Paternal DNA packaging in spermatozoa: more than the sum of its parts? DNA, histones, protamines and epigenetics

David Miller, Martin Brinkworth¹ and David Iles²

Division of Reproduction and Early Development, Leeds Institute of Genetics, Health and Therapeutics, University of Leeds, Clarendon Way, Leeds LS2 9JT, UK, ¹Medical Biosciences Research Focus Group, Division of Biomedical Sciences, University of Bradford, Bradford BD7 1DP, UK and ²Faculty of Biological Sciences, Institute of Integrative and Comparative Biology, University of Leeds, Clarendon Way, Leeds LS2 9JT, UK

Correspondence should be addressed to D Miller; Email: d.miller@leeds.ac.uk

Abstract

Haploid male germ cells package their DNA into a volume that is typically 10% or less that of a somatic cell nucleus. To achieve this remarkable level of compaction, spermatozoa replace most of their histones with smaller, highly basic arginine and (in eutherians) cysteine rich protamines. One reason for such a high level of compaction is that it may help optimise nuclear shape and hence support the gametes’ swimming ability for the long journey across the female reproductive tract to the oocyte. Super-compaction of the genome may confer additional protection from the effects of genotoxic factors. However, many species including the human retain a fraction of their chromatin in the more relaxed nucleosomal configuration that appears to run counter to the ergonomic, toroidal and repackaging of sperm DNA. Recent research suggests that the composition of this ‘residual’ nucleosomal compartment, a generally overlooked feature of the male gamete, is far more significant and important than previously thought. In this respect, the transport and incorporation of modified paternal histones by the spermatozoon to the zygote has been demonstrated and indicates another potential paternal effect in the epigenetic reprogramming of the zygote following fertilisation that is independent of imprinting status. In this review, the most recent research into mammalian spermatozoal chromatin composition is discussed alongside evidence for conserved, non-randomly located nucleosomal domains in spermatozoal nuclei, all supporting the hypothesis that the spermatozoon delivers a novel epigenetic signature to the egg that may be crucial for normal development. We also provide some thoughts on why this signature may be required in early embryogenesis.

Reproduction (2010) 139 287–301

Introduction

All sexually reproducing species depend on the successful fusion between haploid male and female gametes to generate new offspring. In many species, including our own, this process frequently fails for a whole host of reasons that are beyond the scope of this review (Balen 2008). We do know, however, that the paternal nuclear genome in all eutherian mammals is delivered to the oocyte in a far more condensed chromatin state than exists in the nucleus of the oocyte and all somatic cell nuclei. The sperm of many species also deliver a complex RNA package, the role of which is currently unknown but is suspected of carrying epigenetic information for the zygote. Interested readers are directed elsewhere for more comprehensive discussions of this phenomenon (Miller & Ostermeier 2006, Miller 2007).

The greater than tenfold compaction of sperm chromatin is achieved during the final post-meiotic phases of spermatogenesis when the normally nucleosome-bound DNA is almost completely replaced by protamines (Braun 2001). Nucleosomes are octomeric complexes consisting of two copies each of the four core replication-dependent canonical DNA-binding histones, H2A, H2B, H3 and H4. Each nucleosome (sometimes referred to as a core particle) can wrap 146 bp of DNA in a helical pattern around itself, although most studies on chromatin digested by endonucleases reveal an ~200 bp repeating pattern of associated DNA due to the presence of linker DNA strands running between adjacent core particles (Lilley & Pardon 1979). Exhaustive digestion of chromatin with micrococcal nuclease or DNase 1 produces mononucleosomes, without the linker DNA (Sanders 1978, Bryan et al. 1979). Core particles and the way they facilitate the packaging and condensation of DNA in the eukaryotic nucleus is essentially the same from yeast to humans and the iconic electron micrographs of the mid 1970s, showing 10 nm (beads on a string) and 25–30 nm fibres, were the first solid evidence that eukaryotic chromatin exists in different packaging states (Kornberg 1974, Olins
Only unicellular dinoflagellates appear to lack histones and use a completely different process to organise their chromatin (Moreno Diaz de la Espina et al. 2005). In its most condensed state during interphase (higher condensation states are achieved in mitosis and meiosis, the only times when eukaryotic chromosomes are microscopically visible without the aid of dyes or stains. The only exceptions to this general rule are for the giant polytene and lambrush chromosomes of flies and frogs that are visible in some of their interphase nuclei) nucleosomal histones compact genomic DNA some 10\(^3\)-fold by facilitating folding and looping in association with the nuclear scaffold (Razin et al. 2007). This compaction is insufficient for the far smaller sperm nucleus, however, which can package the paternal genome some 10\(^6\)-fold. The repackaging process itself is directed by an elaborate process during the post-meiotic, spermiogenic phase of spermatogenesis, where nucleosomes are destabilised, possibly by massive hyperacetylation of histones (dramatically reducing the core particles’ affinity for DNA by reducing charge difference between them) and then replaced firstly by transition proteins (TPs) and then by protamines (McLay & Clarke 2003, Pivot-Pajot et al. 2003, Kurtz et al. 2007). The generalised substitution of a histone (nucleosome) based chromatin configuration to one based on protamines or protamine-like proteins has arisen many times in evolutionary history (Ausio et al. 2007). In mammalian spermatozoa, chromatin packaging probably undergoes continued ‘maturation’ during facilitated passage along the epididymis, where inter- and intra-protamine disulphide bonds are established to ‘lock’ the chromatin in its final, virtually crystalline state (Huang & Nieschlag 1984, Golan et al. 1996). In the ooplasm, maternal factors rapidly access the paternal chromatin and begin the replacement of protamines with maternal histones before paternal pronuclear formation is possible. To date, one such ooplasmic factor, HIRA, which was identified originally in the yeast Saccharomyces cerevisiae as a repressor of histone expression (Sherwood et al. 1993) has been implicated in this reversal process as its absence in the oocyte blocks male pronuclear formation in flies (Loppin et al. 2005). Its human orthologue, formerly known as TUPLE1, is a candidate gene for DiGeorge syndrome (Wilming et al. 1997). HIRA is a chaperone for histone H3.3 (present in spermatozoa and oocytes) where it is likely to perform equivalent functions in the early mammalian zygote to that proposed in the zygot es of the fly.

Once the remarkable process of sperm chromatin super-condensation is concluded, the casual reader might be forgiven for thinking that maternal factors are left to orchestrate the entire post-zygotic developmental process. However, in mammals at least, we know that gene imprinting gives the paternal genome a measure of autonomy during zygotic gene expression events that can have far reaching developmental effects and consequences (reviewed in Schaefer et al. 2007) and indeed seems also to be necessary for normal spermatogenesis (Marques et al. 2008). These imprinting effects arise through differential methylation of imprinting control regions present at many loci that either suppress or allow gene transcription (Chong & Whitelaw 2004), ensuring that either the paternal or maternal allele is expressed, but not both. Gene imprinting is already established in the fertilising spermatozoon and is not generally affected by the global demethylation of the paternal genome that occurs prior to syngamy. We know that imprinted regions are marked long before they are repackaged during spermiogenesis and that aberrant patterns of methylation in sperm DNA are associated with an infertile phenotype (Kerjean et al. 2000, Housshodaran et al. 2007). However, our understanding of how methylation is established and maintained during the dynamic process of chromatin re-packaging is poor. Moreover, the sperm chromatin of many species is far more highly ordered in the nucleus than would seem necessary for a genome that is simply awaiting delivery to the egg (see below).

This review considers some of the more recent reports examining the unexpected complexity of sperm chromatin. Based on this fresh evidence, we can conclude that the paternal genome is in a ‘frozen’ dynamic state that preserves a novel, post-fertilisation, epigenetic function in the developing embryo that is related somehow to the developmental competence of the new zygote. The focus of this review is on mammalian spermatozoa, although the gametes of other species will be considered in context where appropriate.

### The chromatin of mature spermatozoa

#### From nucleosome to toroid

As indicated in the Introduction, many species, including humans, have solved the problem of how best to fit their paternal DNA into as small a nucleus as possible by switching from a histone to a protamine based DNA packaging configuration (Balhorn 2007). Protamines are arginine and (in eutherians) cysteine-rich proteins that have arisen independently in the genera from different phyla, most likely as a result of a simple frame-shift mutation in an ancestral H1 gene (Ausio et al. 2007). There are two proteins (protamine 1 and protamine 2) involved in repackaging the sperm DNA of most species, although some species, notably bovine and porcine and some primates, while containing genes for the two variants, only express (or use) one of them (Maier et al. 1990, Queralt et al. 1995). The positive selection favouring the formation of a nucleoprotamine based chromatin configuration probably arose because it packages DNA over ten times more efficiently than nucleohistones. In a nucleus that is essentially shut-down, such extreme compaction makes good sense
because the spermatozoon, which should have no need to express its genome during the final stage of its development, does need to ensure that its cargo is packed as efficiently as possible to facilitate safe delivery to the egg. Using biochemical, time lapse and atomic force imaging techniques, the pioneering work of Balhorn showed how protamines (specifically PRM1 (P1)) that associate with ∼10–11 nucleotides per monomer, can complex with and coil 50 kb of naked DNA ex vivo to form toroidal structures of 60–100 nm diameter and a thickness of ∼20 nm in a fashion somewhat analogous to the windings of a copper wire around an electrical transformer core (Brewer et al. 1999, 2003). By stacking these toroids, the sperm nucleus achieves a higher efficiency in packing the paternal genome, reducing the size of the sperm nucleus to an absolute minimum. As head shape and size are known to affect sperm motility and function (Ostermeier et al. 2001, Malo et al. 2006, Ausio et al. 2007, Gillies et al. 2009), it is likely that nuclear dimensions are an important requisite in facilitating optimal head shape and that efficient compaction of the paternal genome facilitates this optimisation. Hence, the evolutionary pressure to substitute histones with protamines in the furtherance of this optimisation process is likely to be strong.

Strategies have evolved to compensate for the inevitable switch-off of gene transcription that accompanies the histone to protamine transition in late spermiogenesis. In general, transcription and translation are temporally uncoupled during spermiogenesis (the post-meiotic stage of spermatogenesis when round cells are extensively remodelled to form mature spermatozoa). The phenomenon of delayed translation is also observed in the oocyte and is known to occur during spermiogenesis in all species studied that substitute histones for protamines, including fruit flies (whose sperm contain a protamine-like protein; McKay et al. 1986, Rathke et al. 2007). There is no detriment associated with expression from just one allele per cell because spermatids are connected by cytoplasmic bridges that allow the sharing of products generated by other spermatids in the testicular syncytium (Dym & Fawcett 1971). Once the replacement of histones commences, spermatids rely solely on translation of existing mRNA to support protein synthesis (including the synthesis of TPs and protamines) as there is no longer any possibility of maintaining transcription (Zakari et al. 1988, Steger 1999). The requirement for TPs in this process is not an absolute, as mice deficient in their expression remain fertile, albeit with smaller litter sizes (Zhao et al. 2001, 2004). Indeed, the sperm from animals heterozygous or homozygous for deletions of the Tp genes show a spectrum of abnormalities ranging from relatively normal to grossly abnormal, but even double knockouts are not completely infertile as offspring can arise following ICSI using sperm from Tp<sup>1<sup>−/−</sup>/Tp2<sup>−/−</sup></sup> mice (Shirley et al. 2004). The primary effect of Tp disruption appears to be a disruption of protamine 2 processing, leading to grossly abnormal levels in affected sperm (Yu et al. 2000). Protamines, however, are essential for normal fertility, as mice do not tolerate haplo-insufficiency of either PRM1 or PRM2 and are severely subfertile or infertile (de Boer et al. 1990, Lee et al. 1995, Cho et al. 2001, Oliva 2006). Interestingly, protamine insufficiency also leads to much higher levels of sperm DNA strand breakage as assessed by the comet assay, suggesting that (irreparable) DNA damage is the main cause of implantation failure in embryos derived from healthy eggs fertilised by protamine compromised sperm (Aoki et al. 2005, Ramos et al. 2008). Together, these data suggest that the additional (super) compaction afforded by sperm protamines confers a measure of protection against DNA damage that may be critical for successful fertilisation.

Are histones replaced or displaced during spermiogenesis?

One of the first signs that DNA repackaging is imminent is a massive increase in the level of acetylation of core histones as determined by immunocytochemistry and western analysis. However, the incorporation of non-canonical, replication-independent testis-specific histone variants into the nucleosomes of developing spermatocytes was recognised many years ago (Marushige & Dixon 1971, Tanphaichitr et al. 1978, Gusse & Chevaillier 1980b, Grimes & Henderson 1983, Trostle-Weige et al. 1984, Meistrich et al. 1985, Nickel et al. 1987, Poccia et al. 1987, Rousseaux-Prevost et al. 1988, Palmer et al. 1990), suggesting that a displacement strategy precedes global replacement. Readers are directed to the excellent review by Churikov et al. (2004a) for a more detailed summary of germline-specific histones in mouse and human spermatogenesis. Here we shall focus on their persistence in the mature spermatozoon.

More recently, H2AX and its phosphorylated form, γ-H2AX have been shown to associate with the sex body of meiotic spermatocytes. The sex body is an interesting structure that accommodates and is closely involved in the inactivation of the sex chromosomes during spermiogenesis. H2AX is essential for sex body formation in that null (male) mice fail to develop this organelle and are infertile (Fernandez-Capetillo et al. 2003). The sex body is not to be confused with the chromatoid body, or ‘nuage’ which is a distinct albeit possibly related structure involved in the post-meiotic control of gene expression (Kotaja et al. 2006).

It is generally agreed (although direct evidence is lacking) that once stripped from the DNA, the canonical nucleosomal histones of condensing spermatids are recycled by the cells’ proteasomal machinery that is intimately involved in the protein turnover kinetics of all cell types, including germ cells (Sutovsky et al. 2003, Haraguchi et al. 2005, Khor et al. 2006).
Indeed, a mutation in one of the testis-specific ubiquitin conjugating enzymes, HR6B, which normally tags proteins for degradation in the proteasomal complex, confers an infertile phenotype resulting from maturation arrest and a general apoptosis of round spermatids (Roest et al. 1996). However, the evidence to date suggests that the replacement of histones by protamines in the mature spermatozoon is incomplete, at least in some mammalian species. The core histones are detected in human (Gatewood et al. 1987, Zalenskaya et al. 2000) and murine (Pittoggi et al. 1999, Govin et al. 2007) spermatozoa, for example by western blot analysis of total acid extracts. Many testis-specific histone variants are also present in human spermatozoa, including: TH2b, which closely associates with telomeres (Churikov et al. 2004a, 2004b); H2AX, so important in sex body formation (Li et al. 2006); and the target for HIRA, H3.3 (van der Heijden et al. 2008; Fig. 1A). The perinuclear histones of bovine sperm may be a special case (see below; Tovich & Oko 2003), but even the heavily condensed sperm chromatin of this species contains a variant of H3, CENPA, which has been shown to decorate centromeres (Palmer et al. 1990; Fig. 1B). Perhaps the most striking of all mammalian sperm investigated so far are the dysmorphic spermatozoa of the dasyurid, Sminthopsis crassicaudata, which have extensive regions in the nuclear periphery that are in a more open and ‘poorly’ condensed configuration compared with the main body of the nucleus (Soon et al. 1997; Fig. 1C). These two compartments or zones, known as C2 and C1, are histone- and protamine-rich respectively, with the presence of typical somatic nucleosomes in the C2 zone only. Interestingly, the C2 zone (and the accompanying dysmorphic nuclear morphology) disappears on the addition of exogenous protamines, showing that the differential packaging exhibited by the sperm of this species is rather plastic.

Studies in the mouse have shown that core histones (and histone-bound DNA) are retained in a peripheral location in mouse sperm nuclei (Pittoggi et al. 1999, Solove’eva et al. 2004), although this location may have been preferentially highlighted by incomplete decondensation of the nuclei prior to the application of antibodies, thus masking their presence deeper within the nucleus. Other reports have localised histones to the annular region of human sperm nuclei (Li et al. 2008). Finally, core histones have been reported to be in association with the perinuclear theca in bovine spermatozoa although again, limited decondensation may contribute to an apparent absence of histones in the nucleus itself (Tovich & Oko 2003). This highlights a limitation of in situ detection methods in this context, whereby incomplete chromatin decondensation may yield false negative results, while full decondensation will distort the architecture that is necessary to inform the location of these proteins.

Reports on the presence of nucleosomal-like structures in human spermatozoa date back to the 1970s and classical biochemical, electron microscopic techniques (Dixon 1972, Puwaravutipanich & Panyim 1975, Philippe & Chevailler 1978, Tanphaichitr et al. 1978, Gusse & Chevailler 1980a, Zalenskaya et al. 1981, Sobhun et al. 1982) and atomic force microscopy (Hud et al. 1993, Allen et al. 1996, Joshi et al. 2000, Nazarov et al. 2008) have been used to demonstrate the presence of toroids and core particles in decondensed human sperm nuclei. Complementary biochemical evidence has shown differences between the sequence composition of histone and protamine-enriched chromatin from human sperm nuclei (Gatewood et al. 1987) and that DNA sequences from the β-globin gene cluster as expressed in the embryo or the adult, are enriched in the histone or protamine compartments of sperm chromatin, respectively (Gardiner-Garden et al. 1998, Wykes & Krawetz 2003). These studies provided the first evidence of a preference for packaging into different domains in sperm chromatin, with more open (nucleosomal) conformations being favoured sites for developmentally regulated genes.

Such reports made use of salt extraction methods to remove histones, followed by restriction digestion of sperm nuclei to release the formerly histone bound DNA. Others used combinations of micrococcal nuclease digestion to release intact nucleosomes from sperm nuclei. Southern blot analysis of the released DNA indicated that human sperm histones package telomeric DNA, arguing that this DNA must have occupied a peripheral location in the nucleus for it to be accessible to the enzyme (Zalenskaya et al. 2000). Similar techniques

Figure 1 Presence and identity of histones in mature spermatozoa. (A) Histone H3 visualised by a pan-H3 antibody (red) and an antibody recognising H3.1/3.2 (green) in two fully decondensed human sperm nuclei. From van der Heijden et al. (2008). Scale bar 10 μm. (B) Localisation of CENPA to centromeric foci of bovine spermatozoa. The upper panel shows immunolabelling of CENPA alone; the lower panel is dually labelled showing the same spermatozoon additionally stained for DNA. From Palmer et al. (1990), with permission. Scale bar 5 μm. (C) Detection of H4 at the periphery of spermatozoa from the marsupial Sminthopsis crassicaudata. The upper panel shows a spermatozoon labelled with antibodies to H4 (green) and stained for DNA (red). The lower panel is stained for DNA only. The H4-bright signal corresponds to the C2 region described by Soon et al. (1997). Scale bar 5 μm.
have been used in numerous reports appearing over the past thirty years, aimed at understanding how conformational changes (reflecting differences in relative accessibility of chromatin domains to salt solutions and/or endonucleases) in somatic and sperm cell chromatin relate to changes in gene expression (Sanders 1978, Davie & Saunders 1981, Zentgraf & Franke 1984, Olivares et al. 1993). Taken together, the evidence suggests that while sperm chromatin may be transcriptionally inactive, it is organised into differentially packaged domains or compartments that resemble the more dynamic configurations normally observed in actively expressed chromatin.

Recent work describing the repackaging process in Drosophila has shown that the post-meiotic replacement of somatic histones by protamine-like proteins also occurs during spermiogenesis in a process that shows striking parallels with the dynamics observed in mammals (Rathke et al. 2007; Fig. 2A). Even post-meiotic gene expression, once thought to be a peculiarity of mammalian spermatogenesis, occurs in the fruit fly when a set of genes encoding proteins involved in sperm individualisation is expressed (Barreau et al. 2008). Fruit flies, and presumably other dipterans that repackage their sperm DNA, use a protamine-like protein to provide highly condensed chromatin and similar strategies must have evolved to circumvent the shutting down of gene transcription that accompanies nuclear condensation. It has been assumed that all histones are replaced, but proteomic data suggest that even the needle-like sperm of Drosophila retain some histones (Dorus et al. 2006 and (unpublished) data from our laboratory has imaged nuclear structures containing an H2A variant by the fluorescence of sperm nuclei from transgenic flies expressing a red-fluorescence protein-H2A fusion protein (Fig. 2B; courtesy of Helen White-Cooper).

In conclusion, nucleosomes are certainly not completely removed from mature ejaculate spermatozoa, although estimates for their contribution to sperm chromatin range from 1% in the mouse (Balhorn et al. 1977), to 15% in the human (Gatewood et al. 1990) and over 50% in some marsupial species (Soon et al. 1997). At first sight, such widely varying contributions suggest that histone retention, if it has some important function, cannot be related to a strictly enforced stoichiometry between nucleoprotamine and nucleosomes in sperm nuclei. And yet, the very existence of species-dependent stoichiometric differences strongly suggests some function. What could that function be?

![Figure 2](https://www.reproduction-online.org)  
**Figure 2** Spermiogenesis in the fruit fly, Drosophila. (A) Late stages in Drosophila spermiogenesis visualised indirectly using an anti-core histone antibody (upper panels) or directly by nascent fluorescence of a protamine-green fluorescing fusion protein (lower panels). The four left-hand panels show young elongating and early canoe stage nuclei. The corresponding right-hand panels show late canoe stage and individualising (separating) nuclei. In these pictures, histones are apparently lost from the later stages of spermatogenesis and replaced by protamines. From Rathke et al. (2007), with permission. Scale bar 5 μm. (B) Isolated Drosophila sperm from transgenic flies harbouring an H2Av-red fluorescing fusion protein. These post-individualised sperm retain histones as indicated by the bright nuclear foci (arrows), suggesting that the complete absence of histones in the later stages of spermatogenesis shown in panel A reflects the inability of antibodies to gain access to the highly condensed chromatin. Pictures with kind permission of Dr Helen White-Cooper, University of Cardiff.

---

**Paternal codes in sperm chromatin**

*Gene imprinting and methylation*

It has been over twenty years since epigenetic differences between parental genomes were suggested to underlie the failure of parthenogenetic mammalian zygotes to survive through embryonic development (Surani & Barton 1983, Banerjee & Smallwood 1995). We now know of at least 80 imprinted gene loci, with the reciprocally imprinted H19/Igf2 and Dlk1-Gtl2 loci (Wu et al. 2006), paternally imprinted Igf2 receptor gene and maternally imprinted Peg1 genes being among the most significant for embryonic development and placentation. Control of expression is mediated by the methylation status of imprint control regions that ensures either the maternal or paternal allele is appropriately expressed. Readers are directed to some excellent reviews on this process for more detail (Trasler 1998, Chong & Whitelaw 2004, Ooi & Henikoff 2007, Rousseaux et al. 2008), but for our purposes, it is sufficient to say that all imprinting signals are erased shortly after primordial germ cells enter the presumptive gonad of the developing embryo and that in males but not females, imprints are probably already established in gonocytes (Luciiero et al. 2002, Delaval et al. 2007).
A more global methylation of parental DNA also occurs during DNA condensation in elongating spermatids and in early germinal vesicle stage oocytes (Marchal et al. 2004). This methylation probably serves to help shut down the genome in both gametes, which both rely on stored mRNAs to support further protein synthesis. In most mammalian species, sperm DNA is rapidly and actively demethylated following entry into the ooplasm and prior to pronuclear formation while the process is delayed in the female genome; only the imprinted genes escape demethylation (Mclay & Clarke 2003, Santos et al. 2005). Recent studies have shown that the patterns of sperm DNA methylation closely resemble those of pluripotent ES cells and that several embryologically important developmental gene sequences including members of the HOX gene clusters are largely unmethylated in both cell types (Weber et al. 2007, Farthing et al. 2008, Arpanahi et al. 2009, Hammad et al. 2009). Imprinting errors in the paternal genome are likely to be a cause of significant morbidity to the developing embryo and there is some evidence for a small increase in imprinting disorders among the offspring of people taking advantage of assisted conception technologies (Li et al. 2005, Marques et al. 2008). Related to this is the finding that an aberrant pattern of DNA methylation in the sperm of some men correlates with known pregnancy failures in their partners (Benchabib et al. 2005). However, methylation is also an indicator of nuclear condensation and as incomplete or aberrant condensation of sperm chromatin is associated with higher levels of DNA damage (Aoki et al. 2005, Aitken & De Iuliis 2007) it is also possible that pregnancy failure arises from the DNA damage.

Is there a sperm-derived histone code?

The histone code is now a recognised additional layer of (epigenetic) information that somatic cells use to help control gene expression. At its simplest level, it involves replacement of many of the canonical, replication-dependent histones with non-canonical and replication-independent variants for various cellular functions (see Sims & Reinberg (2008) and Munshi et al. (2009) for recent reviews). These variants include the testis-specific histones mentioned earlier. However, more complex levels of gene expression control are possible by modification of core histones de novo in various combinations, such that they change their affinity for DNA in subtle ways and become selectively attractive or repellent for chromatin modifiers and transcriptional regulators (Huebert et al. 2006). Disregarding ubiquitination, the three most significant modifications associated with a histone-based epigenome include acetylation (ac), methylation (me) and phosphorylation (P), of core histones at their exposed (within their containing nucleosome) amino-terminal ends. These modifications can come in singlets, pairs (dimethylation, for example) and triplets (triacetylation) and in any combination thereof. Acetylation of lysine 5 on H4 (H4K5ac) and of lysines 9 and 14 on H3 (H3K9K14ac) accompanied by methylation of lysine 4 or 36 on H3 (H3K4me3; H3K36me3) and arginine 3 on H4 (H4R3me) is associated with an openly accessible (to salt and endonucleases), and hence transcriptionally poised or expressing, euchromatin structure. In contrast, inactive, closed and condensed heterochromatin is relatively de-acetylated (except for H4K12ac) and is extensively methylated on H3, lysine 9 (H3K9me3).

There has been renewed interest in both the methylation and histone levels of epigenetic information because of the unique potential for the sperm to deliver to the egg a paternal signature with possible developmental consequences beyond gene imprints. Selective gene imprinting anomalies are certainly one reason for the currently abysmal success rates for somatic cell cloning (Paterson et al. 2003), but the loss of a ‘higher order’ epigenetic signature may be even more difficult to recapitulate or reprogramme in a differentiated and committed somatic cell nucleus.

With respect to testis- and sperm-specific core-histone variants, three proteins, namely H2AL1, H2AL2 and H2BL1 have been discovered in late elongating mouse spermatids and spermatozoa, where they localise to a nucleosome enriched, heterochromatic chromocentre located deep within the murine sperm nucleus (Govin et al. 2007; Fig. 3A). They are interesting because prior to condensation (and the appearance of these new variants), the pericentric heterochromatin of elongating spermatids contains both acetylated H4 (H4K5,8,12ac) and methylated H3 (H3K9me) histones, indicative of more relaxed euchromatin domains (van der Heijden et al. 2006). These modified histones are reported to form slightly smaller nucleosomes in sperm nuclei that apparently lack either H3 or H4 and repackage at least some of the pericentric/chromocentric DNA, providing evidence for novel nucleosome-like complexes in sperm nuclei that are delivered to the egg at fertilisation. Related evidence based on following the fate of modified paternal histones in sperm and in zygotes, suggests that sperm H4Ac is either not lost from the sperm during pericentric condensation or is present in a separate, non-pericentric compartment, possibly located in the posterior of the sperm (Li et al. 2008; Fig. 3C) or peripheral nuclear regions (Pittoggi et al. 1999; Fig. 3D), depending on the species. Either way, H4K8ac and H4K12ac (and H2AX, H3.3) can be detected in the murine zygote prior to full decondensation of the sperm nucleus and prior to any substitutions by maternal factors, indicating that they must have originated from the sperm itself (van der Heijden et al. 2006). Whether these paternally derived histones contribute to and persist in zygotic chromatin, however, is currently unresolved. Both H2AL1/2 rapidly disappear after fertilisation in the mouse (Wu et al. 2008) and while H3.1/H3.2 persists...
in human/mouse zygotes prior to DNA replication, this could be an effect of the heterologous system used (van der Heijden et al. 2008).

**Codes within codes?**

Two recent studies have examined the composition of sperm chromatin from different angles and come to broadly similar conclusions. The first of these studies used a combination of chromatin immuno-precipitation (ChIP) with promoter chip arrays (ChIP-chip) and sequencing (ChIP-seq) to examine the composition of human sperm chromatin more extensively (Hammoud et al. 2009). This study reported that retained nucleosomes of human spermatozoa package sequences with a strong developmental ontology that includes imprinted genes, HOX genes and developmental transcription and signalling factors. Interestingly, methylated histone modifications also showed consistent patterns of enrichment, with H3K4me3 and H3K4me2 (marking more relaxed chromatin domains) present in these developmental loci (Fig. 4A). They also showed that many developmental gene promoters in human sperm are hypomethylated, which contrasts with the situation in the bulk of sperm DNA, which is generally hypermethylated (Table 1).

In the other study (Arpanahi et al. 2009), methods similar to those used by Gatewood et al. (1987) and Zalenskaya et al. (2000) were employed to generate soluble and insoluble DNA fractions from human and murine sperm chromatin that were used to interrogate whole genome comparative genome hybridisation (CGH) microarrays. One striking observation from the CGH data in this study was the tendency for sequences in soluble fractions to align closely with chromosome gene density maps (Fig. 4B). This study also suggests that HOX genes are enriched in soluble DNA fractions (Fig. 4C), which compares favourably with the more relaxed chromatin conformations expected from these regions based on their histone modifications. This relationship was pinpointed to promoter sequences, particularly those associated with or lying close to CTCF-binding sites (which showed highest enrichment in soluble DNA). Intragenic and intergenic sequences lacking CTCF-binding sites showed no clear packaging preference. CTCF is a multi-functional insulator protein that sets boundaries on the actions of enhancer and silencer elements and helps organise the genome into regulatory domains (Ohlsson et al. 2001). This activity extends to the establishment of gene imprints based on differentially methylated imprint control loci, which in both studies, are enriched in soluble DNA. Ablation of the protein in mouse eggs causes severe developmental abnormalities (Wan et al. 2008). Interestingly, soluble promoters also have a strong developmental ontology, while promoters enriched in the insoluble chromatin

---

**Figure 3** Location of some histone variants in mature sperm nuclei. (A) Dual immunodetection of H2AL1/L2 (green) in round (1–3), elongating (4–6), early condensing (7–12), condensing and condensed (13–18) mouse spermatids, co-detected with acetylated histone (red; 1–9) or protamine (red; 10–18). The lower panels correspond to the corresponding DAPI counterstaining. Arrows indicate the distribution of H2AL1/L2 on densely DAPI-stained pericentric heterochromatin. From Govin et al. (2007), with permission. Scale bar 5 μm. (B) Detection of core histones (A, A'), H4K8Ac (B, B') and H4K12 (C, C') by immunofluorescence (green) on fully decondensed mouse sperm nuclei. DNA is pseudo-coloured in red. From van der Heijden et al. (2006). Scale bar 10 μm. (C) Co-detection of core histones (green) in fully decondensed human sperm nuclei alongside protamine 1 (upper panels; red) and protamine 2 (lower panels; red). DNA counterstaining is shown in the rightmost panels. The leftmost panels are merged images from the dual labelling experiments including the counter stain. A clear focus of histone label at the posterior end of these nuclei is evident. From Li et al. (2008). (D) Localisation of H1 (green) to the periphery of mouse sperm nuclei. A merged image from H1-labeled and DNA stained (red) nuclei is shown in the right-hand panel. From Pittoggi et al. (1999), with permission.
Figure 4 (legend continued)
have a strong ontology for sensory stimuli, including olfactory receptor gene clusters. ChIP-chip using an antibody to H4K12ac also pulled down chromatin with promoter sequences representing a wider range of ontologies, including development, but these trends were not as strong as those derived by endonuclease accessibility. Together, the results of this experiment demonstrate that the regulatory sequences of developmentally important genes are strongly enriched in endonuclease-accessible sperm chromatin. Mouse spermatozoal chromatin is also organised in a similar way, indicating that the phenomenon is not restricted to humans and may be evolutionarily conserved.

These complementary studies appear to confirm and extend earlier findings (Gardiner-Garden et al. 1998, Wykes & Krawetz 2003) that sperm contain at least two differentially packaged chromatin domains, of which the minor nucleosomal is enriched in embryologically important developmental gene sequences or their regulatory regions. They also both demonstrate that methylation patterns are important in the two compartments, hinting at an undefined relationship between these two levels of epigenetic regulation in the spermatozoon.

The apparently favoured location of gene sequences in differentially packaged sperm DNA complements the surprising level of chromosomal organisation that is retained by these cells. Various reports have shown that the same chromosomes adopt the same preferential locations from cell to cell and are not just randomly distributed (Zalensky et al. 1995, Hazzouri et al. 2000). A similar, non-random distribution of differentially packaged DNA sequences must also be in place in the millions of sperm cells used in these studies (Arpanahi et al. 2009, Hammoud et al. 2009) otherwise no differences between soluble/insoluble or between

Figure 4 Whole genome findings from human sperm chromatin. (A) Nucleosome enrichment and modifications around and within the human HOXD gene cluster. The upper panel (a) shows clear enrichment of nucleosome-derived DNA sequence tags across this gene cluster from single (D1) and pooled semen donors (boxed). The lower panel (b) shows enrichment for modified histones. Particularly notable is the enrichment of nucleosomes containing H3K4me3 at this locus, indicating a more open euchromatic configuration. From Hammoud et al. (2009). (B) Scatter profile of CGH data obtained from human chromosome 16. Left-hand panels show scatter of log2 ratio values obtained from micrococcal (MND) and salt extracted and restriction-digested (SRD) sperm nuclei. Negative and positive log2 ratio values indicate insoluble (I) and soluble (S) DNA compartments. The gene density profile is shown in the right-hand panel. Note that the clear tendency for S scatter to follow this profile. (C) The same locus shown in A annotated in the UCSC browser (http://genome.ucsc.edu/) with CGH data indicating that it resides in a more openly accessible, endonuclease sensitive DNA fraction of human sperm nuclei (244K_salt_chip; positive log2 ratios; data from Arpanahi et al. (2009)). This locus also contains numerous sites for CTCF binding (CTCF_chip; data from Kim et al. (2007)) and many promoters for the HOXD genes are bound by nucleosomes containing H4K12ac (K4K12_prom; positive log2 ratios). These promoters are also unmethylated (methylation; negative log2 ratios; data from Weber et al. (2007)). From Arpanahi et al. (2009).
Table 1. Histones and their modifications reported in mature spermatozoa.

<table>
<thead>
<tr>
<th>Histone type</th>
<th>Modification</th>
<th>Species</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Core histones</td>
<td>N/A</td>
<td>Human, mouse, marsupial, fruit fly</td>
<td>Gatewood et al. (1990), Soon et al. (1997), Pittoggi et al. (1999) and Dorus et al. (2006)</td>
</tr>
<tr>
<td>H2AX</td>
<td>N/A</td>
<td>Mouse</td>
<td>Li et al. (2006)</td>
</tr>
<tr>
<td>H2A-L1, L2</td>
<td>N/A</td>
<td>Mouse</td>
<td>Govin et al. (2007)</td>
</tr>
<tr>
<td>H2AZ</td>
<td>N/A</td>
<td>Fruit fly</td>
<td>Dorus et al. (2006)</td>
</tr>
<tr>
<td>H2Av</td>
<td>N/A</td>
<td>Fruit</td>
<td>Dorus et al. (2006)</td>
</tr>
<tr>
<td>H3.3B</td>
<td>N/A</td>
<td>Mouse</td>
<td>Govin et al. (2007)</td>
</tr>
<tr>
<td>H2B-L1</td>
<td>N/A</td>
<td>Human</td>
<td>Zalensky et al. (2002)</td>
</tr>
<tr>
<td>TSH2B</td>
<td>N/A</td>
<td>Human</td>
<td>van der Heijden et al. (2008)</td>
</tr>
<tr>
<td>TH2B.175</td>
<td>N/A</td>
<td>Mouse/human</td>
<td>van der Heijden et al. (2008)</td>
</tr>
<tr>
<td>H3.1, 3.2, 3.3</td>
<td>N/A</td>
<td>Fruit fly</td>
<td>Dorus et al. (2006)</td>
</tr>
<tr>
<td>H3 3.1A, B</td>
<td>N/A</td>
<td>Bovine</td>
<td>Palmer et al. (1990)</td>
</tr>
<tr>
<td>H4r</td>
<td>N/A</td>
<td>Fruit-fly</td>
<td>Dorus et al. (2006)</td>
</tr>
<tr>
<td>H3K4 Me2, Me3</td>
<td>Me2, Me3</td>
<td>Human</td>
<td>Hammoud et al. (2009)</td>
</tr>
<tr>
<td>H3K9</td>
<td>Me</td>
<td>Mouse</td>
<td>van der Heijden et al. (2006)</td>
</tr>
<tr>
<td>H4K8, H4K12</td>
<td>Ac</td>
<td>Mouse</td>
<td>Govin et al. (2007)</td>
</tr>
</tbody>
</table>

ChIP-sample/input would have been revealed. Together with differences in the proportion of histones in sperm chromatin between different species, this may also suggest a reason for the species-specific morphology of the sperm of many animals.

Relevance of differential DNA packaging to embryonic development

Ten years ago, a remarkable study showed that it is possible to obtain viable offspring from murine oocytes fertilised merely by sperm nuclei divested of all external components and containing just an intact nuclear matrix and chromatin (Ward et al. 1999). It was later shown that an intact nuclear matrix containing ‘naked’ DNA is able to support male pronuclear formation and the subsequent condensation of morphologically normal paternal chromosomes although no further developmental progress was noted (Mohar et al. 2002). Hence, although maternal histones must be able to package paternal DNA successfully once in the oocyte, these experiments suggest that epigenetically modified paternal chromatin may still be required for normal embryonic development. By inference, a compromised matrix or chromatin configuration is not compatible with development.

The nuclear matrix is thought to act as the ‘skeletal’ framework on which chromatin is organised within the (somatic) nucleus. It is composed of highly insoluble lamin proteins interacting with RNA and the structure probably acts as a scaffold on which the molecular machinery of both DNA replication and transcription can be assembled in a tissue and cell-specific manner. Readers are directed to the following comprehensive reviews of Martelli et al. (1996), Hancock (2000), Barboro et al. (2003) and Albretsen et al. (2009) for more detailed information on the matrix and its controversial role in gene expression control mechanisms. Of course, the scaffold is unlikely to serve these purposes in mature spermatozoa because they do not replicate or transcribe the genome. Yet, the halo-competent nuclei used successfully to generate viable offspring (Ward et al. 1999) are very reminiscent of those that can be prepared from the nuclei of normally active somatic cell nuclei (Nadel et al. 1995, Ma et al. 1999, Iarovaia et al. 2004). The main difference appears to be length of the loop domains that form from stripped spermatozoal nuclei, which generally accommodate ~50 kb of DNA, equivalent to one toroid of nucleoprotamine, compared to the more usual 25–100 kb loops observed in active somatic cell nuclei. Moreover, various reports have given credence to the notion that, despite its ‘frozen’ state, mature spermatozoal chromatin has more openly accessible regions that could act as targets for maternal factors once in the ooplasm. These regions include histone-bound domains that package promoter sequences and support the argument that the organisation of differentially packaged DNA in the mature spermatozoal nucleus is closely tied to its matrix. Most studies into the nuclear dynamics of gene expression support the notion of a physically flexible genome where actively expressed genes not only become more accessible to endonucleases, but also move to new positions within the nucleus, presumably to be in closer proximity to the transcriptional machinery (Chambeyron & Bickmore 2004, Heard & Bickmore 2007). It is therefore tempting to speculate that the spermatozoal nucleus is frozen in a particular dynamic configuration that reflects its pending introduction to the ooplasm. This would emphasise still further the importance of differential DNA packaging as reinforced by the ontological signatures uncovered by the two studies described above (Arpanahi et al. 2009, Hammoud et al. 2009). Deviations from this dynamic due to DNA packaging anomalies are known to lead to infertile phenotypes in...
both human and mouse (Belokopytova et al. 1993, Tomsu et al. 2002, Aoki et al. 2006a, 2006b). One reason why defects in the nuclear matrix component have not come to our attention as a cause of male factor infertility is that such defects are likely to have a wider and more serious range of phenotypes than infertility alone. The lamin deficiency found in precocious senility, for example, is a good indicator that any defects in components of the nuclear matrix are likely to exclude reproductive function by default (Moulson et al. 2007).

The retention of paternal gene sequences having some potentially important embryological function in a more relaxed chromatin configuration begs the question of whether this chromatin is more susceptible to DNA damage than the bulk, protamine-packaged DNA. Sperm chromatin packaging anomalies are closely associated with poor (human) fertility outcomes and higher levels of DNA damage are an accompanying feature of dysfunctional sperm (Aoki et al. 2005). Such reports examined either total protamine content (Bench et al. 1996) or PRM1/PRM2 ratios in sperm populations (Steger et al. 2003, Aoki et al. 2006b) and by inference, the histone/protamine ratio is also probably disturbed in these men. We know that the activation of an endogenous endonuclease in murine sperm releases nucleohistones, confirming the greater accessibility of this compartment to digestion and by logical extension, to DNA damage (Pittoaggi et al. 1999). Experiments aimed at fully determining the sequence composition of this chromatin and/or of DNA released from sperm following exposure to an oxidative insult should help answer this question.

Conclusions and future directions

During spermiogenesis, spermatids repackage their DNA with protamines but in many species examined to date (and probably in all species), a small residue of histone-bound DNA is retained and gains access to the ooplasm. As epigenetic controls of gene expression exist at both DNA (methylation) and histone (methylation, acetylation and phosphorylation) levels, the entry of core histones alongside modified histones and histone variants to the egg leaves room for DNA and histone-based epigenetic signals that may be important for subsequent embryonic development. Recent evidence, arising from a closer examination of the composition of soluble (histone-bound) and insoluble (protamine-bound) domains in human and murine spermatozoa, indicates that the former chromatin is indeed the more significant contributor to a novel epigenetic signal in these cells (Gardiner-Garden et al. 1998, Wykes & Krawetz 2003, Moulson et al. 2007, Arpanahi et al. 2009, Hammoud et al. 2009). This global epigenome differs from the known imprinting processes in that it includes many more genes or their regulatory sequences and carries a very strong developmental ontology. It seems highly unlikely that this is coincidental. BrUTP incorporation in the one cell mouse embryo shows that the paternal pronucleus becomes highly transcriptionally active compared with the female pronucleus before activation of the embryonic genome (Aoki et al. 1997). Moreover, the observations that patterns of hypomethylated sequences packaged by sperm and ES cells are similar and that some of these hypomethylated sperm sequences are carried on modified histones in sperm chromatin that overlap closely with binding to hypomethylated pluripotency gene promoters in ES cells is particularly intriguing. These observations on pre and post (embryonic genome) activated zygotes may connect the differential packaging of sperm chromatin with the earliest developmental commitments in the differentiating embryo. They suggest that without the global spermatozoal epigenome to guide it, developmental progression is likely to be a haphazard affair, prone to many epigenetic errors that lead to non-viable embryos. If the ‘guidance’ information that the paternal chromatin provides is lost once pluripotent ES cells become committed to establishing the earliest cell lineages, the successful totipotent reprogramming of a somatic cell nucleus would be very difficult to achieve, as indicated by the woefully poor efficiency of current cloning strategies (Paterson et al. 2003).

Chromatin research has a long record of uncovering novel new mechanisms for the control of gene expression. It is therefore all the more astonishing that even a terminally differentiated cell like the spermatozoon should retain such complex chromatin architecture. Furthermore, despite the virtually complete replacement of histones by protamines in many species, current research suggests that some modified histones remain and that the strong ontological signature in the DNA that they carry to the egg could have developmental significance. At its simplest, perhaps sperm chromatin acts as a template for establishing the pluripotency of ES cells. Regardless, understanding the nature of those consequences must be one of the next goals for future research as it promises to reveal new insights into early embryonic events that rely on paternal templates. Like gene imprinting, this novel epigenetic marking of sperm chromatin may serve to ensure the continuity and requirement of a paternal contribution to the zygote. If so, it is likely to be a common feature of sexually reproducing metazoans that pre-dates the establishment of mammalian gene imprinting and hence versions of it (whether histone modifications and/or co-existing differential condensation states of sperm chromatin) should be found in lower animals, including non-mammalian vertebrates and invertebrates.
Declaration of interest

There is no conflict of interest that could be perceived as prejudicing the impartiality of the research/work reported.

Funding

The authors’ work as reported herein was supported in part by a grant from the British Biotechnology and Science Research Council (BB/S/B/04900).

Acknowledgements

The authors would like to thank the journals and/or the authors of the articles cited herein for permission to reproduce their work in this review. We thank Helen White-Cooper, University of Cardiff for permission to use the plate shown in Fig. 2B.

References

Aitken RJ & De Iuliis GN 2007 Origins and consequences of DNA damage in male germ cells. Reproductive Biomedicine Online 14 727–733.


Lee K, Haugen HS, Clegg CH & Braun RE 1995 Premature translation of protamine 1 mRNA causes precocious nuclear condensation and arrests spermatid differentiation in mice. PNAS 92 12451–12455.


Maier WM, Nussbaum GA, Domenjoud J, Klemm U & Engel W 1990 The lack of protamine 2 (P2) in boar and bull spermatozoa is due to mutations within the P2 gene. Nucleic Acids Research 18 1249–1254.


Ollshon R, Renkawitz R & Lobanenkov V 2001 CTCF is a uniquely versatile transcription regulator linked to epigenetics and disease. Trends in Genetics 17 520–527.


Rousseaux S, Reynoard N, Escoffer E, Thevenon J, Caron C & Khochbin S 2008 Epigenetic reprogramming of the male genome during gametogenesis and in the zygote. Reproductive Biomedicine Online 16 492–503.


Received 5 July 2009
First decision 4 August 2009
Revised manuscript received 21 August 2009
Accepted 16 September 2009