<table>
<thead>
<tr>
<th><strong>Title</strong></th>
<th>Epigenetic processes in the male germline</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Authors(s)</strong></td>
<td>O'Doherty, Alan; McGettigan, Paul A.</td>
</tr>
<tr>
<td><strong>Publication date</strong></td>
<td>2014-09</td>
</tr>
<tr>
<td><strong>Publication information</strong></td>
<td>Reproduction Fertility and Development, 27 : 725-738</td>
</tr>
<tr>
<td><strong>Publisher</strong></td>
<td>CSIRO Publishing</td>
</tr>
<tr>
<td><strong>Item record/more information</strong></td>
<td><a href="http://hdl.handle.net/10197/6187">http://hdl.handle.net/10197/6187</a></td>
</tr>
<tr>
<td><strong>Publisher's version (DOI)</strong></td>
<td>10.1071/RD14167</td>
</tr>
</tbody>
</table>
Epigenetic processes in the male germline

Alan M. O’Doherty^A and Paul A. McGettigan^B

^A School of Medicine and Medical Sciences, University College Dublin, Belfield, Dublin 4, Ireland.

^B School of Agriculture and Food Science, University College Dublin. Belfield, Dublin 4, Ireland.

**Keywords:** Sperm, Epigenetics, DNA methylation, Chromatin, Gametogenesis, Spermatogenesis, Transgenerational inheritance, non-coding RNA, histone modifications.
Abstract

Sperm undergo some of the most extensive chromatin modifications seen in mammalian biology. During male germline development, paternal DNA methylation marks are erased and established on a global scale through waves of demethylation and de novo methylation. As spermatogenesis progresses the majority of the histones are removed and replaced by protamines enabling a tighter packaging of the DNA and transcriptional shutdown. Following fertilization, the paternal genome is rapidly reactivated, actively demethylated, the protamines are replaced with histones and the embryonic genome is activated. The development of new assays, made possible by high throughput sequencing technology, has resulted in the revisiting of what was considered settled science regarding the state of DNA packaging in mammalian sperm. Researchers have discovered that not all histones are replaced by protamines and in certain experiments various species of RNA have been detected in what was previously considered transcriptionally quiescent sperm. Most controversially, several groups have suggested that environmental modifications of the epigenetic state of sperm may operate as a non-DNA based form of inheritance - a process known as “transgenerational epigenetic inheritance”. Other developments in the field include the increased focus on the involvement of short RNAs such as microRNAs, long non-coding RNAs and piwi-interacting RNAs. There has also been an accumulation of evidence illustrating associations between defects in sperm DNA packaging and disease and fertility. In this paper we review the literature, recent findings and areas of controversy associated with epigenetic processes in the male germline, focusing on DNA methylation dynamics, non-coding RNAs, the biology of sperm chromatin packaging and transgenerational inheritance.
Introduction

While the broad outlines and much of the detail of mammalian sperm packaging has been known for at least 20 years (Ward and Coffey 1991) the area has been receiving much more interest in the recent past. A confluence of factors has been responsible for this sudden peak of interest in sperm biology. These include, development of new technologies (e.g. microarrays and short read sequencing) that enable a more detailed picture of genetics, epigenetics and gene expression, which was previously not possible. The development and improvement of assays (e.g. global methylation profiling and chip-seq/chip-chip of histone modifications) that have allowed researchers to probe past assumptions. Emergence of intriguing patterns (that have received widespread publicity) from epidemiological and experimental biology, that are suggestive of transgenerational and paternal age effects in complex disease. Identification of epigenetic processes in plants and other organisms have led researchers to investigate whether similar processes may also operate in mammalian biology (Bender and Fink 1995; Jacobsen and Meyerowitz 1997).

However, the results of this new enthusiasm have not been without controversy (Ptashne 2013; Maderspacher 2014; Struhl 2014) and critiques are becoming more pointed and public as the claims have received wider publicity. Some of these controversies are inevitable as researchers wrestle with the poorly understood complexities and inherent biases of the new technology and assays (Macarthur 2012). Issues such as poor study design, inappropriate tissue selection, heterogeneity of cellular populations, inadequate controls for population structure and biological variability, insufficient validation of results, hasty inference of causation and the challenges of determining appropriate statistical analysis strategies have already been flagged by others in the field (Michels et al. 2013). Related fields, such as candidate gene studies and genome wide association studies (GWAS), have undergone their own growing pains and have either been abandoned or are now subject to much greater
controls (greater samples sizes and more robust statistical analysis) which has helped to reduce the number of unreplicable results.

Broader criticism of the epigenetics field (Bird 2013) has revolved around over-hyping of findings (Maderspacher 2014) as well as the pursuit of neo-Lamarckian hypotheses (Szyf 2014). The highest profile and most controversial claims in the field of epigenetics relate to the possibility of transgenerational effects (Heard and Martienssen 2014). These claims directly implicate sperm as a mechanism of transmitting heritable non-DNA based information from one generation to the next (Pembrey 2010).

We recognise that even the definition of epigenetics is controversial (Ptashne 2013). The term was first proposed by Conrad Waddington and used in the context of alternate developmental pathways/developmental canalization (Waddington 1942). In the last 15 years the definition has either shifted or been misappropriated (depending on your point of view) and is now more frequently used to describe modifications of DNA or chromatin, especially DNA methylation and histone tail modifications, which may or may not be stably inherited by daughter cells. While acknowledging the controversy, it is this more recent definition that we use in this review.

The influence of an organism’s environment on epigenetic mechanisms is an area of research that has received substantial attention in recent years (Feil and Fraga 2011; Cortessis 2012; Cortessis et al. 2012). Adaptation of epigenetic mechanisms to environmental cues or nutritional status has been demonstrated in a number of species including humans (Heijmans et al. 2008; Waterland et al. 2010), honeybees (Kucharski et al. 2008), mice (Wolff et al. 1998; Morgan et al. 1999b; Ivanova et al. 2012), rats (Burdge et al. 2007; Mychasiuk et al. 2012; Mychasiuk et al. 2013), sea bass (Navarro-Martin et al. 2011), tomato plants (Gonzalez et al. 2013) and dairy cows (O'Doherty et al. 2014).
The contribution of the maternal environment to the development of the offspring has been fairly well defined in recent years. Adverse health outcomes such as obesity, insulin resistance, congenital abnormalities, and cardiovascular disease (Hampton 2004; Stothard et al. 2009; Boerschmann et al. 2010; Maric-Bilkan et al. 2011) have been shown to be more prevalent in children exposed to suboptimal environments (e.g. malnutrition, environmental toxins, and alcohol) prior to pregnancy, in utero and/or during early postpartum development (referred to as developmental origins of human disease (Barouki et al. 2012)). It has been proposed that the underlying mechanisms associated with developmental origins of human disease might be epigenetic in nature, through aberrant DNA methylation, altered expression of non-coding RNAs or inappropriate modification of histones (Soubry et al. 2014). However, the contribution of paternal environmental insults to the outcome of their offspring is less well established.

In this review we will survey the major biological processes during sperm development involving epigenetic change, highlighting the recent findings and areas of existing controversy.

**DNA Methylation in the male germline**

Methylation at cytosine residues (5-methylcytosine/5mC) in the mammalian genome is dynamically controlled throughout development and is a fundamental epigenetic modification. Establishment of DNA methylation patterns is largely accomplished through the action of two members of the DNMT3 de novo methyltransferases family: DNMT3A and DNMT3B. A third member of this family, DNMT3L, is catalytically inactive (Bourc'his et al. 2001; Hata et al. 2002; Kaneda et al. 2004; Borgel et al. 2010; Shirane et al. 2013). Targeted disruption of these genes in mice results in embryonic lethality and sterility (Okano et al. 1999; Bourc'his and Bestor 2004; Webster et al. 2005). In humans, mutations of the
DNMT3 genes are associated with Immunodeficiency, Centromere instability and Facial anomalies syndrome, acute myeloid leukaemia and in developmental growth disorders such as DNMT3A overgrowth syndrome and intellectual disability (Hansen et al. 1999; Ley et al. 2010; Yan et al. 2011; Tatton-Brown et al. 2014).

Mammalian cells generally go through two waves of genome wide DNA methylation programming during germ cell and preimplantation development (Reik et al. 2001; Messerschmidt et al. 2014). The establishment and maintenance of correct DNA methylation patterns is essential for fertility, embryo development and viability of the offspring (Li et al. 1992; Okano et al. 1999; Bourc'his and Bestor 2004). During these developmental reprogramming windows mechanisms that are involved in establishing and removing DNA methylation marks may be susceptible to error. The first reprogramming event occurs in the developing gonad. Murine DNA methylation patterns at most sequences (with the exception of Intracisternal A Particles and some single copy loci (Lees-Murdock and Walsh 2008; Hackett et al. 2012; Seisenberger et al. 2012) are removed during the colonization of the fetal gonad by post-migratory primordial germ cells (PGCs), at around 9.5-12.5 days post coitum (Hajkova et al. 2002; Guibert et al. 2012). Subsequently, establishment of gametic DNA methylation patterns takes place and is strikingly different depending on whether it is occurring in the male or female germline (Smallwood and Kelsey 2012). The most distinctive differences are those at imprinted loci, in which a number of imprint control regions/differentially methylated regions (ICRs/DMRs) are asymmetrically methylated between male and females gametes (Lucifero et al. 2004; Hiura et al. 2006; Trasler 2006; O'Doherty et al. 2012). Approximately 120 imprinted genes have been identified in mammals and it has been postulated that they are regulated by as many as 23 ICR/DMRs (Proudhon et al. 2012; Magee et al. 2014). The majority of the DMRs are methylated in the oocyte and a number of reasons have been offered to explain this sexual dimorphism in
imprint ontogeny, such as the conflict for maternal resources theory (Hurst and McVean 1997; Iwasa 1998), gender-specific stages of DNMT3L expression (Bourc'his and Proudhon 2008) and the central role of maternal ICRs in early development (Schulz et al. 2010). Nevertheless, a number of imprinted loci have been identified in the male germline, such as the H19, Rasgrf1 and Dlk1-Gtl2 DMRs (Tremblay et al. 1997; Shibata et al. 1998; Davis et al. 2000; Takada et al. 2002; Yoon et al. 2002; Lin et al. 2003; Li et al. 2004). RNA sequencing of trophoblast tissue, isolated from reciprocal hybrids of horse and donkey, identified 78 candidate imprinted genes that showed a paternal bias in expression which was lost when compared to transcriptomes of fetal tissue (Wang et al. 2013). Similar to mice, (Barton et al. 1984) these results show an important role for paternally expressed imprinted genes in placental development.

A second reprogramming event occurs following fertilization of the zygote. The mammalian zygote represents a totipotent template harbouring the developmental potential from which an entire organism may be generated. However, since almost all cell-types descended from the zygote (with the exception of mature red blood cells, cornified cells in the skin, hair and nails (Kawane et al. 2001; Eckhart et al. 2013) and certain immune cells (Jung et al. 2006) share its DNA sequence it is widely accepted that epigenetic modifications are associated with defining the developmental potency of the zygote (Ma et al. 2012; Seisenberger et al. 2013). Similar to establishment of germline DNA methylation patterns, the dynamics of reprogramming are asymmetrical depending on the parental origin of the pronuclei/genomic material. For this article reprogramming of the male germline DNA methylation marks will be emphasised. The paternal genome is rapidly depleted of 5-methylcytosine within hours of fertilization (prior to the first cleavage) (Mayer et al. 2000; Oswald et al. 2000; Santos et al. 2002). In contrast, the maternal genome is passively demethylated during subsequent cleavages (and is usually complete by the late morula stage). The maternal genome is
thought to be protected from rapid demethylation through the actions of an essential maternal factor, PGC7/Stella (Nakamura et al. 2007; Nakamura et al. 2012; Wu and Zhang 2014).

While the paternal genome is rapidly demethylated, several loci appear to be protected from this process - these regions include IAP elements, paternally methylated imprinted genes and heterochromatic centromeres (Morgan et al. 2005; Lees-Murdock and Walsh 2008). It has been suggested that the maintenance of methylation patterns at these sequences may be required for successful progression through the early stages of embryogenesis (Seisenberger et al. 2013).

As recently as a decade ago, the mechanisms responsible for active demethylation of the paternal pronuclei were poorly understood. Since that time a number of possibilities have emerged. One of the first demethylation mechanisms to be investigated was the enzymatic removal of cytosine methyl groups by DNA glycosylases (Demeter/repressor of silencing 1 family) and base excision repair (similar to that observed in flowering plants) (Zhu 2009). However, this mechanism of demethylation is unlikely as mammalian orthologs of these enzymes haven’t been discovered. It had also been suggested that the DNMT3 enzymes responsible for establishing DNA methylation marks (Kangaspeska et al. 2008; Metivier et al. 2008) or the methyl-CpG- binding protein MBD2 (Bhattacharya et al. 1999) may also be involved with initiating demethylation, although the involvement of these enzymes remains controversial. The involvement of MBD2 in DNA demethylation is questionable as it has yet to be independently verified (Ooi and Bestor 2008) and MBD2 knockout mice (MBD2−/−) exhibit normal methylation patterns (Hendrich et al. 2001). Another proposed mechanism of active 5mC erasure is through nucleotide excision repair (NER) and the action of the Gad45a family proteins (Barreto et al. 2007; Schmitz et al. 2009). Although some evidence exists to support the possible role of Gadd45 in active demethylation there is also conflicting evidence against this mechanism (Jin et al. 2008; Engel et al. 2009). The latest mechanism under
investigation, for active 5mC removal, is one involving conversion of 5mC to 5-hydroxymethylcytosine (5hmC) by the Ten-Eleven Translocation (TET) family enzymes followed by Thymine DNA Glycosylase (TDG)-mediated Base Excision Repair (BER) (Kohli and Zhang 2013; Zhao and Chen 2013; Wu and Zhang 2014). Immunoflourescence studies, targeting 5mC and 5hmC (as well as two less abundant bases; 5-formylcytosine (5fC) and 5-carboxylyclosine (5caC) (He et al. 2011; Ito et al. 2011)) in the male pronuclei of mammalian zygotes revealed that as the 5mC signal decreased the signal of 5hmC, 5fC and 5caC increased concomitantly, thus providing further evidence for the involvement of this TET mediated pathway in active demethylation (Inoue et al. 2011b; Inoue and Zhang 2011; Iqbal et al. 2011; Wossidlo et al. 2011). Although the mode of active demethylation remains a hotly debated topic, this reprogramming event is vital for establishing pluripotent states in early embryos and for erasure of parent of origin specific methylation marks in primordial germ cells (Feng et al. 2010). As the embryo develops cells differentiate and cell lineages are committed towards distinct cell fates. This process of cellular differentiation begins around the time of embryo implantation. As part of this process the primordial germ cells of the developing gonad are reprogrammed with the appropriate germline DNA methylation marks and the process is cyclically repeated from generation to generation (Seisenberger et al. 2013).

An intriguing challenge to the dogma of demethylation of the paternal pronucleus has been proposed by the work of the Li and O’Neill (Li and O’Neill 2012). Their research has suggested that the apparent loss of methylation is actually a progressive antigenic masking of the methylated bases. They suggest that structural changes to the chromatin, post fertilization, alter the response to immunolocalization assays of 5meC, that the DNA retain its methylation state and that this can be detected by alternate assays e.g. MBD1 immunolocalization or tryptic digestion prior to a 5meC immunolocalization assay. This
model could provide a parsimonious explanation for the lack of demethylation of the paternal pronucleus observed in sheep (Beaujean et al. 2004) and rabbits (Shi et al. 2004). However, genome wide sequencing studies have confirmed that rapid demethylation of the paternal genome does occur in the pronuclei of humans (Guo et al. 2014) and mice (Smith et al. 2012; Guo et al. 2013). (For a review on species specific active DNA demethylation in mammalian embryos see (Ma et al. 2012)).

**Associations with disease and infertility**

Atypical or aberrant male germline DNA methylation has been reported in a number of human conditions such as male factor infertility & pregnancy failure (Benchaib et al. 2005), low semen concentration, reduced sperm motility and poor semen morphology (Houshdaran et al. 2007). In 2007 Kobayashi et al carried out DNA methylation analysis at a panel of imprinted gene DMRs using sperm isolated from 97 infertile males (Kobayashi et al. 2007). The results revealed that methylation defects were frequent in these samples, particularly in samples from patients presenting with oligospermia. They suggested that there could be a higher risk of transmitting inappropriate primary DNA methylation imprints to offspring when using sperm samples from infertile patients in assisted reproductive regimes. In a more recent study, comparing patients with oligozoospermia/asthenozoospermia to fertile normozoospermic males, global analysis of sperm DNA methylation revealed an association between dysregulation of DNA methylation at the H19 DMR and the DAZL promoter in all samples; with the authors suggesting that aberrant H19 DMR methylation is closely associated with oligozoospermia (Li et al. 2013).

In mice, dietary intake of methyl donors (folate) has been implicated in negative pregnancy outcomes through an altered sperm epigenome (Lambrot et al. 2013). In these experiments mice were fed a lifelong diet deficient of folate, beginning in *utero* during the developmental
period of epigenetic programming of germ cells. Paternal folate deficiency was associated with a number of defects in the offspring, such as craniofacial and musculoskeletal malformations. The authors also identified genome wide differential methylation patterns in sperm samples of folate deficient versus folate sufficient mice in regions containing genes related to development, cancer, schizophrenia and diabetes.

Other factors such as heat stress (Rahman et al. 2013) and exposure to environmental toxins (Anway et al. 2005) have also been shown to perturb DNA methylation reprogramming in the male germline, resulting in a reduction in developmental/reproductive competence both in vitro (Rahman et al. 2013) and in vivo (Anway et al. 2005). These epigenetic alterations are believed to persist through subsequent generations through a phenomenon referred to as transgenerational epigenetic inheritance. This controversial phenomenon will be discussed in more detail in a later section.

**Non-coding RNAs (ncRNAs)**

Given that sperm have been generally regarded as being transcriptionally quiescent, the presence of RNA species in the male gamete was originally attributed to carryover from surrounding cells or material left behind following degradation and removal of the residual body, but this source of RNA was ruled out when RNA was specifically identified in sperm samples from several species, including mouse (Wykes et al. 2000), rat (Pessot et al. 1989) and human (Kumar et al. 1993). This repertoire of sperm RNA molecules is now attracting increasing consideration for diagnostic purposes, evaluating sperm function and long-term health of offspring. RNA molecules are not simply messengers between DNA and protein, genome wide transcription studies have identified that a multitude of non-protein coding RNA species are transcribed across eukaryotic genomes (Carninci et al. 2005). These non-coding RNA molecules have been shown to be involved with transcriptional regulation and
have been proposed to influence gene expression through mechanisms such as chromatin configuration, post-transcriptional silencing and post-translational modifications. The ncRNA component of the transcriptome is comprised of an ever increasing list of members (Mukherjee et al. 2014); in this section we will review the role of several of these ncRNAs during male germ cell development.

**Long non-coding RNA (lncRNA)**

Long non-coding RNAs are a large class of non-protein encoding RNA molecules that are generally greater than 200 nucleotides in length. Until recently the expression, and potential biological function, of lncRNAs has remained largely unknown. Initially it was believed that lncRNAs were transcriptional noise caused by transcription from weak promoters facilitated through a low affinity binding of RNA polymerase (Ponjavic et al. 2007; Louro et al. 2009). However, this non-functionality of lncRNAs has been widely disregarded as subsequent targeted functional studies have identified lncRNAs to be involved with many cellular processes such as targeting epigenetic modifications at imprinted loci (Mercer and Mattick 2013; Monnier et al. 2013; Santoro et al. 2013), dosage compensation (Conrad and Akhtar 2012), trans regulation of neighbouring protein coding genes (Rapicavoli et al. 2011) and modulation of mRNA stability & expression in the cytoplasm (Willingham et al. 2005; Carrieri et al. 2012).

Next generation sequencing and microarray technologies have been applied to the analysis of fetal (E12.5 and 15.5) and postnatal mouse testis (days 6-14; encompassing first stages of spermatogenesis), revealing expression profiles of lncRNAs within these tissues (Bao et al. 2013; Laiho et al. 2013; Sun et al. 2013). These studies demonstrated that a large number of lncRNAs (~1000) are temporally up or downregulated during male germ cell development. The presence of lncRNA transcripts in human sperm samples, through RNA-seq profiling,
has also been confirmed (Sendler et al. 2013). Analysis of results from a large scale study carried out comparing RNA-seq data from eight organs and eleven different species (human, chimpanzee, bonobo, gorilla, orangutan, macaque, mouse, opossum, platypus, chicken and frog) led the authors to propose lncRNAs to have potential roles involved with the fundamental processes of spermatogenesis and synaptic transmission (Necsulea et al. 2014). lncRNAs have also been shown to encode micro RNAs (miRNAs, see below) (He et al. 2008). An example of this is the H19 large intergenic non-coding RNA, which encodes a highly conserved miRNA, miR-675 (Cai and Cullen 2007; Smits et al. 2008). This miRNA has been shown to be involved with regulating placental growth towards the end of gestation, through targeting of igf1r (Keniry et al. 2012).

There are several counterpoints to the putative significance of lncRNAs in sperm. Firstly, their expression has been primarily detected in the testes and the very nature of spermatogenesis leads to a very permissive transcriptional environment during the processes that lead to histone removal (see below). Secondly, the expression levels of lncRNAs has been shown to be several orders of magnitude lower than that for protein coding genes this would argue against a significant functional role for these sequences (Hebenstreit et al. 2011; Marinov et al. 2014). In general, the amount of RNA detected in sperm (when detected at all) is extremely low (approximately .01pg per sperm (Ostermeier et al. 2004)) and extreme care needs to be taken to ensure that results are not due to inadvertent genomic DNA contamination (Haas et al. 2012) a problem that may be more prevalent in next generation sequencing RNA-seq studies than is currently appreciated – but of more concern when analysing at the limits of detection. Finally, the viability of bi-maternal embryos (Kawahara et al. 2007) also argue for the relative unimportance of sperm delivered RNAs in fertilization and early embryo development.

**Micro and short interfering RNAs**
miRNA molecules are involved with posttranscriptional regulation of gene expression through interaction with the 3’ UTR (untranslated region) of their target mRNA, leading to degradation/stabilization and repression/activation of translation (Lai 2002; Carrington and Ambros 2003; Meister 2007; Vasudevan et al. 2007). In order for miRNAs to exert their effect they must first be processed from immature primary molecules into mature miRNAs. In the nucleus they are largely processed from primary miRNA complexes into precursor miRNAs through the actions of DROSHA and its co-factor DGCR8/PASHA (Murchison and Hannon 2004; Yang and Lai 2011). Following this, exportin 5 facilitates their transportation into the cytoplasm where they are cleaved by DICER into mature miRNAs (Meister and Tuschl 2004; Zeng and Cullen 2004). DICER is also involved with the siRNA biogenesis pathway by separating double-stranded RNA molecules to generate siRNAs (Carthew and Sontheimer 2009). Several conditional knockout studies, specifically targeting Dicer in the male germline, have defined the role of DICER and the importance of correct small non-coding RNA processing during spermatogenesis (Hayashi et al. 2008; Romero et al. 2011). These studies have shown that mice lacking Dicer present with an infertile or subfertile phenotype. Defective proliferation was observed following targeted disruption of Dicer in primordial germ cells in developing mouse embryos and these mice also had impaired postnatal spermatogenesis. Furthermore, in a comparative study looking at male germline conditional knockouts of both Dicer and Drosha, it was revealed that although DICER and DROSHA are both involved with miRNA biogenesis, testes from these mutants possess morphological and transcriptomic differences (Wu et al. 2012). Results from this investigation identified an essential role for Drosha during male germ cell development, with deficiency in Drosha being associated with impaired spermatogenesis and male infertility; through a lack of DROSHA-dependent miRNAs.

piRNAs
Another class of ncRNAs predominantly expressed in the mammalian germline are the piwi-interacting RNAs (piRNAs). Targeted disruption of genes involved with piRNA biogenesis (MIWI1, MILI2 & MIWI23) resulted in reactivation of retrotransposons and sterility through impaired spermatogenesis, indicating a tentative role for these ncRNAs in regulation of these processes (Kuramochi-Miyagawa et al. 2004; Carmell et al. 2007). Others have also identified that deletion of murine piRNAs results in impaired DNA methylation of transposons in the male germline and mutant mice bear a phenotype resembling that of Dnmt3L deficient animals (Kuramochi-Miyagawa et al. 2008) suggesting that the Piwi family are involved with targeting DNA methylation at transposons during fetal testis development. A role for piRNAs and non-coding RNA in de novo DNA methylation of the murine Rasgrf1 locus has also been suggested (Watanabe et al. 2011). Analysis of DNA methylation in spermatogonia isolated from Mili, MitoPLD, and Miwi2 mutant mice revealed impaired methylation at the DMR of the paternally methylated Rasgrf1 gene, while the other paternally methylated DMRs H19, Dlk1–Gtl2, and Gprl–Zdbf2 had normal levels of methylation. Furthermore, these mice were also infertile through meiotic arrest during spermatogenesis, thus highlighting the importance of the PIWI-piRNA pathway during early development. In addition, this class of non-coding RNAs have been shown to be involved with enforcing retrotransposon silencing during spermatogenesis (Di Giacomo et al. 2013). Molaro et al. have suggested that this silencing occurs in two hierarchically distinct waves (Molaro et al. 2014). They showed that the vast majority of retrotransposon sequences are modified by default de novo methylation. However, some retrotransposon copies evade this wave of methylation and remain transcriptionally active (possibly through mirroring the behaviour of protein coding genes) becoming targets of a second wave of piRNA-mediated methylation.
In humans, genetic variants in Piwi-interacting RNA pathway genes have been identified to contribute to spermatogenic failure (Gu et al. 2010). Genotyping of single nucleotide polymorphisms of several Piwi genes (PIWIL1/HIWI, PIWIL2/HILI, PIWIL3/HIWI3 and PIWIL4/HIWI2) from 490 patients with oligozoospermia or idiopathic azoospermia and 468 control patients revealed that SNPs in HIWI2 and HIWI3 were significantly associated with either an increased or decreased risk of oligozoospermia, respectively. The most recent studies have shown that piRNAs have a major role in mRNA elimination during the later stages of spermiogenesis (Gou et al. 2014). Taken together, the studies discussed in this section have shown essential roles for non-coding RNAs during male germ cell development and in orchestrating DNA methylation dynamics at certain loci.

**DNA and chromatin modifications during spermatogenesis**

The DNA in most vertebrate sperm is dramatically modified in comparison to the structure of DNA found in somatic cells. The most obvious difference is that the removal of most histones and replacement of these by protamines. This change causes a structural switch from the well-recognised (and relatively relaxed) solenoidal/nucleosomal structure of somatic cells to a highly condensed “almost crystalline” (Balhorn 1982) toroidal (doughnut) structure of DNA in mature sperm. While the broad outlines of the process are consistent between mammalian species there are differences, including the number of protamine variants that are present and the percentage of retained histones.

Protamines are unusual proteins having extremely high arginine content (50 - 70%) as well as high cysteine content. Protamines evolved from the H1 histone protein (Eirin-Lopez et al. 2006). Just as with histones, protamines and transition proteins are all highly basic proteins although histones are enriched for lysines rather than arginines. Transition proteins (that replace histones in an intermediate step before themselves getting displaced by protamines)
have approximately equal amounts of lysine and arginine content (20% of TP1 amino acid content is lysine, and 20% is arginine, the figures for TP2 are 10% each) whereas histones are lysine rich and protamines are arginine rich (Meistrich et al. 2003).

The positive charge contributed by each arginine leads to an overall highly positively charged protein (Iso 12.7 for bovine) and enables the protamines to bind strongly to the negatively charged DNA. This interaction is further strengthened by the presence of disulfide bridges facilitated by the cysteines. The resulting structure represents a 10-fold reduction in space occupied by chromatin in the mature sperm compared to the somatic cell when it is at its most compact (during interphase stages of mitosis and meiosis) (Miller et al. 2010). The process of compaction of the paternal DNA is reviewed in detail by others (Ward and Coffey 1991; Kota and Feil 2010; Miller et al. 2010).

Histones are gradually replaced by protamines in a multi-step process. The first step involves the global methylation of DNA followed by hyperacetylation of histones in elongating spermatids (Pivot-Pajot et al. 2003). This global methylation is thought to shutdown the transcription in the gamete (Miller et al. 2010), with hyperacetylation of histones reducing the charge difference between the histone core and DNA and thus reduces the affinity of the nucleosome for DNA.

In fish and birds, histones are directly replaced by protamines. In mammals the histones are first replaced by transition proteins and these are then replaced by protamines. However, despite this process certain regions of DNA retain their histones. The amount of histones retained varies by species, marsupials retain as much as 25% (Soon et al. 1997) while humans retain approximately 15% and mice retain less than 2% (Balhorn et al. 1977). In human and mice, these retained histone loci include regions near housekeeping and developmentally regulated genes (Brykczyńska et al. 2010), as well as regions containing
CTCF motifs that bind the mammalian insulator protein CTCF. Interestingly, while CTCF is not expressed in male germ cells its closely related relative BORIS/CTCFL is (Loukinov et al. 2002), it is only expressed in male germ cells and can bind methylated DNA (unlike its cousin CTCF). It is possible that the function of CTCFL is to protect the histones in the regions containing CTCF motifs from protamine replacement (Arpanahi et al. 2009).

Other DNA regions do not retain their somatic histones but instead of replacing these with protamines the somatic histones are replaced with testis specific histones. These regions include telomeres (Zalensky et al. 2002) and centromeres. Post-fertilization maternal factors are able to access the paternal DNA and begin the replacement of protamines with maternal histones. Key proteins implicated in the process of sperm decondensation are members of the nucleoplasmin/nucleophosmin (NPM) family (Inoue et al. 2011a; Okuwaki et al. 2012). The paternal pro-nucleus is transcriptionally activated before the maternal pronucleus (Aoki et al. 1997). The conserved proteins, HIRA and CHD1, are crucial mediators of nucleosome assembly of the paternal pronucleus in Drosophila (Loppin et al. 2005; Konev et al. 2007).

The successful activation of both paternal and maternal pro-nuclei represents the stage of embryonic genome activation (EGA). The paternal (sperm) DNA is actively demethylated post fertilization however paternally imprinted genes escape this process. As the embryo develops the primordial germ cells enter the gonad the maternal and paternal imprinting signals are erased and in the male they are replaced with paternal imprints and the cycle continues.

Transgenerational and environmental epigenetics

Epigenetic modifications are generally removed and re-established from one generation to the next (Reik et al. 2001). However, studies in some model species have demonstrated that there is a failure to completely remove and re-establish epigenetic marks, at certain genetic
loci, from one generation to the next (Morgan et al. 1999a; Lane et al. 2003; Reinders et al. 2009). The phenomenon is known as transgenerational epigenetic inheritance (Heard and Martienssen 2014). Furthermore, it has been shown that the epigenome can be affected in organisms subjected to certain environmental stressors and that these environmentally induced epigenomic alterations can be transmitted through subsequent generations, which are unexposed to the stressor (Anway et al. 2005). Here we will review recent developments in the understanding of transgenerational inheritance of epigenetic modifications that are transmitted through the male germline.

Early examples of transgenerational epigenetic inheritance in the male germline were demonstrated through experiments investigating the effect of transient exposure of gestating rats to endocrine disruptors (the estrogenic pesticide methoxychlor and antiandrogenic fungicide vinclozolin) on embryonic testis development (Anway et al. 2005; Skinner and Anway 2005). These studies identified a reduced spermatogenic capacity and increased male infertility in the adult F1 population. These phenotypic effects were carried through the male germline to almost all males of subsequent generations (up to the F4 generation). Phenotypic changes were correlated with altered DNA methylation patterns, both hyper- and hypomethylation events, in sperm samples isolated from vinclozolin treated F2 and F3 generation animals. However, whether the changes in methylation were causal of a reduced fertility phenotype in the F1-F4 rats or markers of the transgenerational effect was undetermined. Nevertheless, cause or consequence aside, these early studies and more recent investigations (Anway et al. 2006; Chang et al. 2006; Guerrero-Bosagna et al. 2010; Skinner et al. 2013) have identified that exposure of a gestating mother to endocrine disruptors during the development window encompassing embryonic germ cell reprogramming is capable of altering the epigenome, and that these changes are persistent through the male germ line for several generations. Other environmental toxins such as Bisphenol A, dioxin and DEET have
also been proposed to have similar transgenerational epigenetic effects (Manikkam et al. 2012a; Manikkam et al. 2012b; Singh and Li 2012; Manikkam et al. 2013). Genome wide methylation profiling of male offspring from mice, fed half the calories of controls during their last week of gestation, revealed that over 100 regions (some located in proximity to genes associated with insulin secretion) were hypomethylated in sperm, further demonstrating the plasticity of the epigenome in response to maternal nutrition (Radford et al. 2014). Although methylation differences were observed in the F1 generation, mating of these males (fed a normal diet) with control females revealed that these hypomethylated regions weren’t present in the F2 generation. Some of these previously hypomethylated regions did however exhibit differential expression in the brain and liver of some male F2 fetuses, showing that the grandmother’s diet can impact subsequent generations when transmitted through the male germline. For an recent in depth review of transgenerational epigenetic mechanisms see (Heard and Martienssen 2014).

A recent high profile example of potential transgenerational epigenetic inheritance came from ancestral odour fear conditioning study of Dias and Ressler (Dias and Ressler 2014). While the phenomenon appears very real the epigenetic link is relatively tenuous. The presence of olfactory receptors on sperm and the known chemotaxis of sperm in response to odorant gradients (Spehr et al. 2006) means that direct selection on sperm genotype remains a plausible alternative. Recently, an involvement of sperm RNAs with transgenerational epigenetic inheritance has been suggested (Gapp et al. 2014). The authors of this investigating were aiming to further understand environmental and genetic factors associated with complex neurological diseases and trauma experienced during childhood. Using a mouse model of maternal trauma, that involved separating pups from their mothers, the investigators found a number of behavioural and molecular differences between control and stressed offspring, across several generations. Regarding the observed differences at the
molecular level, using RNA extracted from sperm they detected differences in the abundance of non-coding RNAs (microRNAs and PIWI-interacting RNAs) that have been purported as having a role in epigenetic gene regulation.

Some of the phenomena that are frequently listed under the rubric of transgenerational inheritance are outside the strict definition of a transgenerational epigenetic inheritance as we define it (direct modification of DNA or histones in the germline). The Kit paramutation in mouse (Rassoulzadegan et al. 2006) is the result of the transmission of a mutant RNA via sperm in the absence of the genotype – initially via the cytoplasmic bridges in spermatagonia of a heterozygous male but which apparently is capable of transmission across multiple generations. The paramutation phenomenon is reviewed in depth elsewhere (Chandler and Stam 2004). Many of the validated epigenetic germline inheritance instances occur at loci involving repetitive elements or retrotransposons (Chong and Whitelaw 2004). These elements cluster in regions that are known to avoid histone replacement by protamines during spermatogenesis i.e. telomeres and centromeres as mentioned above.

It remains to be determined whether metastable epialleles that underlie such phenomena as the Agouti viable yellow (coat colour) allele in mouse (Rakyan et al. 2002) are rare phenomena, or whether they are more common and consequently more important for health and disease. Recent research has identified metastable epiallele candidates in humans (Harris et al. 2013), however the further (and more controversial) step of transgenerational transmission of epiallele state has not been demonstrated for these gene candidates.

Confounding issues and controversy

The existence of paternal age effects associated with several conditions has fuelled the speculation about epigenetic transgenerational effects (mediated in part by sperm epigenetic processes). This has become an area of much controversy (Bird 2013) experiencing an
attempted rehabilitation of Lamarckian ideas (Szyf 2014). However, there are issues confounded with the epigenetic explanations that, to date, have received insufficient attention.

In defining inheritance of acquired characteristics, Landman (Landman 1991) noted that in attempting to test for inheritance of acquired characteristics “for individual organisms or cultures of cells” that care must be taken to carry out the experiment “under conditions of little or no growth (thereby ruling out selection of mutants)”[emphasis ours]. This is an under-appreciated consideration for experiments involving sperm epigenetics.

This problem has been belatedly recognised in the context of somatic tissue epigenetic profiling and has recently been tentatively names as “The Houseman Effect”, a name coined for a paper from Houseman et al where they showed that changes in methylation could be used to measure changes in the composition of cell populations (i.e. individual cell methylation levels are not changing but the overall levels are changing due to infiltration/exfiltration of different cell types) (Houseman et al. 2012; Jaffe and Irizarry 2014). This problem, and associated issues with epigenetic profiling, are discussed in more detail in a recent review (Michels et al. 2013).

In the case of sperm, the problem is not (or at least not necessarily) infiltration of different cell types (although immune cells are present in ejaculates (Wolff 1995) and the overall numbers and proportions can be modified by disease) but relates to the individual unique genetic composition of each sperm cell. Individual sperm are haploid organisms with unique genetic complements compared to their cohorts and compared to their parent organism due to meiosis as well as from gene conversion events that may be in the range of 250-800 per sperm cell (Wang et al. 2012). Most analyses implicitly assume that the sperm are genetically identical and have a uniform probability of successful fertilization. However, the
sperm are genetically unique and there are many documented examples of processes which can lead to the sperm genotype influencing fertilization success (Haig 1995). The most easily recognised of these conditions are those which can affect the sex of the offspring. These conditions have been extensively catalogued by William James (James 1996; James 2004). While the actual mechanism of many of these processes remains to be elucidated it is at least as plausible that these types of processes are responsible for some of the phenomena where epigenetic transgenerational effects are currently proposed.

Processes such as transmission ratio distortion have been shown to impact directly on sperm function and lead to non-mendelian inheritance patterns (Herrmann et al. 1999). Other mechanisms include selfish spermatagonial selection; which involves the age related clonal expansion (similar to oncogenic processes) of particular spermatogonial cells ultimately leading to enrichment of mutant sperm over time (Goriely and Wilkie 2012). This process has been identified as the cause of several paternal age effect (PAE) diseases such as achondroplasia and Apert syndrome.

Researchers in this area should be mindful of these problems and look out for the symptoms of these phenomena, such as altered sex ratio in offspring.

**Concluding remarks**

The last decade has seen great growth in the development and application of epigenetics. Some intriguing claims have been made in the field of sperm epigenetics challenging fundamental aspects of human genetics and disease pathogenesis. Now that the epigenetics field is maturing and attracting greater scrutiny and criticism, these claims are moving front and centre. The next decade should determine which of these claims (particularly transgenerational epigenetics) are able to withstand skeptical scrutiny.
References


James, W.H. (1996) Evidence that mammalian sex ratios at birth are partially controlled by parental hormone levels at the time of conception. *J Theor Biol* 180(4), 271-86

James, W.H. (2004) Further evidence that mammalian sex ratios at birth are partially controlled by parental hormone levels around the time of conception. *Hum Reprod* 19(6), 1250-6


O'Doherty, A.M., O'Shea, L.C., and Fair, T. (2012) Bovine DNA methylation imprints are established in an oocyte size-specific manner, which are coordinated with the expression of the DNMT3 family proteins. *Biol Reprod* 86(3), 67


Oswald, J., Engemann, S., Lane, N., Mayer, W., Olek, A., Fundele, R., Dean, W., Reik, W., and Walter, J. (2000) Active demethylation of the paternal genome in the mouse zygote. *Curr Biol* 10(8), 475-8


Reik, W., Dean, W., and Walter, J. (2001) Epigenetic reprogramming in mammalian development. *Science* 293(5532), 1089-93


Santoro, F., Mayer, D., Klement, R.M., Warczok, K.E., Stukalov, A., Barlow, D.P., and Pauler, F.M. (2013) Imprinted Igf2r silencing depends on continuous Airn IncRNA expression and is not restricted to a developmental window. *Development* 140(6), 1184-95


