, Jung SE, Oh YN, Choi A, Yang WI, Shin KJ. Epigenetic age signatures in the forensically relevant body fluid of semen: a preliminary study. *Forensic Sci. Int. Genet.* 19, 28–34 (2015).
- **One of the first studies developing age prediction models from sperm samples; identified loci are the basis for many subsequent studies.**
29. Weidner CI, Lin Q, Koch CM *et al.* Aging of blood can be tracked by DNA methylation changes at just three CpG sites. *Genome Biol.* 15(2), R24 (2014).

30. Pilsner JR, Saddiki H, Whitcomb BW *et al.* Sperm epigenetic clock associates with pregnancy outcomes in the general population. *Hum. Reprod.* 37(7), 1581–1593 (2022).
- **Developed an age prediction model and associated higher age with a lower probability of achieving pregnancy.**
31. Orioli D, Dellambra E. Epigenetic regulation of skin cells in natural aging and premature aging diseases. *Cells* 7(12), 1–30 (2018).
32. Simpson DJ, Chandra T. Epigenetic age prediction. *Aging Cell* 20(9), e13452 (2021).
33. Heidegger A, Pisarek A, de la Puente M *et al.* Development and inter-laboratory validation of the VISAGE enhanced tool for age estimation from semen using quantitative DNA methylation analysis. *Forensic Sci. Int. Genet.* 56, 102596 (2022).
34. Pisarek A, Pośpiech E, Heidegger A *et al.* Epigenetic age prediction in semen – marker selection and model development. *Aging (Albany NY)* 13(15), 19145–19164 (2021).
- **Developed an age prediction model suitable for forensic practice using a sample set larger than 150 samples.**
35. Milekic MH, Xin Y, O'Donnell A *et al.* Age-related sperm DNA methylation changes are transmitted to offspring and associated with abnormal behavior and dysregulated gene expression. *Mol. Psychiatry* 20(8), 995–1001 (2015).
36. Prell A, Sen MO, Potabattula R *et al.* Species-specific paternal age effects and sperm methylation levels of developmentally important genes. *Cells* 11(4), 1–12 (2022).
37. Kayser M, Schneider PM. DNA-based prediction of human externally visible characteristics in forensics: motivations, scientific challenges, and ethical considerations. *Forensic Sci. Int. Genet.* 3(3), 154–161 (2009).
38. Freire-Aradas A, Pośpiech E, Aliferi A *et al.* A comparison of forensic age prediction models using data from four DNA methylation technologies. *Front. Genet.* 11, 932 (2020).
39. Paparazzo E, Lagani V, Geracitano S *et al.* An ELOVL2-based epigenetic clock for forensic age prediction: a systematic review. *Int. J. Mol. Sci.* 24(3), 1–14 (2023).
40. Daunay A, Baudrin LG, Deleuze JF, How-Kit A. Evaluation of six blood-based age prediction models using DNA methylation analysis by pyrosequencing. *Sci. Rep.* 9(1), 8862 (2019).
41. Lee JW, Choung CM, Jung JY, Lee HY, Lim SK. A validation study of DNA methylation-based age prediction using semen in forensic casework samples. *Leg. Med. (Tokyo)* 31, 74–77 (2018).
42. Zbiec-Piekarska R, Spolnicka M, Kupiec T *et al.* Development of a forensically useful age prediction method based on DNA methylation analysis. *Forensic Sci. Int. Genet.* 17, 173–179 (2015).
43. Thong Z, Chan XLS, Tan JYY, Loo ES, Syn CKC. Evaluation of DNA methylation-based age prediction on blood. *Forensic Sci. Int. Genet. Suppl. Ser.* 6, e249–e251 (2017).
44. Cho S, Jung SE, Hong SR *et al.* Independent validation of DNA-based approaches for age prediction in blood. *Forensic Sci. Int. Genet.* 29, 250–256 (2017).
45. Li L, Song F, Lang M *et al.* Methylation-based age prediction using pyrosequencing platform from seminal stains in Han Chinese males. *J. Forensic Sci.* 65(2), 610–619 (2020).
46. Vidaki A, Kayser M. From forensic epigenetics to forensic epigenomics: broadening DNA investigative intelligence. *Genome Biol.* 18(1), 238 (2017).
47. Aliferi A, Ballard D, Gallidabino MD, Thurtle H, Barron L, Syndercombe Court D. DNA methylation-based age prediction using massively parallel sequencing data and multiple machine learning models. *Forensic Sci. Int. Genet.* 37, 215–226 (2018).
48. Bendixen KK, Mindegaard M, Epistolio S *et al.* A qPCR technology for direct quantification of methylation in untreated DNA. *Nat. Commun.* 14(1), 5153 (2023).
49. Cao M, Shao X, Chan P *et al.* High-resolution analyses of human sperm dynamic methylome reveal thousands of novel age-related epigenetic alterations. *Clin. Epigenetics* 12(1), 192 (2020).
50. Jenkins TG, Aston KI, Cairns B, Smith A, Carrell DT. Paternal germ line aging: DNA methylation age prediction from human sperm. *BMC Genomics* 19(1), 763 (2018).