The sperm factor: paternal impact beyond genes

Simone Immler \textsuperscript{1,2}

\textsuperscript{1} School of Biological Sciences, University of East Anglia, Norwich Research Park, Norwich, NR4 7TJ, UK

Corresponding author: S. Immler – s.immler@uea.ac.uk

Short title: Paternal epigenetic effects

Keywords: transgenerational epigenetics, methylation, chromatin modification, RNA, paternal effects, sperm quality, centriole-centrosome complex, protein-based inheritance, small non-coding RNA, transposable elements, prion
Abstract

The fact that sperm carry more than just the paternal DNA has only been discovered just over a decade ago. With this discovery, the idea that the paternal condition may have direct implications for the fitness of the offspring had to be revisited. While this idea is still highly debated, empirical evidence for paternal effects is accumulating. Male condition not only affects male fertility but also offspring early development and performance later in life. Several factors have been identified as possible carriers of non-genetic information, but we still know little about their origin and function and even less about their causation. I consider four possible non-mutually exclusive adaptive and non-adaptive explanations for the existence of paternal effects in an evolutionary context. In addition, I provide a brief overview of the main non-genetic components found in sperm including DNA methylation, chromatin modifications, RNAs and proteins. I discuss their putative functions and present currently available examples for their role in transferring non-genetic information from the father to the offspring. Finally, I identify some of the most important open questions and present possible future research avenues.
Introduction
The importance of non-genetic factors for the transmission of information from parents to offspring is increasingly recognized (Bonduriansky and Day, 2009; Bonduriansky, 2012; Bonduriansky and Day, 2018). In animals, the relatively bigger size of the female gamete – the egg – and the resulting transfer of many different non-genetic components from the mother to her offspring has led to an early recognition of the role of maternal non-genetic effects in determining offspring phenotype (e.g. Dickerson, 1947; Willham, 1963; Legates, 1972; see also Bernardo, 1996; Mousseau and Fox, 1998; Wade, 1998; Marshall and Uller, 2007 for reviews). In contrast, the small compact size and the highly reduced cytoplasm of the animal male gamete – the sperm – was one of the main reasons for the assumption that paternal condition plays little to no role in determining offspring phenotype. This assumption has been overturned just over decade ago and it is now recognised that sperm contribute more than the paternal haploid genome (Krawetz, 2005). In this review, I provide an overview of the potential non-genetic mechanisms and factors transferred via sperm into the zygote. I discuss the evidence for their effects across generations, their putative causes and potential consequences in an evolutionary context. This is by no means a complete account and only provides small insights into a highly complex and fascinating world, but it may stimulate further research into the many processes that can be summarized as “sperm factor”.

3 4 0
Male condition and sperm phenotype

Male condition is affected by environmental factors such as diet, temperature and social interactions and these effects are often reflected in the characteristics of a male’s ejaculate. Nutritional stress is known to negatively affect sperm quality and can lead to an increase in the number of malfunctioning and morphologically abnormal sperm, which in turn may affect male fertilisation success (Gage and Cook, 1994; Merrells et al., 2009; Perry and Rowe, 2010; Tigreros, 2013; Kahrl and Cox, 2015;). Similarly, variation in environmental temperature affects ejaculate traits such as sperm number and sperm morphology in ectotherm insects (Fox et al., 2006) and fish (Breckels and Neff, 2013) but also in endotherm mammals (e.g. Al-Khanaan et al., 2015). Finally, aspects of male social environment such as male:female ratio and the perceived intensity of sperm competition are known to affect sperm numbers (Arnaud et al., 2001; Pilastro et al., 2002; Pizzari et al., 2003), sperm swimming velocity (Burness et al., 2004) and sperm morphology (Crean and Marshall, 2008; Immler et al., 2010). However, while these environmentally induced changes in ejaculate traits are well established, the potential consequences of such changes for the next generation are poorly understood. In order to estimate the importance of paternal effects we need to understand the non-genetic factors carried by sperm and which part of the zygotic development they might affect.

Why do paternal effects exist?

While the evidence for an effect of the paternal condition on the offspring is rapidly mounting (e.g. Curley et al., 2011; Soubry, 2015; Illum et al., 2018 for
review), the evolutionary reason for the existence of paternal effects is less clear. Here below, I discuss four non-mutually exclusive hypotheses that may serve as possible explanations for the transfer of non-genetic information from the father to the offspring.

Paternal effects are non-adaptive

The transfer of non-genetic factors through sperm could be non-adaptive noise caused by physiological processes affecting the epigenetic mechanisms in the male germline in response to changing environmental conditions experienced by the father. Many of the experimental manipulations used to study paternal effects involve a change in the stress level experienced by the male for a defined period during life. Stress generally evokes strong physiological responses, which may negatively affect the germline and with that male reproduction (McGrady, 2009). These negative effects may include an increase in the production of reactive oxygen species (Dickinson and Chang, 2011) and elevated activity of repetitive elements (Capy et al., 2000), both of which jeopardise the integrity of the genome and may increase mutation rates (Maklakov and Immler, 2016). Defense mechanisms of the genome against such mutagenic factors include DNA methylation, chromatin modifications and the production of small RNAs (sRNAs) including Piwi interacting (piRNAs) and microRNAs (miRNAs; Bartel, 2004; Klattenhof and Theurkauf, 2008; Siomi et al., 2011; Ernst et al., 2017). All three factors are known to be involved in mediating the possible effects of selfish genetic elements at the translational and post-translational levels. As a result, relevant epigenetic marks produced in protection of the germline genome may end up
in the mature gametes as relics by chance rather than for adaptive reasons. At this stage, the non-adaptive hypothesis needs careful testing before we can exclude it with certainty.

Paternal effects as an adaptive response to increase offspring fitness

The transfer of information about the environmental conditions encountered by the parents to their offspring may be beneficial and provide an adaptive advantage to the offspring (Bonduriansky and Day, 2009; Turner, 2009). A mechanism that allows for such a transfer of information without modifying the genome may offer a flexible solution particularly in rapidly changing environments. A recent theoretical study described a positive feedback process where the parental phenotype favoured by environmental conditions gets progressively reinforced in the following generations through a learning mechanism (Xue and Leibler, 2016). Empirical evidence for such dynamics have been reported in *C. elegans* where small RNAs have been shown to be inherited for several generations without further additional stimulation with the help of RNA-dependent RNA polymerases (Rechavi *et al.*, 2011; Ashe *et al.*, 2012; Gu *et al.*, 2012; Rechavi *et al.*, 2014). A recent study in *C. elegans* provided direct evidence for such a feedback loop determining the duration of transgenerational inheritance of small RNAs (Houri-Ze’evi *et al.*, 2016).

Similarly, the ability of prions to assume a self-templating fold mechanism (Harvey *et al.*, 2018) suggests that these have the potential to maintain themselves in a self-regulating manner over many generations. Such genome-independent systems could be a way to memorise past conditions and transfer relevant information across generations for swift adjustments to
slow or rapid environmental changes despite the rigidity of the underlying genome.

_Paternal effects to mediate sexual conflict_

The inheritance of a paternal and a maternal genome creates a conflict between males and females over allele expression at heterozygous loci in the offspring (Arnqvist and Rowe, 2005). Epigenetic factors may further contribute to this conflict if they are inherited at an equal rate from both parents, but they may also offer a mechanism to resolve the conflict. Genomic imprinting is an epigenetic mechanism, which determines expression of an allele according to its parental origin (Reik and Walter, 2001). The three main theories proposed for the evolution of genomic imprinting are the kinship theory (Haig, 2000), the sexual antagonism theory (Day and Bonduriansky, 2004; Bonduriansky, 2007) and the maternal-offspring co-adaptation theory (Wolf and Hager, 2006; Wolf and Hager, 2009, all reviewed in Patten et al., 2014). The question at the heart of all three theories is the conflict between the parents over gene expression in their offspring at heterozygous loci. The aspect that varies between the theories is the nature of the involved parties (parent-offspring, male-female or all of them together etc.) and of the resolving mechanism.

Similar to the hypothesis presented for sexual conflict, other genetic conflicts have been proposed as a possible explanation for sperm carrying RNAs (Holman and Price, 2014; Hosken and Hodgson, 2014). These authors suggested that RNAs mediate potential genomic conflicts not only between males and females but also between the diploid male and its haploid sperm,
and among the different sperm within an ejaculate. Given the shear variation of RNAs present within each sperm they possibly cover several of these functions.

Paternal effects to control selfish genetic elements

The genomic conflict arising between the genome and selfish genetic elements may provide another explanation for the evolution of transgenerational epigenetic mechanisms (Holman and Price, 2014). The transfer of defense mechanisms against the detrimental effects of stressful environments from the male germline to the zygote would allow the protection of the zygotic genome during the sensitive stages of early development. The findings of variation in small RNA profiles, methylation patterns and chromatin structure in response to environmental stressors in sperm and the resulting offspring appear to be in line with this idea. However, we still know relatively little about the association between transposable elements (TEs) and epigenetic marks and mechanisms. A recent study in Arabidopsis thaliana showed that changes in methylation patterns and increased levels of gene expression were directly associated with de novo insertions of TEs in the immediate vicinity of affected genes (Stuart et al., 2016). Whether similar associations exist in the male germline and/or in the zygote is currently not known.

Epigenetic factors and RNAs in the sperm may also derive from segregation distorting alleles that involve the incapacitation/killing of sperm or zygotes carrying alternative alleles (Holman and Price, 2014). This suggestion is
purely hypothetical and needs careful testing. But if such a mechanism exists, it would potentially affect male fertility. An association between male fertility and certain RNAs has been shown in humans but the mechanisms involved are unknown (Jodar et al., 2012).

The hypotheses outlined above for the evolution of genetic imprinting and the role of RNAs as signals among different units may apply to any epigenetic factor transmitted via sperm. Males can undoubtedly benefit from transmitting more than just a genome in their gametes, and the idea that these mechanisms are adaptive is enticing. Testing the non-adaptive alternative is therefore even more important and necessary. It will be exciting to examine the different hypotheses and understand more about the evolutionary dynamics involved. This should be increasingly possible with the steadily improving methods available in genomics, transcriptomics and proteomics.

**Which non-genetic components does a sperm transfer to the zygote?**

Beside the nuclear genome, sperm are known to contain a range of epigenetic elements, which are transferred into the zygote upon fertilisation, including chromatin modifications, RNAs and proteins (reviewed in Dadoune, 2009; Carrell, 2012; Casas and Vavouri, 2014; Rando, 2016; Figure 1). Here below, I provide a brief overview of the currently known factors and present examples for the ways these factors might affect processes in the zygote and beyond. I am using the term “epigenetics” in a broad sense and follow Henikoff and Greally’s (2016) definition, where any cellular memory not encoded in the genetic code is included. Genome-carrying cell organelles
such as mitochondria are therefore not included even though these may be inherited through sperm in rare occasions in some organisms, such as insects (Wolff et al., 2012), mammals (Zhao et al., 2004) and birds (Alexander et al., 2015) and regularly in others such as mussels (Sutherland et al., 1998; Zouros, 2000). Even with this relatively restricted definition of the term epigenetic, condition dependent transgenerational effects may be harder to identify than assumed, and some of the aspects that may need further investigation are described in the section Current challenges and future directions below.

DNA methylation/acetylation

DNA methylation is probably the most studied epigenetic mark and is assumed to play a major role in the transfer of non-genetic information across generations. DNA methylation in combination with histone modifications (see section below) plays a key role in regulating gene expression in the germ cells and thereby contributes to three key processes: (I) the specification and formation of primordial germ cells, (II) the genome-wide erasure and re-establishment of germline-specific patterns in the embryo and sex-specific patterns during gametogenesis and (III) the establishment of sex-specific patterns typical for mature male and female gametes (reviewed in Allegrucci et al., 2005). Given their key role in governing gene expression throughout development, it is not surprising that paternal condition affects methylation patterns in the offspring. Fathers kept on a high fat diet in Sprague-Dawley rats for example sired daughters with impaired insulin secretion and glucose tolerance. Their female offspring exhibited altered expression in 642
pancreatic islet genes with some of the key genes being hypomethylated (Ng et al., 2010). More generally, environmental changes during early developmental stages seem to have a major impact on germline methylation patterns (see Faulk and Dolinoy, 2011 for review).

The molecular mechanism is based on the binding of a methyl/acetyl group to a DNA molecule, which may affect the transcriptional activity of the underlying gene without changing the genetic code. The percentage of methylation inherited from the father through sperm varies markedly across taxa and may range from fully maternally inherited to largely paternally inherited patterns. In house mice *Mus musculus* (and other mammals), the methylation structure in the developing zygote is re-structured during early embryogenesis following the maternal template and paternal marks are mostly removed (see Daxinger and Whitelaw, 2012 for review). In contrast, in zebrafish *Danio rerio*, the paternal methylation pattern forms the template and the maternal methylation pattern is largely restructured according to the information coming from the father (Potok et al., 2013; Jiang et al., 2013). These taxonomic differences in methylation inheritance are currently not explained and possible evolutionary reasons need to be tested.

**RNA families**

Sperm contain many families of RNAs, which may be transferred into the zygote during fertilisation and may therefore affect processes involved during early embryogenesis (Dadoune, 2009). These RNA families include messenger RNAs (mRNA; Alcivar et al., 1989; Ostermeier et al., 2002; Yang
et al., 2009; Bonache et al., 2012), micro RNAs (miRNAs; e.g. Krawetz et al., 2011), Piwi interacting RNAs (piRNAs; e.g. Krawetz et al., 2011), transfer RNA derived small RNAs (tRNAs; e.g. Peng et al., 2012) and a number of other to date un-specified RNA families. mRNAs are a large group of different molecules that are the direct result of gene transcription and are therefore also known as “coding” RNAs. The mRNA content in sperm is relatively low compared to any other cell type, and their origin (i.e. pre- versus postmeiotic) and role need further investigation.

The three remaining families (i.e. miRNAs, piRNAs and tRNAs) belong to the group of “small non-coding” RNAs (sRNAs) as they are transcribed from non-coding regions of the genome, and for many, their origin and function is still unknown. miRNAs are short (about 22-nucleotides) molecules that are involved in RNA silencing and regulation of gene expression at the transcriptional and post-transcriptional stages (Bartel, 2004). They may mediate the activity of selfish genetic elements by triggering small interfering RNAs (siRNAs) in a highly specialised and pathway specific manner (Creasey et al., 2014). Similarly, piRNAs (21-32 nucleotides) in the germline are involved in the silencing of selfish DNA elements and the maintenance of DNA integrity through the formation of RNA-protein complexes that act at the transcriptional and post-transcriptional levels (Klattenhof and Theurkauf, 2008; Siomi et al., 2011; Ernst et al., 2017). However, the exact mechanisms and origins of piRNAs are currently elusive. tRNAs (sometimes also referred to as tsRNAs) may vary in length (from 20 nucleotides into the range of piRNAs) and have been assumed to be the result of transmitter RNA
degradation until they were clearly identified as a distinct group of small non-coding RNAs (Lee et al., 2009). Observations in house mice *M. musculus* suggested that in testicular sperm, the tRNA content is low but increases with maturation through the fusion with epidydosomes (Sharma et al., 2016). The same study also reported that the function of these tRNAs is to repress genes associated with the selfish element MERVL active in preimplantation embryos.

The total amount of RNA molecules transferred through sperm is vanishingly small compared to the RNAs present in the egg. Nevertheless, several families of RNAs have been reported to be involved in non-genetic inheritance of paternal conditions across generations. miRNAs and piRNAs were differentially expressed in the sperm of male house mice *M. musculus* exposed to traumatic stress during the juvenile life stage compared to sperm of control male mice (Gapp 2014). The injection of these differentially expressed RNAs into early zygotes lead to similar offspring phenotypes as those observed in the experiments using traumatised males as fathers. Furthermore, miRNAs were involved in the transmission of chronic stress responses experimentally evoked in adult male mice to their offspring (Rodgers et al., 2013). The precise role of tRNAs needs further investigation but they seem to affect gene expression during early embryo development (Sharma et al., 2016).

*Proteins*
Sperm are composed of a wide range of proteins located on the sperm surface, in the acrosome (where present), in and around the nucleus and even in the flagellum. The sperm proteosome as a whole has been analysed with respect to human infertility and 20 proteins have been identified to be associated with fertility issues (Lefievre et al., 2003; Pixton et al., 2004; Rawe et al., 2008). A similar study in the house mouse *M. musculus* shortlisted 132 proteins that may affect fertility, some of which seem to be evolutionarily preserved across taxonomic groups (Chu et al., 2006). These findings suggest a potential major role for proteins in transgenerational epigenetics.

In fact, in non-rodent mammalian fertilisation, the centriole-centrosome is inherited through the sperm and acts as a template for all subsequent cell divisions from early embryogenesis into adulthood. Any malformations of this complex result in severe infertility due to disruption or insufficiency during mitotic divisions and may hence cause developmental problems anywhere from interrupting the first mitotic divisions to causing embryonic malformations (Schatten and Sun, 2013). The centriole-centrosome complex likely varies in its shape and therefore function also among fertile males, and these more subtle variations may contribute to the fitness and performance of the offspring in the next generation.

In a recent review, Harvey et al. (2018) proposed that prions are ideal candidates for non-genetic transgenerational inheritance due to their conformational flexibility and their ability to transform into self-templating folds, which allows them to proliferate independently even across generations.
Prions are considerably more stable during meiotic processes compared to other epigenetic factors experiencing major re-structuring (Cox, 1965; Young and Cox, 1971). The independence and stability of prions may imply that protein-based transgenerational inheritance could be important but the idea needs careful testing.

Histone modifications
Although histone modifications could be regarded as part of the sperm proteome, I discuss them separately as they have received a lot of attention in the context of trans-generational epigenetics. Modifications of the histones are assumed to affect gene expression and therefore may play a key role in gene regulation (e.g. Kouzarides, 2007). Gene regulation is particularly important during the early stages of development and any marks inherited from the father may contribute to embryonic gene expression – with potential effects later on in life. In mammalian sperm, 90 (in humans) to 95% (in house mice) of histones are replaced by protamines during spermatogenesis, and the remaining histones may undergo post-translational modifications affecting gene expression at these loci (Luense et al., 2016). These post-translational modifications may regulate gene expression during spermatogenesis and during early embryo development (Brykczyńska et al., 2010; Hammoud et al., 2011; Erkek et al., 2013; Brunner et al., 2014). In human sperm, histone modifications appear to be particularly enriched around developmental loci. Dimethylated lysine 4 on histone H3 (H3K4me2) for example, is found at promoter loci, whereas H3K4me3 is found in large clusters of paternally expressed imprinted genes, miRNAs and HOX genes (Hammoud et al.,
These patterns also seem to hold for the zebrafish *Danio rerio*, where sperm retain the histones and lack protamines altogether, but chromatin markers such as permissive H3K4me3 with or without repressive H3K9me3 or H3K27me3 are associated with developmental loci (Lindeman et al., 2011). A study manipulating the dietary conditions in male house mice observed differential gene expression in the next generation and found a consistent decrease in H3K27me3 at the promoter of monoamine oxidase in sperm of low-protein diet males compared to control males (Carone et al., 2010).

**Current challenges and future directions**

The study of paternal epigenetic effects inherited across generations is still in its early days and many fundamental questions are currently unanswered. The many unfilled gaps and fundamental unknowns put limitations to our ability to summarise the relative importance, prevalence, and/or impact of each of the factors discussed. It may be worth identifying some of the key aspects that we should focus on in the near future.

The term “sperm factor” may be somewhat misleading in being an oversimplification of what is clearly a varied set of highly complex factors. One of the questions is therefore: How are the different mechanisms linked? Understanding whether the different epigenetic components act independently, complementarily, additively, or interactively and how these interactions and the resulting effects may be context-dependent are some of the challenges we are currently facing. The interaction between some of the factors such as the tight linkage between DNA methylation and histone
modifications for the regulation of gene expression during proliferation and differentiation of the germline is relatively well understood. In contrast, other factors are still largely a black box (piRNAs), and many have not even been properly identified yet (other small RNAs). Carefully designed experiments combined with the latest –omics technology may be a valuable way to gain insights into what are clearly highly complex processes.

Another currently open question is whether the non-genetic transfer of information in sperm is truly “non-genetic” or whether there is a causal connection between the non-genetic information and the underlying genome. Non-genetic factors may fall into one of three possible categories: (I) independent of sequence variation, (II) partially dependent on sequence variation, and (III) completely dependent on sequence variation (based on epiallelic variation as proposed by Richards, 2006). An additional aspect that needs to be considered is whether the transfer of information is based (A) purely on transmitted genes or (B) on a combination of transmitted genes and non-genetic material. In case III, all the observed variation should be explained by focusing exclusively on sequence variation and the distinction between scenario A and B is not necessary. However, in cases I and II, sequence variation will not explain everything as non-genetic material may be generated independently and add variation through non-genetic mutations occurring between transcription events. Performing experimental manipulations of paternally experienced environmental conditions in combination with long-read DNA sequencing, RNA sequencing of different
RNA families, ChIP sequencing and bisulfite sequencing is not an easy but a promising way forward to answer these questions.

The importance of the relative timing and duration of changes in environmental conditions experienced by a male to affect the following generation(s) is still poorly understood. In mammals (and probably most other taxa), early embryo development is a particularly sensitive period and methylation patterns and histone modifications are strongly affected by environmental conditions during this time (reviewed in Faulk and Dolinoy, 2011). However, effects across generations have also been shown in studies where males were exposed to stressful environments as juveniles before sexual maturity (e.g. Gapp et al., 2014), during adulthood (e.g. Carone et al., 2010) or both (e.g. Rodgers et al., 2013). It would be interesting to understand, which epigenetic factors are mostly affected by environmental conditions in the male germline during each of these life stages and how strong the observed transgenerational effects are relative to each other.

Of particular relevance for the fields of ecology and evolution is the question about the stability of epigenetic alterations. While some epigenetic marks are stable and conserved even across taxa (Provataris et al., 2018), others are seemingly more apt to change. Having said that, even sRNAs can be transferred across many generations without further stimulation in a self-regulating process (Rechavi et al., 2014) suggesting that such systems may provide a reliable way to memorise environmental conditions. Understanding
the flexibility and stability of epigenetic mechanisms is important to fully assess their relative contribution to inheritance.

Finally, ejaculates generally consist of more than just sperm, and we know that the content of seminal fluids may have severe effects on female fitness (Chapman et al., 1995; Wolfner, 2002), and also on their offspring (Chapman et al., 2001; Crean et al., 2014; Crean et al., 2016). Controlling for such effects and disentangling factors carried by sperm from factors in the seminal fluid will be imperative when studying the various mechanisms.

In summary, non-genetic factors transferred through the sperm into the zygote are very likely to affect the resulting generation(s) and this in itself is a very important insight. We now need to understand, which mechanisms contribute to this transfer of information and how and what the true purpose of non-genetic information transferred in sperm across generations is. With a great range of novel tools becoming available and increasingly affordable we should be able to address these important questions.

**Acknowledgements**

I thank F. Spagopoulou and M. Lind for giving me the opportunity to write this review and to A. S.-J. Suh for inspiring discussions and four anonymous reviewers for valuable comments on earlier versions of this manuscript. I was supported by grants from the European Research Council (HAPSELA336633), Human Frontier Science Program (R0025/2015) and the Knut and Alice Wallenberg Foundation.
Conflict of interest

The authors declare no conflict of interest.

References


Figure 1: Illustration of non-genetic components transferred via sperm from the father to the offspring and their putative effects in the offspring. The description of the effects is very general as many of them are currently still poorly understood.
Figure 1

Sperm factors

Proteins
Cell divisions, other effects?

Histone modifications
Gene regulation, TE activity

RNAs
Gene regulation, TE activity

DNA methylation
Gene regulation, TE activity