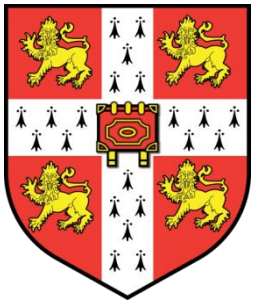
Programming of the paternal nucleus for embryonic development



Marta Teperek

Clare Hall College

University of Cambridge

March 2014

This dissertation is submitted for the degree of Doctor of Philosophy

Supervisor: Dr Jerome Jullien

Co-supervisor: Dr John Gurdon

Preface

The work described in this dissertation has been conducted in the Gurdon Institute, University of Cambridge, Cambridge, United Kingdom, under the supervision of Dr Jerome Jullien and Dr John Gurdon.

This dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration except where specifically indicated in the text.

The work described here has not been submitted to any other university.

Marta Teperek

Statement of length

This dissertation does not exceed the prescribed word limit (i.e., it is below 60 000 words, excluding bibliography, figures and appendices).

Marta Teperek

Acknowledgements

I am extremely grateful to Dr Jerome Jullien, my supervisor, who provided me with his constant, day-to-day support during my rotation in the Gurdon laboratory and then later on through the whole period of my PhD. I am grateful not only for his scientific input, creative ideas, but also for being a patient, understanding and motivating supervisor. I am also immensely grateful for Dr Jullien's sense of humour, which made my days in the lab enjoyable and entertaining.

I am very grateful to Dr John Gurdon, my co-supervisor, who allowed me to do my rotation project as well as my PhD in his laboratory. I am grateful for his input into my work and all his valuable advice. I am also extremely grateful to Dr Gurdon for giving me a lot of freedom in my research, which stimulated my development as an independent scientist.

I would like to thank Dr Vincent Gaggioli, Dr Angela Simeone and Dr Kei Miyamoto for their great help with performing the experiments as well as with the experimental design and analysis. Their input was vital for the progress of this project. I am also grateful to them for being wonderful friends, on top of being excellent scientific collaborators.

I would like to thank all my colleagues from the Gurdon laboratory for creating a nice working atmosphere in the lab, and who were never tired of my questions and who were also fantastic colleagues.

I am thankful to Dr Charles Bradshaw and Dr George Allen for performing vital bioinformatic analyses for this project.

I am extremely grateful to all my friends from the Gurdon Institute, and particularly to Beata Wyspianska, for their endless support and who made my time in the Gurdon Institute such a great experience.

I would like to thank my collaborators from outside the Institute for all the scientific help and input into this project. I am especially grateful to Dr Antoine Peters for hosting me in his laboratory and for teaching me some of his expert techniques.

I would like to thank David Simpson for taking care of our frogs; Renata Feret and other member of the proteomics facility for performing 2-DIGE and mass spectrometry analyses of the data; and Dr Rachael Walker for performing the flow cytometry analysis for me.

I am grateful to my dear friend, Jerzy Wilczynski, for his expert advice on the statistical analysis conducted in Chapters 3 and 5 of this thesis. I would also like to thank him for being a never-ending source of motivation and inspiration for my career and life choices.

I am immensely grateful to my parents and to my sister for their enormous support during all my time in Cambridge and beyond. I would like to thank my parents for their guidance in life and for always telling me (and being a wonderful example of this) that education and self-development is important in one's life. I would like to thank them and my sister for providing me with constant motivation and for never giving up on me. I could have never gone so far without them.

Finally, I would like to thank the funding bodies: the Wellcome Trust and the Medical Research Council UK for their financial support.

Summary

Historically, sperm has been considered merely as a carrier of genetic material at fertilisation. However, it is known that sperm supports embryonic development better than other cell types, suggesting that it might also have additional important, non-genetic contributions to embryonic development. The work described in this dissertation focuses on identifying the molecular determinants of developmental programming of sperm.

First, the development of embryos derived from sperm and spermatids, immature precursors of sperm was compared. Sperm-derived embryos developed significantly better than spermatid-derived embryos. Further research aiming to identify the reasons for the developmental advantage of sperm led to the identification of proteins that are present specifically in sperm and not in spermatids. Moreover, egg factors which are preferentially incorporated into the sperm, but not into the spermatid chromatin were identified with the use of egg extracts, suggesting that the chromatin of sperm could be programmed to interact with the components of the egg.

Subsequently, the reasons for developmental failure of spermatid-derived embryos were investigated. By comparing the sperm with spermatids it was shown that the programming of sperm to support efficient development is linked to its special ability to regulate expression of developmentally-important embryonic genes, and not to its ability to support DNA replication or rRNA production. Further characterisation of the sperm and spermatid chromatin with the use of genome-wide sequencing allowed me to link the correct regulation of gene expression in the embryo with a certain combination of epigenetic marks in the sperm, but not in the spermatid chromatin. Finally, it is shown that enzymatic removal of epigenetic modifications at fertilisation leads to misregulation of gene expression. This therefore suggests that epigenetic information contained in parental genomes at fertilisation is required for a proper regulation of embryonic transcription.

My results support the hypothesis that the sperm is not only a carrier of genetic material, but also provides the embryo with epigenetic information for regulation of transcription after fertilisation. I believe that these findings advance our current understanding of the nature and mechanisms of sperm programming for embryonic development, and are important contributions to the emerging field of transgenerational inheritance of epigenetic traits in general.

Table of Contents

Chapter 1
Introduction

1.1. Stabilisation of the differentiated state by epigenetic mechanisms

Vertebrate development starts when the sperm fertilises the egg. This leads to the formation of a totipotent zygote, which subsequently replicates its DNA and undergoes series of cell divisions. Later on embryonic cells start to differentiate to form multiple tissues and organs constituting the adult body. It is very important that such differentiated cells forming particular tissues do not de-differentiate and do not change their fates to other cell types, as this could lead to tissue malfunction and tumour formation. That is why a fully differentiated state is thought to be safe-guarded by several layers of stabilising mechanisms. Virtually all the cells in the body have the same genetic content, as evidenced by nuclear transfer experiments (see below) (Gurdon 1962, Gurdon & Uehlinger 1966, Hochedlinger & Jaenisch 2002, Eggan *et al.* 2004, Sung *et al.* 2006). This means that differential expression of genes as cells specialise in their developmental pathways is not the result of gene loss or gain. In other words, features that make cells in a body different from each other are epigenetic. The word 'epigenetic' is derived from the Greek preposition 'epi' meaning 'above, outside, besides' combined with the word 'genetic', therefore meaning 'outside genetic'. Currently this definition is used to describe heritable, non-genetic changes in cellular states (Bonasio *et al.* 2010). Interestingly, it has been observed that the DNA on chromosome (and in the nucleus) tends to be clustered according to its transcriptional activity: regions of chromatin which are transcriptionally active and decondensed are termed 'euchromatin' and regions which are transcriptionally repressed and densely packed are termed 'heterochromatin'. Epigenetic mechanisms were proposed to be involved in the stabilisation and separation of

these two different states (Lamond & Earnshaw 1998, Noma *et al.* 2001). There are various epigenetic mechanisms that can contribute to transcriptional activation and silencing. The best characterised epigenetic mechanisms stabilising the expression states of genes are DNA methylation, post-translational histone modifications, presence of histone variants or density of nucleosomes itself. Such mechanisms can act cooperatively to achieve multiple layers of epigenetic mechanisms ensuring stable maintenance of chromatin states. All these mechanisms are briefly reviewed below.

1.1.1. DNA methylation

Methylation of cytosine is the best described epigenetic modification occurring on the DNA itself. Cytosines in DNA can be unmethylated or methylated (or be an intermediate between the two states, for example 5-hydroxymethylcytosine, which is however not discussed here). Methylation occurs by covalent modification of cytosine: by an addition of a methyl group to its fifth carbon, which creates 5-methyl-cytosine (5meC). Methylation can occur on cytosines in various contexts (Lister *et al.* 2009), but 5meC in CpG dinucleotides is the best described (Doerfler 2008). CpG dinucleotides are usually overrepresented in gene promoters. Methylation of cytosines in such CpG-rich promoters often leads to gene silencing (Boyes & Bird 1991, Boyes & Bird 1992, Hsieh 1994, Schubeler *et al.* 2000, Chen *et al.* 2001a, Song *et al.* 2005). Conversely, the removal of methyl group from 5meC (DNA demethylation), leads to gene expression (Benvenuto *et al.* 1996, Papageorgis *et al.* 2010, Stengel *et al.* 2010). DNA methylation is implied in the establishment and/or the maintenance of correct gene expression patterns during development, differentiation and tissue specification (Maatouk *et al.* 2006, Song *et al.* 2009). Interestingly, comparison of DNA methylation profiles from 17 distinct adult mouse tissues led to identification of tissue-

specific DNA methylation sites (Hon *et al.* 2013). This finding points towards important roles of DNA methylation during organism development and specialisation. It is therefore not surprising that alteration of these methylation patterns, on a genome-wide, or on a loci-specific level, is associated with genomic instability and with cancers (Gaudet *et al.* 2003, Reddington *et al.* 2013). In support of this hypothesis, it has been well documented that during the germline specification in mouse, DNA methylation patterns are globally erased to allow reprogramming of the germ cells and to allow them to re-gain the pluripotency (Surani *et al.* 2007, Hajkova *et al.* 2010, Popp *et al.* 2010).

1.1.2. Post-translational histone modifications

Post-translational histone modifications are on the other hand not related to changes to DNA bases itself, but to histones – small basic proteins packing the DNA within the nucleus. There are four major types of histones which form the nucleosome – H2A, H2B, H3 and H4. Furthermore, inter-nucleosomal regions of DNA are bound by linker histone H1. All these histones can be subjected to post-translational modifications, which can affect the chromatin compaction and in turn lead to gene expression or repression or to stabilisation/destabilisation of these states (Strahl & Allis 2000, Jenuwein & Allis 2001, Godde & Ura 2008). For example, it is known that dimethylation or trimethylation of lysine 4 on histone 3 (H3K4me2/3), methylation of argininie 17 or arginine 26 of histone H3 or histone acetylation are associated with transcriptional activation (Simpson 1978, Chahal *et al.* 1980, Turner *et al.* 1992, Jeppesen & Turner 1993, Lee *et al.* 1993, Ogryzko *et al.* 1996, Noma *et al.* 2001, Bernstein *et al.* 2002, Santos-Rosa *et al.* 2002, Torres-Padilla *et al.* 2007, Wu *et al.* 2009). On the other hand, ubiquitination of lysine 119 of histone H2A (H2AK119Ub), di- or trimethylation of lysine 9 of histone H3 (H3K9me2/3) or trimethylation of lysine 27 of

histone H3 (H3K27me3) are associated with transcriptional repression (Lachner *et al.* 2001, Nakayama *et al.* 2001, Noma *et al.* 2001, Cao *et al.* 2002, Czermin *et al.* 2002, Muller *et al.* 2002, Snowden *et al.* 2002, Plath *et al.* 2003, Wang *et al.* 2004, Richly *et al.* 2010). The effect of post-translational histone modifications on the chromatin state can be direct or indirect. It has been demonstrated that acetylation directly decreases the affinity of histones for DNA and therefore makes DNA more accessible to transcription factors (Cary *et al.* 1982). Non-modified histones are positively charged, which facilitates their binding to negatively charged DNA. Acetylation brings a negative charge to the histone, which therefore decreases its affinity to DNA. Interestingly, post-translational histone marks can also indirectly affect the chromatin structure, by the recruitment of proteins that recognize such marks, known as chromatin 'readers'. Acetylated histones are recognized by proteins containing bromodomains, for example Brd2, Brd3 or Brd4 (Bromodomain proteins 2, 3, 4), which, by various mechanisms, can in turn facilitate transcription (Dey *et al.* 2003, Kanno *et al.* 2004, LeRoy *et al.* 2008, Dey *et al.* 2009, Umehara *et al.* 2010a, Umehara *et al.* 2010b, Zhao *et al.* 2011, Draker *et al.* 2012). It has been shown that Brd4-bound acetylated histones can recruit transcription elongation factor P-TEFb, which leads to transcriptional elongation and gene activation (Jang *et al.* 2005, Yang *et al.* 2005, Hargreaves *et al.* 2009, Zippo *et al.* 2009, Liu *et al.* 2013). Conversely, recognition and binding of repressive H3K9me2/3 mark by heterochromatin protein 1 (HP1) leads to heterochromatin formation and transcriptional repression (Bannister *et al.* 2001, Lachner *et al.* 2001, Motamedi *et al.* 2008).

1.1.3. Histone variants

There are also numerous histone variants that can compose the chromatin and presence of these variants instead of canonical histones also affects the chromatin

accessibility (Talbert *et al.* 2012). For example, H3 variant, histone H3.3 is associated with open chromatin state and active transcription and is preferentially incorporated into the paternal pronucleus after fertilisation (McKittrick *et al.* 2004, Loppin *et al.* 2005, Torres-Padilla *et al.* 2006, Jullien *et al.* 2012). Interestingly, genome-wide analysis of H3.3 binding revealed that on top of being enriched at actively transcribed genes (transcriptional start sites and gene bodies), H3.3 was also enriched at transcriptional start sites (but not gene bodies) of repressed genes and was also shown to be important for the establishment of the pericentromeric heterochromatin mouse embryos (Goldberg *et al.* 2010, Santenard *et al.* 2010). Another histone variant, H2A.Z, was shown to occupy developmentally-important gene promoters in embryonic stem cells (ES cells) and its downregulation led to overexpression of some of its target genes, suggesting that it is important for gene repression (Creyghton *et al.* 2008). Interestingly, presence of the very same histone variant H2A.Z in a combination with acetylated forms of histone H4, leads to recruitment of Brd2 and, in turn, to transcriptional activation (Draker *et al.* 2012). It has been also shown that presence of H2A.Z histone variant is essential for the removal of nucleosomes from the transcriptional start sites (TSSs) of tissue-specific genes during ES cells differentiation (Li *et al.* 2012) and that it generally facilitates transcription-coupled nucleosome removal from gene TSSs (Schones *et al.* 2008). These suggest that the presence of the very same histone variant can have different effects on transcription, depending on the chromatin context. Furthermore, combination of different histone variants can have additive effects on chromatin structure. It was shown that nucleosomes assembled with histone variants H3.3 and H2A.Z were more unstable than nucleosomes containing H3.3 and H2A or H3 and H2A.Z alone (Jin & Felsenfeld 2007). Instability of nucleosomes can then positively affect transcription, as discussed above. Besides, not only core histones have their variants. Linker histone H1 also has multiple variants. For example, there are 11 variants of histone H1 in humans and mice (Izzo *et al.*

2008), some of which are oocyte/early embryo-specific (H1foo) or testis-specific (H1t, H1t2 or Hils1) (Seyedin & Kistler 1980, Tanaka *et al.* 2001, Yan *et al.* 2003, Martianov *et al.* 2005, Tanaka *et al.* 2005). Interestingly, it has been shown that early embryonic variant of histone H1 (dBigH1) in *Drosophila* is crucial for regulation of zygotic gene activation (ZGA): deletion of this histone variant resulted in pre-mature ZGA and embryo death (Perez-Montero *et al.* 2013). Oocyte and embryo-specific linker histone variant, B4, a homologue of the mammalian H1foo, has been also identified in frogs (Smith *et al.* 1988, Tanaka *et al.* 2001). It was shown that the affinity of B4 histone for DNA is 6 times lower than that of a canonical, somatic-type H1. The incorporation of B4 instead of canonical H1 into chromatin facilitated transcription (Ura *et al.* 1996). Interestingly, the very same histone B4 was shown to be incorporated into chromatin of transplanted nuclei during the nuclear transfer procedure and to be necessary for transcriptional reprogramming (Teranishi *et al.* 2004, Jullien *et al.* 2010). It is therefore not surprising that the embryonic-like linker histone variants are lost as embryos develop, probably in order to prevent promiscuous transcription and enable tissue specification (Smith *et al.* 1988). All these suggest that the mere composition of the nucleosomes can affect the susceptibility of chromatin to transcription.

1.1.4. Nucleosome density

The positioning of nucleosomes on DNA and, specifically, the density of nucleosomes at particular genomic sites, is also known to affect the accessibility of the chromatin. In general, the higher the nucleosome density, the more compact the chromatin structure and the less accessible for transcriptional machinery (Kwon *et al.* 1994, Bai & Morozov 2010). Comparison of genome-wide profiles of nucleosome occupancy in seven different human cell lines revealed the presence of nucleosome-free regions at TSSs of

transcriptionally active genes (Ozsolak *et al.* 2007). It was also shown that during ES cell differentiation, the activation of genes correlated with nucleosome depletion around their transcriptional start sites (TSSs) (Li *et al.* 2012). Furthermore, in yeast nucleosome occupancy is inversely correlated with transcriptional activity not only at gene promoters, but also across the gene body: nucleosomes are depleted at the promoters of highly transcribed genes and are more abundant at promoters and at gene bodies of repressed genes (Lee *et al.* 2004). Interestingly, it was shown in zebrafish embryos that the canonical organisation of nucleosomes on chromatin is detected only from around the time of zygotic genome activation, suggesting that proper positioning of nucleosomes on chromatin is important for the regulation of gene expression (Zhang *et al.* 2014). All these results support the model that the mere presence of nucleosome modulates the accessibility of the underlying DNA sequence.

1.1.5. Cooperative action of different epigenetic mechanism to achieve stable maintenance of a differentiated state

It was demonstrated that various epigenetic mechanisms mentioned before can act cooperatively to achieve multiple layers of epigenetic stabilisation of chromatin states. For example, it has been shown that cytosine methylation is recognized by proteins that have methylated-cytosine binding domains (MBD): MeCP2, Mbd1, Mbd2, Mbd3 and Mbd4 (Bogdanovic & Veenstra 2009). These proteins then recruit other factors, which can further induce heterochromatinization of the DNA-methylated region. For example, it has been shown that MBD domain-containing proteins can recruit histone deacetylases (HDACs) to chromatin. These are responsible for removing the acetyl groups from histones and therefore lead to chromatin compaction and transcriptional repression (Jones *et al.* 1998, Wade *et al.*

1999, Zhang *et al.* 1999, Le Guezennec *et al.* 2006). DNA methylation can be also reciprocally induced by the presence of certain histone marks. For example, HP1 protein that binds to H3K9me2/3 interacts with Dnmt1 and Dnmt3b, which are maintenance and *de novo* DNA methylases, respectively. This interaction is important for the induction of DNA methylation and gene repression (Lehnertz *et al.* 2003, Smallwood *et al.* 2007). Presence of DNA methylation can regulate the distribution of repressive histone marks also by restricting them. It has been shown that DNA methylation is mutually exclusive to H3K27me3 mark (Lindroth *et al.* 2008). Removal of DNA methylation leads to widespread accumulation of H3K27me3 at illegitimate regions (Hagarman *et al.* 2013, Reddington *et al.* 2013). In contrast, the removal of H3K27me3 results only in mild changes in 5meC distribution across the genome (Hagarman *et al.* 2013). Interestingly, it was shown that unmethylated CpG islands can be recognized and bound by Kdm2b, which is a histone H3 lysine 36 demethylase. Kdm2b then recruits Polycomb repressive complex 1 (PRC1), which deposits H2AK119Ub, leading to gene repression (Farcas *et al.* 2012, He *et al.* 2013). Presence of certain histone variants in the nucleosome can also influence the deposition of epigenetic marks. For example, in mouse ES cells, the presence of H3.3 is necessary for proper establishment of H3K27me3 mark at promoters of developmentally-important genes (Banaszynski *et al.* 2013).

Concluding, various epigenetic mechanisms can act independently or synergistically in several layers of regulation to ensure a stable maintenance of a differentiated state (Fig. 1).

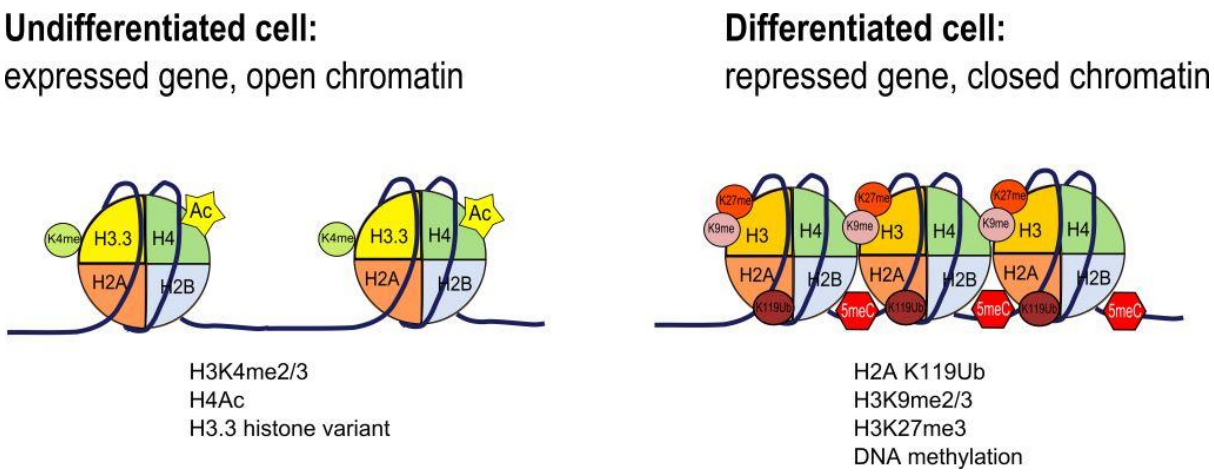


Fig. 1. Epigenetic mechanisms stabilising a differentiated state.

Example of a gene which is expressed in an undifferentiated cell and repressed in a differentiated cell. Repression of this gene is achieved and stabilised by several layers of epigenetic mechanisms: DNA becomes methylated, nucleosomes occupancy increases, histone variants associated with active transcription are replaced with canonical histone variants, activating post-translational histone marks are removed and repressive marks established.

1.2. Reprogramming as the reversal of differentiation

Interestingly, epigenetically-stabilised differentiated states can be experimentally reversed back to embryonic-like pluripotent states. It has been shown that if a nucleus of a somatic cell is transplanted to an unfertilised, enucleated egg, this can lead to a development of a new embryo, which can subsequently differentiate into a an adult organism (Gurdon *et al.* 1958, Gurdon & Uehlinger 1966). It was also proved with the use of nuclear markers that the genetic material in such cloned adult organisms was derived exclusively from the donor cell nucleus and not from the recipient egg (Elsdale *et al.* 1960, Hochedlinger & Jaenisch 2002). It means that a differentiated cell was able to de-differentiate to a totipotent state. However, it has been also shown that this process of de-differentiation is limited: the more differentiated the donor cell nucleus, the less frequent is the normal development of resulting nuclear transfer embryos. For example, it was shown in frogs that if donor cells for the nuclear transfer were derived from blastula or early gastrula stage embryos, up to 36% of successfully reconstructed embryos could reach a feeding tadpole stage, whereas if intestinal epithelium cell nuclei were used as donors, only 1.5% of embryos reached a feeding tadpole stage (Gurdon 1962). It has been further explained that the decreased susceptibility of differentiated cells to undergo reprogramming is related to the fact that multiple layers of epigenetic stabilisation of the differentiated state are acquired sequentially as the development progresses: the more differentiated the cell, the more epigenetic mechanisms safeguard the differentiated state and the more difficult it is to successfully de-differentiate such cells (Pasque *et al.* 2010, Pasque *et al.* 2011b). It was reported in mammalian systems that nuclear transfer-derived embryos exhibit abnormally high DNA methylation levels as compared with control, fertilised embryos. Moreover, these non-reprogrammable methylation abnormalities were memories of the methylation states in the donor cells used for the nuclear transfer: similarly high methylation pattern were also detected in the donor cells (Bourc'his *et*

al. 2001, Kang *et al.* 2001, Chan *et al.* 2012). This suggests that repressive marks acquired during differentiation were a barrier for reprogramming. It has been also shown that histone variants can be a barrier for reprogramming. Experiments using the transcriptional reprogramming system of nuclear transfer to *Xenopus* prophase-arrested oocytes (Fig. 2) demonstrated that mouse embryonic fibroblasts (MEFs) reprogram worse than less differentiated epiblast stem cells due to the presence of repressive histone variant macroH2A in MEFs (Pasque *et al.* 2011a). Post-translational histone marks themselves were also shown to be refractory to correct reprogramming after nuclear transfer and to be responsible for the retention of the donor cell-like characteristics in nuclear transfer embryos. For example, cloned bovine embryos 'remembered' the histone H4 lysine 5 acetylation levels from the donor cells, which was also correlated with levels of gene expression in such embryos. Interestingly, this memory of histone acetylation level was retained even when blastomeres derived from the first round of nuclear transfers were used as donors for the second round of nuclear transfers, suggesting persistence of the memory through generations (Wee *et al.* 2006). The best documented and molecularly characterised phenomenon of the epigenetic memory in nuclear transfer embryos comes from experiments in frogs. It was shown that embryos derived from nuclear transfer with somite cells (muscle precursors) showed expression of MyoD gene (which is a gene normally expressed in somite tissues) in illegitimate regions of the embryo, in which this gene is not normally expressed (Ng & Gurdon 2008). This memory was dependent on the presence of lysine 4 of H3.3 histone variant at MyoD gene promoter (Ng & Gurdon 2008). This suggests that problems with resetting the epigenetic landscape apply not only to genes which are repressed and cannot be properly activated in the embryo, but also to those for which activating epigenetic environment inherited from the donor cell causes illegitimate gene expression in the resulting embryo.

Concluding, even though the process of cell differentiation can be reversed experimentally, nuclear reprogramming has its limitations and in general is inefficient; especially when highly differentiated cells are used as nuclear donors.

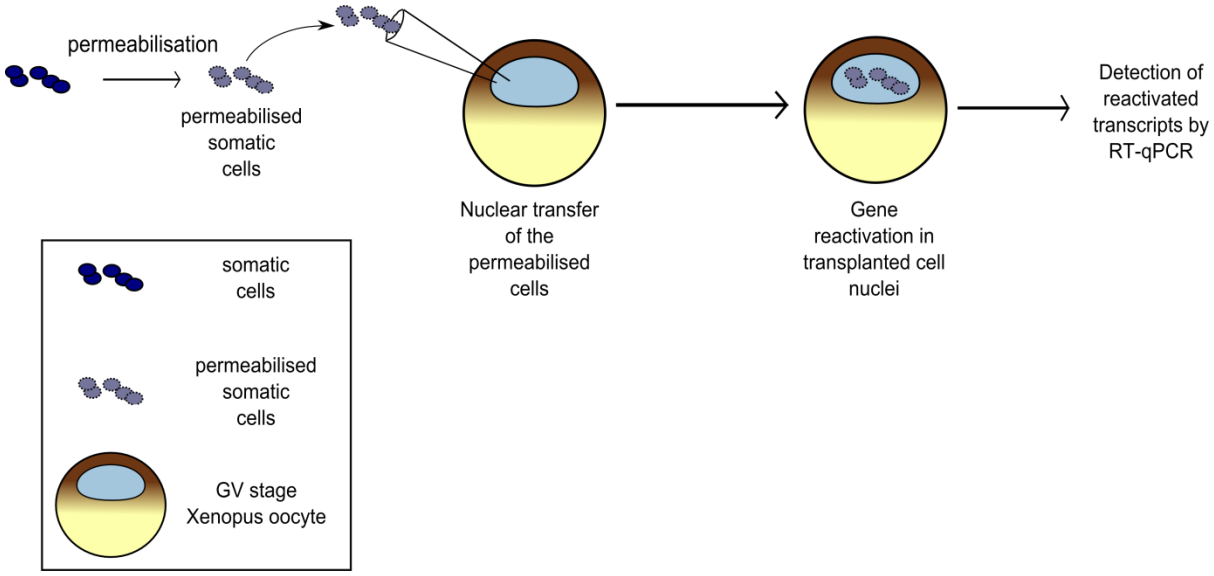


Fig. 2. Transcriptional reprogramming system of the nuclear transfer to *Xenopus* prophase-arrested oocytes.

Somatic cells are first permeabilised and subsequently transplanted into the nucleus (germinal vesicle – GV) of a prophase-arrested Xenopus laevis oocyte. After the transplantation, nuclei of somatic cells reactivate gene transcription. This transcription can be measured by RT-qPCR.

1.3. Sperm is a highly specialised cell

1.3.1. Global changes occurring during spermiogenesis

The sperm cell is highly specialised - designed to deliver genetic material to the egg at fertilisation. Spermatogenesis is the process in which a pluripotent germ cell differentiates into a mature spermatozoon. It starts when the germ cell precursors of the sperm, spermatogonia undergo a cell division, which can be either proliferative or differentiative. The first one produces more spermatogonial cells, the second one results in a formation of a spermatogonial cell and a more differentiated precursor – a primary spermatocyte (de Rooij 2001) (Fig.3). Primary spermatocytes reduce their chromosome and DNA content in a meiotic division and eventually, as a result of meiosis, give rise to four haploid spermatids (Roosen-Runge 1969). Spermatids are immediate precursors of sperm, but in order to be transformed into mature spermatozoa, they have to complete a series of substantial structural and morphological changes, called spermiogenesis (Toshimori 2003) (Fig. 3).

One of the most dramatic changes occurring during spermiogenesis is a global compaction of the chromatin. The molecular basis behind the compaction of the sperm nucleus is the best described in mammalian systems, so in this section I will review what is known about this process in mammalian systems. It has been calculated that the volume of chromatin in mature sperm in mammals is around six times smaller than the volume of mitotic chromosomes (Ward & Coffey 1991). Such a high compaction of the sperm nucleus is possible due to the presence of protamines. Protamines are small basic proteins, which have higher affinity for DNA than histones, allowing a tight packaging of sperm DNA (Balhorn 2007). They are incorporated into the sperm chromatin in place of core histones in a multi-step process. Initially, the chromatin in the sperm is globally hyperacetylated (Hazzouri

et al. 2000, Govin *et al.* 2004), which is thought to play at least two roles in promoting chromatin compaction. First, acetylated lysine residues of histones are recognized and bound by a protein called Brdt (Bromodomain-testis specific), which is suggested to be involved in histone eviction (Pivot-Pajot *et al.* 2003, Govin *et al.* 2006, Moriniere *et al.* 2009, Gaucher *et al.* 2012). Second, it has been recently shown that acetylation of core histones is also directly involved in their proteasome-mediated degradation during sperm maturation (Qian *et al.* 2013). Core histones are replaced with testis-specific histone variants, such as hTSH2B, H2AL2 or H3t, which are thought to further increase the instability of the nucleosome structure (Li *et al.* 2005, Syed *et al.* 2009, Tachiwana *et al.* 2010). Subsequently, histones are replaced with transition proteins (Yu *et al.* 2000, Meistrich *et al.* 2003, Zhao *et al.* 2004), which are ultimately exchanged for protamines in the final stages of sperm maturation (Chen *et al.* 2001b, Cho *et al.* 2003, Balhorn 2007, Ravel *et al.* 2007, Steger *et al.* 2008). These nuclear changes are accompanied by cessation of transcription and disappearance of the basal transcriptional machinery – as opposed to spermatids, mature spermatozoa are transcriptionally silent (Monesi *et al.* 1978, Martianov *et al.* 2001, Zheng *et al.* 2008) (Fig.4).

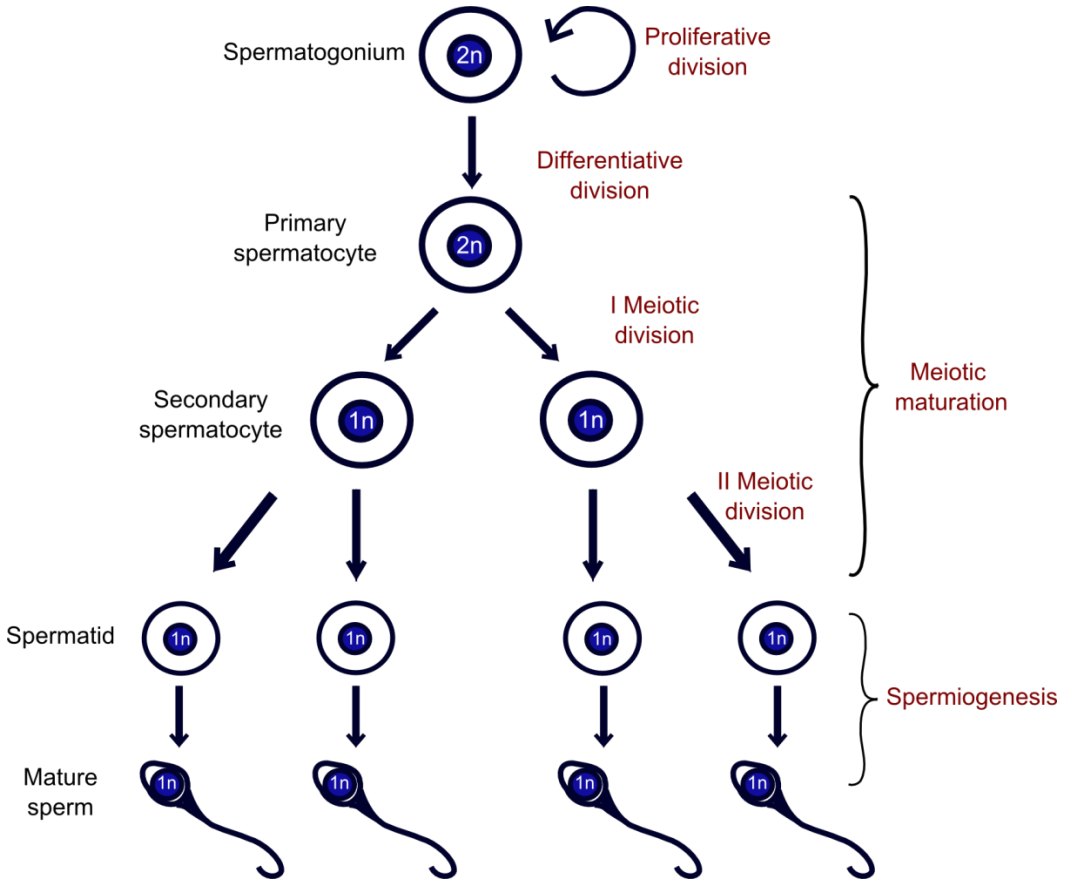


Fig. 3. Spermatogenesis.

The diagram depicts sequential stages of spermatogenesis. Spermatogenesis consists of meiotic maturation, in which the chromosome and DNA content is reduced and of spermiogenesis, in which immature spermatid maturates into spermatozoon.

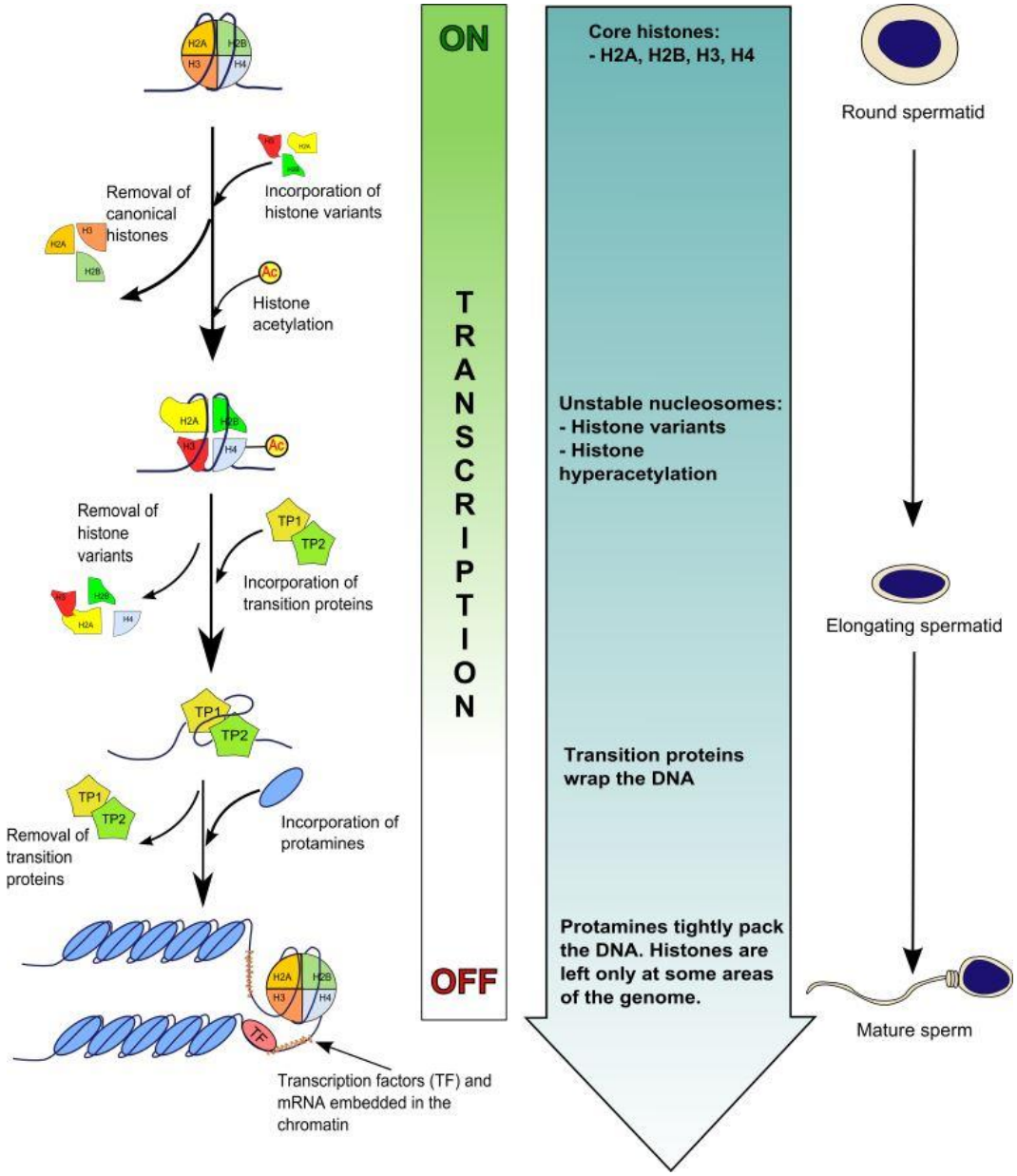


Fig. 4. Major nuclear changes during spermiogenesis (as described in mammalian systems) - adapted from (Teperek & Miyamoto 2013).

The round spermatid undergoes a series of chromatin remodelling events that lead to the formation of a fully mature spermatozoon. First, canonical core histones packing the chromatin in the round spermatid are replaced by histone variants, which, together with global histone acetylation, leads to instability of the nucleosome structure. Subsequently, transition proteins are incorporated in place of unstable nucleosomes in the elongating spermatid. Finally, transition proteins are replaced with protamines. The mature sperm chromatin is mainly composed of protamines, with interspersed histones and with tightly associated mRNAs and transcription factors (see chapter 1.4.2). All these processes are occurring in parallel with the cessation of transcription – round spermatids are transcriptionally active, whereas no transcription is detected in the mature sperm.

1.3.2. Epigenetic marks are retained in the chromatin of mature sperm

Interestingly, it has been shown that not all the regions in the sperm chromatin are subjected to histone to protamine replacement. Regions which retain histones in the chromatin of mature sperm in human were shown to be endonuclease-sensitive and to be enriched for gene promoters and other regulatory sequences of the genome, such as enhancers (Arpanahi *et al.* 2009). It was also shown that among the genes that retain histones at their promoters in mature sperm in mouse and human, there is a significant overrepresentation of genes that are developmentally-important (Hammoud *et al.* 2009, Brykczynska *et al.* 2010). Furthermore, histones retained in sperm turned out to be post-translationally modified at specific regions of the chromatin in mouse and human. Some of the histone modifications, like acetylation of H4K12, (Paradowska *et al.* 2012), histone crotonylation (Tan *et al.* 2011) or H3K4me2 and H3K4me3 (Hammoud *et al.* 2009, Brykczynska *et al.* 2010) localise to genes which are highly active during spermiogenesis, suggesting that they are important for the regulation of expression of spermiogenesis-related genes. Therefore, the presence of these marks at the promoters of these genes in the chromatin of mature sperm likely reflects their expression at earlier stages. Strikingly however, other marks like H3K27me3, but also H3K4me2 and H3K4me3 (and a combination of H3K4me2/3 with H3K27me3), also marked promoters of genes that are developmentally-important, for example Hox genes, that are involved in embryo patterning (Hammoud *et al.* 2009, Brykczynska *et al.* 2010). Furthermore, comparison of sperm epigenetic landscapes with embryonic transcriptome revealed the following correlation: genes that have activating epigenetic marks in sperm tend to be expressed early in development, whereas those with repressive marks are activated late or silenced in early embryonic stages (Hammoud *et al.* 2009, Brykczynska *et al.* 2010). Surprisingly, similar observations were also made in zebrafish, which does not have

protamines packing the DNA in the mature sperm. In zebrafish sperm, developmentally-important genes were enriched for histone variant H2AFV, for H3K4me2/3, H3K27me3 and for H3K36me3 at their promoter sequences (Wu *et al.* 2011). Genes bearing H3K4me3 together with H3K14 acetylation (H3K14ac) in sperm were enriched at promoters of genes related to spermiogenesis and also for housekeeping genes (Wu *et al.* 2011). Analogically to what was observed in mammals, genes with activating marks at their promoters in sperm (H3K4me2/3 and H3K14ac) tend to be expressed early in embryonic development, whereas genes with repressive mark H3K27me3 or enriched for histone variant H2AFV, tend to be repressed at the earliest developmental stages in the embryo (Wu *et al.* 2011). Interestingly, a mass-spectrometry based study of post-translational histone and protamine marks in mouse sperm, revealed that not only histones, but also the protamines, are subjected to numerous post-translational modifications (Brunner *et al.* 2014).

Similarly to histone marks, DNA methylation patterns also show uneven distribution in the genome of mature sperm. Overall, cytosine methylation patterns in gene promoters of sperm are very similar to those of embryonic stem cells (Farthing *et al.* 2008). Interestingly, it was shown that developmentally-important genes were hypomethylated in sperm in mammals, as well as in zebrafish (Hammoud *et al.* 2009, Brykczynska *et al.* 2010, Wu *et al.* 2011). Importantly, experiments in zebrafish demonstrated that DNA methylation pattern of sperm, but not of oocytes, is recapitulated in the embryo at the time of zygotic genome activation. The oocyte methylation pattern had to be adjusted to match the sperm methylation pattern by a passive loss of DNA methylation (at loci methylated in an oocyte-specific manner) or by undergoing *de novo* methylation (for these loci, that were unmethylated in oocytes, but methylated in sperm) (Jiang *et al.* 2013, Potok *et al.* 2013). Experiments performed using *in vitro* egg extracts in *Xenopus laevis* showed that maintenance of the sperm-derived DNA methylation after replication is dependent on the recruitment of the

maintenance DNA methylase via UHRF1-mediated ubiquitination of lysine 23 on histone H3 (Nishiyama *et al.* 2013).

1.3.3. Chromatin composition of the mature sperm in *Xenopus laevis*

Global condensation of the mature sperm nucleus is also observed in *Xenopus laevis*. Six protamine-like sperm basic proteins (Sps) have been identified in *Xenopus*: Sp1-6 (Abe & Hiyoshi 1991, Hiyoshi *et al.* 1991, Yokota 1991, Frehlick *et al.* 2007). Furthermore, it has been shown that mature *Xenopus* sperm cells are devoid of somatic type of linker histone H1, and instead have linker histone variants composing their chromatin, like H1fx and H1.sp (Abe & Hiyoshi 1991, Shechter *et al.* 2009). Interestingly, it has been also shown that the levels of somatic histone variants H2A and H2B, but not H3 and H4, are significantly reduced in *Xenopus* sperm (Abe & Hiyoshi 1991, Yokota 1991, Shechter *et al.* 2009). Furthermore, mass spectrometry and immunoblotting analyses in *Xenopus laevis* sperm identified the presence of various post-translational histone marks on the sperm chromatin (Nicklay *et al.* 2009, Shechter *et al.* 2009) (Table 1). However, nothing is known about the positioning of these marks on the sperm chromatin or about the function (if any) of these marks in the future embryo development.

Table1. Histone variants and post-translational histone marks identified in *Xenopus laevis* sperm*

Histone variants	H1fx, H1.sp, H2A.Z,
Modifications on H3	*Methylation:* K4me1/2, K9me1/3, K27me1/2/3, K36me1/2, K79me1/2, R2me2a, R17me2a *Acetylation:* K18ac1, K23ac2,
Modifications on H4	*Acetylation:* K5ac2/3, K8ac1/2/3, K12ac1/2/3, K16ac1/2/3 *Phosphorylation:* S1ph *Methylation:* K20me1/2/3, K79me2, R3me1, R3me2s

* - information from (Nicklay *et al.* 2009, Shechter *et al.* 2009).

1.4. Why is the sperm better than a somatic cell at supporting embryonic development?

As mentioned before, the differentiation state of a somatic cell can be reversed back to pluripotency by reprogramming, for example by transferring the nucleus of such a differentiated cell to an unfertilised, enucleated egg, which in some cases can lead to a full-term embryonic development. However, it was also mentioned that the nuclear transfer procedure is generally highly inefficient. This inefficiency remains in contrast to the high rates of successful embryonic development when the sperm fertilises the egg. There are definitely some similarities between events occurring after somatic cell nuclear transfer and after fertilisation. First, upon introduction into the egg cytoplasm, both the chromatin of the sperm and the chromatin of the somatic cell nucleus decondense and increase in volume (Gurdon 1976, Lassalle & Testart 1991, McLay & Clarke 2003). Global chromatin changes accompany the decondensation process. During fertilisation, sperm-specific protamines are replaced with histones deposited in the egg (McLay & Clarke 2003). In nuclear transfer, the oocyte-specific histone variants are also incorporated into the chromatin of somatic cells (Jullien *et al.* 2010, Jullien *et al.* 2012). Interestingly, it was shown that in mice the same egg-specific linker histone variant – H1foo is incorporated into sperm chromatin upon fertilisation and into the nucleus of a somatic cell after somatic cell nuclear transfer (Gao *et al.* 2004, Becker *et al.* 2005). Moreover, it was shown that in *Xenopus* both processes are dependent on a protein called Nucleoplasmin, which is maternally deposited in the egg (Philpott *et al.* 1991, Philpott & Leno 1992, Gillespie & Blow 2000, Tamada *et al.* 2006, Ramos *et al.* 2010, Inoue *et al.* 2011, Okuwaki *et al.* 2012). After reprogramming, the DNA of the sperm or of the somatic cell can be replicated and used as a template for RNA synthesis (Aoki *et al.* 1997, Bouniol-Baly *et al.* 1997). These observations would suggest that the embryos generated by both ways undergo similar changes leading to the same outcome – embryonic development.

However, embryos develop much better to the adulthood after fertilisation than after a somatic cells nuclear transfer. Does it mean that the sperm is somehow specialised to support embryonic development?

There are numerous events occurring during spermiogenesis that transform the immature germ cell precursor into a mature sperm that is ready to fertilise the egg. It is possible that these events are important to somehow program the sperm to support embryonic development. Interestingly, ICSI (Intra Cytoplasmic Sperm Injection) experiments in mice showed that a spermatid (Fig. 3), when injected into an unfertilized egg, is a much less efficient donor cell at supporting normal embryonic development than a mature sperm cell (Kishigami *et al.* 2004). Moreover, experiments in mice also proved that when injected into eggs, spermatids that are more advanced in spermiogenesis are more efficient in producing normal embryos than the less advanced ones (Ohta *et al.* 2009). Both results could suggest that acquisition of specific sets of proteins and/or depletion of others may be responsible for the high efficiency of reprogramming and embryo generation upon fertilisation. Despite the fact that no factors directly responsible for differences in the efficiency of distinct sperm cell progenitors have been identified so far, there are some, which are more likely than the others to be involved in the sperm programming. Those include protamines, sperm-derived transcriptional regulators and sperm epigenetic marking. Each one of those is briefly discussed below in the context of its potential ability to confer to the sperm developmental advantage when compared with a somatic cell.

1.4.1. Protamines

A striking difference between the sperm and the somatic cell is the packaging of their DNA. DNA in sperm is wrapped around protamines (at least in the majority of the vertebrate species), whereas histones are the core unit organising chromatin in somatic cells. Having the protamines instead of canonical histones as the main components of chromatin can provide the sperm with at least two developmental advantages. Firstly, upon fertilisation, protamines from the sperm are removed and subsequently, maternally-derived histones are assembled onto the paternal DNA to allow its chromatinisation (McLay & Clarke 2003, van der Heijden *et al.* 2005). The chromatin of a somatic cell is packed with histones instead of protamines, and therefore there may be no need to replace the somatic cell histones with the oocyte-stored histones. As mentioned before, somatic cells are differentiated and during the course of specialisation they acquire many epigenetic mechanisms stabilising the expression states of their own differentiation-specific genes, for example in the form of post-translational histone marks. If such epigenetic marks are not correctly erased at fertilisation (as could be the case if the oocyte would not remove the somatic cell-specific histones), the epigenetic memory of a donor cell can be carried over to the embryos, as demonstrated in the examples above (chapter 1.2). In support of this hypothesis it was shown that embryos generated by the injection of round spermatids, which do not have protamines on their DNA, had elevated levels of DNA methylation and of H3K9me3, as compared to embryos generated by sperm injection (Kishigami *et al.* 2006).

Secondly, protamines on sperm DNA ensure the tight packaging of its chromatin. This can protect the DNA from any physical damage. Interestingly, in rabbits no offspring could be derived after the round spermatid injection and the developmental arrest of the reconstructed embryos was correlated with their abnormal ploidy (Ogonuki *et al.* 2005). The

nuclear transfer procedure involves numerous micromanipulations to the somatic cell nucleus. Since somatic cells do not have protamines, such procedures could lead to DNA damage. Indeed, it was suggested that a major cause of developmental failure of nuclear transfer embryos was DNA loss, which could be a result of DNA damage induced during the nuclear transfer procedure (Mizutani *et al.* 2012).

On the other hand, experiments in mouse in which spermatids at different stages of spermiogenesis were used as the paternal DNA donors demonstrated that the highest difference in the ability of spermatids to support embryonic development was acquired during spermatid maturation from step 7-8 to step 9-10. Interestingly, at this stages protamines are not yet present in the spermatid nuclei, which argues against necessary roles of protamines in supporting embryonic development (Ohta *et al.* 2009).

1.4.2. Sperm-derived transcriptional regulators and RNAs

Interestingly, even though mature sperm is transcriptionally silent and devoid of basal components of transcriptional machinery, some transcription factors, like Oct-1, Ets-1, C/EBP and TBP are retained in the chromatin of the mature sperm, and are associated with the hypersensitive regions of the sperm chromatin (Fig. 4) (Pittoggi *et al.* 2001, Zheng *et al.* 2008). Proteomic analysis of the mature sperm led to identification of several proteins involved in transcriptional regulation, for example, Bromodomain-containing protein 7 (Brd7) or Polycomb protein Suz12 (de Mateo *et al.* 2011). If such factors were delivered to the oocytes at fertilisation, they could be involved in the regulation of embryonic gene expression. On the other hand, a somatic cell has its own transcriptional regulators important for the maintenance of its differentiated state. When the nucleus of such a cell is transplanted

to the oocyte, some of these differentiation-specific factors characteristic for a somatic cell transcriptional program would be delivered to the cytoplasm of the oocyte. This could be detrimental for the embryonic development, since it could interfere with the proper regulation of the embryonic gene expression. Indeed, it was shown that if fertilised embryos were injected with a cytoplasm derived from somatic cells, their development was impaired and such embryos had a decreased expression of pluripotency gene Oct4 (Van Thuan *et al.* 2006), suggesting that some components of a somatic cell cytoplasm may be indeed toxic for development.

It was also hypothesised that sperm-derived RNAs may be needed for the future embryonic development. Despite the fact that mature sperm is transcriptionally silent and that it has almost no cytoplasm, it has been estimated that it contains about 10-100fg of RNA compared to 10-50pg of RNA typically found in a somatic cell (Pessot *et al.* 1989, Krawetz 2005), some of which were shown to be delivered to the oocyte at fertilisation (Ostermeier *et al.* 2004). Of the RNAs delivered by the sperm to the oocyte at fertilisation, microRNA miR-34c was shown to be required for the first embryonic cleavage division in mouse (Liu *et al.* 2012). This led to the hypothesis that sperm-derived RNAs may be important for embryo development and that lack of such RNAs could explain the low developmental potential of somatic cells as donors in the nuclear transfer procedures (Miller 2007). However, the results of recent genome-wide analyses of sperm RNAs argue against this hypothesis. First, the great majority of the RNA contained in the sperm nucleus turned out to be fragmented to a variable extent (Sendler *et al.* 2013). Second, the comparison of RNA-seq profiles obtained from the purified sperm, from the whole testis and from various human tissues, revealed that there are only 102 transcripts (out of 37974 transcripts tested) that were expressed exclusively in the sperm and testis (or detected in other tissues at very low levels). All these transcripts were present throughout spermatogenesis and retained until the final stages of sperm maturation,

so no sperm-specific transcripts were detected (Sendler *et al.* 2013). This is not surprising taken into account the fact that sperm is transcriptionally silent, and thus all transcripts present in mature sperm must be a result of transcription occurring at earlier stages of spermiogenesis. Therefore, the presence of transcripts specific for testicular cells could perhaps explain the developmental advantage of sperm as compared to somatic cells; however, it cannot explain the fact that the mature sperm is better than a spermatid in supporting embryonic development.

1.4.3. Sperm-derived epigenetic marks

As mentioned before, it has been demonstrated in several species that mature sperm may also deliver to the oocyte information about the regulation of embryonic gene expression in the form of epigenetic marks on its chromatin. Promoters of developmentally-important genes in the sperm were shown to bear various post-translational epigenetic marks, which, if delivered to the oocyte at fertilisation, could instruct the embryo about the regulation of the future gene expression (Hammoud *et al.* 2009). Furthermore, it was also shown that the pattern of DNA methylation of the zebrafish sperm is retained in the embryo at the time of zygotic gene activation (Potok *et al.* 2013). Such epigenetic marking could therefore provide the sperm with a developmental advantage over a somatic cell nucleus. As discussed before, the chromatin of a somatic cell acquires various epigenetic marks in order to stabilise its gene expression. These multiple layers of epigenetic marks could be refractory to reprogramming after nuclear transfer and would result in the retention of the epigenetic memory of the previous state and interfere with embryonic development (Pasque *et al.* 2011b, Chan *et al.* 2012). Furthermore, in support of the hypothesis for the importance of sperm-derived epigenetic marks for successful embryonic development, it was shown that abnormal histone

retention profiles or incorrect DNA methylation patterns in sperm correlated with idiopathic infertility cases in humans (Hammoud *et al.* 2010, Hammoud *et al.* 2011). However, due to the fact that the highly condensed sperm nucleus is not an easy material for experimental manipulations, so far there were no experiments proving this hypothesis.

1.5. Evidence for the sperm developmental advantages

After fertilisation, the embryo has to successfully accomplish several major developmental processes. First, the DNA from both parents needs to be faithfully replicated. Second, the embryo has to undergo successful zygotic genome activation (ZGA) and initiate the synthesis of both zygotic rRNA and mRNA. All these events need to be tightly regulated, as otherwise they may create a barrier for successful embryonic development (Newport & Kirschner 1982). Interestingly, it was suggested that the sperm may be advantageous as compared with a somatic cell in undergoing these processes (see below).

1.5.1. DNA replication

It was shown that a poor survival of nuclear transfer-derived embryos is correlated with chromosome loss during early embryonic division stages (Mizutani *et al.* 2012). It has been suggested that chromosome loss during cell division may result from incomplete DNA replication at the time of division (Laskey 2005). The use of egg extracts of *Xenopus laevis*, which are able to replicate DNA *in vitro*, demonstrated that the sperm, as opposed to a differentiated somatic cell (erythrocyte), can replicate DNA efficiently. DNA replication initiates from origins of replication. It was shown that during sperm replication, the median inter-origin distances were around 23.4kbp, whereas the median inter-origin distances for

replicating erythrocytes were around 120.9kbp (Lemaitre *et al.* 2005). Since, at least in early *Xenopus* development, early cell cycle phases are very rapid, and the first cell cycle last for only around 120mins (at 18°C), too sparsely positioned origins of replication in somatic cells could prevent the replication from being timely completed before the onset of the first cell division. Incompletely replicated chromosomes could then be either stretched and broken at the division or pushed towards one of the two resulting blastomeres, resulting in DNA loss (Gurdon & Laskey 1970). Interestingly, it was shown that pre-treatment of mouse embryonic fibroblasts (MEF) nuclei with *Xenopus* egg extracts before the nuclear transfer procedure increased the efficiency of replication (making it similar to what is observed for sperm nuclei) (Ganier *et al.* 2011). This pre-treatment of MEFs with *Xenopus* egg extracts also increased the frequency of normal nuclear transfer embryo development in mouse (Ganier *et al.* 2011). Therefore, the ability of undergoing efficient DNA replication was suggested to be one of the developmental advantages of sperm. Conversely, inefficient replication was suggested to be the cause of frequent developmental failures of embryos derived from a somatic cell nuclear transfer (Laskey 2005).

1.5.2. rRNA synthesis

Another challenge that the developing embryo has to face is a timely activation of rRNA synthesis. rRNA is a necessary component of ribosomes, which mediate protein translation. Series of experiments in mouse demonstrated that sperm-derived embryos, are significantly better at timely activating zygotic rRNA transcription than nuclear transfer-derived embryos. First, reverse transcription, quantitative PCR (RT-qPCR) measurements of ribosomal RNA content revealed that nuclear transfer embryos (NT-embryos) have significantly lower levels of 18S, 28S and 5.8S rRNA than control sperm-derived embryos

(Suzuki *et al.* 2007). Second, inefficient rRNA activation, was also evidenced by a smaller amount of labelled RNA precursor incorporated into rDNA and by a significant delay in the appearance of markers of active nucleolar organising regions in NT-embryos, as compared to sperm-derived embryos (Bui *et al.* 2011). Finally, inefficient activation of rRNA synthesis was correlated with poor developmental outcomes of NT-embryos. First, NT-embryos developed worse than sperm-derived embryos, but also, NT-embryos which were more efficient at rRNA synthesis, developed better to adulthood than embryos less efficient at rRNA synthesis (Zheng *et al.* 2012). Interestingly, the ability to activate rRNA synthesis efficiently in NT-embryos was correlated with the levels of rDNA activity in the donor cells: the more active the donor cells at synthesising rRNA, the more efficient the resulting NT-embryos at activating rRNA synthesis and the better the embryonic development. Furthermore, this phenomenon was also related to the DNA methylation levels of the rDNA loci in the donor cells, which were shown to be the most methylated in cells which were the less active at rRNA synthesis, and the less methylated in cells which were the most active in the rRNA transcription (Zheng *et al.* 2012). This is an important finding, as it again points towards the epigenetic memory of a previous state being a barrier for a successful reprogramming, and suggests that the inability of NT-embryos to correctly activate rRNA transcription may be the cause of their developmental failures.

1.5.3. mRNA synthesis

Finally, embryos need to initiate the synthesis of zygotic mRNAs in order to succeed in development (Newport & Kirschner 1982). There are reports indicating that NT-derived embryos do not properly initiate embryonic gene expression, as compared to sperm derived-embryos. It was shown in mouse that NT-embryos often fail to induce correct expression of a

key pluripotency factor Oct4 and even if they succeed to induce its expression, they often do so abnormally and express it in ectopic tissues, in which Oct4 is not expressed in control sperm-derived embryos (Boiani *et al.* 2002). Again, correct activation of Oct4 correlated with successful development of NT-embryos (Bortvin *et al.* 2003). Microarray studies revealed that abnormal expression levels of 1633 genes was already detected at 2-cell stage in NT-embryos in mouse (as compared with fertilised embryos). Furthermore, the largest group within these early abnormally expressed transcripts were transcription factors (Vassena *et al.* 2007). Another microarray study in mouse in which 87 single blastocyst stage embryos derived from nuclear transfers with different donor cells were compared with control fertilised embryos revealed that gene expression profiles differed between nuclear transfer embryos and fertilised embryos and also that nuclear transfer embryos coming from the same donor cell type, were more similar to each other, than to the embryos coming from nuclear transfer of different donor cell types (Fukuda *et al.* 2010). The same conclusion has been drawn in an independent study, also using mouse nuclear transfer-derived embryos from different types of donor cells (Hirasawa *et al.* 2013). In the latter paper it was also showed that the donor-cell dependent aberrant gene expression was more pronounced in embryonic than in extraembryonic tissues (Hirasawa *et al.* 2013). All these results point towards the same conclusion, that the NT-derived embryos experience problems with initiation of embryonic gene expression and that the abnormalities in gene expression profiles are often correlated with the origin of the donor cell, suggesting persistent epigenetic memory in NT embryos.

Concluding, the sperm, as opposed to a somatic cell, may be programmed for several aspects of embryo development: DNA replication, initiation of rRNA synthesis and zygotic mRNA transcription.

1.6. Work described in this thesis

The experiments described in this thesis aim to answer the question of what is the nature of developmental programming of sperm. I started from establishing a model system in which the developmental potential of sperm can be compared to the developmental potential of another cell type. I concluded that a suitable cell to compare with sperm is a spermatid. A spermatid is an immediate precursor of sperm, has already completed meiotic division, and differs from sperm only in the degree of specialisation. Furthermore, a spermatid can be compared to sperm in the same assay. I then showed that sperm is better at supporting embryonic development than spermatids, suggesting that sperm is programmed to support proper embryonic development. Subsequently, I aimed to identify the molecular basis for the sperm programming. I started from the identification of proteins present in sperm and spermatids by mass spectrometry analysis, reasoning that sperm-specific factors could be responsible for the developmental advantage of sperm. I also used the mass spectrometry-based approach to identify egg proteins which bind specifically to sperm or spermatid chromatin, since sperm programming could be also reflected in the ability to interact with the egg factors. Mass spectrometry analysis allowed me to identify several interesting candidate proteins, which I then functionally tested for their roles in the programming of the paternal nucleus. Unfortunately, none of these factors turned out to be beneficial for embryonic development. I have then changed my strategy for investigating the nature of sperm programming by characterising the developmental defects of spermatid-derived embryos. I showed that these defects are unlikely to be explained by inefficient DNA replication or by their inability to initiate zygotic rRNA transcription. Interestingly, RNA-seq analyses allowed me to identify 100 developmentally-important mRNAs which were misexpressed in spermatid-derived embryos as compared with sperm-derived embryos. Interestingly, the majority of them (82/100) were upregulated in spermatid-derived embryos. Further analysis

revealed that orthologues of these genes are modified by a repressive H3K27me3 mark in human sperm. These results encouraged me to test the hypothesis that perhaps the sperm is epigenetically programmed to regulate embryonic gene expression after fertilisation. Indeed, the enzymatic removal of H3K27me3 marks from the parental chromatin in embryos at the time of fertilisation resulted in gene misexpression. This led to the hypothesis that perhaps spermatids lack the repressive H3K27me3, as compared with sperm, which causes upregulation of genes in spermatid-derived embryos. Surprisingly, ChIP-seq analyses on the chromatin isolated from sperm and spermatids revealed that spermatids, similarly to sperm, already have repressive H3K27me3 marks at misregulated genes. Therefore, H3K27me3 itself could not explain the difference in gene expression between sperm- and spermatid-derived embryos. Interestingly, further ChIP-seq analyses for activating histone marks revealed that spermatids have more of H3K4me2/3 marks than sperm. Interestingly, the presence of these activating marks at misregulated genes in spermatids correlated with their upregulation in spermatid-derived embryos. All these results support the hypothesis that the sperm, as opposed to the spermatid, is epigenetically programmed to regulate embryonic gene expression after fertilisation.

The results presented in this thesis are divided into 5 separate chapters (Chapters 3-7). Each of these chapters has a brief introduction that explains the rationale for the experiments conducted, followed by a detailed report of the results and a subsequent short summary/discussion. Since all the chapters have their own short summary/discussion sections, when writing the final discussion of the thesis (Chapter 8) I took the advantage of the writer's freedom and I do not discuss the results one by one in the order of appearance in the thesis. Instead, I took the pleasure of talking about the most exciting results; therefore I sometimes shuffle the results and discuss them according to the common connecting theme. In the discussion section I also propose new experiments, which are based on the results

described, that are either currently conducted in the laboratory or are in the future plans. Finally, I put the obtained results in the broader context and I explain how they advance the current knowledge of the subject.

In the appendices section of the thesis I attach publications that arose as a result of this work. There are two review articles, one book chapter and a research article by a colleague from my laboratory, Dr Kei Miyamoto, with whom I collaborated on some parts of his project (which is not connected to the work described in this thesis and therefore is not separately described here). The manuscript describing the work presented in this thesis is currently in preparation.

Supplementary tables S1-S6 are Excel files with long protein/gene lists, and they are therefore provided electronically as separate attachments to this thesis.

Chapter 2
Experimental procedures

All the experiments involving the use of animals comply with the Home Office regulations as set out in the Animals (Scientific Procedures) act 1986 (Establishment Code 8002802).

All the animal care (tadpole and frog husbandry, hormonal injections (PMSG/hCG), sacrifice and dissections) was done by David Simpson at the Frog Facility at the Gurdon Institute.

All the bioinformatic analyses described are written and kindly shared by Dr Angela Simeone, Dr George Allen and Dr Charles Bradshaw.

Experimental procedures are described in order of their appearance in the 'Results' section of the thesis (Chapters 3-7), apart from all the bioinformatic analyses, which are described altogether as the last section.

2.1. Separation of sperm and spermatids

Testes from 6 adult *Xenopus laevis* males were isolated and manually cleaned from blood vessels and fat in 1XMMR (100mM NaCl, 2mM KCl, 2mM $CaCl_2$, 1mM $MgCl_2$, 5mM HEPES pH 7.4) using forceps and paper tissues. It is crucial to clean the testes well from any non-testicular tissues, as otherwise the cells released from the tissues may negatively affect the final purity of the isolated cells. Subsequently, testes were torn into small pieces with

forceps and homogenised with 2-3 strokes of a Dounce homogenizer (tissue from 1 testis at a time). The cell suspension was then filtered (50um pore size, CellTrics, cat. 04-0042-2317) to remove tissue debris and cell clumps and spun down at 800rcf, 4°C, 20 minutes. Supernatant was discarded and the cell pellet was resuspended in 12ml of 1XMMR. If any red blood cells were visible at the bottom of the pellet (a result of incomplete removal of blood vessels), only the uncontaminated part of the pellet was recovered, taking extreme care not to disturb the red blood cells. Subsequently, six (1 per testes from each frog) step gradients of iodixanol (Optiprep – Sigma, D1556, is 60% iodixanol in water) in 1XMMR were manually prepared in pre-chilled 14ml ultra-clear centrifuge tubes (Beckman Coulter, 344060) in the following order from the bottom to the top of the tube: 4ml of 30% iodixanol, 1ml of 20% iodixanol, 5ml of 12% iodixanol and 2ml of cell suspension in 1XMMR. Gradients were spun down in pre-chilled SW40Ti rotor at 7500rpm (10000g), 4°C, 15 minutes, deceleration without break (Beckman Coulter Ultra-centrifuge, Optima L-100XP). The top interface fraction (between 1XMMR and 12% iodixanol), containing spermatids, and the pelleted fraction, containing mature sperm, were collected. Collected fractions were diluted six times with 1XMMR and collected by spinning first at 805 rcf, 4°C, 20 minutes and re-spinning at 3220 rcf, 4°C, 20 minutes to pellet the remaining cells. Pelleted cells were subjected to nuclei preparation (see below).

2.2. FACS analysis

All the cell cycle analyses were performed by Dr Rachel Walker at the Flow Cytometry facility of Cambridge Stem Cell Institute, Cambridge, UK (samples were provided to Dr Walker after fixation and labelling).

Cells were fixed in cold 70% EtOH for 30mins at 4°C, washed twice in 1XMMR, spun at 850g and resuspended in 50ul of 1XMMR supplemented with 100ug/ml of DNase-free RNase. DNA was stained by the addition of 200ul of 50ug/ml propidium iodide solution.

2.3. Sperm and spermatid nuclei preparation, intra-cytoplasmic sperm injections (ICSI) to non-enucleated and to enucleated eggs, and embryo culture

Sperm and spermatids nuclei have been permeabilised with digitonin (Smith *et al.* 2006) and stored at -80°C. Injections were performed using Drummond Nanoject microinjector (NanojectII Auto Nanolitre Injector, Biohit, 3-00-206A) and glass caplillaries (Biohit, 3-00-203-G/XL) pulled using a Flaming-Brown micropipette puller (settings: heat 700, pull 100, velocity 100, time 10). Cell suspension in sperm dilution buffer (SDB) (Smith *et al.* 2006) was sucked into the injection needle filled with mineral oil. Cell concentration was adjusted by doing mock injections on a microscope slide to deliver 1 cell per 4.6nl injection. The eggs were placed in batches of 20-25 on a blotting paper. If they were to be enucleated, they were placed with animal pole facing upwards, whereas if they were not subjected to enucleation, they were placed on a side (with the marginal zone upwards). For enucleation, eggs were treated for 30s with a UV mineralite lamp (Gurdon 1960) (this step was omitted for non-enucleated eggs). The jelly of the egg was removed by a minimal length of the Hanovia lamp treatment that allows for the needle penetration. The eggs were immediately injected with sperm or spermatid solution and moved to 1XMBS (Gurdon 1976) supplemented with 0.2% bovine serum albumin (BSA). Cell suspension in the needle was replaced every 20-25 eggs injected. At 4-cell stage embryos were sorted (all the non-cleaved embryos or those with irregular cleavage furrows were discarded) and the culture media was

replaced with 0.1MBS, 0.2% BSA. Embryos were cultured in 0.1MBS, 0.2% BSA (changed daily) at 16-18°C incubator. Assessment of developmental stages was performed according to Nieuwkoop and Faber (Nieuwkoop & Faber 1994).

2.4. Nuclear transfer

Nuclear transfers were performed as described (Elsdale *et al.* 1960, Gurdon 1960), with a modification: donor cells were dissociated in 2mM EDTA. Donor cells were derived from animal caps of early gastrula (stage 10.5) embryos. Eggs were enucleated by a 30s treatment of Mineralite UV lamp and dejellied by a treatment with a Hanovia UV lamp (minimal length required to allow for the needle penetration).

2.5. Protein preparation for the mass spectrometry analysis

All the methods were obtained from the Cambridge Centre for Proteomics (CCP), at the Biochemistry Department, University of Cambridge. I have performed all the steps of protein isolation, quantification and labelling under the supervision of Renata Feret from the CCP. All the subsequent steps of the analysis were carried by Renata Feret and colleagues at the CCP.

Proteins from pelleted sperm and spermatids were extracted in Urea/Thiourea buffer (4% Chaps, 2M Thiourea, 6M Urea; Sigma: C9426, T7875, respectively and Fisher Chemicals, U/0500/53) supplemented with protease inhibitors (Roche, 11873580001) by

keeping on ice for 1h and vortexing from time to time until the pellet disappeared. Subsequently, the solution was sonicated with a probe sonicator (3s, Amplitude 40, sonication over an ice bath) in four separate rounds. The protein concentration was quantified with a Quick Start Bradford Protein Assay (Bio-Rad, 500-0202), following manufacturer's recommendations. Fifty micrograms of protein lysate was subsequently labelled with Cy3 or Cy5 dye (Cy dyes label the lysine residues of proteins) by adding 1ul of 0.2mM dye. Labelling reaction was carried at 4ºC for 30mins in the dark and quenched by the addition of 1ul of 1mM lysine. Quenching was performed for 10mins at 4ºC in the dark. After the labelling was completed, labelled protein lysates were mixed. Reciprocal labelling was performed to rule out any abnormalities or biases in labelling (none were noticed). Further processing of the samples (1^{st} and 2^{nd} dimension electrophoresis, scanning, staining, spot excision, protein digestion and the mass spectrometry analysis) were performed by the staff at the CCP according to the methods used at the CCP.

2.6. Interphase egg extract preparation

Eggs were collected in 1XMMR, dejellied with 0.2XMBS, 2% cysteine (pH 7.8-7.9) (Sigma, W326305) and washed with 0.2XMMR. Subsequently, eggs were activated for 3 minutes at RT with 0.2XMMR supplemented with 0.2ug/ml calcium ionophore (Sigma, C7522). Eggs were rinsed with 0.2XMMR and subsequently all abnormal or not activated eggs were removed. Eggs were washed with 50ml of ice-cold extraction buffer (EB) (5mM KCl, 0.5mM $MgCl_2$, 0.2mM DTT, 5mM Hepes pH 7.5) supplemented with protease inhibtiors (PI) (Roche, 11873580001), transferred into centrifugation tube (Thinwall, Ultra-Clear™, 5mL, 13 x 51 mm tubes, Beckman, 344057) and supplemented with 1ml of EB buffer with PI and 100ug/ml of cytochalasin B (Sigma, C2743) and placed on ice for 10

minutes. Subsequently, eggs were spun briefly at 350g for 1 minute at 4°C (SW55Ti rotor, Beckman Coulter Ultra-centrifuge, Optima L-100XP) and excess buffer was discarded. Eggs were then spun at 18000g for 10 minutes at 1°C, the extract was collected with a needle, transferred to a fresh, pre-chilled tube, supplemented with PI and 10ug/ml of cytochalasin B and re-spun using the same conditions. Extract was collected with a needle and used: frozen (on liquid nitrogen, in 100ul aliquots) for the analysis of proteins bound to the chromatin followed by the proteomic analysis; or fresh, for replication assays (see below).

2.7. Egg extract treatment and protein isolation for mass spectrometry analysis of egg proteins incorporated into sperm or spermatid chromatin.

Sperm and spermatids were collected and permeabilised as described above. Egg extracts were prepared as described above. Twenty millions of permeabilised sperm or spermatids were treated with 3650ul of egg extract, supplemented with the 1X final energy regeneration mix (20X energy regeneration mix is prepared and stored in aliquots at -80ºC: 2mg/ml Creatine Kinase (Roche, 10127566001), 150mM Creatine Phosphate (Roche, 10621714001), 20mM ATP (Roche, 10519979001), 2mM EGTA, 20mM MgCl2). Control permeabilised cells were treated with EB buffer and the energy regeneration mix alone. Cells were incubated in the extract/buffer for 1h at room temperature with frequent tapping. After that, cells were washed with 15 volumes of Buffer D (10mM Hepes pH 7.7, 10mM KCl, 1.5mM $MgCl_2$, 1mM DTT) (spin at 3220 rcf, 4°C, 20mins). Pellets of black colour were observed in the egg extract-treated samples. Subsequently, pellets were resuspended in 15ml of Buffer E (250mM Sucrose, 10mM Tris-HCl pH 8.0, 5mM $MgCl_2$, 1mM DTT, 0.1% Triton X-100) and spun at 3220 rcf, 4°C, 20mins. Subsequently, pellets were resuspended in 1ml of Buffer E each and transferred to a 1.5ml tubes and washed 6 more times with Buffer E (each

time spun at 1000rcf, 5mins, 4°C). The final pellet contained proteins bound to chromatin, which were isolated and processed for 2-DIGE and proteomics analysis as described in chapter 2.5.

2.8. Immunoblotting analysis.

All the immunoblotting analyses were performed according to standard immunoblotting protocols (Green & Sambrook 2012). For the validation of mass spectrometry results, equal amount of proteins isolated in Urea/Thiourea buffer were loaded on each blot (see chapter 2.5). For the immunoblotting on oocytes, embryos or GVs, equivalents of 1 oocyte, cytoplasm or embryo or an indicated number of GVs were loaded on each lane of a gel. Isolations of cytoplasmic proteins from oocytes/embryos were performed using a mild lysis buffer (10mM Tris pH 8.0, 150mM NaCl, 1%NP-40, 1mM EDTA and protease inhibitors): oocytes/embryos were disrupted in the lysis buffer, kept on ice for 10mins, vortexed from time to time, spun at 16100 rcf, 4ºC, 1min and the supernatant was used for immunoblotting. Isolations of nuclear proteins (from somatic cells, sperm, spermatids or from embryos) were performed using Emilie's Buffer (500mM Tris pH 6.8, 500mM NaCl, 1% NP40, 0.1% SDS, 1% b-Mercapthoethanol and protease inhibitors). If nuclear proteins were isolated from embryos, then first the nuclei were isolated (to remove contaminating yolk, see below) and only then the Emilie's buffer was used to extract the proteins from the pelleted nuclei. Nuclei were isolated from embryos with the following method: up to 5 embryos were homogenised by pipetting up and down in 50ul of buffer E1 (50mM Hepes-KOH pH 7.5, 140mM NaCl, 1mM EDTA pH 8.0, 10% glycerol, 0.5% Igepal CA-630, 0.25% Triton X-100, 1mM DTT and protease inhibitors) and spun at 1100rcf, 2mins, 4ºC (all the following centrifugations were performed in the same conditions). The

supernatant and the lipids attached to the walls of the tube were discarded and the pellet was resuspended in 2ml of buffer E1. The spin was repeated and the pellet was again resuspended in 2ml of buffer E1 and incubated on ice for 10mins. After another spin, the pellet was resuspended in 2ml of buffer E2 (10mM Tris pH 8.0, 200mM NaCl, 1mM EDTA pH 8.0, 0.5mM EGTA pH 8.0 and protease inhibitors) and centrifuged. This procedure was repeated twice (the last wash in E2 was preceded by 10mins incubation in E2 on ice). Finally, the pellet (containing the isolated nuclei) was resuspended in the desired volume of Emilie's buffer).

Polyacrylamide gels were cast at the percentage appropriate for the separation of the desired protein size (Green & Sambrook 2012). Gel electrophoresis, transfer of proteins, immunoblotting and washes were performed according to the standard protocols (Green & Sambrook 2012). PVDF membranes with 0.45uM pore size were used (Immobilon, Millipore, IPFL00010) with a semi-dry transfer system (Trans-Blot® SD Semi-Dry Transfer Cell, BioRad, 170-3940) transferring for 30mins, at room temperature (25V).

All the protein detections were performed using immunofluorescence with the use of the LI-COR Imaging System. Primary antibodies were used at 1:1000 dilutions, unless stated otherwise and blots were incubated with the primary antibodies overnight at 4°C. Primary antibodies against the following proteins were used in this thesis: Hdac1 (rabbit polyclonal; Abcam, ab33278), Hdac2 (rabbit polyclonal; Epitomics S2398), Mbd3 (1:250; mouse monoclonal; Abcam, ab45027), Hp1γ (goat polyclonal; Abcam, ab40827), HA-tag (1:2500, mouse monoclonal; Sigma, H9658), Rbbp4 (rabbit polyclonal; Abcam, ab1765), Rbbp7 (rabbit polyclonal; Abcam, ab3535), Lsf (rabbit polyclonal; Abcam, ab80445), actin (1:2500, rabbit polyclonal; Sigma, A2103), beta-actin (1:2500; mouse monoclonal; Abcam, ab6276), H3K27me3 (rabbit monoclonal; Cell Signalling Technology, 9733), H2A (rabbit polyclonal; Millipore, 07-146), H2B (1:2000, rabbit polyclonal; Abcam, ab1790), H3 (rabbit polyclonal;

Abcam, ab18521) and H4 (1:500, rabbit polyclonal; Abcam, ab10158). To detect the primary antibodies, the following secondary antibodies were used (all at 1:25000 dilution): goat anti-mouse IRDye 800 (Licor, 926-32210), Goat anti-Rabbit IRDye 800 (Licor, 926-32211), goat anti-rabbit Alexa 680 (Invitrogen, A-21109), goat anti-mouse Alexa 680 (Invitrogen, A-21057), donkey anti-goat Alexa 680 (Invitrogen, A-21084). Blots were incubated with the secondary antibody solution for 1h at RT. To reveal the proteins, blots were scanned using a LI-COR detection system (Odyssey laser scanner, LI-COR Biosciences).

2.9. Molecular cloning of candidate sperm factors and mRNA synthesis

cDNA from testis isolated from *Xenopus laevis* was generated with the use of oligodT(15) primer. NCBI-deposited sequences were then used to design primers to amplify the sequences of candidate sperm factors: Sp1 (NM_001137586.1), Sp4 (NM_001087761.1), Sp5 (S71764.1), H1fx (BC041758.1), Mlf1 (NM_001095375.1). Primers used to amplify the candidate factors cDNA sequences and to allow their cloning into the pEntry vector are listed in Table 2. Clonings were performed using pENTR™/D-TOPO® Cloning Kit (Life Technologies, K2400) and Gateway® LR Clonase® II enzyme mix (Life Technologies, 11791-020) according to the manufacturer's instructions. Directional cloning was performed into the C-terminus (destination vector used was pCS2+ with a C-terminal HA-tag). Clones were first checked by a directional colony PCR screen (using M13 forward primer: GTAAAACGACGGCCAGT and the insert-specific reverse primer) and second, by restriction enzyme digestion and Sanger sequencing (Sanger sequencing reactions were performed at the sequencing facility at the Department of Biochemistry, University of Cambridge). mRNA was synthesized in vitro using MEGAscript® SP6 Kit (Ambion, AM1330M) following the manufacturer's instructions. mRNA concentration was measured

with a Nanodrop, adjusted to a final concentration of 1mg/ml with DEPC H2O, aliquoted (aliquots of 2ul) and frozen at -80ºC. Clones of mouse K6B (amino acids 1025-1642) and its catalytically inactive mutant (K6B-mut) were in pCS2+ destination vector with a C-terminal HA-tag and NLS-tag and were a kind gift from Dr Jerome Jullien.

Table 2. Sequences of primers used for the cloning of candidate sperm factors

Name of primer	Sequence
XL_SP1_Entry_F	CACCATGGCACTGCCCTCCGAGACC
XL_SP1_Entry_R	CACTATCATGGTTCTGGGAACCCTGCGCTTG
XL_SP4_Entry_F	CACCATGAGCAAAGTGAGTGGCGGG
XL_SP4_Entry_R	ACTGCGATAATCTGAGCCATAGTCTCTTGCC
XL_SP5_Entry_F	CACCATGAGCAAAATGAGAGGCGGG
XL_SP5_Entry_R	ACTGCGATATTCTGACCCATAGGC
XL_H1fx_Entry_F	CACCATGGCTCTAGAGCTGGAAGAGAATTTACACAGC
XL_H1fx_Entry_R	CGCTTTCTTGGATTTAGGCGCTTTGCGGACGC
XL_MLF1_Entry_F	CACCATGTTCCGCAGTTTGCTGAGAGACTTTGACG
XL_MLF1_Entry_R	TTTCTCCCTTGCCGGCAACTGCAGCTG

2.10. Cell culture, transfection, immunostaining and microscopic analysis

Mouse C2C12 cells were cultured in DMEM (GIBCO, 41965-062) supplemented with 10% FBS and 100u/ml of penicillin/streptomycin at 37ºC, 5% CO_2. *Xenopus laevis* XL 177 cells were cultured in the same medium, but diluted to 60% with a ddH_2O at 23ºC. Cells were split whenever the confluence was reached. Transfections were performed with lipofectamine 2000 (Invitrogen, 11668027), following the manufacturer's protocol. For the immunostaining procedure, cells were fixed with 4% PFA in1XPBS, permeabilised with 0.1% Triton X-100 in 1XPBS and blocked with 5% BSA in 1XPBS. Antibody against the HA-tag (mouse monoclonal; Sigma, H9658) was diluted 1:500 and detected with a secondary antibody conjugated with Alexa 488, diluted 1:500 (donkey anti-mouse, Invitrogen, A21202). Samples were mounted on a microscope slide in Vectashield mounting medium with DAPI

(Vector Labs, H-1200) and sealed with nail polish. Microscopic analyses were performed using Zeiss 510 META confocal LSM microscope.

2.11. Cell squashing for ploidy assessment

Early tadpoles (stage 30-36) were anaesthetised and decapitated. Subsequently, tadpoles were transferred on a microscope slide and the yolky tissues were removed. The remaining tissues of the tadpole were then squashed with the coverslip and observed under a phase contrast microscope.

2.12. Injection of 1-cell embryos with mRNA

Eggs were *in vitro* fertilised and dejellied using 2% cysteine solution in 0.1XMMR. Injections into 1-cell stage embryos were performed in injection solution (Smith *et al.* 2006) using a Drummond Nanoject microinjector, delivering 9.2ng of mRNA per injection (mRNA at 1mg/ml in DEPC H_2O). Embryos were cultured at 18°C and collected for qRT-PCR analysis (chapter 2.14) at the stage indicated in the text (see Chapter 3.7). Expression of K6B and K6B-mut proteins in embryos, as well as the removal of H3K27me3 by K6B were confirmed by immunoblotting.

2.13. Injection of mRNA or oligonucleotides into the oocytes

Oligonucleotides and qPCR primers for HP1γ were designed by Matthew Jones, a 4-year PhD rotation student during his rotation in the Gurdon laboratory.

Ovaries were isolated from frogs pre-primed with PMSG at least 48h before use. Oocyte defolliculation was performed enzymatically with a liberase treatment (Halley-Stott *et al.* 2010). For *in vitro* maturation experiments, oocytes were liberated for a maximum time of 1h 15mins and the remaining defolliculation was done manually (Miyamoto *et al.* 2013). Oocytes were cultured in 1XMBS medium (88mM NaCl, 1mM KCl, 2.4mM $NaHCO_3$, 0.82mM $MgSO_4$, 0.41mM $CaCl_2$, 0.33mM $Ca(NO_3)_2$, 10mM HEPES pH7.4, 10ug/ml streptomycin sulfate and 10μg/ml penicillin). mRNA was injected as described above, with the difference that the use of injection solution was omitted (oocytes were injected directly in the culture medium). Oligonucleotide design was performed according to the guidelines published (Hulstrand *et al.* 2010). Sequences of the oligonucleotides used are listed in Table 3. The oligonucleotide injection and *in vitro* maturation and ICSI (for the K6B/K6B-mut) experiments were conducted as described before (Miyamoto *et al.* 2013). Sequences of primers used for the assessment of the knockdown are listed in Table 4.

Table 3. Sequences of oligonucleotides used for the knockdown of candidate egg factors

Name of oligonucleotide	Sequence
HP1g #1	C*T*T*CCTCCACCTTCT*T*G*C
HP1g #2	A*C*C*TTTCCATTTACTAC*A*C*G
HP1g #3	T*T*C*AATCAACTCTGGACA*G*T*C
Hdac1 #1	C*C*A*ACATCACCATCA*T*A*G
Hdac1 #2	C*C*A*TAGTTGAGCAGC*A*G*G
Hdac1 #3	T*G*T*CTGGTCGTATGG*A*G*C
Hdac2 #1	G*A*C*CTTCTTCTTGGC*A*C*C
Hdac2 #2	C*A*C*CATCATAATAGT*A*G*C
Hdac2 #3	C*G*G*ATTCTGTGAGGC*T*T*C
Mbd2/3 #1	C*T*G*AGGCTTACTGCG*G*A*A
Mbd2/3 #2	T*C*C*TAAGTAACGAGC*A*A*G
Mbd2/3 #3	C*C*G*CATTCGCTGTCT*G*T*T
LSF #1	T*G*C*CAGCGGTAACTT*C*A*G
LSF #2	T*C*C*ACCTCCATCCTT*C*T*A
LSF #3	T*A*T*TCTGTCTCCAGG*T*C*T
Rbbp4 #1	T*A*T*GACTCGTTCCTC*C*A*C
Rbbp4 #2	G*A*G*CATCATCATTAG*G*A*A
Rbbp4 #3	A*G*G*TTGCCACTTAGG*T*T*G
Rbbp7 #1	C*T*T*CCACTGTATCCT*C*A*A
Rbbp7 #2	C*A*T*AACCAGGTCATA*C*A*G
Rbbp7 #3	G*G*A*ACCTGGACACGA*G*C*C

* denotes a phosphorothioate bond

Table 4. Sequences of primers used for qRT-PCR assays

Name of the primer	Sequence
Hdac1_F1	CGCTCCATACGACCAGACA
Hdac1_R1	GCCATCAAACACAGGACAGT
Hdac2_F1	TCTGTAGCTGGTGCTGTAAAACTCA
Hdac2_R1	CCTGCCCAGTTAACAGCCATA
HP1g_F1	GGGAGCCTGAGGAAAACTTAG
HP1g_R1	CAAATCCCCGTGGTTTATCA
Rbbp4_F1	GGAGAGTTTGGAGGCTTTGG
Rbbp4_R1	TGGAGTTTTGGTGGCAATAA
Rbbp7_F2	GTGTCCAGGTTCCCAATGAT
Rbbp7_R2	CAAAGCCACCAAACTCTCCT
PWP1_F	GACTTCGAAAATCTGGCATCTCA
PWP1_R	GGGACTTTCACCATTGACTTAAACA
GATA3_F2	CACAGGATCTCCATTGGCATT
GATA3_R2	CCTGTGCAAACTGTCAAACCA
SFRP2_F1	GGAATAAGAAGAGACAGGCCCAAT

SFRP2_R1	TTACCAAAGCCACCCCAGAA
C19ORF26_tr_F2	GCCATCAACCCCTTCTTCATC
C19ORF26_tr_R2	ACACGTTACCACAGCACTTTGC
PLOD2_F2	CACATTCTTTATTCTGCCGACAGAT
PLOD2_R2	AAGAGCCAATCACGCAAGCT
NA_F1	GATGCTCAGCTTTGGATCTTGA
NA_R1	CCACACGGGCCTGATCTG
MN1_F2	TGCCTTCAGCTAGGGACACA
MN1_R2	CACCCAGTCGTGATAAAGCAGTAG
XL_HES1_F1	TGAGCAATACCCCGGATAAG
XL_HES1_R1	TCCAGGATGAGGGTTTTGAG
WAVE1_F2	AGGAATCCAGCTTCGCAAAG
WAVE1_R2	ACGCGCTCGTGTTTTGCT
HOXB1_F	CCCCACAAAATTGCAACCA
HOXB1_R	TCTGCTTCTTGGCTGGCATA
HES7_F1	TCCTCTCCCTCCGCCTTTT
HES7_R1	CCATGGAAACCCATAGAAAGCT
DOLPP1_F	GGGCATTCGCTATGCTCTCT
DOLPP1_R	GCCTGAAATCCCTCAACCAA
ZNF33A_F2	GGTCTGTCTCATCCTGAATGCTT
ZNF33A_R2	AATAGGTGTGGATTCTGCTGTTGA
SOX21_F1	CCCACATTGGGTTCCAACTG
SOX21_R1	GGCATGACAGCCCGACTAAG
XL_GJB1_F1	GCATCAGCAAAGAGCATCAA
XL_GJB1_R1	CAGGGAGCCGTGAGAGTTAG
FOSL2_F1	TGTGTGATAAAGTAGACCAGAGGATTTT
FOSL2_R1	GTCGCTTGTTTCCTTTTCAACA
FOS_F2	AGTCCTGGATCGCCGAGTT
FOS_R2	TCACAGTAACCGCAACGATCTATT
CHD3_F2	GTTCCCACGCACGTTTGTT
CHD3_R2	TGGCTGCTGCATCCATAATG
MIX1b_F2	AGGATGGAGTGGAGGATCTGAA
MIX1b_R2	GGTTCTCCGGGAAGGTAAAGG
XL_18srRNA_F1	ATGGCCGTTCTTAGTTGGTG
XL_18srRNA_R1	TATTGCTCGATCTCGTGTGG
XL_28srRNA_F1	TCATCAGACCCCAGAAAAGG
XL_28srRNA_R1	GATTCGGCAGGTGAGTTGTT

2.14. RNA extraction and qRT-PCR analysis

Oocytes or embryos were collected and frozen at -80°C. RNA extractions were performed using Qiagen RNeasy Mini kit (Qiagen, 74106) according to the manufacturer's protocol, unless stated otherwise. RNA was eluted in 50ul DEPC H2O and was used for cDNA synthesis and for RT-qPCRs, as described before (Halley-Stott *et al.* 2010). For all the experiments described, apart from rRNA levels assessment, cDNA synthesis was performed using oligo dT(15) primer (0.5ul of 100uM primer/reaction). For rRNA levels measurement, gene-specific primers were used for cDNA synthesis. qPCRs were performed with gene specific primers (Table 4) using a SybrGreen detection system (Sigma, S9194) and ABI 7300 machine (Applied Biosystems) as detailed before (Halley-Stott *et al.* 2010). Results were exported to Microsoft Excel for analysis. Gene expression was normalized to *pwp1 or dolpp1* transcripts. Subsequently, a Grubb's test was used to identify and exclude any potential outliers in the datasets with a p-value cut-off < 0.05: 1 out of 15 samples for spermatid-derived embryos for rRNA expression analysis was identified as outlying, 1 out of 6 sperm/spermatid embryos was outlying for *mn1* and 1 out of 6 for *chd3* transcript for RNA-seq validation analysis. Statistical significance was assessed using a t-test.

2.15. Analysis of DNA replication

The protocol described here is a modification of the original protocol kindly provided by Dr Vincent Gaggioli.

Replication on single DNA fibres was performed as described before (Gaggioli *et al.* 2013) with slight modifications. Freshly prepared egg extracts were supplemented with energy regeneration mix (components as mentioned in chapter 2.7) and with 20uM biotin-16-dUTP (Roche, 11093070910). Permeabilised cells were added to a final concentration of 200 nuclei/ul of extract and incubated at RT for 30mins, 40mins or for 2h (tapping every 10 minutes). Reaction was stopped by adding 10 volumes of ice-cold 1XPBS (Phosphate Buffer Saline: 137mM NaCl, 2.7mM KCl, 10mM $Na_2HPO_4x2H_2O$, 2mM KH_2PO_4) and cells were spun down at 1000g, 4°C, 7 minutes. Cells were resuspended in 50ul of 1XPBS and mixed immediately with 50ul of melted (at 65°C) 2% low melting point agarose (Invitrogen, 16520050) in 1XPBS. After solidification, the agarose plug was incubated overnight (O/N) at 50°C with 1ml 0.5M EDTA pH8.0, 100uL 10% sarkosyl (Sigma, L5125), 1mg/mL Proteinase K (New England Biolabs, P8102S) followed by three washes in TE pH 6.5. Subsequently, the plug was incubated twice in TE supplemented with 0.1mM PMSF (Sigma, 93482) for 30 minutes at 50°C and washed four times with 1ml of 50mM MES (Sigma, 69889) pH 6.35, 1mM EDTA (1h at RT each wash). Then the solution was removed; the plug was melted in 400ul of MES pH 6.35, 1mM EDTA at 68°C for 20 minutes and the agarose was digested with 2 units of β-agarase (New England Biolabs, M0392S) O/N at 42°C.

Silanised coverslips were prepared as described before (Labit *et al.* 2008). Thirty microliters of replicated DNA solution was pipetted onto a silanised coverslip, covered with a

non-silanised coverslip and incubated for 5 minutes at RT. Subsequently, the top coverslip was slid away to stretch DNA fibres and the silanised coverslip with stretched fibres was fixed in 3:1 solution of Methanol:Glacial Acetid Acid for 10 minutes, RT. The fibres were then denatured with 2.5M HCl (1h, RT) and dehydrated by washes in 70% ethanol, 90% ethanol and 100% ethanol (1 minute for each wash). Subsequently, the coverslip was dried, washed 3 times in 1XPBS, 0.1% Tween (Sigma, P5927) (5 minutes for each wash) and blocked in 3%BSA in 1XPBS (1h, RT). All antibodies were diluted in 1XPBS, 3%BSA, 0.1% Tween. Total DNA and replicated DNA were simultaneously detected using the following primary antibodies: anti-DNA antibody (Millipore, MAB3034) 1:300 dilution and streptavidin-Alexa 594 antibody 1:50 to detect biotin (Invitrogen, S-11227) for 30 minutes at 37°C. Slides were washed with 1XPBS, 0.1% Tween (4 washes) and secondary antibodies (diluted 1:50) were added: chicken anti-mouse Alexa 488 (Invitrogen, A-21200) and biotinylated antibody anti-streptavidin (Vector Labs, BA-0500) (incubation for 30 minutes, 37°C). After four washes in 1XPBS, 0.1% Tween, a tertiary detection was performed with antibodies diluted 1:50: goat anti-chicken Alexa 488 (Invitrogen, A-11039) and streptavidin-Alexa 594 (for 30 minutes, 37°C). The coverslip was washed three times with 1XPBS 0.1% Tween, three times in 1XPBS, mounted on a microscope slide with a mounting medium (50% glycerol in 1XPBS), and sealed with nail polish. Images were acquired with a Zeiss 510 META confocal LSM microscope. Image analysis was performed in ImageJ; the amount of replicated DNA and total DNA was measured individually on single DNA fibres. The data was analysed in Microsoft Excel with a macro kindly provided by Dr Vincent Gaggioli.

2.16 Pulldown of BrUTP-labelled RNA from haploid sperm- and spermatid-derived embryos

Sperm- and spermatid-derived haploid embryos were generated as described above, but using SDB buffer containing 100mM BrUTP (Sigma, B7166). Single embryos were harvested at stage 10.5-11.5 and total RNA was isolated with Qiagen RNeasy Mini kit (Qiagen, 74106) according to the manufacturer's protocol (elution in 50ul DEPC H_2O). For each embryo 20ul of agarose-conjugated beads with antibody against BrdUTP (Santa Cruz Bioechnology, sc-32323) were washed twice in 1ml of 0.5XSSPE (wash buffer) (1XSSPE: 300mM NaCl, 20mM $NaH_2PO_4xH_2O$, 2mM EDTA), 0.05% Tween 20, 0.1% PVP (Sigma, P5288) by centrifugation for 2 minutes at 2000g, RT. Subsequently, beads were blocked O/N on a rotating wheel at 4°C in wash buffer supplemented with 1mg/ml of RNase-free BSA final (New England Biolabs, B9000S). Next, 20ul of beads were resuspended in 500ul of 0.5XSSPE supplemented with 0.05% Tween and RNase Inhibitor (New England Biolabs, M0314S), mixed with 30ul of purified RNA and incubated for 4 hours on a rotating wheel at 4°C. Subsequently, beads were washed for 10 minutes on a rotating wheel at 4°C with the following solutions: once with 0.2XSSPE, 0.05% Tween; twice with 0.5XSSPE, 0.05% Tween, 150mM NaCl and once with 1XTE (10mM Tris, 1mM EDTA, pH 8.0), 0.05% Tween (spins 2000g, RT for 2 minutes). RNA was eluted four times for 5 minutes at RT with 100ul of elution buffer (300mM NaCl, 5mM Tris, 1mM EDTA, 0.1%SDS, 2mM DTT); the eluates were pooled, extracted with phenol:chloroform, precipitated by adding 1ml of EtOH, 40ul of 3M NaAcetate and 1ul of 10mg/ml tRNA (Sigma, R8759) and resuspended in 15ul DEPC H_2O.

2.17. Preparation of cDNA library for sequencing

Embryos were collected and frozen at -80°C (pools of 5 embryos) at stage 10.5-11.5. RNA extractions were performed using Qiagen RNeasy Mini kit (Qiagen, 74106) according to the manufacturer's protocol. RNA was eluted in 50ul DEPC H2O and was used to generate a cDNA sequencing library with Illumina TrueSeq kit (RS-122-2001), according to the manufacturer's protocol. Quality and the size of libraries obtained were validated using Tapestation equipment and software (Agilent).

2.18. Preparation of ChIP-seq samples

Sperm and spermatids were separated as described above. Chromatin fractionation and chromatin immunoprecipitation (ChIP) were performed as described before (Erkek *et al.* 2013, Hisano *et al.* 2013) with slight modifications. Pre-treatment of cells with DTT was omitted and chromatin was digested with 2.5U of MNase/1 million of cells (Roche, 12533700) for 30 minutes at 37°C. Following the digestion, antibodies against histone marks were added: anti-H3K4me2 (Millipore, 07-030), anti-H3K4me3 (Abcam, ab8580), anti-H3K4me3 (Millipore CS200580), anti-H3K27me3 (Millipore, 07-449), anti-H3K27me3 (kind gift from Dr Thomas Jenuwein). Before ChIP, primary antibodies were bound to magnetic beads conjugated with secondary antibodies (Invitrogen, 11204D) according to the manufacturer's protocol and all wash steps in the protocol were performed using a magnet, instead of centrifugation. Bound DNA was isolated, separated by electrophoresis and mononucleosomal bands from sperm and spermatids were excised (Hisano *et al.* 2013) and subjected to library preparation with TruSeq DNA kit (Illumina, FC-121-2001), according to

the manufacturer's protocol. Quality and the size of libraries obtained were validated with the use of Tapestation equipment and software (Agilent).

2.19. Bioinformatic analyses

All the methods in this section were written and kindly shared by Dr Angela Simeone, Dr George Allen and Dr Charles Bradshaw.

2.19.1. Sequencing of libraries

RNA-Seq and ChIP-Seq libraries were sequenced on an Illumina HiSeq 2000 instrument in single read mode at 36 base length or paired read mode at 50 or 100 base length. The resulting fastq files were filtered and mapped against the *Xenopus laevis* genome (JGI version 6.1) using BWA 0.6.2 (ChIP-Seq) or TopHat 2.0.6 (RNA-Seq) (Li & Durbin 2009, Trapnell *et al.* 2009, Kim *et al.* 2013).

2.19.2. Xenopus laevis transcriptome

The 553,960 assembled transcripts were provided by the International Xenopus Genome Project (http://www.marcottelab.org/index.php/Xenopus_Genome_Project) in October 2012. This assembly was augmented with *Xenopus laevis* sequences from the NCBI RefSeq database downloaded in Feburary 2012 (30,611 sequences). The combined transcript sequences were filtered with cd-hit-est 4.5.7 (Li & Godzik 2006) with a similarity score of

95% to remove redundant sequences. This resulted in a final set of 39,384 transcripts. To provide gene names, orthologues were found against the *M. musculus* proteome (downloaded in January 2013 – NCBI RefSeq) using Inparanoid 4 (Alexeyenko *et al.* 2006) on predicted ORFs from the Trinity Suite (Grabherr *et al.* 2011). The sequences were further annotated using InterProScan 4.8 (Zdobnov & Apweiler 2001) to provide both InterPro Domains (Release 35) and Panther 7.2 ontology terms (Thomas *et al.* 2003). *Xenopus laevis* NCBI descriptions were provided for transcripts that originated from the NCBI.

2.19.3. Filtering sequencing data

Fastq files were filtered for low quality reads (<Q20) and low quality bases were trimmed from the ends of the reads (<Q20). Reads of good quality where a paired read was of low quality were kept. Adapter sequences were removed from both pairs using cutadapt (Martin 2011).

2.19.4. Genome based RNA-Seq mapping

Xenopus laevis draft genome (JGI version 6.1) from the International Xenopus Genome Project was used as a reference genome (ftp://ftp.xenbase.org/pub/Genomics/JGI/Xenla6.1/). Transcript sequences obtained from the assembly were assigned to genome using BLAT (Kent 2002). The resulting mappings were filtered by a mismatch threshold (2%) as well as requiring 90% of the transcript to match to the genome and all exons to match to a single scaffold. To prevent spurious matching, the genome was filtered to only include scaffolds with length > 100kb. This resulted in 34,373 transcripts mapping to the genome. This mapping was used as a junction file for Tophat 2.0.6

(Trapnell *et al.* 2009, Kim *et al.* 2013) which was used to map the RNA-Seq reads to the genome. Read counts were then generated for each of the transcripts.

2.19.5. Differential expression

RPKMs were calculated by normalizing read counts for each transcript by the transcript length and the total number of reads in the corresponding sample. Zeros were replaced with values obtained by randomly sampling from all RPKM values greater than zero and less than 0.2. These were converted back to raw counts, rounding up to the nearest integer, and then normalized using the Bioconductor package EdgeR (Robinson *et al.* 2010). Transcripts were kept in the analysis if they had at least one count per million in all of the sperm-derived embryo samples or all of the spermatid-derived embryo samples, leaving 18,340 transcripts post-filter. Differentially expressed transcripts were then called using EdgeR, taking into account the pairing of sperm- and spermatid-embryos in the design matrix of the model. Gene ontology terms over-represented among the differentially expressed genes were found using topGO (Alexa *et al.* 2006).

2.19.6. Heatmaps for differentially expressed genes

For each differentially expressed gene, $\log_2$ fold changes were calculated pairwise for each spermatid/sperm-derived embryo pair. Genes were filtered out if they were not consistently upregulated in at least 6 of 7 pairs or consistently downregulated in 6 of 7 pairs.

These $\log_2$ fold change values were then plotted for each remaining gene, ordered by mean fold change, using heatmap.2 from the gplots library in R.

2.19.7. Genome-wide correlation analysis of ChIP-seq data

For each ChIP experiment, reads in the bound (IP) and in the input samples (input) were normalised to the total number of reads aligned and scaled by a factor of 10^6 (i.e. values represent count per million, cpm). The entire genome was binned into 200bp wide windows. The coverage was computed as the number of reads in each window normalised by the total number of reads in the experiment. For each mark in each cell type the reproducibility was evaluated by estimating the Pearson correlation coefficient between ChIP-Seq replicates.

2.19.8. Histone methylation level analysis

The methylation level was computed as:

$$Methylation\ level = \frac{Coverage_{IP}}{N_{IP}} 10^6 - \frac{Coverage_{input}}{N_{input}} 10^6$$

where N_{IP} is the total number of aligned reads in the IP experiment and N_{input} is the total number of aligned reads in the input experiment.

For H3K4me2/3 the methylation level was computed in a window around the TSS [TSS-10kb, TSS+2kb]. For H3K27me3 the window considered for estimating the methylation level included the 10 kb upstream region together with the gene body. Methylation levels across replicates were averaged. Heatmaps for methylation levels at misregulated genes were generated in the same way as for differentially expressed genes.

2.19.9. Peak calling for histone marks

The detection of highly methylated histone regions (peaks) was performed with MACS2 2.0.9 (Zhang *et al.* 2008) using the broad-region option and a q-value of < 0.01. The list of confirmed peaks for each histone mark analyzed consisted of the peaks with a p-value of < 0.01 detected in at least two out of three replicates.

2.19.10. Statistical testing of ChIP-seq data

Statistical analysis was conducted in R (http://www.R-project.org). The comparison of the methylation levels between promoter regions of 100 misregulated genes and promoter regions in the entire genome was performed using Kolmogorov-Smirnov test (R function ks.test()). The difference between methylation levels of promoter regions between cell types (sperm, spermatids) was tested in the same way.

The enrichment in the proportion of misregulated genes positive for H3K27me3 in sperm and spermatid was tested by the non-parametric Chi-squared test for proportions (R function prop.test()).

Chapter 3

Sperm-derived embryos develop better than nuclear transfer-derived and spermatid-derived embryos

3.1. Introduction

The goal of my PhD is to understand the mechanisms of sperm programming for embryonic development. To be able to investigate these mechanisms, firstly a good system is needed in which the developmental capacity of sperm, and of other cells, could be tested and compared. One would ideally want to compare the developmental potential of embryos generated with sperm or with a somatic cell. However, a typical somatic cell is not appropriate for such comparison, for the reason that it is diploid. Therefore, if embryos are to be generated with a diploid cell (by a nuclear transfer to unfertilised egg), the egg would need to be enucleated. By contrast, fertilisation of an egg by a sperm naturally gives rise to a diploid embryo with no need for egg enucleation. Thus, comparison of somatic cell nuclear transfer with fertilisation is limited by technical differences in embryo generation; one would need another haploid cell for fair comparisons with sperm. Therefore, it was reasoned that spermatids, immature precursors of sperm, could be appropriate for such comparisons. First, the same ploidy would enable their comparison in the same assay and second, the fact that they come from the same lineage, reduces the differences between the two cells. Too many differences, for example between sperm and somatic cells, would make it difficult to identify those differences, which are developmentally-relevant.

In this chapter I first described how I isolated sperm and spermatids from *Xenopus laevis* testicular cells. Second, I described the results of the comparison of the developmental potential of sperm- and of spermatid-derived embryos. Last, I described the results of the comparison of the developmental ability of spermatid-derived embryos and nuclear transfer-derived embryos.

3.2. Separation of *Xenopus laevis* sperm and spermatids

To compare the abilities of sperm and spermatids to support embryonic development I first needed to establish a method allowing their separation from *Xenopus leavis* testes. I have achieved this by adapting and modifying a previously described method using a density gradient centrifugation in a Metrizamide gradient (Risley & Eckhardt 1979). Metrizamide was unfortunately no longer commercially available, therefore I have found another non-ionic solution, Optiprep (Sigma, D1556), with a similar density. The density of 60% Metrizamide solution used in (Risley & Eckhardt 1979) was 1.33g/cm^3, whereas the density of Optiprep solution (which is 60% solution of iodixanol in water) was 1.32g/cm^3. I have therefore used Optiprep in my attempts to separate *Xenopus* testicular cells.

Briefly, a mixture of testicular cells was isolated from testes (see Experimental procedures) and overlaid on an Optiprep step gradient prepared in a centrifuge tube. A step gradient was prepared as described (Risley & Eckhardt 1979), with a slight modification: an additional 1ml of 20% step was introduced between the 30% and 12% steps (Fig. 5). Centrifugation was performed in the same conditions as described in (Risley & Eckhardt 1979) and this allowed separation of testicular cells into two different fractions: spermatids (collected from the top interface fraction) and mature sperm (pelleted at the bottom of the tube) (Fig. 5).

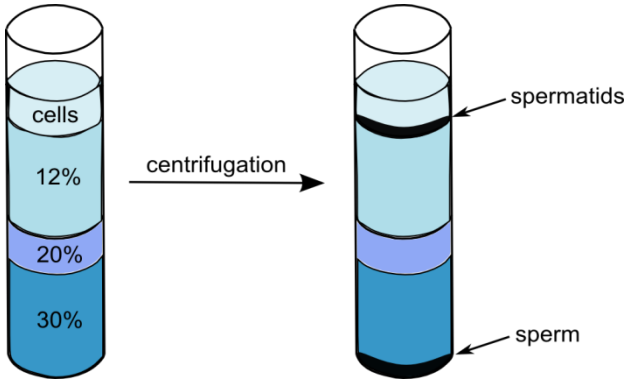


Fig. 5. Separation of sperm and spermatid by density gradient centrifugation

The diagram explains the procedure of separating sperm and spermatids from *Xenopus laevis* testis. First, step gradients of Optiprep are prepared in a centrifuge tube. Percentages of Optiprep in 1XMMR in each step are indicated in the diagram (see also Experimental procedures). Testicular cells isolated from the testis are overlaid on top of the prepared gradient. After centrifugation, spermatids are recovered from the interface between the 1XMMR and 12% Optiprep fraction, and the sperm is recovered from the bottom of the tube (pelleted cells).

Morphological observations under a phase contrast microscope, as well as DAPI staining of these two cell populations confirmed their successful separation (Fig. 6A and 6B). The mature sperm population was 95-99% pure, and it was possible to assess the purity of the sperm by microscopic observations, due to the fact that mature sperm in *Xenopus laevis* have a characteristic, snake-like shape (Risley & Eckhardt 1979, Risley *et al.* 1982). It is however not possible to judge the purity of the spermatid population by simple microscopic observations, due to the fact that spermatids at many stages of their maturation, have a round or elongated shape. Spermatids can therefore be misassigned as another cell types by an inexperienced researcher (Abe 1988, Abe & Hiyoshi 1991). To circumvent this problem, the purity of the spermatid fraction was assessed by flow cytometry analysis. Spermatids have already completed meiosis and are haploid with a reduced DNA content, so they can be

distinguished from other round-shape diploid and tetraploid testicular cells by DNA staining and ploidy assessment. Flow cytometry assessment estimated the spermatid content to be around 80% (Fig. 6C).

These two purified populations were used in all the subsequent experiments in which sperm and spermatids are compared.

3.3. Sperm-derived embryos develop better than spermatid-derived embryos

Next, I needed to develop an experimental setup in which the developmental potential of sperm and spermatids can be compared. As opposed to sperm, spermatids are not motile and therefore cannot swim to the egg for fertilisation. I therefore reasoned that a fair comparison between the two cells would be to inject them directly into the cytoplasm of the egg. For that, a protocol for intra-cytoplasmic sperm injection (ICSI) was adapted from (Smith *et al.* 2006). The technique relays on injecting single permeabilised sperm cells into unfertilised eggs (Fig. 7A). I confirmed that embryos obtained in such way can develop into healthy adult organisms (Fig. 7B).

For the purpose of comparing the developmental capacity of sperm and spermatids, sperm and spermatids were permeabilised and subsequently injected into unfertilised *Xenopus* eggs. The embryo development was assessed at two different stages: at an early gastrula stage and at a swimming tadpole stage and the embryos were scored as the percentage of those reaching a gastrula/swimming tadpole stage to the total number of cleaved embryos (Fig. 7C).

Both sperm- and spermatid-derived embryos reached the gastrula stage with a similar efficiency. However, sperm-derived embryos developed significantly better to the swimming tadpole stage than spermatid-derived embryos (p-value < 0.05) (Fig. 8A and 8B).

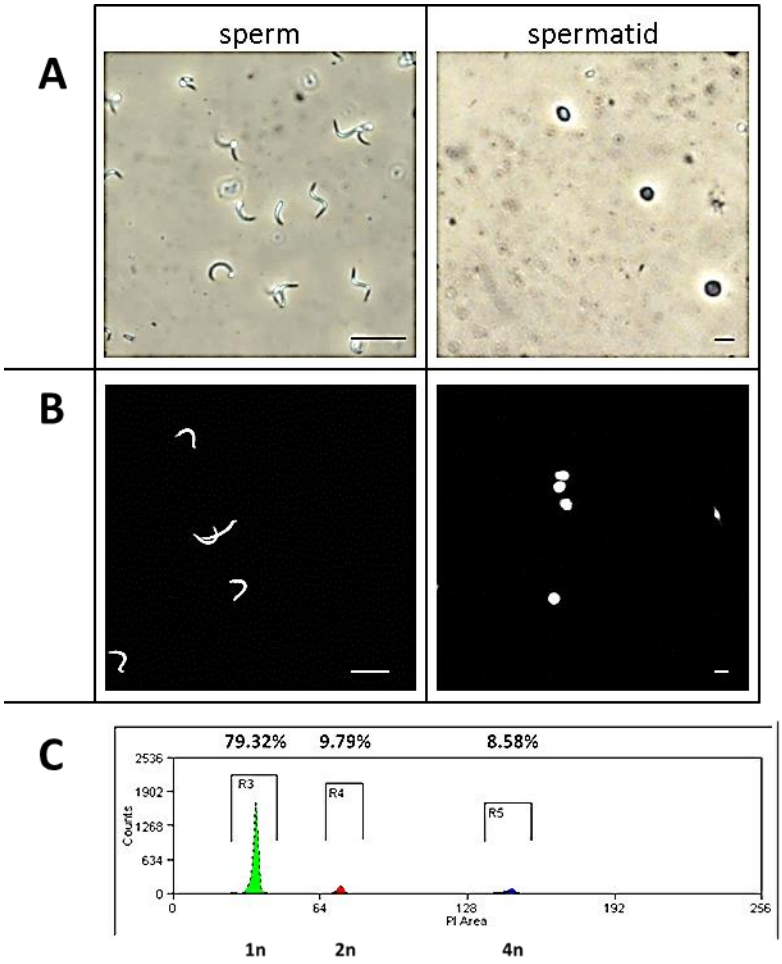


Fig. 6. Purity assessment of the spermatids and sperm populations

(A) Observations of sperm and spermatids population with a phase contrast microscope reveals that the two populations isolated are morphologically different. Sperm cells have a characteristic snake-like shape (left panel), whereas spermatids are round cells with a dense interior structures (right panel). Scale bar = 10um. (B) DNA staining with DAPI of sperm and spermatids confirms their morphological (nuclear shape) differences. Scale bar = 10um. Due to the fact that the sperm cells have a distinct morphology, microscopic observations (A and B) allow the estimation of the purity of sperm population to be around 95-99%. (C) Flow cytometry-based assessment of the purity of the spermatid population. Spermatids (haploid cells) consist around 80% of all the cells.

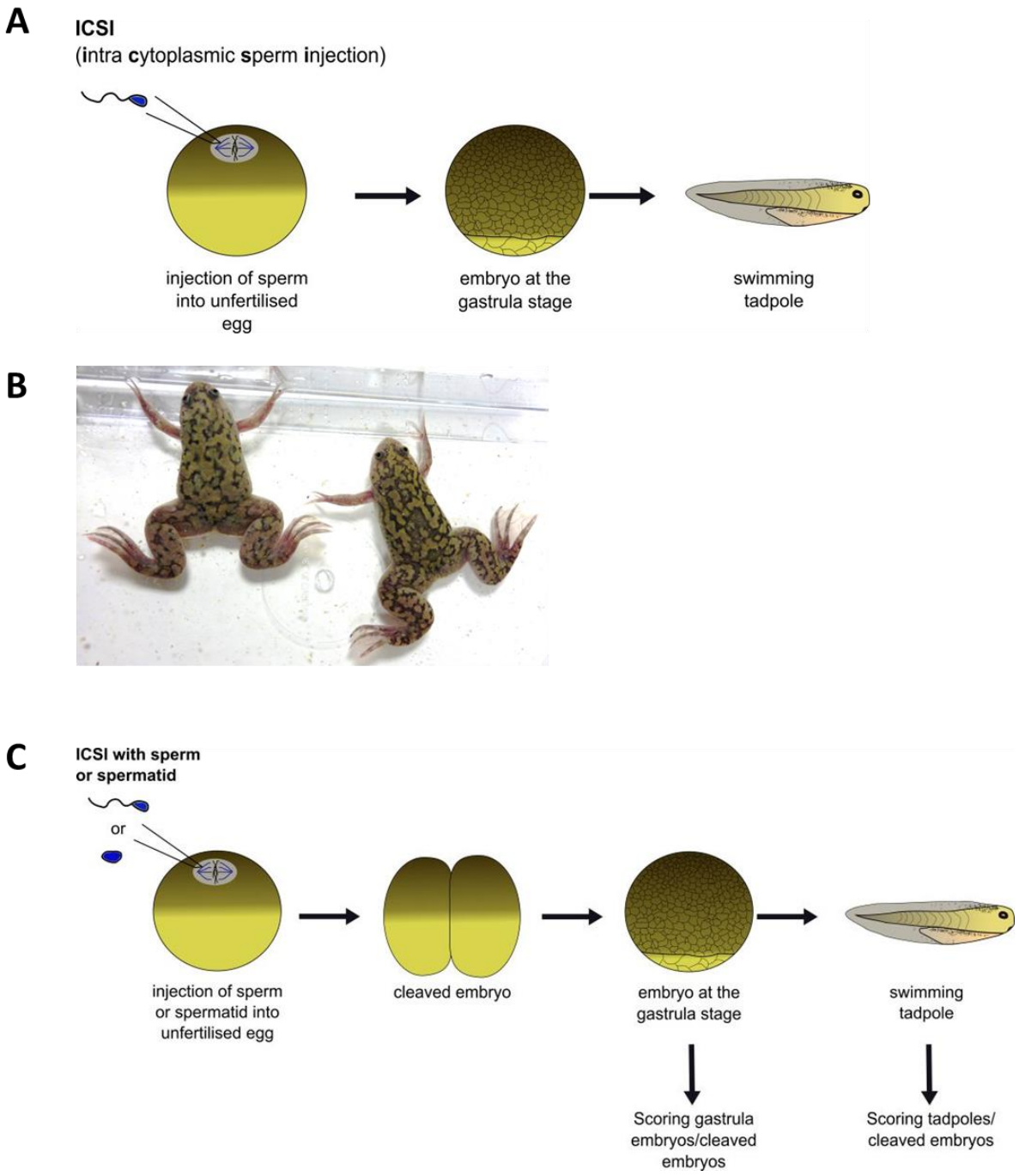


Fig. 7. ICSI with sperm and spermatids

(A) Diagram explaining the principles of the ICSI procedure. The sperm is injected directly into the cytoplasm of the unfertilised egg. Successfully injected eggs can subsequently develop into tadpoles. (B) Healthy frogs obtained by ICSI procedure. (C) Diagram explaining the design of ICSI experiment to compare the developmental potential of sperm and spermatids. Sperm or spermatids are injected into unfertilised egg. Such egg develops and embryos are scored at two stages: at an early gastrula stage and at a swimming tadpole stage.

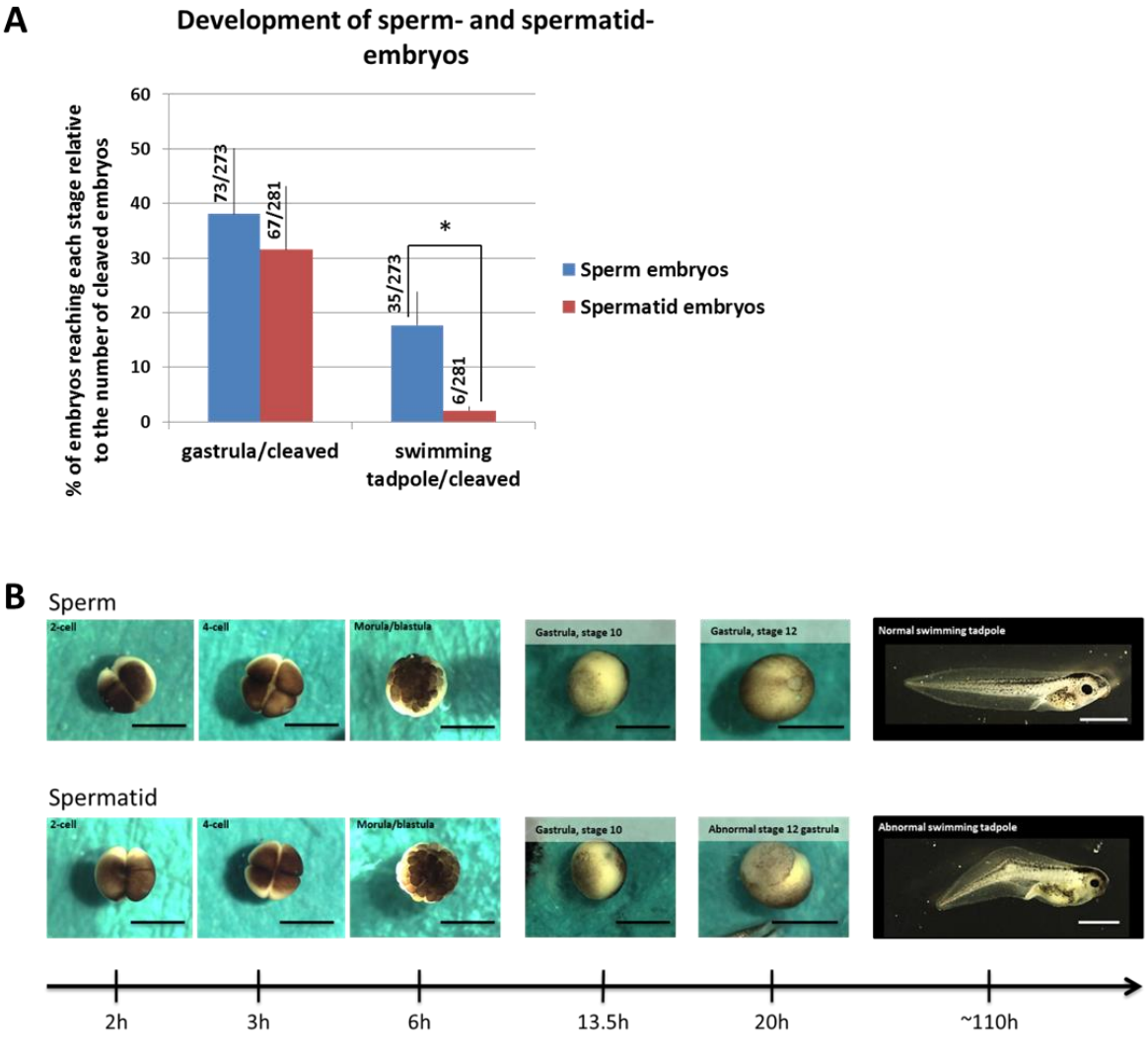


Fig. 8. Sperm-derived embryos develop better than spermatid-derived embryos

(A) Graph summarising the results of ICSI experiments with sperm and spermatids. Embryos were scored at two different developmental stages: at a gastrula stage and at a swimming tadpole stage. Sperm- and spermatid-derived embryos developed similarly to the gastrula stage, but the sperm-derived embryos developed significantly better to the swimming tadpole stage. * - p-value = 0.000002 (z-test). N = 6 independent experiments. Error bars show ±SEM. Numbers above each bar represent the number of embryos tested. (B) Shows representative images of sperm-derived and spermatid-derived embryos. Note that developmental abnormalities of spermatid-derived embryos are not visible before stage 12. Timescale below the images indicates average developmental time for embryos cultured at 18ºC. Scale bar = 1mm.

3.4. Nuclear transfer-derived embryos develop with a similarly low efficiency as spermatid-derived embryos and worse than sperm-derived embryos.

Developmental potential of somatic cells cannot be directly compared to that of sperm and spermatids, due to differences in ploidy of these cells and therefore different technical manipulations required to generate the embryos. However, to get an idea as to whether development of somatic cell-derived embryos is more similar to the development of sperm-derived or spermatid-derived embryos, nuclear transfer experiments were performed and compared to the results of ICSI experiments obtained above.

Late blastula/early gastrula stage cells were used as nuclear donors for the somatic cell nuclear transfer experiments, as these cells proved to be efficient at supporting embryonic development (Gurdon 1962). Optimisation of the nuclear transfer procedure is described in the chapter 5.2.3. Cell membranes were mechanically disrupted and single nuclei were then transferred into enucleated eggs (Fig. 9A). Resulting embryos were scored as before: at an early gastrula and at a swimming tadpole stage (as a percentage of the cleaved embryos). Subsequently, the results were compared to the results obtained with ICSI with sperm and spermatids.

Nuclear transfer-derived embryos developed to the gastrula stage with a similar frequency to sperm- and spermatid-derived embryos. Interestingly, nuclear transfer-derived embryos developed to a swimming tadpole stage less efficiently than sperm-derived embryos (p-value < 0.05) and with a similar efficiency to spermatid-derived embryos (p-value > 0.05) (Fig. 9B).

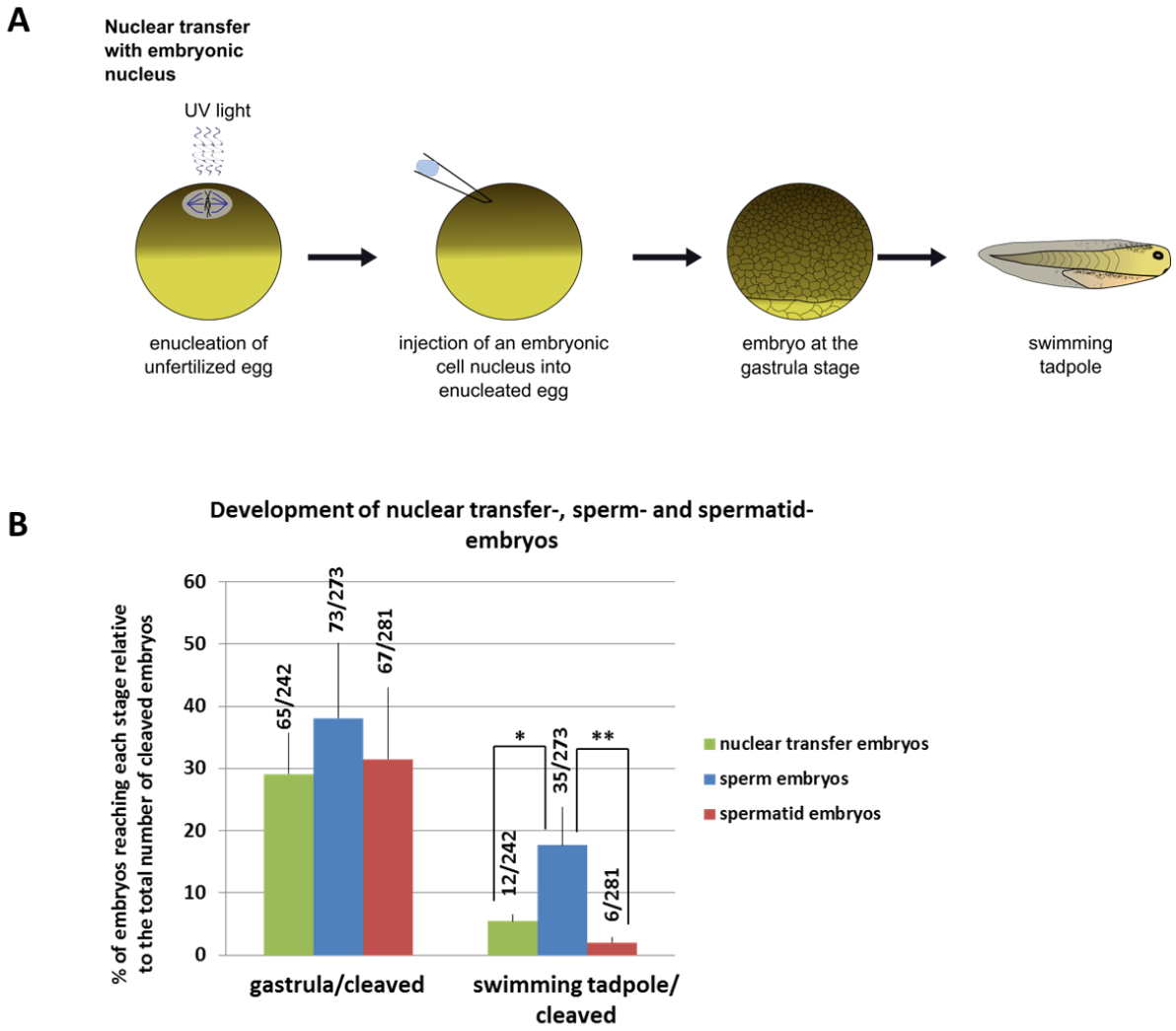


Fig. 9. Sperm-derived embryos develop better than nuclear transfer-derived embryos

(A) Diagram explaining the procedure of nuclear transfer to *Xenopus laevis* eggs. The egg is enucleated by UV radiation and subsequently injected with an embryonic cell nucleus. Successfully reconstructed embryos can develop to a swimming tadpole stage. (B) Comparison of the developmental potential of sperm-, spermatid- and nuclear transfer-derived embryos. Scoring was performed as described in Fig. 8. Note that the data for sperm- and spermatid-derived embryos come from the experiments described in Fig. 8. Nuclear transfer embryos are from 4 independent experiments. Error bars show ±SEM. Numbers above each bar represent the number of embryos tested. * indicates p-value=0.002; ** indicates p-value=0.000002 (z-test).

3.5. Summary

The results obtained so far show that sperm-derived embryos developed significantly better than spermatid-derived embryos. Furthermore, the developmental potential of nuclear transfer-derived embryos was as low as that of spermatid-derived embryos (as compared with sperm-derived emrbyos). This suggests that the sperm is better suited to support embryonic development than a spermatid and a somatic cell.

Since a direct comparison between sperm and somatic cells is not possible and since spermatids were similarly inefficient at supporting development as somatic cells, in all subsequent analysis the sperm is compared with spermatids.

Chapter 4

Identification of proteins present in sperm, spermatids and incorporated into sperm and spermatids from the egg extract

4.1. Introduction

Results obtained so far suggest that the sperm, as opposed to the spermatid, is programmed to support efficient embryonic development. Spermiogenesis is a complex, multistep process, involving numerous molecular changes to the maturing spermatid nucleus. Many proteins are lost and gained during sperm maturation (Gaucher *et al.* 2010). It is therefore possible, that the loss or gain of particular proteins in the course of spermiogenesis is responsible for the acquisition of the developmental advantage of sperm, as compared to the spermatid. Such proteins could have a direct or indirect effect on embryonic development. For example, if sperm was delivering transcription factors to the embryo, they could directly influence the embryonic development. On the other hand, sperm factors could also have indirect effect on development if they were recognised by egg-derived effector proteins. For example, sperm-derived protamines are recognised and processed by egg-derived Nucleoplasmin (Philpott *et al.* 1991, Philpott & Leno 1992), which could help the sperm nucleus to acquire a chromatin state compatible with early development. I therefore aimed to identify proteins that: 1) are present in the sperm nucleus itself, or 2) are egg factors that are specifically attracted to the sperm chromatin. To identify the first factors, I have compared proteins present in sperm and spermatids. To identify the second type of factors, I have incubated sperm and spermatids in egg extracts and compared the proteins bound to each type

of chromatin. In both cases the identification of differences in protein composition between the two samples was performed with the use of 2-DIGE electrophoresis (2-D Fluorescence Difference Gel Electrophoresis) followed by mass-spectrometry analysis of selected protein spots. This approach led to the identification of 51 sperm-specific proteins, 47 spermatid-specific proteins and also 107 egg proteins binding specifically to sperm upon egg-extract treatment and 20 egg proteins incorporated specifically into spermatid chromatin upon egg extract incubation.

4.2. Sperm and spermatids differ in their nuclear protein composition

First, the nuclear composition of sperm and spermatids was compared. For that, the cells were prepared in the same way as for ICSI experiments (see Experimental procedures). Proteins were extracted with Urea/Thiourea buffer (see Experimental procedures). Subsequently, equal amount of proteins isolated from sperm and spermatids were labelled with fluorescent Cy3 and Cy5 dyes and separated in two dimensions (Fig. 10A and B). During spermiogenesis, the protein composition of the nucleus of the maturing spermatid undergoes numerous changes (Gaucher *et al.* 2010). For example many proteins that become incoportated into the sperm chromatin in *Xenopus laevis*, are highly basic (Abe & Hiyoshi 1991, Hiyoshi *et al.* 1991). Therefore, in order to allow an appropriate separation of all proteins and of nuclear proteins, amongst which many are highly basic, the first dimension electrophoresis (separating proteins according to their isoelectric point) was performed separately in two different pH ranges: 3-10 (to better separate the majority of the proteins) and in pH range 7-11 (to specifically separate the basic proteins). Subsequently, all the proteins were separated by electrophoresis in the second dimension, according to their molecular mass. Gels were imaged using a laser scanner to identify protein spots which were

present specifically in sperm, spermatids and those which were common between the two samples (manual identification) (Fig. 11A and B). Afterwards, gels were silver-stained in order to allow visualisation and excision of selected protein spots (Fig. 12A and B). Subsequently, proteins isolated from each spot were subjected to mass spectrometry analysis. Only those proteins which had an overall protein probability score (calculated by Mascot: http://www.matrixscience.com/help/scoring_help.html) above 100 (the higher the score, the more probable the correct identification of the protein), or a score below 100, but at least two different peptides confirming their identity, were included in the final list of identified proteins. In total 51 sperm-specific, 47 spermatid-specific and 38 proteins present in both cell types were identified (Table S1). Amongst the sperm-specific proteins identified, proteins previously reported to be present in *Xenopus laevis* sperm were found, for example sperm basic protein 1 (Sp1), sperm basic protein 4 (Sp4) (Sp1-6 proteins are functional orthologues of mammalian protamines) and histone 1 variant H1fx (Shechter *et al.* 2009), which confirms that the approach used to identify these proteins is valid. Similarly, in the list of spermatid-specific proteins a homologue of a human spermatid-specific protein, Rsb-66, was found (Yang *et al.* 2003, Chen *et al.* 2008), which confirms a successful separation of a spermatid population and also indicates that the approach used can successfully identify spermatid-specific proteins. Proteins that were found in both cell types contained mainly basic metabolism and structural proteins, for example actin, tubulin or ATP synthase (Table S1).

Within the sperm-specific proteins there are several proteins which can be implicated in rendering the sperm susceptible to egg reprogramming. For example, I have identified Wdr5 protein in the sperm nuclei. Wdr5 has been shown to recognise and bind to dimethylated and trimethylated lysine 4 of histone H3 (H3K4me2/3). Wdr5 then recruits histone H3K4 methylase (via a direct interaction with the methylase), which results in further spreading of the activating H3K4me2/3 epigenetic mark and leads to a transcriptional

activation (Wysocka *et al.* 2005, Wysocka *et al.* 2006, Zhang *et al.* 2012). It has been shown that in *Xenopus laevis* embryos, knockdown of Wdr5 leads to abnormal expression patterns of developmentally important *Hox* genes. Furthermore, Wdr5 was shown to be required for the self-renewal of embryonic stem cells (ES cells) and also for the induction of pluripotency during the derivation of induced pluripotent stem cells (iPS cells) (Ang *et al.* 2011). Therefore, the presence of Wdr5 protein in the sperm, but not in the spermatid nucleus, could be advantageous for the sperm, as Wdr5 protein was directly shown to be involved in the regulation of transcription. Another example of a sperm-specific protein identified in this study, which could also explain the developmental advantage of the sperm is Mlf1. Mlf1 is a transcription factor that has been shown to regulate both gene transcription and cell cycle progression (Winteringham *et al.* 2004, Yoneda-Kato *et al.* 2005, Yoneda-Kato & Kato 2008). Therefore, Mlf1 could be important for the early phases of the embryonic development, and could potentially explain the developmental advantage of sperm over the spermatids. Conversely, proteins named Prohibitin and Prohibitin 2 were identified as being present specifically in spermatids. They are highly conserved proteins, with 90% and 88% amino acids identity between *Mus musculus* and *Xenopus laevis*, respectively. They were shown to have anti-proliferative functions, but also to be involved in differentiation and morphogenesis (Chowdhury *et al.* 2013). It is thus also possible that the presence of certain proteins in the spermatids, but not in sperm, such as anti-proliferative Prohibitins, could impair the development of spermatid-derived embryos.

I also tested whether any particular types of proteins are enriched in sperm and spermatids. For that I have performed gene ontology (GO) analysis using the online Database for Annotation, Visualization and Integrated Discovery (DAVID; http://david.abcc.ncifcrf.gov/). Gene ontology analysis indeed indicated that distinct classes of proteins are overrepresented in sperm and spermatids (see Table S2 for terms enriched

with a p-value < 0.05). GO analysis for biological processes (BP) for sperm-specific proteins showed a significant enrichment for terms associated with chromatin and nuclear changes, for example nucleosome organisation, nucleosome assembly or DNA packaging. This likely reflects the high degree of sperm nucleus specialisation that occurs during spermiogenesis (incorporation of protamines and global chromatin remodelling). In agreement with that, such terms were not enriched in the list of the spermatid proteins. Among the spermatid proteins the only significantly enriched BP terms were related to cellular homeostasis. Similarly, the BP terms enriched among the proteins present in both cell types were related to basic metabolism, such as glycolysis, oxidation reduction or anion transport.

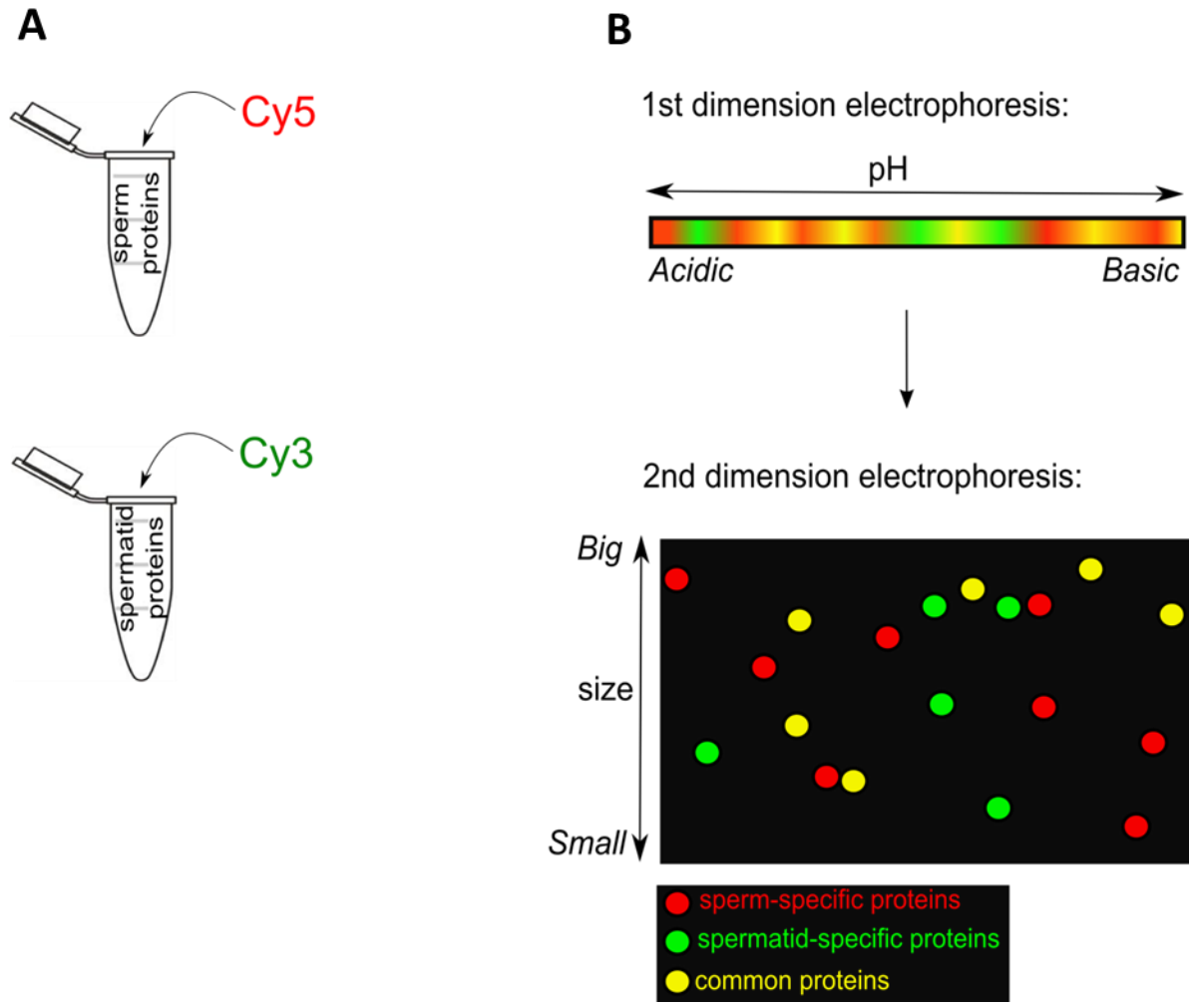


Fig. 10. Experimental design for 2-D Fluorescence Difference Gel Electrophoresis (2-DIGE) analysis

(A) The same quantity of proteins isolated from sperm or spermatids was labelled with Cy5 (red) or Cy3 (green) dye. (B) After labelling, proteins were mixed and separated in the first dimension electrophoresis in the pH range, according to isoelectric points of proteins. Subsequently, the proteins were run in the second dimension electrophoresis, according to the molecular weight of proteins. These two runs allowed separation of proteins into spots of three colours: in this example the red spots were sperm-specific; the green ones spermatid-specific and the yellow ones were proteins present in both cell types.

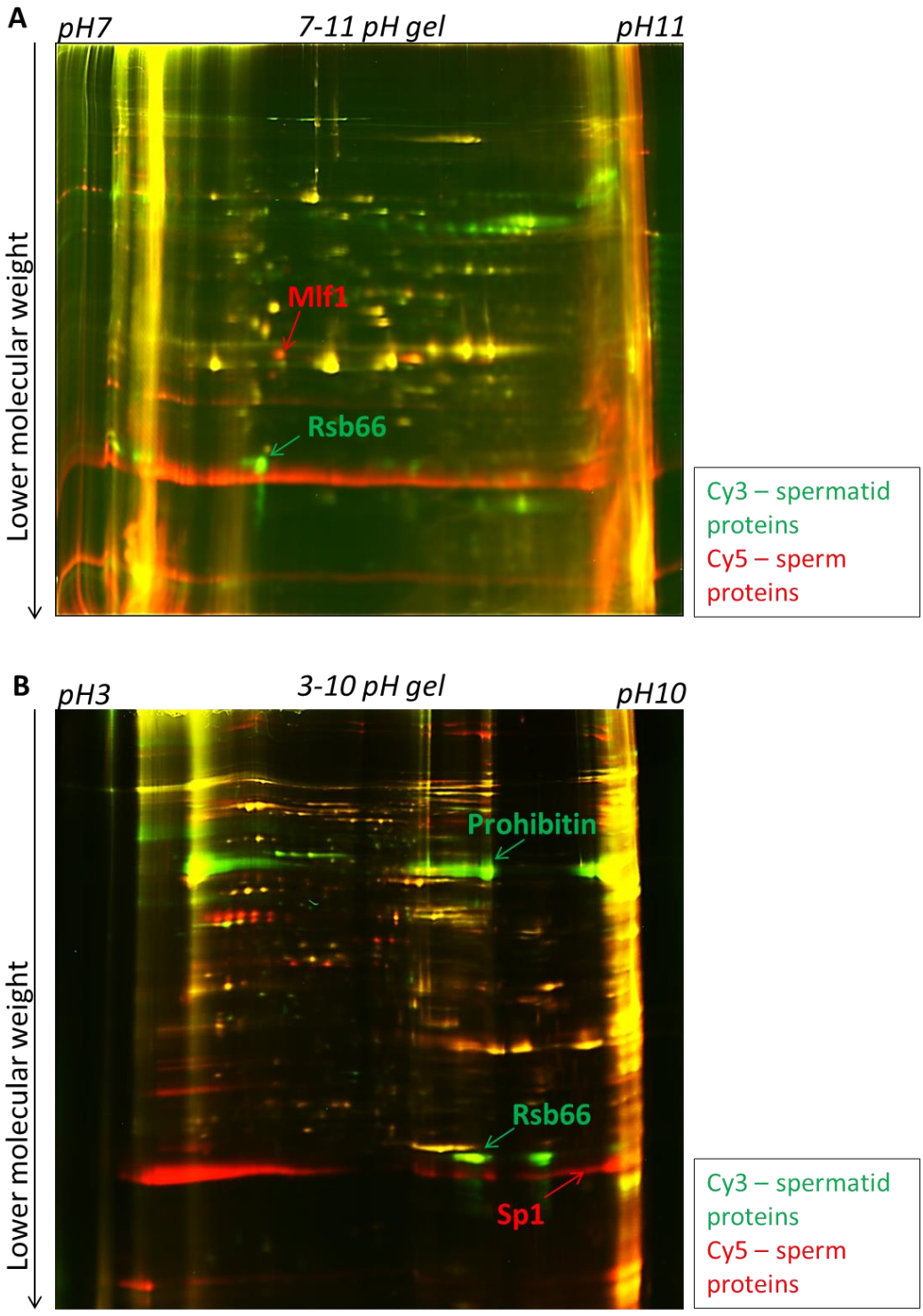


Fig. 11. 2-DIGE electrophoresis of proteins isolated from sperm and spermatids

Proteins isolated from spermatids and sperm were labelled with Cy3 (green) and Cy5 (red) dyes, respectively, and subsequently separated in two dimensions. Examples of sperm- and spermatid-specific proteins identified are indicated with arrows: red arrows for sperm-specific proteins and green arrows for spermatid-specific proteins. (A) Laser scanned image of a gel with the proteins separated in pH range 7-11 (basic range). (B) Laser scanned image of gel with proteins separated in the pH range 3-10 (broad range).

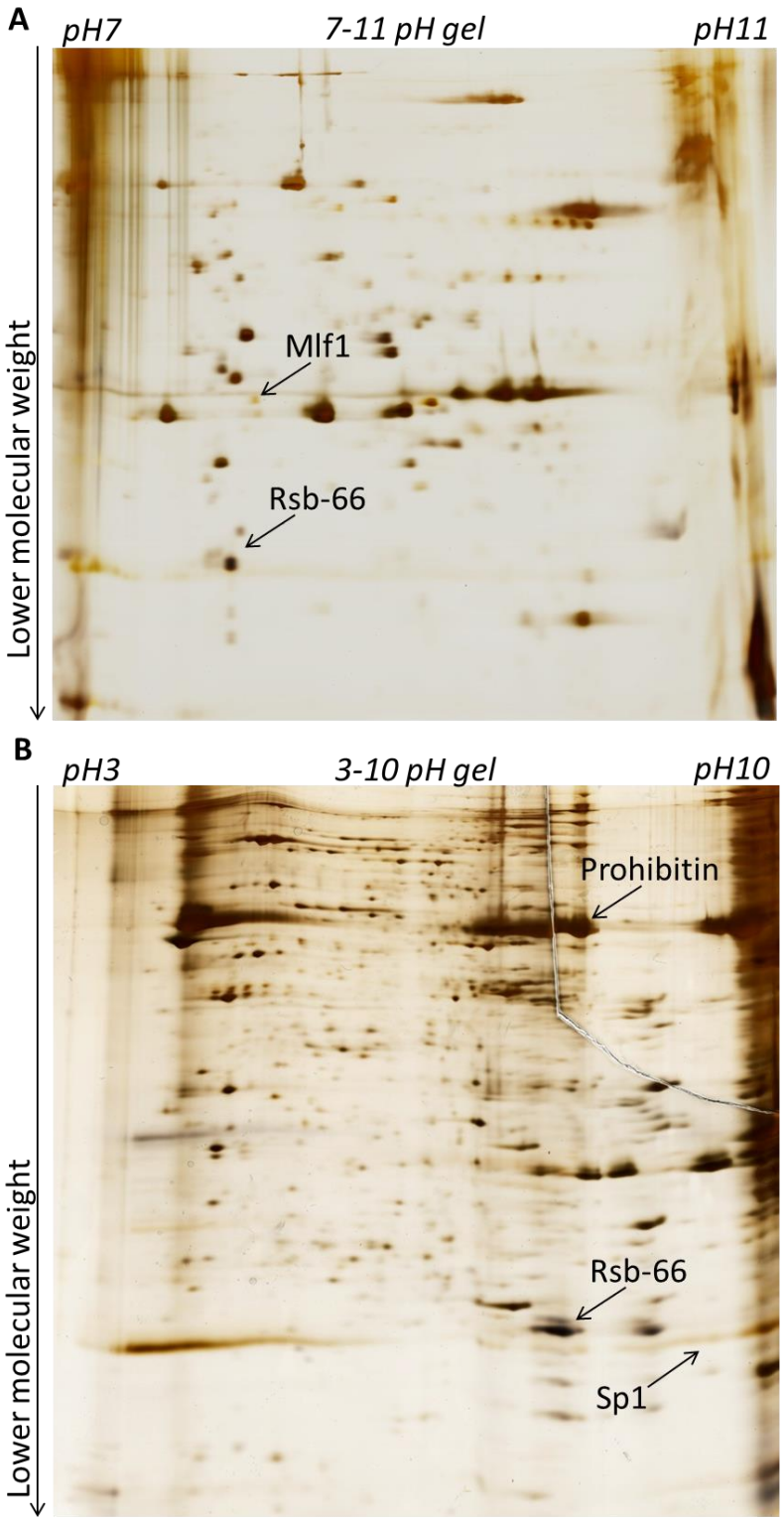


Fig. 12. Silver staining of 2-DIGE gels for spot excision

Gels from figure 11 were silver-stained to allow protein visualisation and spot excision. Protein examples from Fig. 11 are also indicated with arrows. (A) proteins separated in pH 7-11; (B) proteins separated in pH 3-10.

4.3. Sperm and spermatids bind distinct egg factors

I next tested whether differences in the reprogramming potential between sperm and spermatids could be a result of their differential ability to attract and bind egg factors after fertilisation. One cell embryo contains only one set of maternal and paternal chromosomes, making the analysis of chromatin-bound proteins challenging for two reasons. First, there is an equimolar amount of paternal and maternal chromosomes, which therefore does not allow the analysis of the factors bound only to the paternal chromosomes. Second, just one set of chromosomes per embryo limits the material availability. Therefore, I took advantage of the availability of egg extracts in *Xenopus laevis*. Extracts from activated eggs can recapitulate the whole first embryonic cell cycle: global protamine to histone exchange in the sperm nucleus, DNA synthesis and chromosome condensation for mitosis (Lemaitre *et al.* 2001, Gillespie *et al.* 2012). Incubating large number of sperm/spermatids in extracts prepared from activated egg overcomes the problem of limited amount of material in 1-cell stage embryos and mimics the events happening after fertilisation. Therefore, to test whether sperm and the spermatids attract different egg factors, I incubated permeabilised sperm and spermatids in egg extracts. Subsequently, chromatin and chromatin-bound proteins were isolated, followed by extensive washes to enrich for proteins bound to chromatin (Fig. 13 and Experimental procedures). Proteins were then isolated, labelled with Cy3 and Cy5 dyes and subjected to 2-DIGE analysis. Since only two samples can be simultaneously compared on one gel, and there were four samples to be compared (sperm, sperm-extract treated, spermatid, spermatid-extract treated), I first compared sperm with sperm-extract treated and spermatid with spermatid-extract treated (Fig. 14A and 14B, respectively). Interestingly, it turned out that virtually all the detectable proteins changed after the egg extract treatment (almost no 'yellow' spots, see fig. 14A and 14B). This can be explained by two possibilities: 1. the great majority of the donor nuclear proteins are removed (and/or modified post-translationally)

upon the incubation with egg extract or 2. egg-derived proteins are in such excess over the donor cell-derived proteins that the donor cell-derived proteins become undetectable after the egg extract treatment. Therefore, I decided to perform a direct comparison of sperm-extract treated sample with the spermatid-extract treated sample to identify egg proteins that are specifically incorporated from extracts to the sperm or spermatid nuclei. 2-DIGE analysis identified numerous protein spots present specifically in the sperm extract-treated or in the spermatids extract-treated sample (Fig. 15A). Selected spots were excised and subjected to mass spectrometry-based identification (Fig. 15B). Mass-spectrometry based analysis of peptides isolated from the selected spots, led to the identification of 107 proteins bound specifically to sperm nuclei, 20 proteins bound to spermatid nuclei and 108 proteins incorporated from the egg into both cell types (Table S3).

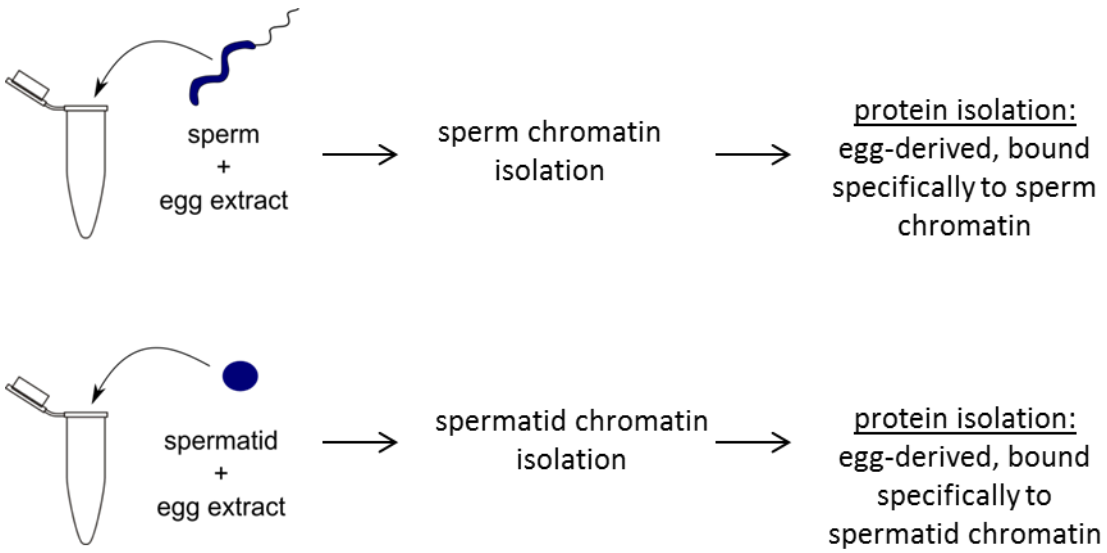


Fig. 13. Experimental design for mass spectrometry analysis of extract-treated sperm or spermatids

Sperm or spermatids are separately treated with egg extracts. Subsequently, sperm or spermatid chromatin is purified and chromatin-bound proteins are isolated. Isolated proteins are then subjected to 2-DIGE and mass spectrometry identification.

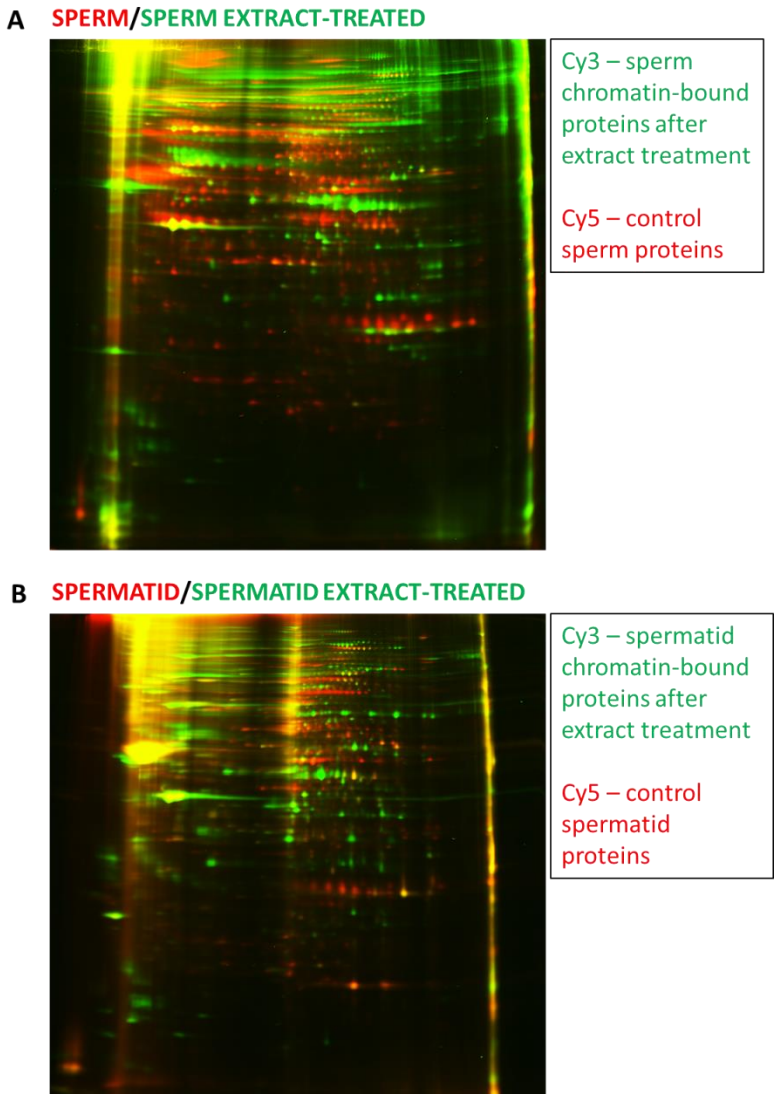


Fig. 14. 2-DIGE electrophoresis of proteins from sperm and sperm-extract treated and of spermatid and spermatid-extract treated.

(A) Proteins isolated from sperm (red) were run on 2-D gel together with proteins bound to the sperm chromatin after egg extract treatment (green). (B) Proteins isolated from spermatid (red) were run on 2-D gel together with proteins bound to the spermatid chromatin after egg extract treatment (green). Note the presence of numerous red or green spots on both gels (A and B) and the low number of yellow spots. The first dimension electrophoresis for both gels shown was carried in the pH 3-10 (broad range).

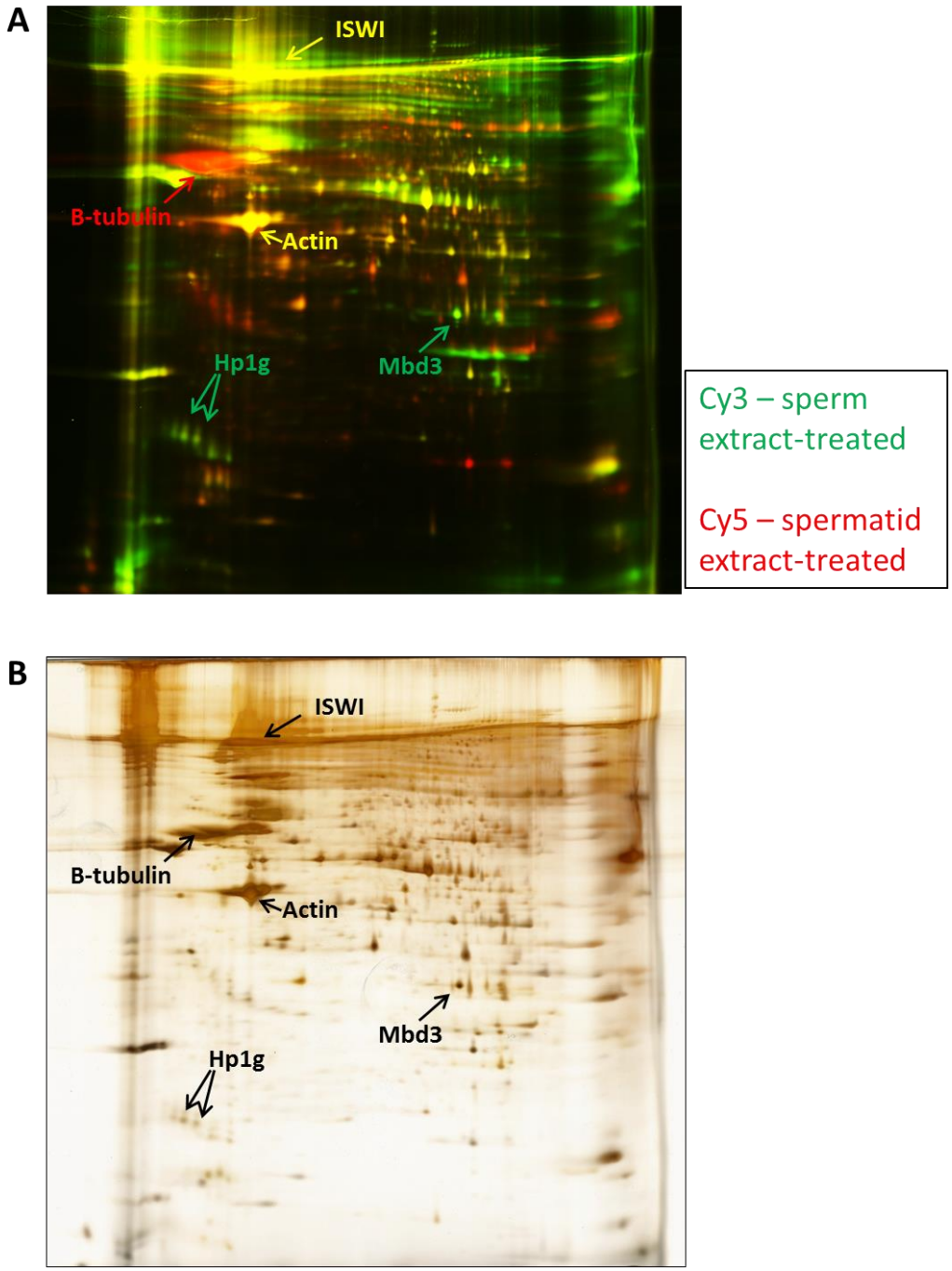


Fig. 15. 2-DIGE electrophoresis of sperm-extract treated and spermatid-extract treated.

(A) Proteins bound to sperm chromatin after egg extract treatment (green) were run on a pH 3-10 gel together with proteins bound to spermatid chromatin after egg extract treatment (red). Examples of proteins binding specifically to sperm, spermatid or to both types of cells are indicated with arrows. (B) The same gel as in (A) silver-stained to allow protein spot visualisation and excision. Examples of proteins are indicated with arrows.

Among the egg proteins that were bound to both types of nuclei, many proteins involved in global chromatin remodelling, for example, chromodomain helicase DNA binding protein 4 (Chd4) and Imitation Switch (ISWI) were identified. Both Chd4 and ISWI are ATP-dependent chromatin remodelling complexes that move nucleosomes. Nucleosome sliding can change the accessibility of the DNA and plays important roles in multiple biological process, like regulation of transcription mediated by polymerases I-III, DNA replication or DNA repair (Tong *et al.* 1998, Poot *et al.* 2005, Erdel & Rippe 2011). ISWI has been already reported to be recruited to chromatin of somatic cell nuclei incubated in egg extracts (Kikyo *et al.* 2000), thus its identification here validates the experimental setup applied. Another class of proteins abundantly detected in both types of cells after extract treatment are structural proteins that are important for the maintenance of the cell cytoskeleton and of the nucleus and nuclear structures, like actin, nucleoporin or nuclear pore complex proteins Nup98-Nup96 (Fontoura *et al.* 1999, Fontoura *et al.* 2001, Enninga *et al.* 2003, Loiodice *et al.* 2004), which likely reflect pronuclear formation activities induced by the extract. An example of another highly represented group of proteins incorporated from the egg extract into sperm and spermatid nuclei are proteins directly involved in DNA replication: origin recognition proteins (Orc1), protein involved in unwinding and remodelling of the DNA (topoisomerase I and II, FACT complex subunit Spt16, DNA replication helicase Mcm2), but also proteins involved in stabilising the single stranded DNA, necessary for replication, like Replication protein A (RPA) (Henricksen *et al.* 1996, Bochkarev *et al.* 1997, Rowles & Blow 1997, Sible *et al.* 1998, Wang *et al.* 2002, VanDemark *et al.* 2006, Han *et al.* 2010). Identification of the types of proteins mentioned above confirms that the extracts used were functional and that they were able to recapitulate at least some of the events occurring in the first embryonic cell cycle. Last, a protein named Sal-like protein 4 (Sall4) was also identified as bound to both cell types after egg extract

treatment. This finding is a bit surprising, as Sall4 is a transcription factor, and transcription factors are usually not present at very high concentrations. Sall4 was highly abundant among the proteins incorporated into chromatin after the egg extract treatment, as it was identified as the fourth most abundant protein in all the groups analysed. Interestingly, Sall4 has been reported as a master regulator of the core pluripotency network, necessary for the early embryonic development (Elling *et al.* 2006, Zhang *et al.* 2006, Tan *et al.* 2013). This finding could suggest that factors important for the pluripotency in the early embryo are incorporated into the paternal chromatin immediately after fertilisation.

Interestingly, some proteins that were identified in spots originating uniquely from extract-treated sperm or from extract-treated spermatids, turned out to be isoforms sharing a high degree of similarity, for example, a protein identified within sperm-extract treated (gi|27881711) was just 3 amino acid different from protein identified in spermatids-extract treated (gi|639691) and both of them were isoforms of High mobility group protein X (HMG-X) (Kinoshita *et al.* 1994). Another such example is a protein name Hira (histone cell cycle regulation defective homolog A) (Ray-Gallet *et al.* 2002). One isoform of this protein (Hira-A) was identified as specific for sperm-extract treated (gi|50416397) and another one (Hira) as present in both cell types after egg extract treatment (gi|14330670). There were also some cases in which the same protein was identified separately in sperm- and spermatid-extract treated. For example, Metastasis associated 1-like protein (mta2) (gi|5901733) was identified independently in sperm- and in spermatid-specific protein spots. This is likely a result of post-translational modification of the protein; however, due to the fact that the analysis performed did not discriminate between different post-translational modifications, such proteins were classified as present in both cell types after the egg extract treatment.

There were however also egg proteins that were bound exclusively to one chromatin type and not to the other. Many of the protein identified exclusively in sperm extract-treated were structural chromatin proteins, for example core histones: H2A, H2B, H3 and H4. Presence of core histones incorporated into the sperm chromatin likely reflects the remodelling of the paternal chromatin after fertilisation and the exchange of sperm-derived protamine-like proteins (sp1-6) to canonical type of histones derived from the egg. Another class of egg proteins binding specifically to the sperm chromatin are transcriptional repressors, for example heterochromatin protein 1 gamma (HP1γ), methyl-CpG binding domain protein 3 (Mbd3), histone deacetylase 1 and histone deactylase 2 (Hdac1 and Hdac2, respectively). HP1γ was shown to recognise and bind methylated lysine 9 of histone H3 (Lachner *et al.* 2001). This binding is important for the regulation of gene expression (Kwon & Workman 2011, Smallwood *et al.* 2012) and also for cell reprogramming to pluripotency (Sridharan *et al.* 2013). Mbd3 does not recognise post-translational marks on histones, but binds to methylated DNA (Wade *et al.* 1999). Mbd3 was shown to be necessary for embryonic development in *Xenopus* (Iwano *et al.* 2004) and to be a roadblock for reprogramming to pluripotency (Rais *et al.* 2013). Hdac1 and Hdac2 are enzymes responsible for removal of acetyl marks from histones, which were shown to be involved in transcriptional repression (Laherty *et al.* 1997, Hassig *et al.* 1998). Furthermore, both Hdac1 and Hdac2 are involved in DNA replication, for example by stabilising newly formed nucleosomes and also by directly interacting with topoisomerase II (Tsai *et al.* 2000, Bhaskara *et al.* 2013). Identification of many repressive egg proteins binding specifically to the sperm chromatin may seem somewhat surprising. However, during the rapid cell cycle phases of early *Xenopus* development, no transcription is observed (Newport & Kirschner 1982). Therefore, the ability to recruit all the repressive proteins from the egg may reflect

programming of sperm to participate in the earliest phases of embryonic development – to support efficient replication and prevent premature transcription.

Such a wide variety of chromatin remodelling proteins were not identified in spermatid-extract treated, with the exception of Baf57/Smarce1. Baf57 was shown to be important for cell cycle progression via transcriptional regulation of cell-cycle related genes (Hah *et al.* 2010). Furthermore, a couple of unique isoforms of structural proteins – tubulin and vimentin were identified as binding specifically to spermatid extract-treated (Table S3).

GO analysis of egg proteins binding specifically to sperm, to spermatids and to both cell types confirmed the observations made by looking at examples of proteins (Table S4). First, among the egg proteins incorporated specifically to the sperm chromatin, there was a significant overrepresentation of those belonging to BP (biological process) terms connected with chromosome and chromatin organization, chromatin modification, chromatin assembly and disassembly. Such BP terms were not enriched among proteins incorporated into spermatid-nuclei, and instead terms related to protein polymerisation and protein complex assembly were identified as significantly enriched. As expected, many cell cycle related BP terms were enriched among egg proteins incorporated into both cell types, such as DNA replication, mitosis, cell division, spindle assembly etc. This likely reflects the functionality of the extracts used for the experiments, and their ability to support the events happening during the first embryonic cell cycle.

4.4. Validation of mass spectrometry results by immunobloting.

Next I wanted to validate the mass spectrometry approach. One possibility to validate the mass spectrometry results is to perform immunoblotting analysis for candidate, mass-

spectrometry-identified proteins. For that I have chosen four proteins, Hdac1, Hdac2, Hp1γ and Mbd3, identified as binding preferentially from the egg to the sperm chromatin. Those particular proteins were chosen due to the availability of antibodies that recognise the *Xenopus laevis* proteins. Immunoblotting analysis confirmed that those proteins are preferentially incorporated into the sperm chromatin upon egg extract treatment (Fig. 16A and B), therefore validating the use of mass spectrometry approach.

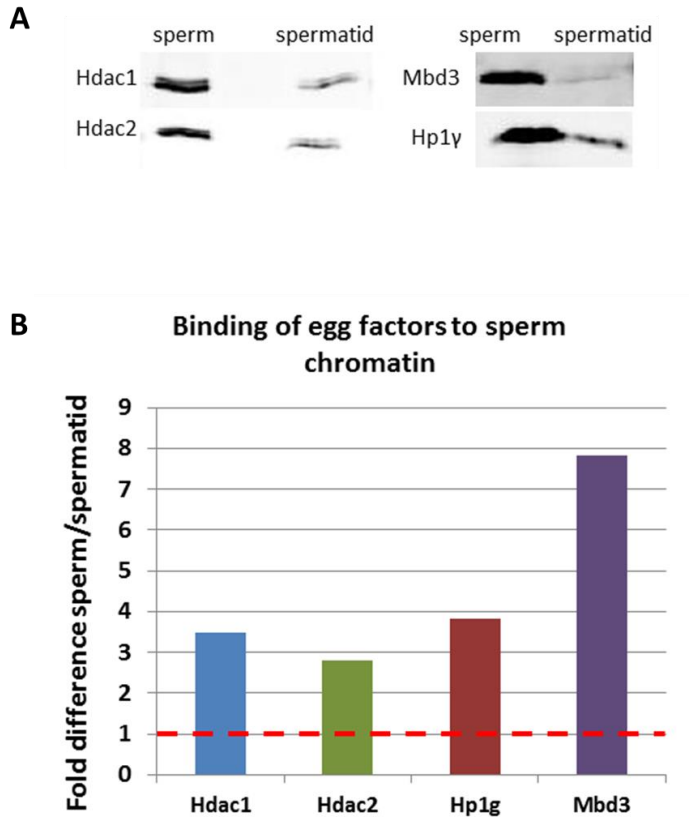


Fig. 16. Validation of mass spectrometry results by immunoblotting.

(A) Immunoblotting results on proteins isolated from sperm and spermatids after extract treatment. 'sperm' - proteins bound to sperm chromatin after egg extract treatment; 'spermatid' - proteins bound to spermatid chromatin after egg extract treatment. The antibody used is indicated on the left hand side of the blot inset. The same quantity of proteins was loaded on each lane. (B) Quantification of the blots shown in (A). Results are shown as fold difference between the band intensity in sperm-extract treated to spermatid extract-treated. Names of the proteins are indicated below the x axis.

4.5. Summary

To conclude, mass spectrometry analysis of sperm, spermatids and of egg extract-treated sperm and spermatids allowed the identification of numerous proteins present specifically in these cell types. Some of these proteins were previously identified as present in these cell types, therefore validating the approach used, for example the presence of Sp1 basic protein in sperm, Rsb-66 in spermatids or the incorporation of ISWI into chromatin after egg extract treatment. On the other hand, interesting novel findings were also made. For example, Wdr5 protein was found as present exclusively in sperm nuclei, which could indicate sperm programming for efficient embryonic development. Also, several proteins were identified as incorporated specifically to the sperm chromatin, for example HP1γ. This may reflect interplay between the sperm chromatin and the egg cytoplasm – unique features of sperm, but not of spermatid chromatin, may allow the binding of specific egg factors.

It is also important to note that the approach used here to identify the proteins has some limitations. First, proteins need to be sufficiently abundant in order to be identified by 2-DIGE/mass spectrometry approach. Therefore, less abundant proteins may be missed in this approach. Second, only some of the spots (not all of them) were excised and analysed. This means that not all the proteins that are different between sperm and spermatids were analysed and also, not all the egg proteins preferentially binding to the sperm or spermatid chromatin were identified. Therefore, even though there are many interesting candidate proteins identified in the analysis (see above), one has to remember that this is only a subset of all the proteins changing between these different conditions.

Chapter 5

Functional assessment of candidate reprogramming factors

5.1. Introduction

Mass spectrometry based approach allowed the identification of several factors which could potentially explain the developmental advantage of sperm over spermatids. My next aim was to functionally validate these factors. To do this, I first selected several sperm-specific factors identified by mass spectrometry, as well as several egg factors specifically incorporated from the egg extract into the sperm chromatin.

For the functional validation of the sperm-specific factors I have selected these factors which were abundantly detected in sperm, and which are implicated in the most dramatic nuclear changes between the spermatid and sperm nucleus. These were: sperm-specific basic nuclear proteins Sp1, Sp4 and Sp5 and sperm linker histone variant H1fx. Additionally, I also included a transcription factor Mlf1 for the following reasons. First, it was very abundant specifically in the sperm nucleus, and such high abundance is somewhat surprising for a transcription factor, especially in the transcriptionally silent sperm nucleus. Second, recent proteomic studies of human sperm also identified MLF1 as present specifically in the sperm (Wang *et al.* 2013), and the conservation of the presence of this protein in the sperm nucleus between *Xenopus* and human suggests that it could be functionally relevant. Third, microarray analysis identified mRNA encoding an interacting partner of Mlf1, *Mlf1IP* (Mlf1 Interacting Protein) as present in the oocytes/eggs of three different animal species: mouse,

bovine and *Xenopus laevis* (Vallee *et al.* 2005). All these suggest that the presence of Mlf1 in the sperm could be important for the embryonic development, perhaps due to interaction of Mlf1 with its oocyte counterpart. I have additionally chosen another sperm factor, Brdt, based on published research indicating that Brdt was necessary during spermiogenesis to remodel the maturing sperm nucleus (Gaucher *et al.* 2012). Furthermore, chemical inhibition of Brdt caused infertility in mouse, suggesting that the presence of Brdt could be important for the developmental potential of sperm (Matzuk *et al.* 2012).

For the functional validation of egg factors identified as binding specifically to the sperm chromatin I have selected 7 proteins. Four of them were previously reported as transcriptional repressors (Laherty *et al.* 1997, Hassig *et al.* 1998, Jiang *et al.* 2004, du Chene *et al.* 2007): Hdac1, Hdac2, Mbd3 and Hp1γ and their recruitment to the sperm chromatin at fertilisation could be important to maintain a transcriptionally silent state during the earliest phases of embryonic development, before the onsets of zygotic genome activation. Two other proteins incorporated to the sperm chromatin from the egg extract: Rbbp4 and Rbbp7, were selected for a functional validation due to their presence in repressive complexes with Hdac1 and Hdac2 proteins, but also with Polycomb group proteins (they are both present in Polycomb Repressive Complex 2) and also because of their reported interactions with histones and roles in chromatin assembly (Vermaak *et al.* 1999, Nicolas *et al.* 2000, Kuzmichev *et al.* 2002, Yao & Yang 2003). Lastly, Lsf protein (Late SV40 Protein, also known as Cp2 or Tfcp2) was also selected for a functional validation, as it is an egg-derived transcription factor incorporated specifically to the sperm chromatin and it was reported to have oncogenic properties and to be important for the cell cycle entry and progression (Saxena *et al.* 2009, Yoo *et al.* 2010). These therefore suggest that binding of Lsf1 to the sperm chromatin could potentially facilitate the rapid cell cycles of early *Xenopus laevis* embryos and thus explain the developmental advantage of sperm over spermatids.

The functional tests of sperm-specific proteins were performed by a somatic cell nuclear transfer of donor cells ectopically overexpressing these factors, whereas tests of the functional importance of egg-derived factors were attempted by a knockdown of the egg factors.

Unfortunately, none of the sperm-specific factors exogenously expressed in the donor cells led to an increase in the efficiency of nuclear transfer. Similarly unsuccessful was the attempt to knock down the egg factors, as even though the strategy used allowed to downregulate the expression measured at the mRNA levels, the protein levels of none of these factors were reproducibly downregulated.

5.2. Functional assessment of sperm-specific proteins

5.2.1. Experimental design

High condensation of the sperm nucleus makes it very inaccessible for any technical manipulations, for example for selective protein depletion. In order to test whether the presence of candidate sperm-specific factors makes the sperm better at supporting development, I have therefore chosen to use a different strategy than depleting these factors in the sperm itself and instead I decided to ectopically overexpress these factors in somatic cells and examine the ability of such cells to support embryonic development. To achieve this, mRNA encoding a factor of interest is first injected into a 1-cell stage *Xenopus* embryo. Such embryo is then allowed to develop and as the embryo develops, the injected mRNA becomes translated into the corresponding protein. Subsequently, when the embryo reaches a late blastula/early gastrula stage it is collected and disaggregated to obtain single cells that

overexpress the protein of interest and which can be used as donors for nuclear transfer experiments (Fig. 17). Late blastula/early gastrula stage embryos were used as cell donors, since cells coming from such early embryos proved before to be efficient donors in nuclear transfer experiments (Gurdon 1962). Successfully reconstructed embryos can develop into tadpoles and their developmental potential can be compared to the developmental potential of control reconstructed embryos. I have therefore used this experimental setup to functionally test the candidate sperm-specific factors.

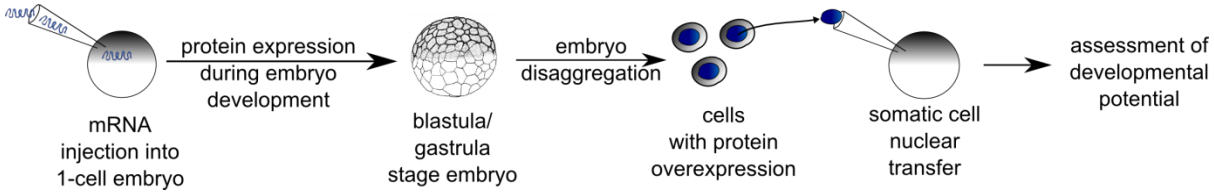


Fig. 17. Overexpression of sperm factors in donor cells for a nuclear transfer experiment.

To test whether a sperm factor can increase the developmental potential of somatic cells, mRNA encoding the factor is first injected into a 1-cell stage embryo. During embryonic development, the mRNA is translated into protein. The resulting embryo expressing the protein of interest is then disaggregated and the cells, pre-loaded with the factor of interest are used as donors for nuclear transfer experiments.

5.2.2. Cloning and ectopic expression of sperm-specific factors

Candidate sperm-specific factors (Sp1, Sp4, Sp5, H1fx and Mlf1) were first cloned from cDNA into pCS2 vectors that allow *in vitro* mRNA synthesis. They were additionally tagged with a hemagglutinin tag (HA-tag) to allow monitoring of protein expression in the absence of available antibodies. The construct encoding Brdt was tagged with a green fluorescent protein (GFP) and was a kind gift of Dr Saadi Khochbin and was sub-cloned into pCS2 vectors. At this stage, pre-testing of some of the constructs (Sp1, Sp4, Sp5 and Brdt) was performed by transiently transfecting cultured cells and checking whether the proteins

are targeted to the nucleus, since the nuclear localisation of these proteins was reported previously (Risley & Eckhardt 1981, Abe & Hiyoshi 1991, Pivot-Pajot *et al.* 2003). Microscopic observations of the transfected cells confirmed that all the tested proteins localised to the nuclei (Fig. 18A - D). Interestingly, ectopic expression of Brdt combined with TSA (Trichostatin A) treatment led to chromatin compaction (Fig. 19). TSA is an inhibitor of histone deacetylases, and therefore TSA treatment leads to an increase of histone acetylation levels (Yoshida *et al.* 1990). It was reported previously that Brdt protein can recognise and bind to acetylated histones via its bromodomains and that this leads to chromatin compaction (Pivot-Pajot *et al.* 2003, Govin *et al.* 2006, Moriniere *et al.* 2009). Therefore, my observation that expression of Brdt combined with TSA treatment led to the chromatin compaction validates the functionality of Brdt exogenously expressed in the transfected cells.

Subsequently, mRNAs encoding the candidate proteins were *in vitro* transcribed. The size and the purity of the synthesised mRNA was confirmed by an agarose gel electrophoresis and mRNAs were subsequently injected into 1-cell stage embryos to test whether they can be efficiently translated into proteins. Expression of Sp1, Sp4, Sp5, H1fx and Mlf1 was tested by immunoblotting (staining against the HA tag of the proteins), whereas the expression of Brdt was assessed by microscopic observations of whole embryos (Brdt protein was tagged with GFP). All the mRNAs tested allowed efficient protein synthesis in the embryos (Fig. 20).

5.2.3. Validation of UV treatment length required for the nuclear transfer procedure

Since the embryonic cells used as donors for the nuclear transfer procedure are diploid, the recipient egg needs to be enucleated to allow the development of a diploid embryo. Enucleation is performed by first placing the eggs on a small piece of a blotting paper soaked in water and mounted on a microscope slide. The eggs are oriented with their

animal poles and the white spot (which indicates the position of the meiotic spindle and egg chromosomes) upwards. Subsequently, the eggs are placed for 30s under a Mineralite UV lamp for enucleation and finally, for 3-6s under a Hanovia UV lamp for dejellination (Gurdon 1962) (Fig. 21). The latter treatment is used to soften the jelly that coats the egg and to make it penetrable by the injection needle. The length of the Hanovia lamp treatment has to be optimized every time a new batch of eggs is used. In general, too short a treatment does not allow the insertion of a needle into the egg, whereas too long a treatment is detrimental for the development of embryos (Fig. 22), therefore each time the nuclear transfer procedure is performed, the researcher has to determine the shortest length of Hanovia lamp treatment that allows the needle to penetrate the egg.

Next I validated whether 30s treatment of eggs with Mineralite UV lamp, which has been used in the nuclear transfer procedure in the past (Gurdon 1962), is sufficient to enucleate the eggs. If the eggs are successfully enucleated, the resulting embryos should be haploid (since the sperm that fertilises the egg is haploid and the maternal genetic content is destroyed). Haploid embryos in *Xenopus laevis* are viable, but differ from the diploid ones morphologically: haploid tadpoles are more vegetalised (shorter and thicker) than the diploid ones. It is also possible to assess the ploidy of the embryo by squashing its cells on a microscope slide and looking at their nuclei with a phase contrast microscope. In *Xenopus laevis* each set of parental chromosomes give rise to one nucleolus, visible as a black dot in the nucleus. Therefore, diploid cells have two nucleoli in each nucleus (two black dots), whereas haploid cell have just one nucleolus in each nucleus (one black dot). To test whether 30s treatment with Mineralite UV lamp is sufficient to enucleate the eggs, I first enucleated them as described above (Fig. 21) and then fertilised them. The control embryos were fertilised without enucleation. Morphological observations of the embryos revealed that the Mineralite-treated eggs gave rise to haploid embryos (they were vegetalised as compared to

control, diploid embryos) (Fig. 23A). Furthermore, microscopic observations of squashed cells also confirmed that embryos obtained from Mineralite-treated eggs were haploid (evidenced by the presence of only one nucleolus/nucleus) (Fig. 23B).

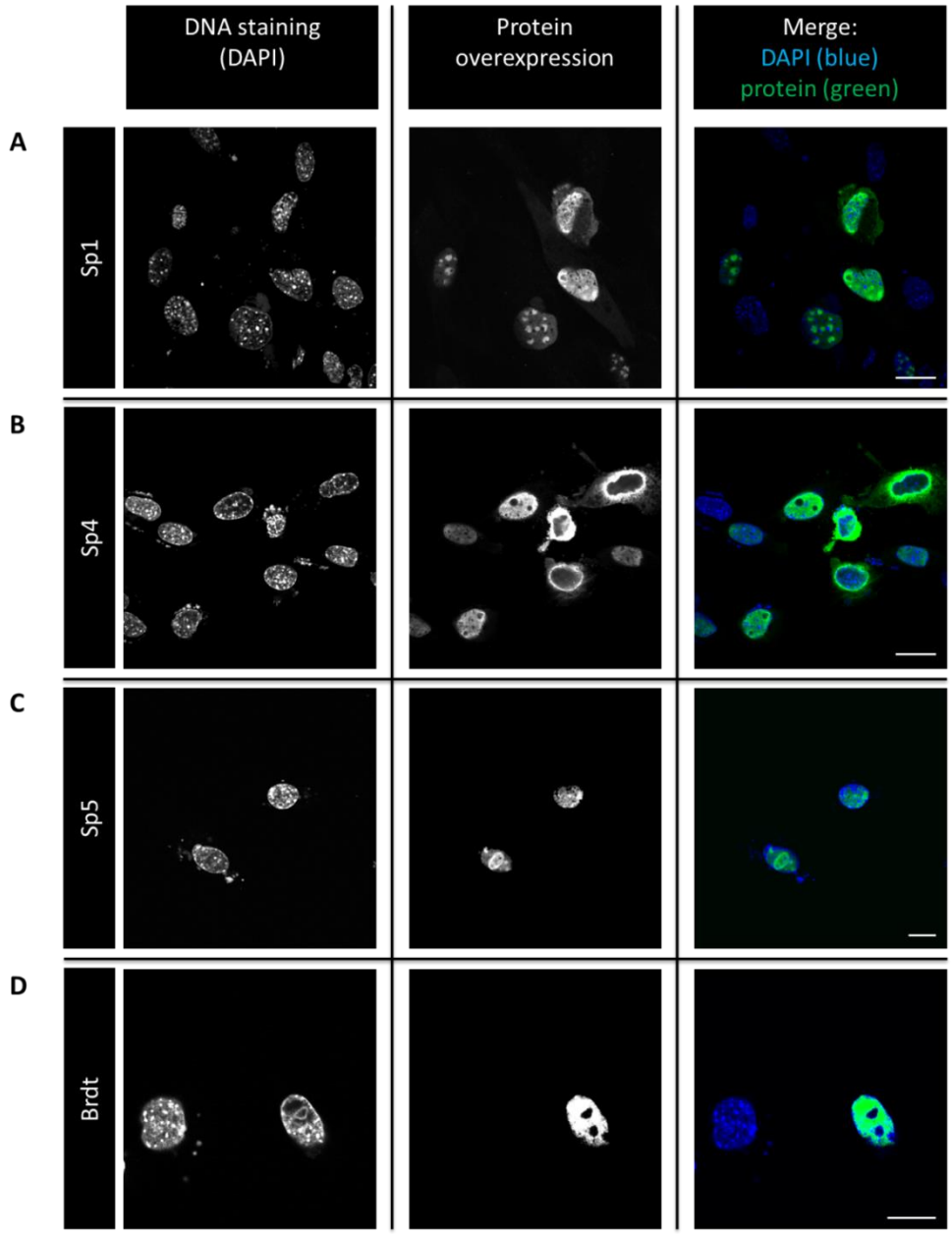


Fig. 18. Nuclear localisation of Sp1, Sp4, Sp5 and Brdt in transfected C2C12 cells.

C2C12 myoblast cells were transfected with plasmids encoding selected sperm factors: Sp1 (A), Sp4 (B), Sp5 (C) and Brdt (D). 48h after transfection cells were fixed and subjected to immunostaining revealing DNA (staining with DAPI, left column) and the overexpressed sperm factor (middle panel). Merge images (right column) show that the overexpressed proteins (green) localise to the nuclei (blue). Scale bars = 10um.

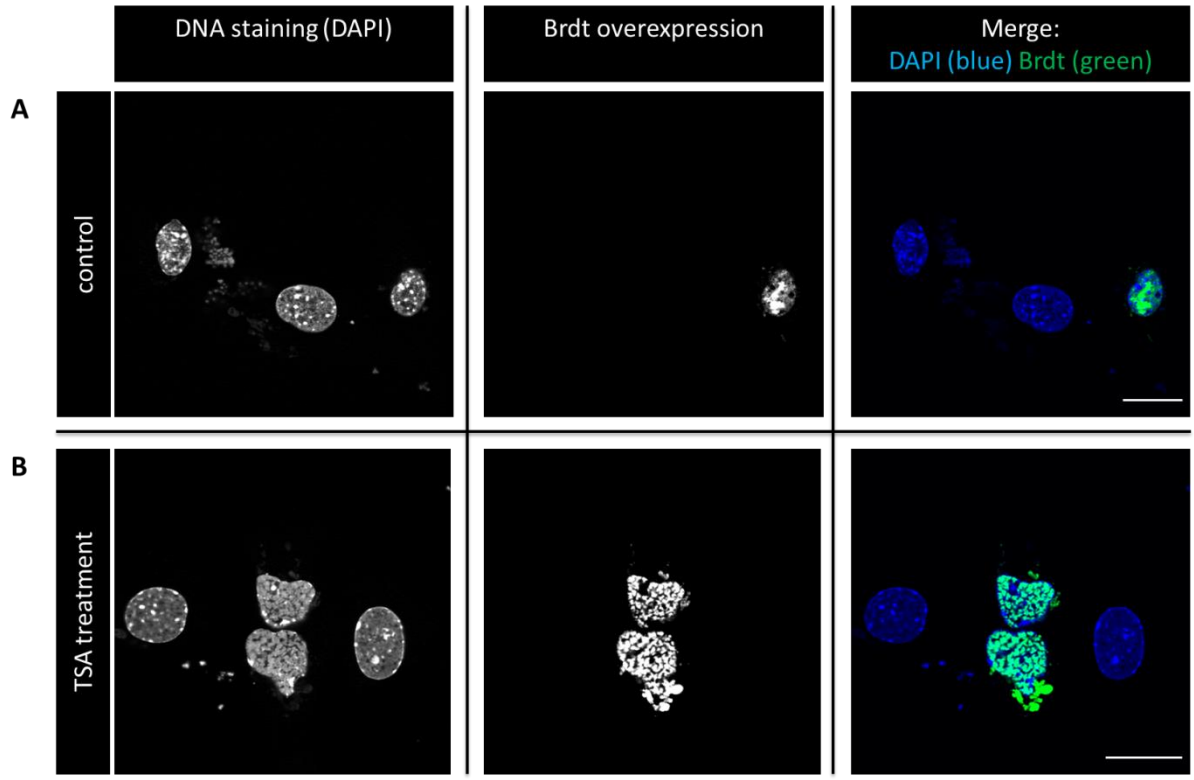


Fig. 19. Brdt-dependent compaction of chromatin upon TSA treatment.

C2C12 myoblast cells were transfected with a plasmid encoding Brdt and treated with 120ng/ml TSA or with an equivalent concentration of DMSO (control). 48h after transfection cells were fixed and subjected to immunostaining revealing DNA (staining with DAPI, left column) and Brdt (middle panel). Merge images (right column) show that control, DMSO-treated cells do not compact chromatin upon Brdt overexpression (A), but cells treated with TSA and overexpressing Brdt, do compact the chromatin (B). Scale bars = 10um.

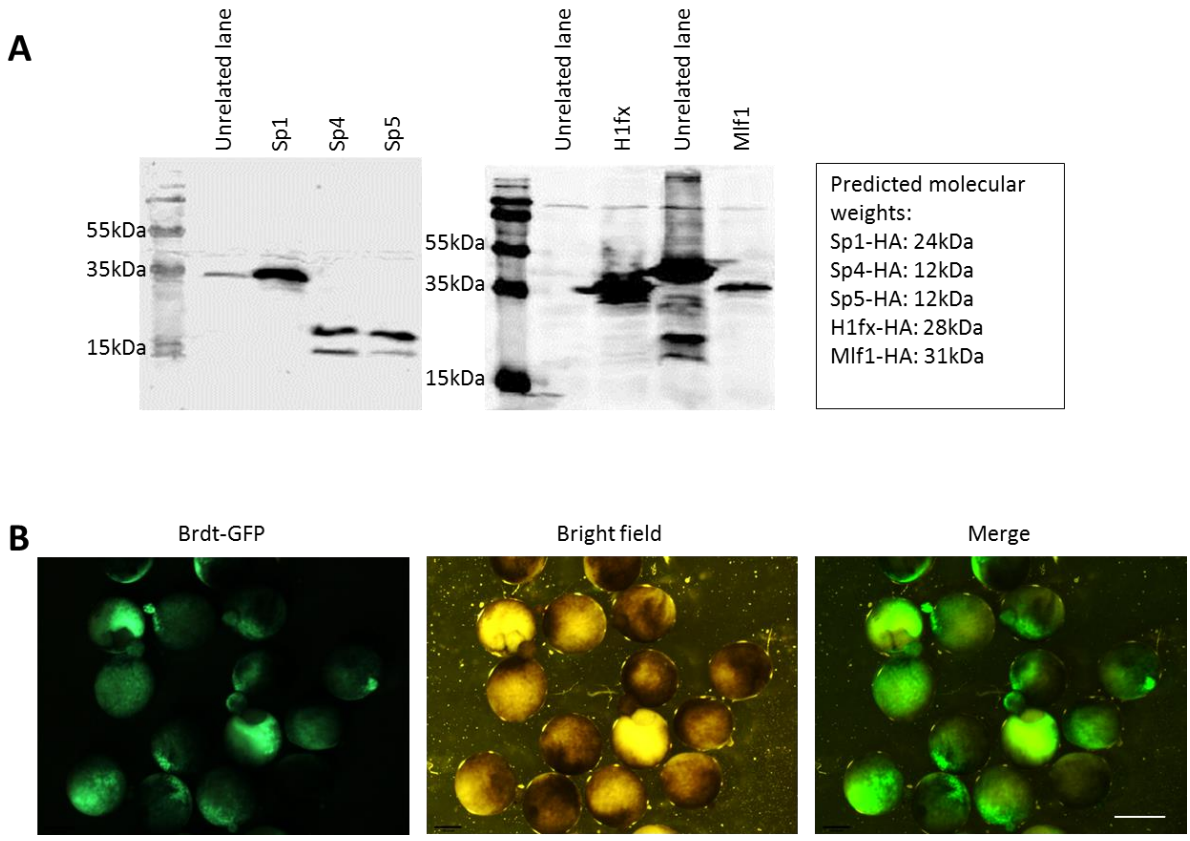


Fig. 20. mRNA injection into 1-cell stage embryos allows protein synthesis.

One cell stage embryos were injected with mRNAs encoding various sperm-specific proteins. At gastrula stage the embryos were either collected for immunoblotting analysis against the HA-tag (A) or were photographed under the microscope equipped with a fluorescent lamp (B). (A) Immunoblotting analysis confirmed that all proteins encoded by the injected mRNAs (Sp1, Sp4, Sp5, H1fx and Mlf1) were translated in the embryos. Note that the predicted molecular weights are slightly different from the observed molecular weights of proteins, which could be due to their post-translational modifications. (B) Microscopic observations under fluorescent light of embryos injected with mRNA encoding Brdt-GFP, revealed that the protein is expressed in the embryos (merge image). Scale bar = 1mm.

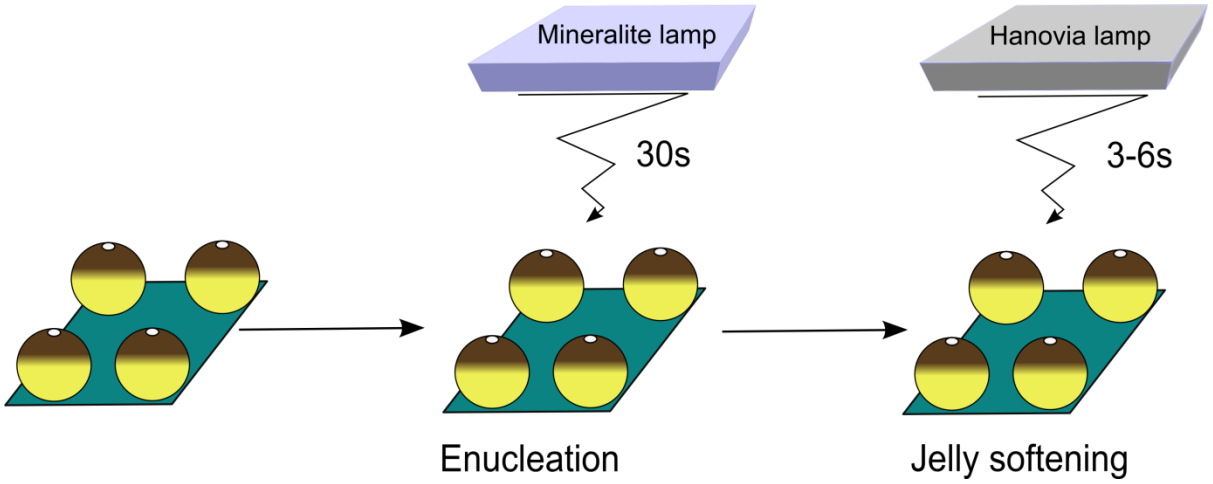


Fig. 21. Diagram explaining recipient egg preparation for the nuclear transfer procedure.

Eggs are first immobilised with the animal pole (white spot) facing upwards on a wet blotting paper. Subsequently, eggs are enucleated with a 30s treatment with a Mineralite lamp UV light and then the jelly coat of the eggs is softened by a 3-6s treatment with a Hanovia lamp UV light.

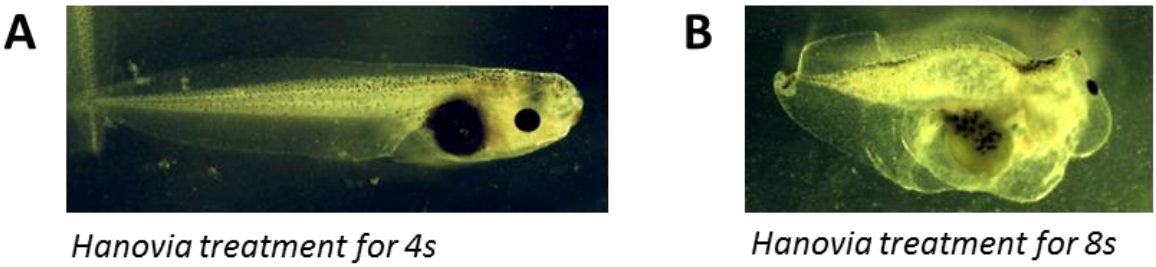


Fig. 22. Prolonged Hanovia lamp treatment leads to abnormal development.

Control fertilised embryos were treated with a Hanovia lamp for 4s - the shortest time allowing the needle penetration (A) or for a prolonged time – 8s (B). Note that a prolonged treatment with Hanovia lamp results in abnormal development of the embryos.

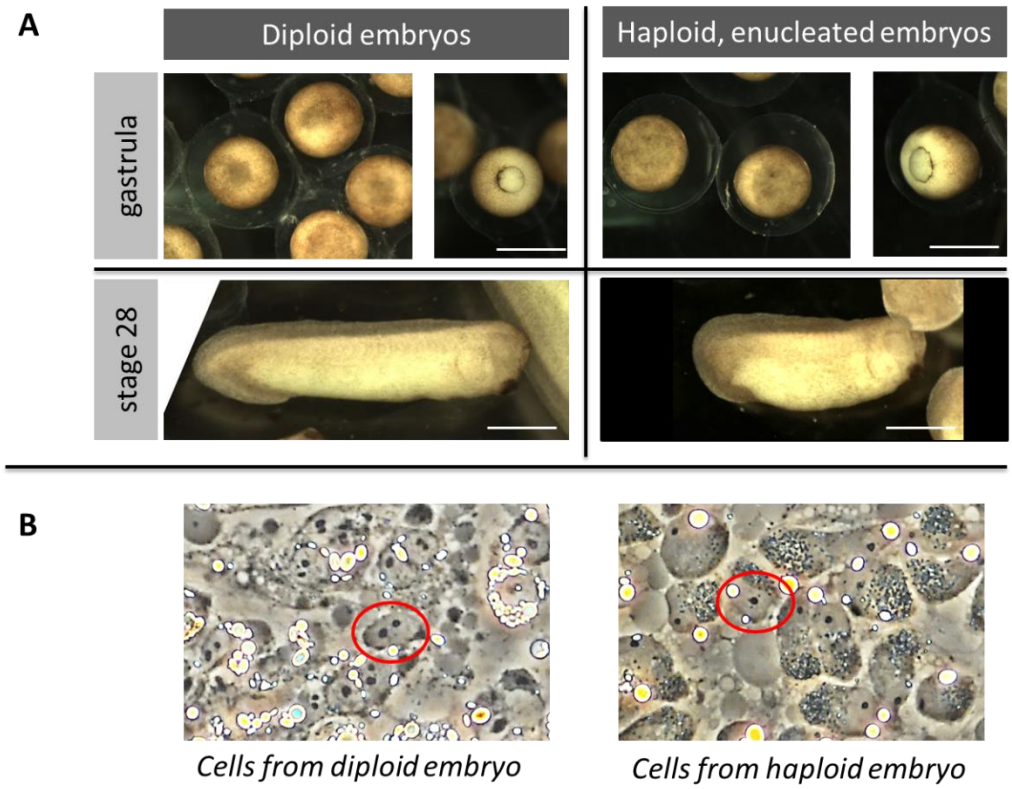


Fig. 23. Mineralite UV lamp treatment for 30s destroys the genetic material of the egg.

Eggs were subjected to 30s Mineralite UV lamp treatment for enucleation and subsequently fertilised. Control eggs were directly fertilised, omitting the Mineralite treatment. (A) Microscopic observations of control embryos (non-enucleated) (left panel, 'diploid embryos') and of embryos obtained from Mineralite-treated eggs (enucleated) (right panel, 'haploid, enucleated embryos') reveals that haploid embryos are vegetalised and therefore confirms the successful egg enucleation. Note that the vegetalised phenotype is not apparent at the gastrula stage (upper panel) and only becomes visible when the embryo elongates (bottom panel). Scale bars = 1mm. (B) Microscopic observations of cells from control embryos (left panel) and from embryos obtained from Mineralite-treated eggs (right panel) confirms a successful egg enucleation. Note that in the nuclei of cells from diploid embryos two nucleoli (black dots) are visible (left panel), whereas in the nuclei of cells from haploid embryos only one nucleolus per nucleus can be detected (single black dots). Examples of nuclei in each cell preparation are inside the red circles.

5.2.4. Overexpression of candidate sperm factors does not increase the efficiency of nuclear transfer.

I next tested the effect of overexpression of candidate sperm factors in the donor cells on the efficiency of nuclear transfer. In this experimental setup the developmental capacity of embryos reconstructed with cells overexpressing the factor of interest was compared with the developmental capacity of embryos reconstructed with control cells, which did not overexpress the sperm factor. Reconstructed embryos were scored as the number of swimming tadpoles obtained to the total number of cleaved embryos. The reasoning behind this experimental design was that if the presence of some sperm-specific factors is beneficial for the sperm to support the embryonic development, its ectopic expression in a donor cell should increase the efficiency of the nuclear transfer. Some of the candidate sperm factors were therefore expressed as single factors: Brdt, Mlf1, H1fx. A combination of factors was also used for those which have similar functions: Sp1, Sp4 and Sp5 or Sp1, Sp4, Sp5 and H1fx altogether (as those factors are structuring the chromatin in the mature sperm). These experiments, which initially gave promising outcomes (tendency of sperm factor-overexpressing cells to support higher efficiency of nuclear transfer than control cells) were independently repeated to validate whether initially promising outcomes are reproducible.

Unfortunately, none of the sperm factors tested, alone or in combinations, reproducibly increased the efficiency of the nuclear transfer. Even those factors which initially gave a promising outcome (Brdt and the mixture of Sp1, Sp4, Sp5 and H1fx) did not show reproducible effects in the following experiments (Table 5).

Table 5. Summary of nuclear transfer experiments using cells overexpressing candidate sperm factors.

Factor/combination of factors tested	Number of embryos tested: total number of swimming tadpoles obtained (ST)/cleaved embryos (CE)	
	Control cells (% ST/CE)	Overexpressing cells (% ST/CE)
Brdt – experiment 1	2/42 (4.8%)	3/46 (6.5%)
Brdt – experiment 2	10/36 (27.8%)	8/31 (25.8%)
Brdt – experiment 3	5/48 (10.4%)	1/49 (2.0%)
Brdt – experiment 4	7/61 (11.5%)	1/30 (3.3%)
Brdt – total from 4 experiments	*24/187 (12.9%)*	*13/156 (8.3%)*
Mlf1	6/35 (17.1%)	5/42 (11.9%)
H1fx	2/18 (11.1%)	2/19 (10.5%)
Sp1, Sp4, Sp5	3/31 (9.7%)	2/28 (7.1%)
Sp1, Sp4, Sp5 and H1fx – experiment 1	0/44 (0%)	4/44 (9.1%)
Sp1, Sp4, Sp5 and H1fx – experiment 2	1/20 (5%)	6/27 (22.2%)
Sp1, Sp4, Sp5 and H1fx – experiment 3	2/21 (9.5%)	1/22 (4.5%)
Sp1, Sp4, Sp5 and H1fx – experiment 4	3/26 (11.5%)	2/38 (5.3%)
Sp1, Sp4, Sp5 and H1fx – total from 4 experiments	*6/111 (5.4%)*	*13/131 (9.9%)*

5.3. Functional assessment of egg factors preferentially associating with the sperm chromatin

5.3.1. Experimental design

In order to test whether egg factors preferentially associating with the sperm chromatin have a function in early development, one would ideally remove such factors from the early embryos. The fact that the egg is a much more accessible cell for any type of manipulations than sperm makes it possible to try to downregulate the selected factors. The best characterised way of downregulating proteins in *Xenopus laevis* oocytes is by the injection of antisense deoxy-oligonucleotides (Hulstrand *et al.* 2010). In this approach, oligonucleotides complimentary to mRNA encoding the protein of interest are designed and injected into the GV stage oocyte. Such oligonucleotides form DNA-RNA heteroduplexes with the target mRNA, which are recognised and cleaved by endogenous oocyte-derived RNase-H activity. Cleaved mRNAs are subsequently degraded by oocyte-derived exonucleases. If the protein of interest is sufficiently unstable, then downregulation of the mRNA can lead to the reduction of the desired protein level (Fig. 24A). Such oocytes in which the protein is downregulated can be subsequently *in vitro* matured to eggs and injected with sperm (in ICSI procedure), which allows the assessment of the effects of the protein downregulation on embryonic development (Fig. 24B). Here I aimed to use this approach to assess the developmental function of egg-derived candidate factors binding specifically to the sperm chromatin after the egg extract treatment.

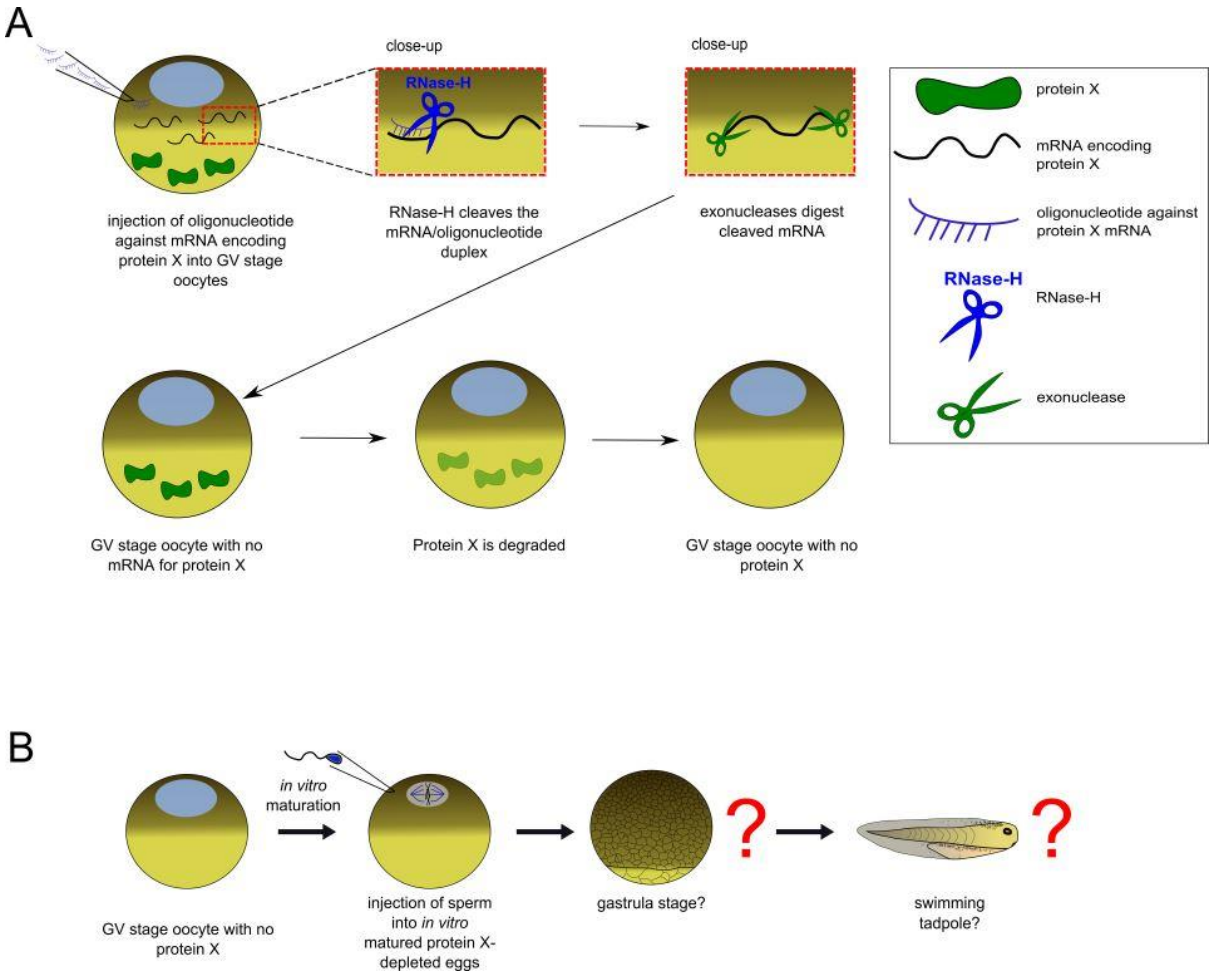


Fig. 24. Diagram explaining oligonucleotide-mediated knockdown of proteins in *Xenopus laevis* oocytes.

(A) Antisense deoxy-oligonucleotides against mRNA encoding the protein X are injected into the GV stage oocyte. These oligonucleotides form DNA-RNA heteroduplexes with the mRNA in the region of the base pair complementarity. DNA-RNA heteroduplexes are recognised and cleaved by endogenous, oocyte-derived RNase-H. The cleaved mRNA is then degraded by endogenous exonucleases. In the absence of mRNA, protein X becomes degraded. (B) GV stage oocytes depleted for protein X can be *in vitro* matured into eggs and injected with sperm (by ICSI procedure) to generate embryos depleted for the protein X. The effect of protein X depletion on embryonic development can be subsequently assessed.

5.3.2. Validation of the antibodies

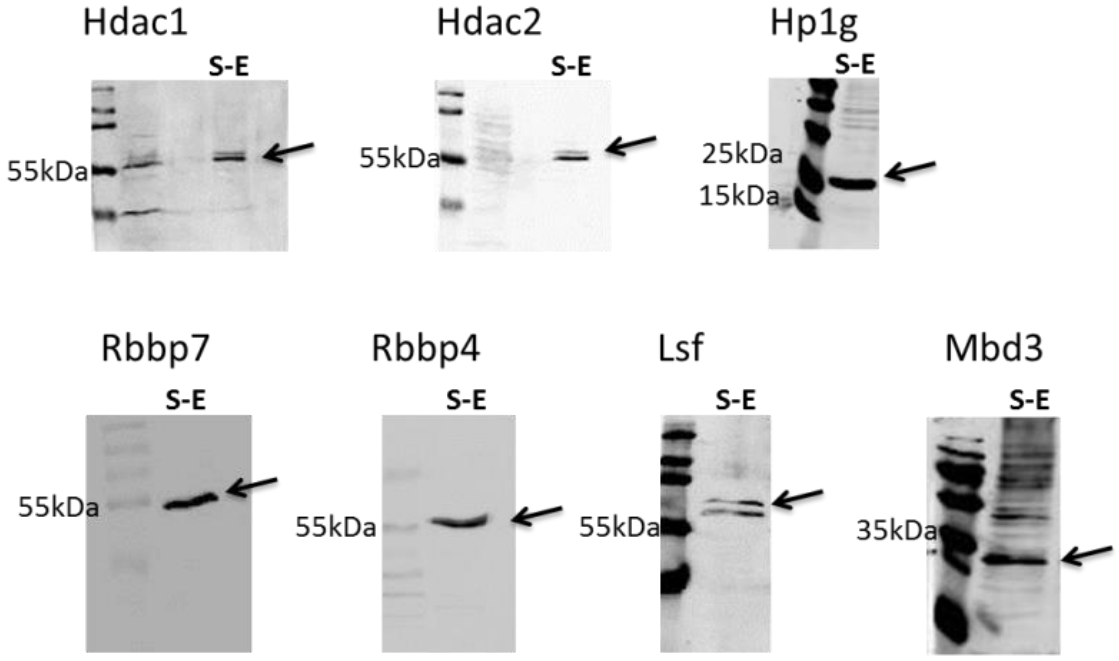
I have first tested whether the proteins I want to downregulate (Hdac1, Hdac2, Rbbp4, Rbbp7, Mbd3, Hp1γ and Lsf) are detectable with commercially available antibodies. To do this I have treated sperm with egg extracts to test whether bands of correct sizes are detected. All antibodies recognised bands of approximately the expected size (Fig. 25), therefore validating their use for assessing the knockdown efficiency of the oligonucleotides. Next, I tested whether the proteins I want to downregulate are also detectable by immunoblotting in the oocytes, as the knockdown itself is performed in the oocytes. All proteins apart from Mbd3 and Lsf were detected in the oocyte lysates. Due to the fact that the oocyte is pre-loaded with a lot of proteins, it is difficult to load more than one oocyte per a gel lane. The ability to detect the protein in the oocyte is crucial for the assessment of knockdown efficiency. Therefore, I tested whether I could detect these proteins if instead of the whole oocyte lysate I would use germinal vesicles (nuclei) isolated from the oocytes. The advantage of using the nuclei is that one can load many nuclei per one lane and in this way focus the analysis on nuclear proteins. I have therefore tried to detect Mbd3 and Lsf proteins in lysates from 20 germinal vesicles. Unfortunately, even the use of this approach did not allow me to detect the proteins of the correct size (Fig. 26). The anti-Mbd3 antibody recognised a band of around 45kDa instead of 33kDa. This antibody was reported to also recognise Mbd2, which was reported to migrate at 45kDa (Zhu *et al.* 2011), therefore the observed band could be Mbd2 and not Mbd3. Alternatively; the 45kDa band could be a post-translationally modified form of Mbd3. For Lsf protein a band of around 35kDa was detected instead of expected 57kDa. This again could be the effect of post-translational clipping of the protein in the oocyte and not in the egg, or of unspecific recognition of another protein by anti-Lsf antibody. Due to the uncertainty about the detection of Mbd3 and Lsf proteins in the oocyte, which could impede the validation of the knockdown effect, I have decided to exclude these

two proteins from the list of candidate egg factors to be downregulated and for the further steps I focused on the downregulation of the five remaining proteins: Hdac1, Hdac2, Hp1γ, Rbbp4 and Rbbp7.

5.3.3. Downregulation of the mRNA encoding the selected factors

I next tested whether deoxy-oligonucleotides injected into the oocytes can degrade the mRNAs encoding the target proteins. I have designed antisense deoxy-oligonucleotides complimentary to Hdac1, Hdac2, Hp1γ, Rbbp4 and Rbbp7, following the guidelines described before (Hulstrand *et al.* 2010). Three different oligonucleotides targeting each of the mRNAs were designed, whereas scrambled oligonucleotides were designed for control experiments. Oligonucleotides were injected into GV stage oocytes. Injected oocytes were collected 48h after the oligonucleotide injection and processed for reverse-transcription quantitative PCR analysis (RT-qPCR). Amount of transcripts present in the oligonucleotide-injected oocytes were normalised to the amount of transcript present in the control, scrambled-oligonucleotide injected oocytes. qPCR results showed that all the oligonucleotides allowed a significant downregulation of the target mRNAs (Fig. 27). Subsequently, the most efficient oligonucleotides at degrading the mRNAs (oligonucleotide 3 for Hdac1, Hdac2 and Rbbp7, oligonucleotide 1 for Hp1γ and oligonucleotide 2 for Rbbp4), were chosen to assay the degradation of the target proteins.

A



B

Protein	Expected molecular weight (kDa)
Hdac1	55
Hdac2	55
Hp1g	20
Rbbp7	48
Rbbp4	48
Lsf	57
Mbd3	33

Fig. 25. Validation of the antibodies on sperm-extract treated samples.

(A) Immunoblotting for Hdac1, Hdac2, Hp1γ (Hp1g), Rbbp7, Rbbp4, Lsf and Mbd3. 'S-E' indicates the position of sperm-extract treated sample and the arrow indicates the protein band of interest. (B) Table presenting the expected molecular weights of the proteins tested. Note that the protein bands detected in (A) are of similar weight to the expected molecular weight (B), suggesting that the antibodies recognised the correct proteins.

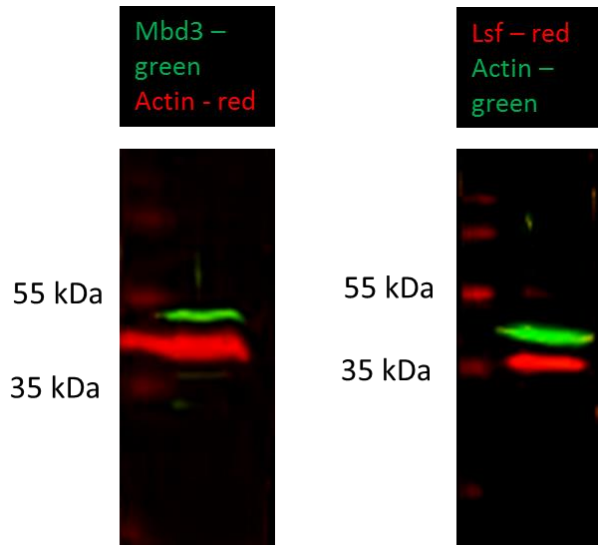


Fig. 26. Mbd3 and Lsf proteins are not detected in the oocyte.

Proteins isolated from 20 germinal vesicles (oocyte nuclei) were loaded on each gel. Subsequently gels were stained for Mbd3 (left panel, green) and for Lsf (right panel, red). Detected proteins were not migrating at the expected molecular weights (observed before in the lysate from egg extract-treated sperm chromatin): Mbd3 should be detected at 33kDa (not at 45kDa) and the expected molecular weight of Lsf is 57kDa (not 35kDa).

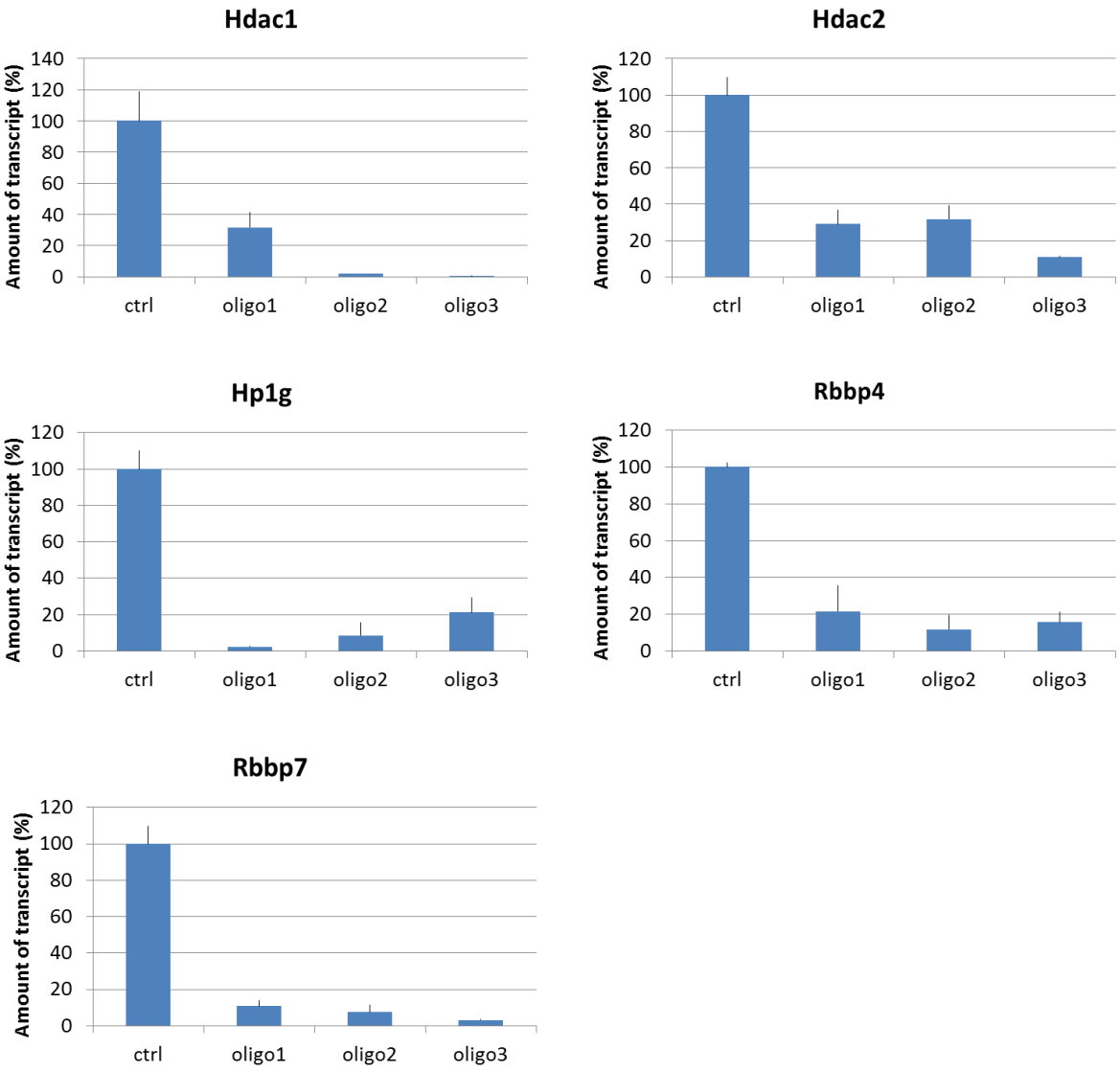


Fig. 27. Assessment of mRNA degradation by oligonucleotide-mediated knockdown.

Oocytes were injected either with scrambled oligonucleotides (control) or with three different oligonucleotides (oligo1, oligo2, oligo3) designed to target the mRNA of interest (name of the target mRNA indicated above the graph). All samples were normalised against a housekeeping mRNA *pwp1* and the amount of mRNA in the control sample was set to 100% (the other samples were normalised accordingly). All oligonucleotides tested led to a significant downregulation of the mRNA levels ($p < 0.05$, t-tests). Error bars show ±SEM.

5.3.4. Protein levels of the candidate egg factors are not reproducibly downregulated by the antisense oligonucleotides

I have subsequently tested whether the injection of the most efficient oligonucleotide affects the target protein levels in the oocyte. For each candidate factor oocytes were injected either with the oligonucleotide against the mRNA or with the control scrambled oligonucleotide. Oocytes were collected for protein extraction and immunoblotting analysis 48h, 96h and 144h after the injection. Only the injection of the oligonucleotide against Hp1γ led to a downregulation of the protein level in the oocyte, whereas the expression level of other proteins tested was not affected (Fig. 28A). I have therefore repeated the experiments with the oligonucleotide against Hp1γ. Unfortunately; the initial downregulation of the protein level observed after injection of the oligonucleotide 1 was not reproduced in independently repeated experiments. One of the reasons for that could be a sequence polymorphism between different frogs (animals in our frog colony are not from an inbred line and can therefore have slightly different DNA sequence which could prevent targeting by the oligonucleotide). I have therefore tried to inject the two other oligonucleotides against Hp1γ: oligonucleotide 2 and 3; which however did not result in a protein knockdown (Fig. 28B). I have then also tried a higher dose of the oligonucleotide 1: 2 times more than initially, the same amount and 2 times less; however, none of the concentrations led to the protein downregulation (Fig. 28C). The lack of knockdown could have been also caused by the fact that the oocytes used in the initial experiments were fully grown and therefore already ceased RNA transcription, whereas the oocytes used for the repeat experiments were still transcribing mRNAs and therefore the knockdown was not efficient. Alternatively, oligonucleotides could have degraded during the storage period and could have become less efficient at the mRNA degradation. To shed light on the discrepancy in the results obtained, I have treated non-injected oocytes with cycloheximide, which inhibits protein translation.

Treatment with cycloheximide therefore allows monitoring of the half-life of proteins. Immunoblotting of proteins from cycloheximide-treated oocytes revealed that the level of Hp1γ protein was not reduced during 48h of cycloheximide treatment (Fig. 28D) (longer treatment with cycloheximide is toxic to the oocytes and therefore cannot be applied). This result is contradictory to the initial result that demonstrated a knockdown of Hp1γ already after 48h from the oligonucleotide injection. The reason for the discrepancy between the results remains therefore unclear; however, since all the follow-up experiments suggest no knockdown of Hp1γ, it is unlikely that the oligonucleotide-mediated route can lead to a reproducible depletion of this protein.

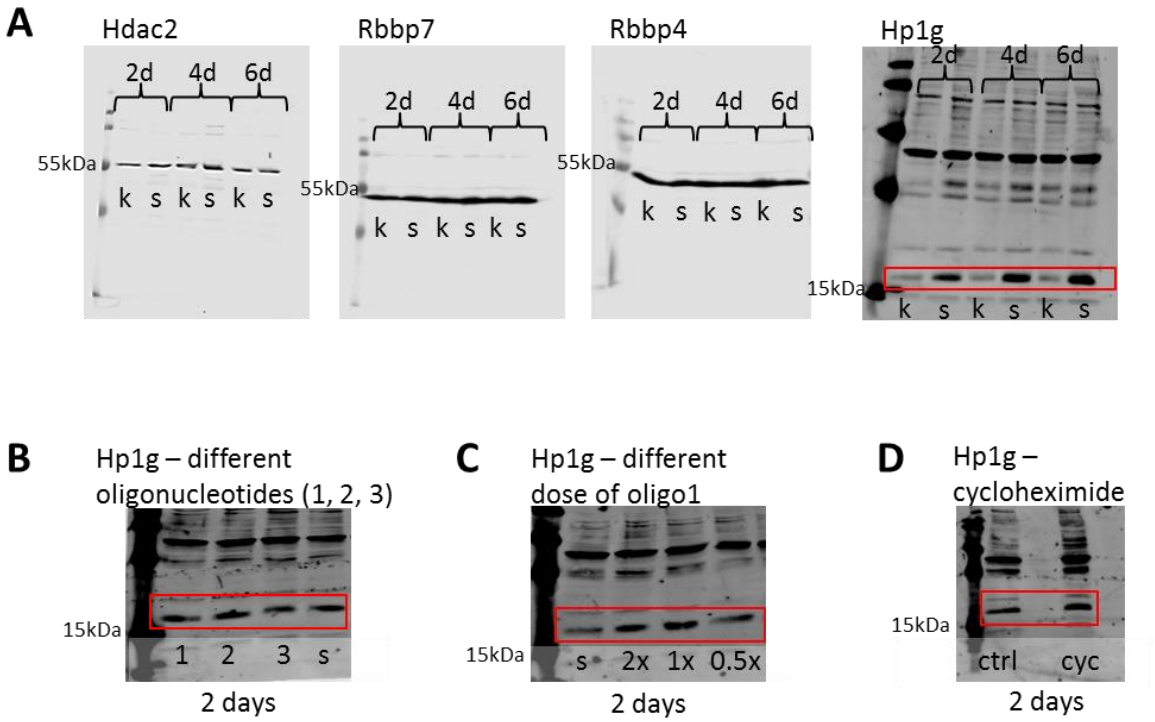


Fig. 28. Assessment of protein degradation by oligonucleotide-mediated knockdown.

(A) Oocytes were injected either with scrambled oligonucleotides (s) or with oligonucleotides against the mRNA of interest to achieve a protein knockdown (k) (indicated below the lanes). Oocytes were collected 48h (2d), 96h (4d) or 144h (6d) after injection for immunoblotting analysis. The protein detected on each blot is indicated above the blot. Red rectangle in the Hp1γ blot indicates the position of the Hp1γ protein band. (B) Assessment of Hp1γ level upon injection of various oligonucleotides. Oocytes were injected with three different oligonucleotides against Hp1γ (lanes 1, 2, 3) or with scrambled oligonucleotides (s). Samples were collected 48h after injection. Red rectangle indicates the position of the Hp1γ protein band. (C) Assessment of Hp1γ level upon injection of various doses of oligonucleotide 1. Oocytes were injected with three different doses of oligonucleotide 1 against Hp1γ mRNA (2x, 1x and 0.5x times of the amount injected in panel A) or with scrambled oligonucleotides (s). Samples were collected 48h after injection. Red rectangle indicates the position of the Hp1γ protein band. (D). Assessment of Hp1γ level upon cycloheximide treatment of oocytes. Cycloheximide-treated oocytes are labelled with 'cyc' below the lane, whereas the control oocytes, are labelled 'ctrl'. Samples were collected 48h after the addition of cycloheximide to the media. Red rectangle indicates the position of the Hp1γ protein band.

5.4. Summary and discussion

In this section I described the effect of the overexpression of candidate sperm factors on the efficiency of nuclear transfer and my attempts to downregulate the egg factors that bind specifically to the sperm chromatin upon egg extract treatment.

In the first part I have tested the effect of several candidate sperm factors and several combinations of sperm factors on the efficiency of nuclear transfer. However, none of the factors/combinations of factors led to a reproducible improvement of the nuclear transfer efficiency. There could be several explanations for that, for example: wrong factors were selected, essential co-factors were not co-expressed with the candidate factors, over-expressed factors were not functional or they were expressed at inappropriate levels. Also, it is likely that introduction of sperm factors could have rescued some aspects of the development, but not all. In order to address this in the future a more detailed molecular analysis of the nuclear transfer embryos (for example, qPCR analysis for candidate gene expression) would be needed.

I have also attempted to downregulate several candidate factors which were identified as binding specifically to the sperm chromatin upon egg extract treatment. Knockdown was performed with the use of oligonucleotides antisense to mRNA encoding the protein of interest. All the oligonucleotides tested led to a significant downregulation of mRNA targets, as evidenced by a qPCR. However, none of the candidates tested was reproducibly downregulated at the protein level. This suggests that even though mRNA levels decreased, the target proteins were stable and did not degrade. In the future it would be worth performing a simple test with cycloheximide treatment before choosing the candidate

proteins to be downregulated. In this way one could eliminate those proteins, whose levels are unlikely to be affected by the oligonucleotide-mediated knockdown.

Chapter 6

Characterisation of the developmental defects of spermatid-derived embryos

Replication assays described in this chapter were performed in collaboration with Dr Vincent Gaggioli. Preparation of libraries and sequencing of two RNA-seq samples was done in collaboration with Dr Taejoon Kwon and Dr Edward Marcotte. Bioinformatic analyses described in this chapter were performed by Dr Angela Simeone, Dr George Allen and Dr Charles Bradshaw.

6.1. Introduction

Comparison of the developmental potential of sperm- and spermatid-derived embryos revealed that sperm-derived embryos develop significantly better than spermatid-derived embryos to a swimming tadpole stage. Since the mass spectrometry analysis and further functional testing of sperm proteins and of sperm-binding factors failed to unravel the source of the developmental advantage of sperm, I modified my strategy. Instead of directly looking for potential factors conferring developmental benefits to sperm, I have decided to first characterise the developmental defects of spermatid-derived embryos. Understanding what processes occur abnormally in spermatid-derived embryos could help to identify in what respects sperm is better at supporting embryonic development.

Early *Xenopus* embryo development starts with rapid cell divisions (fast DNA replication cycles) in the absence of transcription from the zygotic genome. Only around the 12th cell cycle division, at the time called mid-blastula transition (MBT), cell divisions slow down and the zygotic genome activation occurs (Newport & Kirschner 1982, Kimelman *et al.* 1987) (Fig. 29). I have therefore tested whether these major events during embryonic development: DNA replication accompanying cell divisions and gene transcription needed for further embryonic development occur normally in spermatid-derived embryos. My results demonstrate that spermatids are equally good as sperm at supporting DNA replication. Furthermore, I showed that developmental failure of spermatid-derived embryos is not due to RNA carried over to the embryo. I also showed that zygotic rRNA transcription is initiated normally in spermatid-derived embryos. Interestingly, RNA sequencing (RNA-seq) analysis of spermatid- and of sperm-derived embryos revealed that one hundred developmentally-important mRNAs are misregulated in spermatid-derived embryos, which can reflect developmental programming of sperm to support correct regulation of gene expression in the embryo.

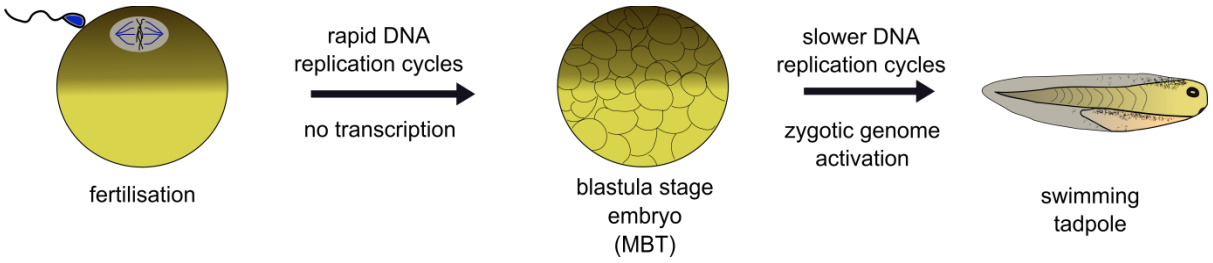


Fig. 29. Rapid cell divisions precede zygotic genome activation in *Xenopus*.

In the very early stages of embryonic development in *Xenopus* rapid cell divisions (fast DNA replication cycles) occur in the absence of zygotic transcription. Only when the embryo reaches the mid-blastula transition stage, the zygotic genome is activated, which is concurrent with slowing down of cell divisions.

6.2. Spermatids replicate their DNA as efficiently as sperm

All replication studies described below were performed with great help and supervision from Dr Vincent Gaggioli from the Gurdon Institute, at the University of Cambridge, UK.

Cell divisions occurred normally in spermatid-derived embryos (Fig. 8); however, it is not known whether accompanying DNA replication was also normal. Experiments in which sperm and somatic cells (erythrocytes) were incubated in *Xenopus* egg extracts showed that the sperm chromatin was better suited for supporting DNA replication that the chromatin of somatic cells (Lemaitre *et al.* 2005). DNA replication starts from the origins of replication and spreads in both directions from the origins (Fig. 30A). The first cell cycle in *Xenopus* lasts for only about 2h (at 18ºC). During this 2h the sperm needs to remodel its chromatin (exchange the protamines for histones) and initiate and timely complete DNA replication before the onsets of the cell division. It was shown that DNA replication in the sperm chromatin is initiated from multiple origins of replication spaced on average every 23.4kb of DNA. This was in sharp contrast to what was seen in erythrocyte chromatin in which the origins of replication were much sparser and positioned on average every 120.9kb of DNA (Lemaitre *et al.* 2005). This could prevent the erythrocytes from completing the replication before the onset of the cell division (Fig. 30B). Indeed, it was shown that after 2h of egg extract treatment, erythrocytes replicated less than 10% of DNA, whereas 80% of sperm DNA was replicated in this time (Lemaitre *et al.* 2005). This led to the hypothesis that inefficient replication is a roadblock to embryo development after somatic cell nuclear transfer in frogs (Laskey 2005). It was also shown that pre-treatment of mammalian donor

cells with *Xenopus* egg extracts increased the efficiency of mouse nuclear transfer and that it was correlated with an increased replication efficiency (Ganier *et al.* 2011). I have therefore hypothesised that the developmental failure of spermatid-derived embryos might be related to inefficient DNA replication. I have therefore decided to test the ability of sperm and spermatids to undergo replication with the use of *in vitro Xenopus* egg extracts.

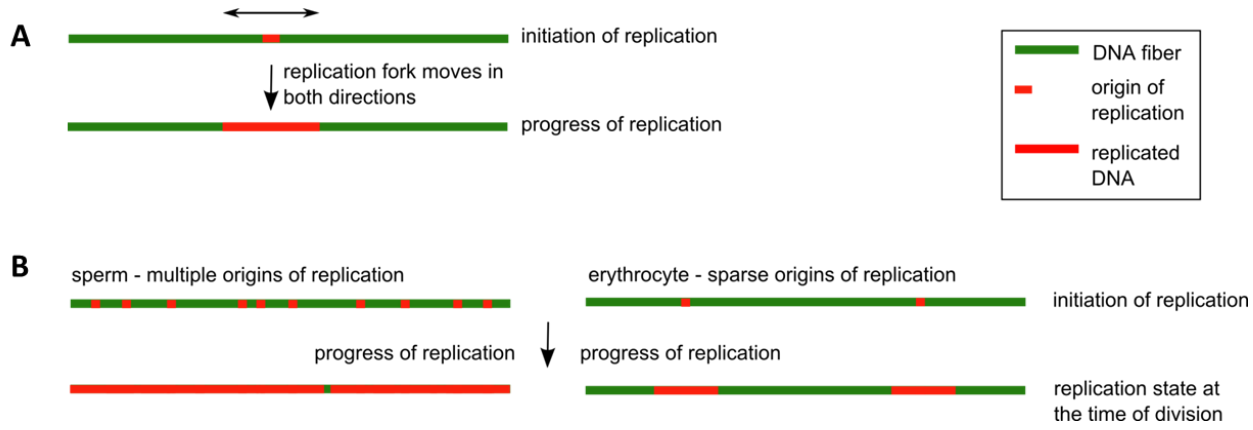


Fig. 30. Multiple origins of replication allow timely finishing of DNA replication in egg extracts.

Figure based on results from (Lemaitre *et al.* 2005). (A) DNA replication occurs in both directions from the origin of replication. (B) Sperm has multiple origins of replication, densely positioned on the DNA fiber (left panel), whereas the erythrocyte has sparse origins of replication (right panel). Multiple origins of replication ensure timely finishing of replication after the egg extract treatment in sperm, but not in erythrocytes.

To be able to assess the replication efficiency and at the same time to look into the initiation of replication I decided to use the technique of molecular combing. Help with this technique was kindly provided by Dr Vincent Gaggioli. I have prepared extracts from activated eggs, which are able to replicate DNA (Blow & Laskey 1986). Permeabilised sperm or spermatids nuclei were incubated in extracts at a concentration of 200 nuclei per a microliter of extract. Extracts were supplemented with biotin-dUTP, which is a DNA precursor incorporated into DNA upon replication (Fig. 31A). Newly synthesised (replicated)

DNA can be then revealed by biotin detection with fluorescently tagged streptavidin. In molecular combing procedure DNA fibers are isolated, spread on a microscope slide and then stained for DNA, to visualise all DNA fibers, and for biotin, to reveal those fibers/fragments of fibers that were replicated (Fig. 31B).

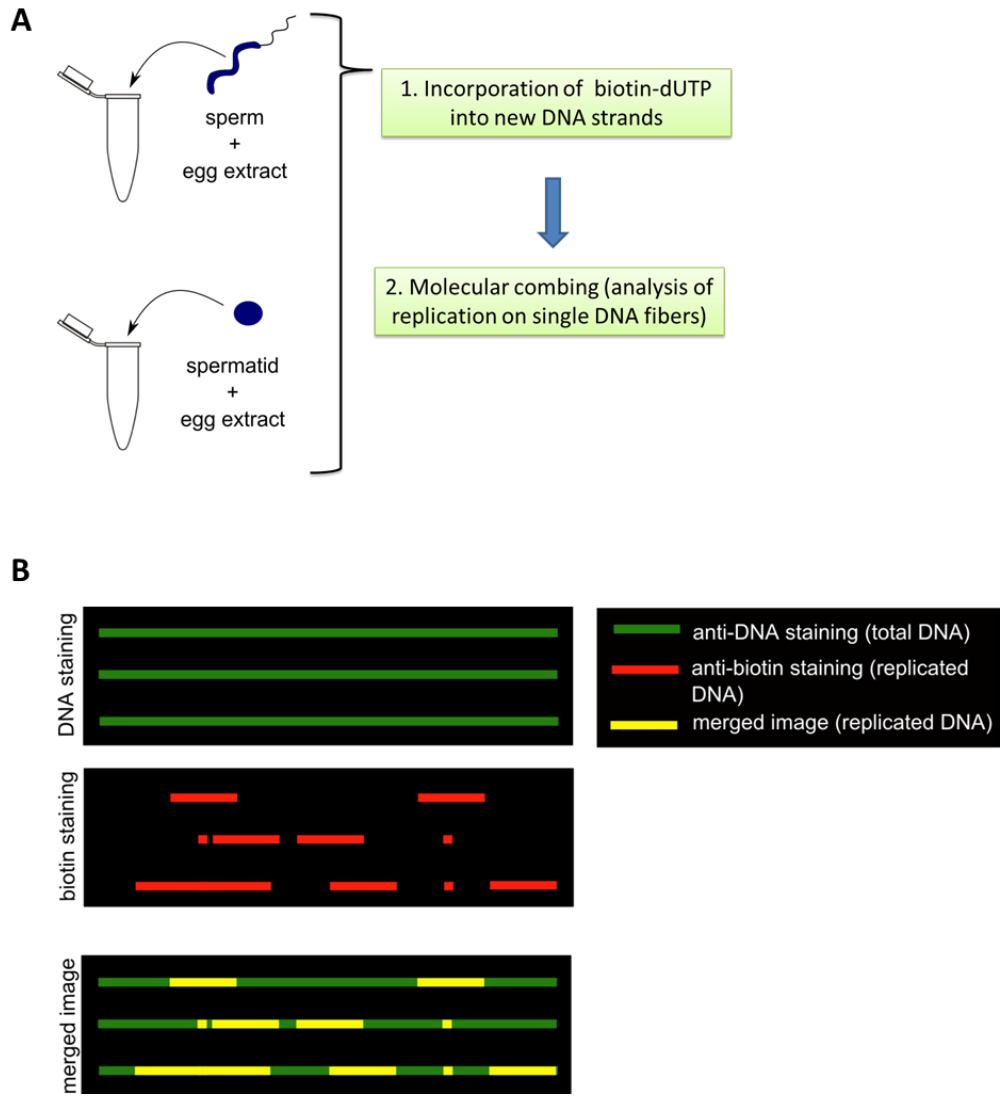


Fig. 31. Diagram explaining the experimental design for the replication assessment in sperm and spermatids.

(A) Permeabilised sperm or spermatids are incubated in egg extracts supplemented with biotin-dUTP. Subsequently, replicated DNA is revealed by molecular combing. (B) Diagram explaining the results of molecular combing. DNA fibers are stretched on the microscope slide and revealed with anti-DNA staining (shown in green). Replicated DNA is revealed with anti-biotin staining (shown in red). Replicated DNA in the merged image is shown in yellow.

The initial experiment was designed to get an idea as to whether the cells can replicate and how they initiate replication (what is the spacing of the origins of replication in each cell type). For that I have incubated sperm or spermatids in freshly prepared egg extracts, supplemented with biotin-dUTP, for short time periods: 30mins and 40mins. Then the DNA fibers were isolated and subjected to molecular combing and immunostaining. Interestingly, immunostaining results revealed that both samples were able to initiate replication with very similar spacing of origins of replication; however, initiation of replication in sperm was approximately 10mins delayed as compared with spermatids. At 30mins from the start of extract treatment some fibers in spermatids already started to initiate replication (Fig. 32). At the same time the replication was not yet detectable in sperm. At 40mins most of the fibers in spermatids were replicating, whereas only some fibers in the sperm sample initiated replication (Fig. 32). This suggests that sperm is not better than spermatids at initiating DNA replication. If anything, the sperm is delayed as compared with spermatids, which is likely due to the fact that it has to remodel its chromatin (replace protamines for egg-derived histones) before the onset of replication.

No major differences between sperm and spermatids were found at the initiation of replication, which however does not mean that both cell types are capable of timely completing the replication. It could be that the speed of the replication fork is different in both cell types (for example due to differences in the chromatin structure, which could interfere with the replication progress). If that was the case it could be that even though the cells started the replication similarly, they would not be able to complete replication at the same time. I have therefore incubated sperm or spermatids in egg extracts, supplemented with biotin-dUTP) for 2h – the time which is equivalent to the length of the first embryonic cell cycle. Also, not to miss any potential differences between the samples, this time I have precisely quantified the extent of replication. Quantification of the replication extent was

performed by measuring with ImageJ the length of replicated DNA in each fiber (staining anti-biotin) to the total length of the fiber (staining against DNA). Measurements were performed on at least 125 independent DNA fibres (22000kb of DNA for each sample) (Fig. 33A). Data acquired from all the measurements were then exported to Microsoft Excel for calculations with the use of a macro created and shared by Dr Vincent Gaggioli. At 120mins from the start of the egg extract treatment both cell types replicated more than 80% of the total DNA length. There was no difference in the replication extent between sperm and spermatid fibers, suggesting that replication problems are unlikely to be the explanation of the developmental defects of spermatid-derived embryos (Fig. 33B).

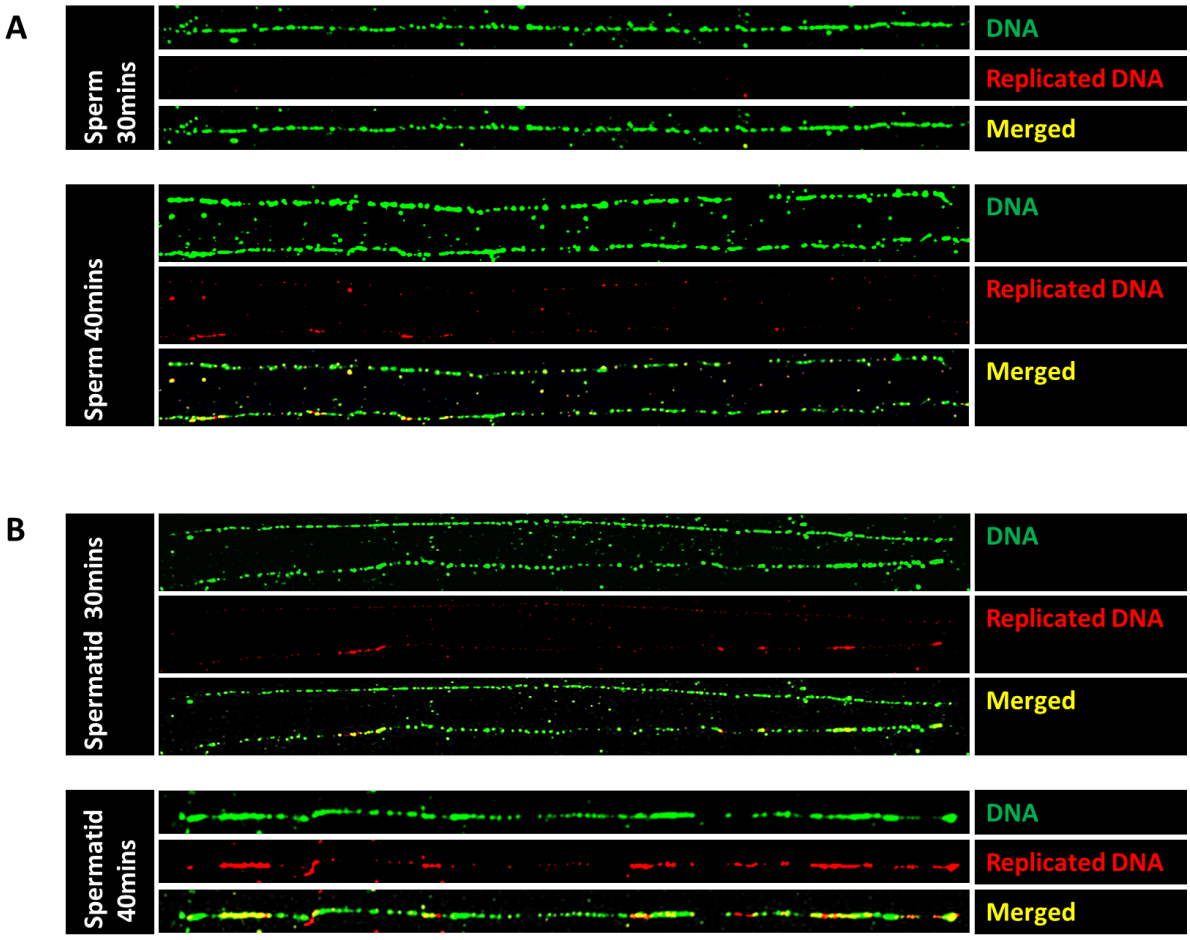


Fig. 32. Spermatids initiate replication earlier than sperm and show similar spacing of origin of replication to that of sperm.

Sperm or spermatids were incubated in egg extracts supplemented with biotin-dUTP for 30 or for 40mins. Antibody staining against DNA reveals the total length of the fiber (green) and antibody staining against biotin reveals the replicated DNA (red). Replicated regions of the fibers are yellow in the merged images. (A) Examples of fibers isolated from the sperm sample. (B) Examples of fibers isolated from the spermatid sample. Note that spermatids initiate replication earlier than sperm (30mins for spermatids and 40mins for sperm) and that the spacing of origins is similar in both samples.

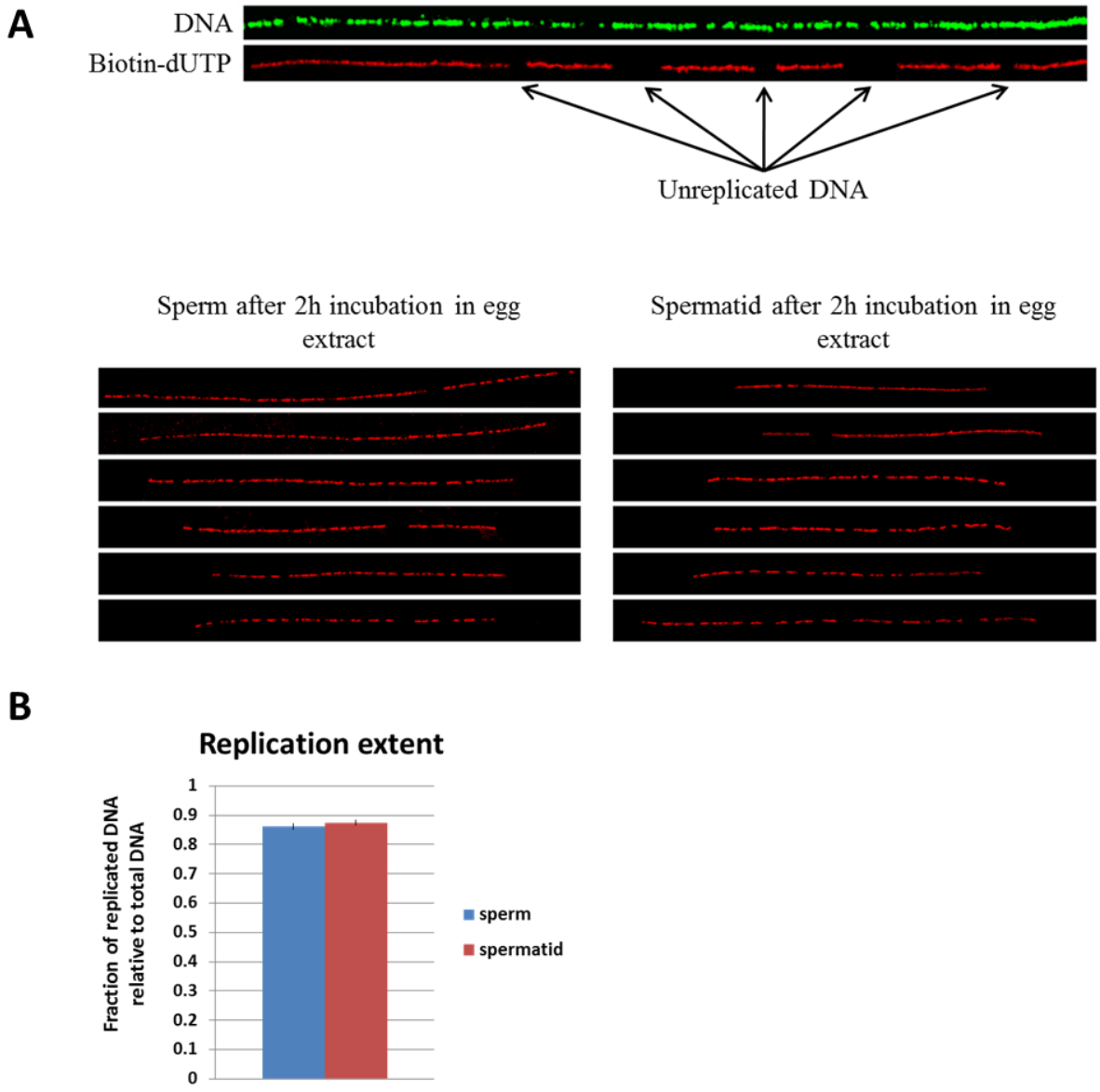


Fig. 33. Spermatids replicate DNA as efficiently as sperm

(A) Examples of DNA fibers after immunostaining procedure. Antibody staining against DNA reveals the total length of the fibre (green) and antibody staining against biotin reveals the replicated DNA (red). The bottom panels show representative examples of replication staining from sperm and from spermatids incubated in egg extracts. (B) Replication extent measured as the proportion of DNA that incorporated biotin-dUTP to the total fiber length. Results are from at least 125 independent DNA fibers (22000kb of DNA for each sample). Error bars show ±SEM. Samples were not significantly different (p-value = 0.37, t-test). *Panel 'A' of this figure was created and kindly shared by Dr Vincent Gaggioli.*

6.3. Spermatid-derived RNA is not deleterious for embryonic development

Since my results showed that problems with DNA replication are unlikely to be the cause of developmental defects of spermatid-derived embryos, I next investigated other possible reasons which could explain their defects. I first tested whether carried-over RNA can be a problem for development of spermatid-derived embryos. All the RNAs present in sperm are also present in spermatids; however, during spermiogenesis the maturing spermatozoon reduces its cytoplasmic contents together with the vast majority of RNAs. It has been estimated that mature sperm contains only about 10-100fg of RNA compared to 10-50pg of RNA typically found in a somatic cell (Pessot *et al.* 1989, Krawetz 2005), which is about 1000 times reduction in the RNA content. Since there is no transcription in the nucleus of the mature sperm, spermatids contain all RNAs which are required for sperm maturation, for example mRNAs encoding sperm basic proteins 1-6, which are the functional equivalents of mammalian protamines in *Xenopus laevis* (Abe & Hiyoshi 1991, Hiyoshi *et al.* 1991). Translation of these mRNAs into proteins and their further incorporation into chromatin is thought to be important for the acquisition of the highly specialised, almost crystalline structure of the sperm nucleus. One could imagine that delivery of all these spermiogenesis-specific RNAs by the spermatid to the egg at fertilisation could lead to the illegitimate translation of spermiogenesis-specific mRNAs in the embryo after fertilisation. Presence of such translated proteins could then interfere with the embryonic development, for example by altering the chromatin architecture of the spermatid-derived embryos. Interestingly, when mRNAs encoding the sperm-specific factors Sp4 or Sp5 were injected into 1-cell stage embryos at high concentrations (9.2ng of mRNA per embryo), embryos died around the gastrulation stage, suggesting that illegitimate expression of spermiogenesis-related proteins can indeed be deleterious for embryonic development (Fig. 34).

I have therefore tested whether the potential carried-over RNA from spermatids can have any effects on embryonic development. I have injected fertilised embryos (at 1-cell stage) with 50pg (corresponding to the maximum amount of RNA found in a typical somatic cell) of either total RNA isolated from testis (isolated with Trizol, therefore recovering all different RNA types) or with 50pg of the mRNA encoding the mixture of sperm basic proteins (Sp1, Sp4 and Sp5 – mRNAs which were shown to be toxic for embryonic development when injected at high doses). None of the injections had any detrimental effects on embryonic development (Fig. 35) suggesting that carried-over RNA from spermatids is not the cause of developmental failure of spermatid-derived embryos.

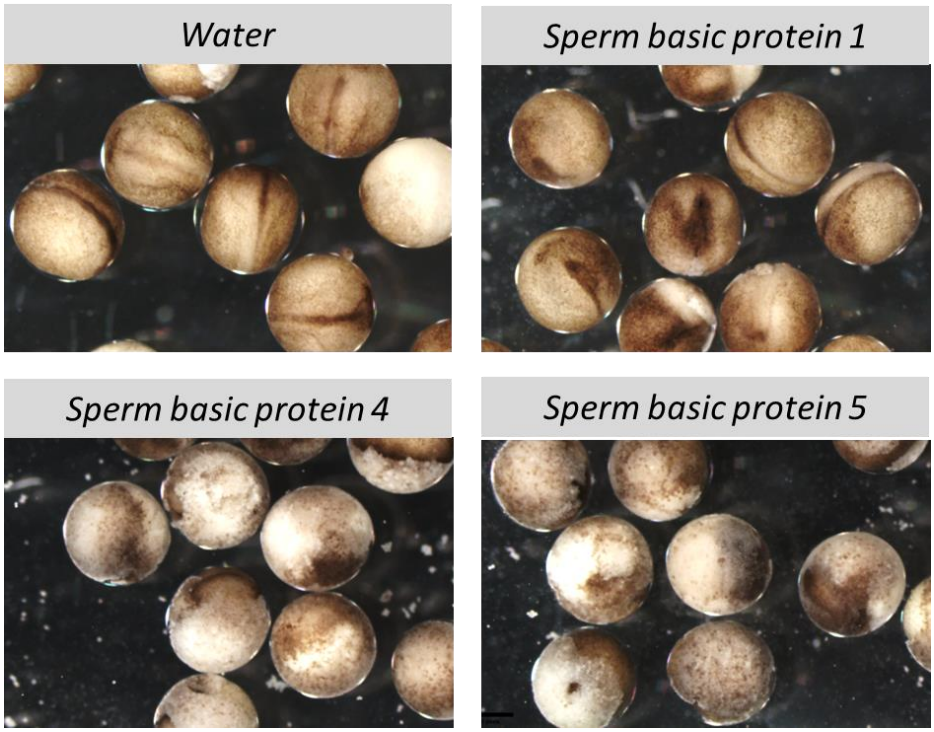


Fig. 34. Injection of high doses of mRNAs encoding spermiogenesis-related proteins is toxic for embryos

Embryos were injected at 1-cell stage with water or with 9.2ng of mRNAs encoding Sperm basic protein 1 (Sp1), Sperm basic protein 4 (Sp4) or Sperm basic protein 5 (Sp5). Injection of Sp4 or Sp5 was toxic to the embryos.

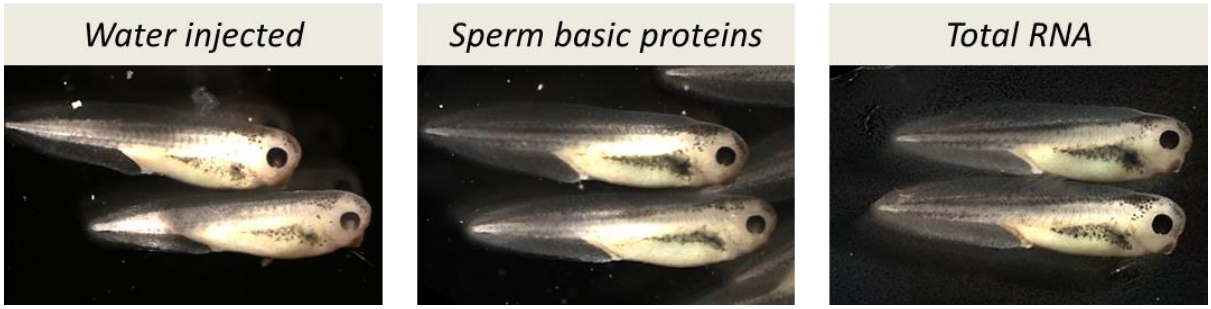


Fig. 35. RNA carry-over is not the cause of developmental defects of spermatid-derived embryos

Fertilised embryos (at 1-cell stage) were injected either with water, with 50pg of mRNAs encoding sperm basic proteins (Sp1, Sp4 and Sp5) or with 50pg of total testicular RNA ('Total RNA'). Injections did not affect the normality of embryonic development, since all embryos developed into normal swimming tadpoles.

6.4. Haploid paternal embryos as a tool for a specific assessment of transcription from the paternally-derived chromatin

Another major challenge that the embryo needs to accomplish to develop successfully is the zygotic genome activation (Newport & Kirschner 1982). Therefore, I decided to assess the ability of sperm- and spermatid-derived embryos to support embryonic transcription. After ICSI, embryonic transcription occurs from both the paternal and the maternal genome. In order to specifically assess the ability of the paternally-inherited chromatin to drive transcription, I decided to use haploid paternally-derived embryos. Haploid paternally-derived embryos are typically generated by first enucleating the egg with a Mineralite UV lamp treatment, followed by *in vitro* fertilisation. Developing embryos are haploid and their genetic material is inherited solely paternally (Gurdon 1960, Hamilton 1963) (Fig. 23). Since here I wanted to generate haploid sperm- or spermatid-derived embryos, I had to slightly modify this protocol to compensate for the fact that a spermatid cannot fertilise the egg: sperm or spermatids were injected into enucleated eggs (instead of performing the *in vitro* fertilisation) (Fig. 36A).

I have first tested whether haploid embryos recapitulate the developmental phenotypes of diploid sperm- and spermatid-derived embryos. I have allowed the generated haploid embryos to develop and scored them in the same way as diploid embryos were scored: as a number of gastrula embryos and as a number of swimming tadpoles to the total number of cleaved embryos. The results obtained with haploid embryos agreed with the findings obtained with diploid embryos – there was no significant difference in the embryo development to the gastrula stage, but sperm-derived embryos developed significantly better to the swimming tadpole stage compared to spermatid-derived embryos (p-value < 0.05) (Fig. 36B). This therefore validates the use of haploid embryos for the assessment of transcription

originating specifically from sperm- or spermatid-derived chromatin at the time of embryonic gene activation.

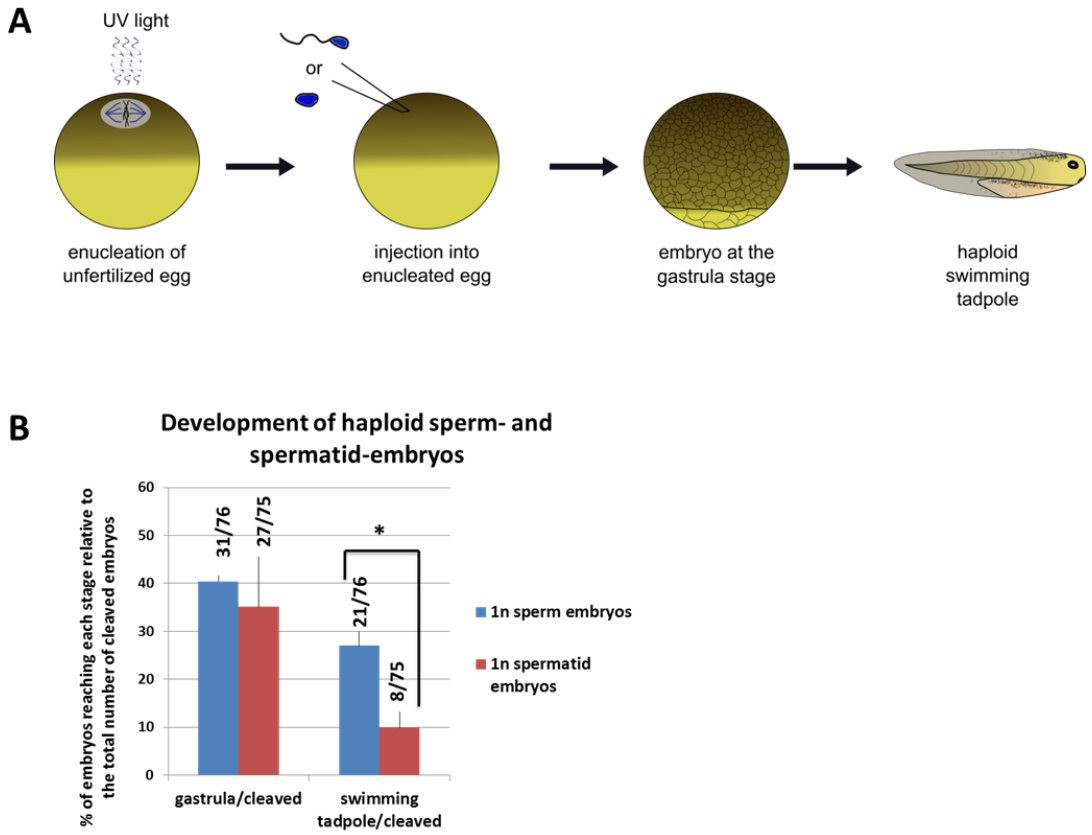


Fig. 36. Developmental advantage of sperm over spermatid is maintained in haploid embryos

(A) Diagram explaining ICSI into enucleated eggs. Eggs are first enucleated with the Mineralite UV lamp treatment and subsequently injected with sperm or spermatids. (B). Haploid sperm-derived embryos developed better than haploid spermatid-derived embryos. Embryos were scored as the % of embryos reaching a gastrula stage and a swimming tadpole stage to the total number of cleaved embryos. Numbers of embryos analysed are indicated above the bars. N = 3 independent experiments. Error bars show ± SEM. * indicates p-value = 0.008 (z-test).

6.5. rRNA synthesis occurs normally in spermatid-derived embryos

I next tested whether spermatid-derived embryos are equally suited to support embryonic transcription of rRNA as sperm-derived embryos. It was reported that mouse nuclear transfer-derived embryos aberrantly expressed rRNAs when compared to *in vitro* fertilised embryos (Suzuki *et al.* 2007) and that this correlated with their poor developmental outcomes (Zheng *et al.* 2012). I have therefore hypothesised that sperm is programmed to support efficient rRNA synthesis, whereas the spermatid is not. To test this hypothesis, enucleated eggs were injected with sperm or spermatids and with BrUTP. BrUTP was co-injected with sperm or spermatids in order to label only the newly synthesised transcripts (Core *et al.* 2008) and to distinguish them from rRNA maternally accumulated during oogenesis (Roger *et al.* 2002). This procedure allowed generation of haploid sperm- or spermatid-embryos which were collected at the gastrula stage. Subsequently, newly synthesised RNA was pulled down and quantified by reverse transcription quantitative PCR (RT-qPCR) for 18S and 28S rRNA (Fig. 37A). The results of the RT-qPCR analysis revealed that there are no significant differences in the amount of 18S or 28S rRNA synthesised between the sperm- and spermatid-derived haploid embryos (Fig. 37B). This suggests that problems with correct activation of embryonic rRNA synthesis are unlikely to be the cause of developmental defects of spermatid-derived embryos.

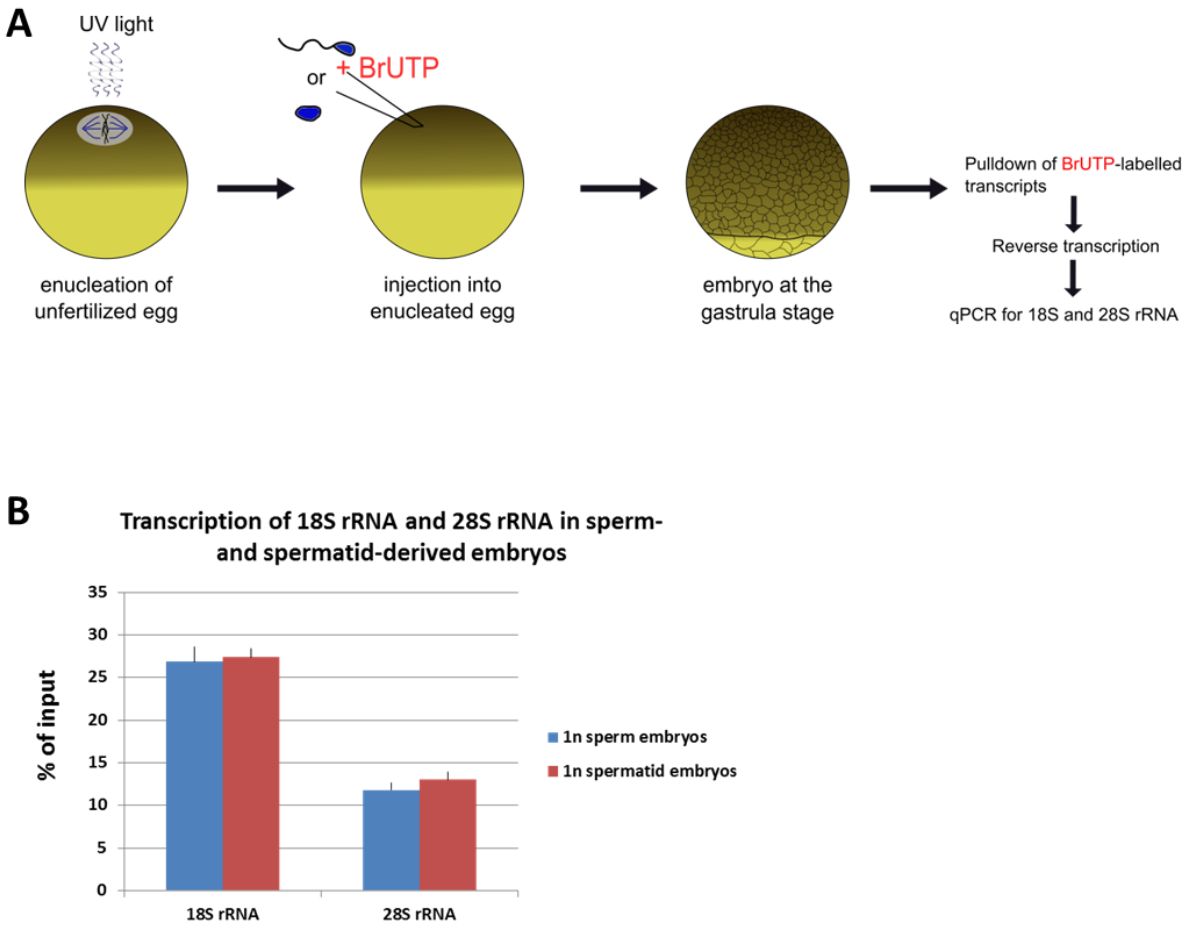


Figure 37. Spermatid-derived embryos are as good as sperm-derived embryos at synthesising rRNAs.

(A) Diagram explaining newly synthesised RNA isolation from sperm- and spermatid-embryos. Haploid sperm- and spermatid-embryos are obtained by ICSI to enucleated eggs and are co-injected with BrUTP to label newly synthesised RNA. Embryos develop to a gastrula stage when they are collected. Subsequently, BrUTP-labelled, newly synthesised RNA is pulled down. (B) Spermatid-derived embryos synthesised rRNA as efficiently as sperm-derived embryos, as evidenced by RT-qPCR quantification of 18S and 28S rRNAs. Values are shown as a percentage of pulled down RNA to the total input RNA. Error bars show ± SEM. N=20 sperm-derived embryos and N=14 spermatid-derived embryos. Samples were not significantly different (p-value = 0.82 for 18S rRNA and p-value = 0.36 for 28S rRNA, t-test).

6.6. Developmentally-important mRNAs are misexpressed in spermatid-derived embryos

Zygotic rRNA activation was not different between sperm- and spermatid-derived embryos, therefore I next hypothesised that spermatid-derived embryos may show aberrant mRNA transcription. In order to investigate the potential mRNA expression changes in a global way, I decided to perform RNA sequencing analysis of sperm- and spermatid-derived embryos. To focus my analysis on the transcription originating from the paternal chromatin, I again used haploid sperm- and spermatid-derived embryos. In order to eliminate technical variation between experiments and to facilitate the identification of the biologically meaningful differences, I collected the embryos (pools of 5 sperm- or spermatid-derived embryos) in seven independent experiments: experiments were conducted on different days, with eggs obtained from seven independent frogs and from three independent sperm and spermatid cell preparations. I generated haploid embryos as described above (Fig. 36A) and I collected them at the gastrula stage, before the onset of developmental defects. Subsequently, I isolated RNA from the embryos and generated sequencing libraries for five out of seven experimental replicates and sequenced them at the sequencing facility at the Cambridge Research Institute. Two remaining RNA samples were sent to our collaborators, Dr Taejoon Kwon and Dr Edward Marcotte at the University of Texas, USA, for independent library preparation and sequencing.

Bioinformatic analyses performed by Dr Angela Simeone, Dr Charles Bradshaw and Dr George Allen identified 255 out of 18,340 transcripts as abnormally expressed in spermatid-derived embryos (compared to sperm-derived embryos) with a false discovery rate (FDR) below 0.05 (Table S5). When applying more stringent filtering criteria (selecting only those transcripts which were consistently up- or down-regulated in at least 6 out of 7 separate

experiments), a final list of 100 transcripts differentially expressed in spermatid-derived embryos was obtained. From now on I refer to these 100 transcripts as 'misregulated' (Fig. 38A and Table S5). The majority of these misregulated transcripts (82 out of 100) were found to be upregulated, while only 18 out of 100 were downregulated in spermatid-derived embryos, as compared with sperm-derived embryos. The RNA-seq results were confirmed by RT-qPCR analysis (Fig. 38B).

In order to further characterise the misregulated genes, gene ontology (GO) enrichment analysis was performed by Dr George Allen. The analysis revealed that several developmentally-important terms are significantly enriched in the misregulated gene list (p-value < 0.05) (Fig. 39A). Indeed, more than 25% of the misregulated transcripts are known transcriptional regulators essential for embryonic development, for example *gata2*, *gata3*, *hes1* and *fos* (Zon *et al.* 1991, Kelley *et al.* 1994, Maeno *et al.* 1996, Kim *et al.* 1998, Read *et al.* 1998, Nardelli *et al.* 1999, Jouve *et al.* 2000, Friedle & Knochel 2002, Nakazaki *et al.* 2008, Lee *et al.* 2011). Furthermore, other misregulated transcripts, such as *bmp2*, *bmp7* or *dhh*, are morphogens with crucial roles in the induction of germ layers and cell signalling (Bitgood & McMahon 1995, Reversade & De Robertis 2005, Reversade *et al.* 2005, Wills *et al.* 2008).

6.7. mRNAs misexpressed in spermatid-derived embryos are Polycomb targets in human sperm

Interestingly, I noticed that the GO terms enriched in the misregulated set of transcripts (Fig. 39A) are very similar to those enriched for genes bearing trimethylated lysine 27 on histone H3 (H3K27me3) in the mature sperm in human (Brykczynska *et al.* 2010) (Fig. 39B). I therefore tested whether human orthologues of the *Xenopus* misregulated transcripts identified here may also be H3K27me3-modified in human sperm. Orthology search was performed by Dr Charles Bradshaw and the cross comparison of H3K27me3-modifed genes in human sperm to their *Xenopus* orthologues was performed by Dr Angela Simeone. It was found that among the *Xenopus* misregulated genes that have human orthologues, 41% were enriched for H3K27me3 in human sperm (Fig. 39C). This is a significant overrepresentation of H3K27me3-modified genes (p-value < 0.05), since of all human orthologues of *Xenopus* genes, only 16% are enriched for H3K27me3 in sperm (Fig. 39C and Table S6).

Concluding, 100 developmentally-important transcripts were identified as misregulated in spermatid-derived embryos, of which the majority of transcripts was upregulated. Furthermore, orthologues of these misregulated genes are enriched for the H3K27me3 mark in human sperm. This result supports the hypothesis that the nucleus of the sperm, but not of a spermatid, can be a subject of epigenetic programming to regulate the transcription of developmentally-important genes in the future embryo.

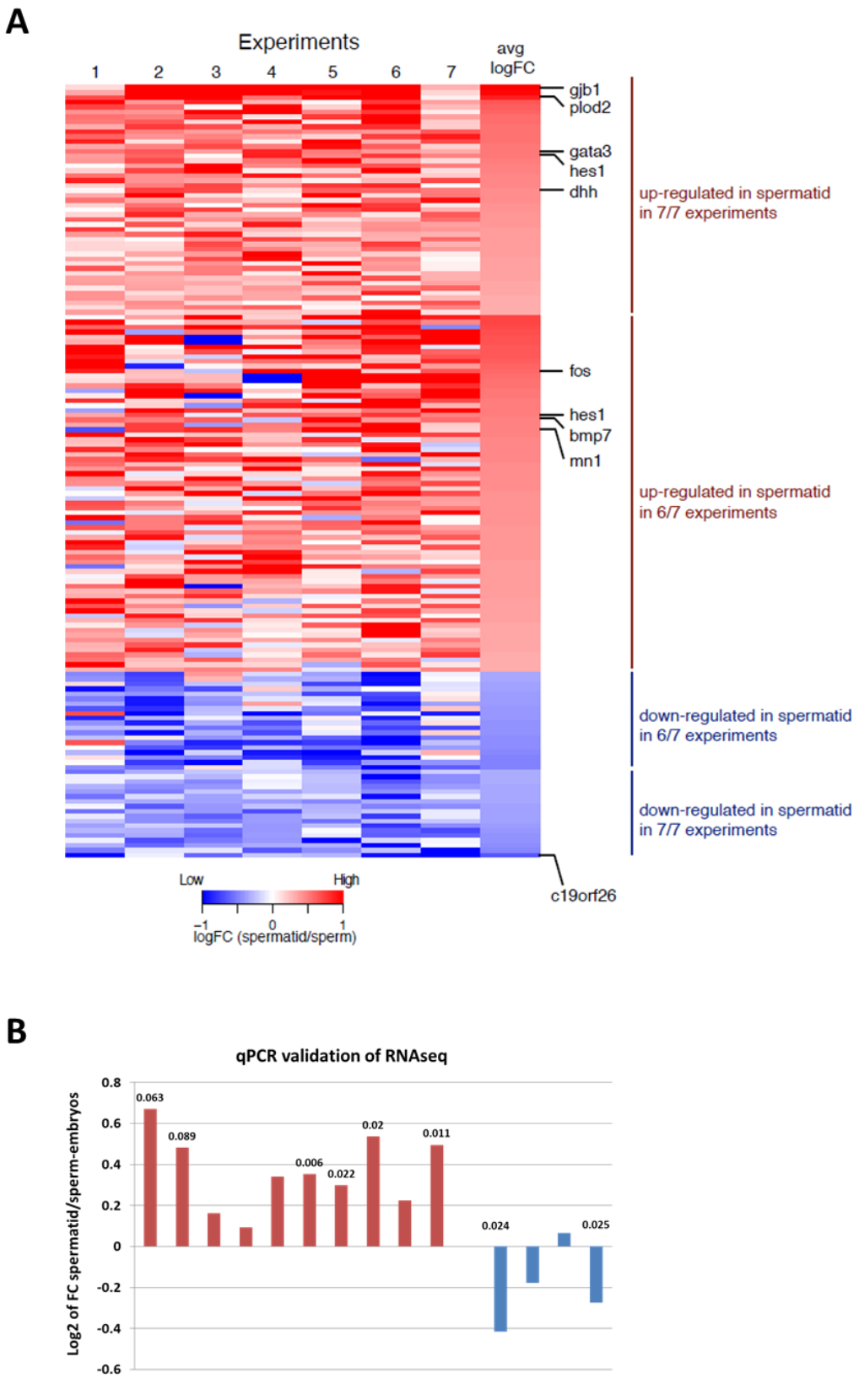


Fig. 38. Legend on the subsequent page

Fig. 38. Identification of 100 transcripts misregulated in spermatid-derived embryos

(A) Heatmap of transcripts misregulated in spermatid-embryos. Haploid sperm- and spermatid-embryos were collected in seven experimental replicates and subjected to RNA-seq analysis. Heatmap shows 100 misregulated transcripts in spermatid-embryos: transcripts in red were upregulated, whereas transcripts in blue were downregulated (FDR < 0.05). Columns represent expression values in counts per million of reads (cpm) for each transcript (rows) obtained in seven independent experiments (columns 1-7). Transcripts are sorted by average $\log_2$ of fold difference in expression levels between spermatid- to sperm-embryos. Examples of interesting misregulated transcripts are indicated on the right hand side of the heatmap (B) RT-qPCR validation of misregulated transcripts identified by RNA-seq analysis. Ten randomly selected transcripts from the fifty lowest FDR, upregulated transcripts, and four randomly selected out of eighteen downregulated transcripts were selected for RT-qPCR validation. Expression values for each gene were normalised to the housekeeping gene *pwp1*. Bars show average $\log_2$ of fold change (FC) of expression values obtained for spermatid-derived embryos to values obtained for sperm-derived embryos. Red bars show transcripts which were identified as upregulated in spermatid-derived embryos in RNA-seq, and blue bars show transcripts identified as downregulated in RNA-seq. N=6 independent experiments for all transcripts, apart from *Mn1* and *Chd3* for which N=5 independent experiments. P-values below 0.1 are shown above the bars (t-test).

Panel 'A' of this figure was created and kindly shared by Dr Angela Simeone and Dr George Allen.

A

Enriched GO terms – 'misregulated genes'	
GO:0007275	Multicellular organismal development
GO:0009888	Tissue development
GO:0048731	System development
GO:0032502	Developmental process
GO:0048856	Anatomical structure development
GO:0007498	Mesoderm development

B

Top 6 enriched GO terms in *Brykczynska et al., 2010*
Multicellular organismal development
System development
Anatomical structure development
Developmental process
Multicellular organismal process
Organ development

C

Number of *Xenopus* orthologues of human genes with H3K27me3 and H3K4me2 marks in sperm

	Number of genes	Number (%) of H3K27me3-positive genes	Number (%) of H3K4me2-positive genes
All orthologues (X. leavis to Human)	8812	1140 (16%)	2469 (35%)
Misregulated orthologues	54	18 (41%)*	12 (27%)

Fig. 39. Human orthologues of *Xenopus* misregulated genes are marked by H3K27me3 in sperm.

(A) Developmentally-important gene ontology terms are enriched in the list of misregulated genes in spermatid-derived embryos (p-value < 0.05). (B) Top 6 gene ontology terms enriched within genes having H3K27me3 mark in human sperm (Brykczynska *et al.* 2010). (C) Number of *Xenopus laevis* orthologues of genes enriched for H3K27me3 and H3K4me2 in human sperm (Brykczynska *et al.* 2010). * - p-value=0.000009 and p-value=0.00006 (proportion test and hypergeometric test, respectively), demonstrating that H3K27me3-positive genes are significantly overrepresented in the list of human orthologues of *Xenopus* misregulated genes. Statistical analysis of the enrichment was performed by Dr Angela Simeone.

6.8. Summary

Results described in this chapter aim at the identification of abnormalities observed in spermatid-derived embryos, as compared with sperm-derived embryos. I showed that DNA replication problems, carry-over mRNA or problems with the activation of the zygotic rRNA transcription are unlikely to be the cause of developmental defects of spermatid-derived embryos. On the other hand, RNA-seq analysis of sperm- and spermatid-derived embryos allowed the identification of 100 developmentally-important mRNAs which are misexpressed in spermatid-derived embryos, as compared with sperm-derived embryos. Interestingly, the majority of these mRNAs (82/100) turned out to be upregulated in spermatid-derived embryos. Misregulation of these mRNAs is a plausible explanation for the developmental defects of spermatid-derived embryos, especially since many of them are important transcriptional regulators of embryonic development.

Another interesting and unexpected finding is that the gene ontology terms enriched in the misregulated gene list in spermatid-derived embryos were strikingly similar to the gene ontology terms enriched within H3K27me3-positive genes in mature human sperm. More detailed analysis of human orthologues of *Xenopus* genes revealed a significant enrichment for genes positive for H3K27me3 in the list of human orthologues of the *Xenopus* misregulated genes. This finding suggests that perhaps sperm, but not the spermatid, is epigenetically programmed to regulate transcription of embryonic genes.

Chapter 7

Epigenetic profiling of sperm and spermatids

Experiments with enzymatic removal of H3K27me3 mark in in vitro matured/ICSI embryos were performed under the supervision and with a great help from Dr Kei Miyamoto and Dr Jerome Jullien. I was trained and supervised on how to perform MNase digestions and ChIP-seq analyses on the chromatin from sperm and spermatids by Dr Serap Erkek and Dr Antoine Peters (Friedrich Miescher Institute for Biomedical Research, Basel, Switzerland), who collaborated with our laboratory on this project and who allowed me to visit their laboratory to obtain a technical training vital for this project. Bioinformatic analyses described in this chapter were performed by Dr Angela Simeone.

7.1. Introduction

The results discussed in the previous chapter show that the expression of one hundred developmentally-important genes is misregulated in spermatid-derived embryos, as compared with sperm-derived embryos. Interestingly, human orthologues of these genes are enriched for H3K27me3 mark in the mature sperm. This result prompted me to hypothesise that perhaps sperm, as opposed to spermatids, is epigenetically programmed to regulate expression of embryonic genes in the embryo after fertilisation. Indeed, it was shown in many species that mature sperm retains post-translationally modified histones on chromatin

(Hammoud *et al.* 2009, Brykczynska *et al.* 2010, Wu *et al.* 2011). It has been suggested that such histones can be delivered to the embryo at fertilisation (van der Heijden *et al.* 2008). If the epigenetic marks on the chromatin could be delivered to the oocyte at fertilisation, then they could pattern the gene expression in the future embryo. Any differences in the chromatin structure between sperm and spermatids could potentially result in differences in the gene expression patterns. I have therefore decided to test the hypothesis that sperm, but not spermatids, may be epigenetically suitable to support embryonic transcription.

In this chapter I describe my results that characterise the epigenetic status of the chromatin of sperm and spermatids in *Xenopus laevis.* I first tested whether the presence of H3K27me3 mark on the paternal chromatin at fertilisation is important for the regulation of the gene expression in the embryo. The results obtained showed that this is indeed the case – enzymatic removal of H3K27me3 marks from the parental chromatin at fertilisation led to gene misexpression in the embryo. This confirmed that the presence of epigenetic marks on the parental chromatin is important for the future regulation of embryonic gene expression. I therefore hypothesised that differential gene expression between sperm- and spermatid-derived embryos could result from differences in epigenetic marks between sperm and spermatids. To interrogate the possible epigenetic differences between sperm and spermatids I first performed a brief general characterisation of sperm and spermatids chromatin by micrococcal nuclease digestion, which allowed the identification of unique chromatin structures in sperm, but not in spermatids. Subsequently, I performed ChIP-seq analyses for H3K27me3 and for H3K4me2 and H3K4me3. Interestingly, my results suggest that the overexpression of genes in spermatid-derived embryos is not explained by the lack of H3K27me3 in spermatids, but instead by a higher abundance of H3K4me2/3 marks in spermatids than in sperm.

7.2. Parentally-derived H3K27me3 is necessary for correct gene expression in embryos

Cross-comparison of genes modified by H3K27me3 in human sperm with the misregulated genes suggested that sperm might be epigenetically programmed by H3K27me3 to support a proper embryonic gene transcription. Before embarking on a detailed characterisation of the epigenetic marks in sperm or spermatids, I wanted to test whether epigenetic marks on the parental chromatin are of any importance for the regulation of gene expression in the early embryo. In other words, I first wanted to test the functional importance of parentally-derived H3K27me3 in the regulation of embryonic gene expression. To interfere with H3K27me3, I used Kdm6b (K6B), an enzyme that specifically demethylates H3K27me3, and as a control I used K6B mutant, which is catalytically inactive (K6B-mut). Constructs encoding these enzymes were kindly provided by Dr Jerome Jullien. First, by injecting mRNA encoding K6B or K6B-mut into 1-cell stage embryos I confirmed that K6B enzyme, but not its mutant version, removed H3K27me3 from embryonic chromatin (Fig. 40).

I designed my experiments in a way that allows H3K27me3 removal from both parental chromatin sets immediately at fertilisation. For that purpose I adapted the technique of *in vitro* maturation of prophase-arrested oocytes (Miyamoto *et al.* 2013) and I was supervised and assisted in performing these experiments by Dr Kei Miyamoto and Dr Jerome Jullien. Firstly, mRNA encoding K6B is injected into immature, GV stage oocytes to allow for protein overexpression. Such oocytes, pre-loaded with K6B or K6B-mut enzyme, are subsequently *in vitro* matured into eggs and injected with sperm (via ICSI procedure) to generate embryos (Figure 41A). In this experimental setup, with the currently available protocols, it is not technically possible to enucleate the oocyte and to obtain haploid embryos

derived from paternal genome only. Instead, in the conditions used here, both maternal and paternal chromatins contribute to development of the embryo and the enzyme accumulated in the oocyte prior to fertilisation acts on both parental chromatin sets. Embryos derived in such way were allowed to develop and were collected at the gastrula stage for RT-qPCR analysis. I assessed the expression levels of five genes misregulated in spermatid-derived embryos (*gata3, plod2, hes1, mn1* and *c19orf26*), and of two control genes: *hoxb1* and *wasf1*. Expression of *hoxb1* is regulated by H3K27me3 (Agger *et al.* 2007). Therefore, *hoxb1* serves as a positive control for the experiment, since the expression level of this gene should be affected by the removal of H3K27me3. *Wasf1* does not have H3K27me3 around its genomic region (Akkers *et al.* 2009), so its expression levels should not change upon H3K27me3 removal. As expected, in this experimental setup K6B overexpression had pronounced effects on gene expression (Figure 41B). *Hoxb1* expression was upregulated 20-fold. Importantly, expression of 4 out of 5 misregulated genes tested was also upregulated (statistically significant upregulation was observed for *plod2* and *gata3*, p-values < 0.05, t-test). I concluded from this experiment that the presence of H3K27me3 on parental chromatin is required for the proper regulation of embryonic gene expression.

To determine whether the upregulation of gene expression observed was a result of H3K27me3 removal immediately at fertilisation or at a later point during development, I tested the effect of H3K27me3 removal during embryogenesis (after fertilisation). For that I injected fertilised 1-cell embryos with mRNA encoding K6B (Fig. 41C). In that way the enzyme is absent at fertilisation and becomes translated and modifies the chromatin only as the embryo develops. I again collected the embryos for RT-qPCR analysis at the gastrula stage and performed the analysis for the same genes as described above. Interestingly, in this experimental setup I observed that control genes, as well as genes identified as misregulated in spermatid-derived embryos were similarly transcribed in embryos expressing K6B or

K6B-mut (Fig. 41D). I therefore concluded that H3K27me3 removal during embryogenesis does not affect early embryonic gene expression.

Summarising, these two sets of experiments (the removal of H3K27me3 at fertilisation and the removal of H3K27me3 later during embryonic development) demonstrate that the presence of H3K27me3 on parental chromatin at fertilisation is required for the proper regulation of expression of embryonic genes. Removal of the H3K27me3 marks during embryogenesis does not affect the expression levels of the tested genes.

Results of these experiments support the hypothesis that the epigenetic marking in the sperm may be required for a proper embryonic development.

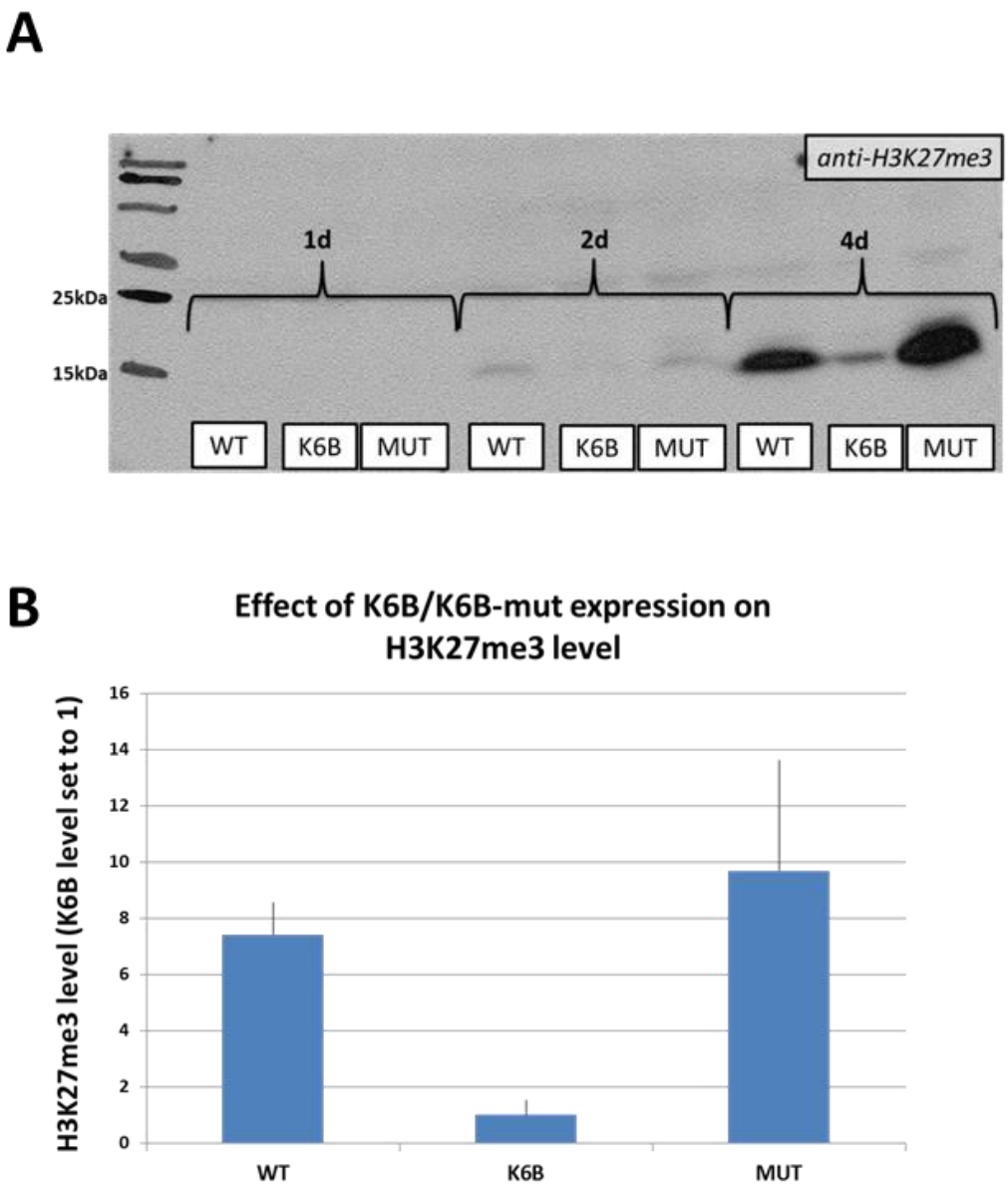


Fig. 40. Overexpression of K6B leads to a removal of H3K27me3 mark.

(A) One-cell embryos were injected either with mRNA encoding K6B or K6B-mut (MUT). Non-injected embryos were used as a control (WT). Embryos were collected for a western blot analysis 1 day (1d, gastrula stage), 2 days (2d, tailbud stage) or 4 days (4d, early tadpole) after mRNA injection. Membrane is stained with an antibody against H3K27me3. H3K27me3 is removed upon overexpression of K6B, but not upon overexpression of K6B-mut. (B) Quantification of H3K27me3 removal based on the results of immunoblots from three independent experiments. Error bars show ±SEM.

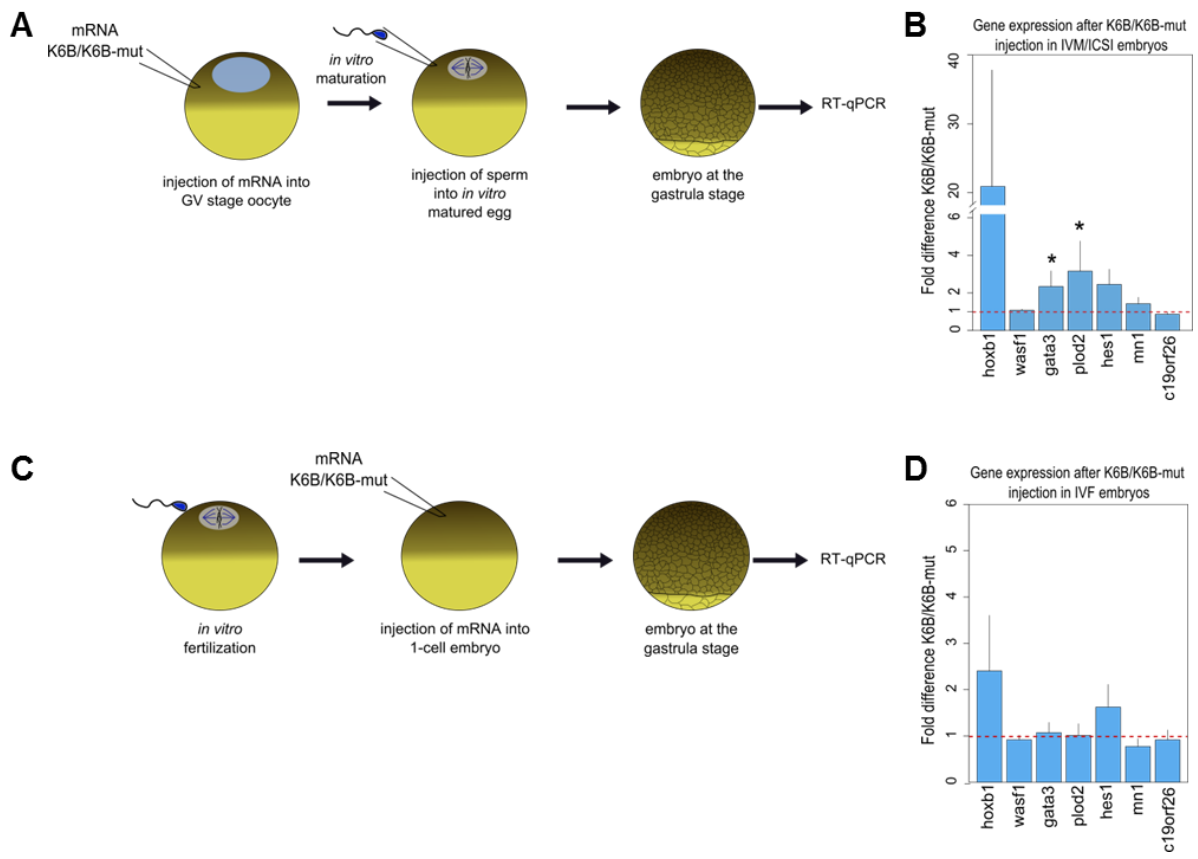


Fig. 41. Removal of H3K27me3 at fertilisation leads to gene misexpression.

(A) H3K27me3 mark removal at fertilisation. Immature, prophase-arrested germinal vesicle (GV) stage oocytes are injected with mRNA encoding K6B or K6B-mut. mRNA is translated into protein and subsequently the oocytes, pre-loaded with K6B or K6B-mut, are *in vitro* matured into eggs. Sperm is then injected into such eggs and the resulting embryos are collected for RT-qPCR analysis at the gastrula stage. (B) Gene expression levels in *in vitro* matured and sperm injected (IVM/ICSI) embryos expressing K6B or K6B-mut. Gene expression was normalised to the housekeeping gene *pwp1*. Expression levels are shown as fold differences in gene expression between embryos expressing K6B to embryos expressing K6B-mut. Error bars show ± SEM. N=5 experimental replicates. * indicates p-value < 0.05 (t-test). (C) H3K27me3 mark removal during embryonic development. 1-cell stage embryos are injected with mRNA encoding K6B or K6B-mut. In this way the proteins are absent at fertilisation and they become expressed only as the embryo develops. Embryos are collected for RT-qPCR analysis at the gastrula stage. (D) Gene expression in *in vitro* fertilised (IVF) embryos expressing K6B or K6B-mut. Gene expression was normalised to the housekeeping gene *pwp1*. Expression levels are shown as fold differences in gene expression between embryos expressing K6B to embryos expressing K6B-mut. Error bars show ±SEM. N=3 experimental replicates.

7.3. General characterisation of chromatin structure in sperm and spermatids

Removal of H3K27me3 marks from the parental chromatin at the time of fertilisation showed that they are important for the correct regulation of embryonic gene expression. I have therefore decided to further characterise the structure of chromatin in sperm and spermatids. I initially decided to see whether the chromatin is any different in sperm or spermatids in terms of its accessibility or its global structure. One simple way to test this is to treat the isolated chromatin with micrococcal nuclease (MNase). MNase is an endo-exonuclease enzyme that digests DNA at regions which are not protected from digestion. Protection from digestion can be mediated for example by the presence of nucleosomes on DNA. The nuclesome protects DNA fragment of around 150bp, therefore after the MNase digestion of chromatin isolated from somatic cells, one usually obtains a so called 'nucleosome ladder' – a DNA ladder of band sizes which are a multiplication of 150bp, for mono-, di- trinucleosomes etc. (Fig. 42). To establish a protocol for *Xenopus laevis* sperm and spermatids chromatin preparation and MNase digestion, a collaboration with the lab of Dr Antoine Peters (FMI, Basel) has been initiated, in order to benefit from the expertise of Dr Peters in this methodology (Brykczynska *et al.* 2010, Erkek *et al.* 2013, Hisano *et al.* 2013).

Chromatin from sperm and spermatids was isolated and subjected to MNase digestion for 30mins at 37ºC (using 2.5 units of MNase per chromatin isolated from 1 million of cells). As a control for MNase digestion, I have also prepared and digested chromatin from *Xenopus laevis* cell line XL177. Subsequently, DNA has been isolated and run on an agarose gel. Somatic cells and spermatids displayed a typical 'nucleosome ladder' digestion pattern (Fig. 43). Spermatids were more digested than XL177 cells (more mononucleosomes and less of the higher order chromatin structures observed after digestion). Unexpectedly, it turned out that the sperm has a very unique pattern of nuclease-protected DNA regions. Three fragments

of different sizes were observed after MNase digestion: ~75bp, ~110bp and ~150bp long fragments (Fig. 43). Interestingly, mouse sperm digested with MNase displays a digestion pattern typical for any other somatic cells (150bp size nucleosomes) (Brykczynska *et al.* 2010), which therefore suggests that *Xenopus* sperm is very unique in its chromatin composition.

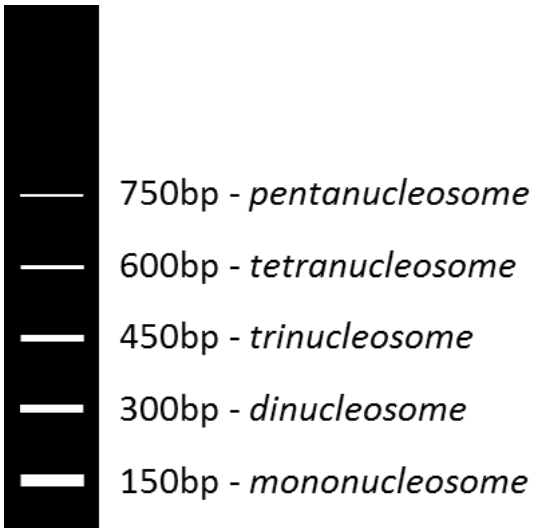


Fig. 42. 'Nucleosome ladder' after MNase digestion

Drawing showing an example of a 'nucleosome ladder' when DNA isolated from chromatin from somatic cells digested with MNase is run on a gel. Note that the size of the bands of the 'ladder' is approximately a multiplication of 150bp, which is a size of DNA protected by a single nucleosome.

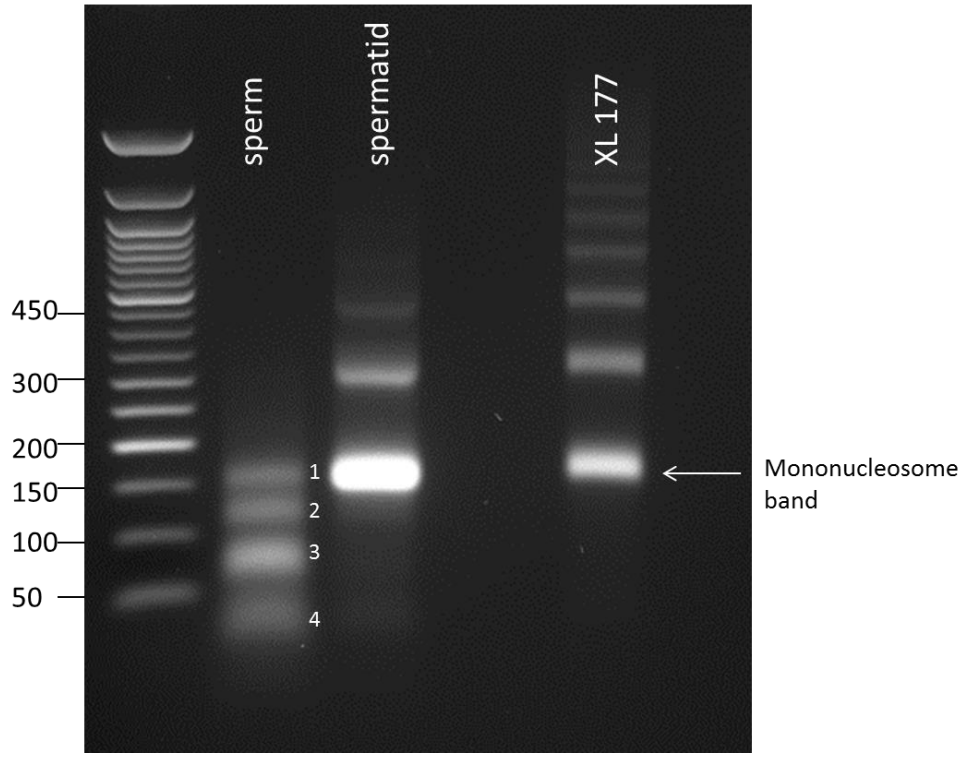


Fig. 43. MNase digestion reveals that sperm chromatin has a unique structure

Chromatin isolated from 1 million of sperm, spermatids or XL177 cells was digested with 2.5 units of MNase for 30mins at 37ºC. Subsequently, DNA was isolated and run on the agarose gel. Spermatids and XL177 cells digestion pattern displays a typical 'nucleosome ladder' (mononucleosome band size indicated), whereas digestion of sperm revealed a unique structure of sperm chromatin, consisting of three bands of approximately 150bp, 110bp and 75bp (bands are respectively numbered 1, 2, 3 and indicated on the gel). The lowest band (4) is digested DNA, which was also observed when mouse or human sperm is digested with MNase (Brykczynska *et al.* 2010).

It is known that mature sperm in *Xenopus laevis* retains core histones H3 and H4, but has a reduced amount of histones H2A and H2B (Abe & Hiyoshi 1991, Yokota 1991, Shechter *et al.* 2009). My immunoblotting results on *Xenopus leavis* sperm and spermatids agreed with the published observations – I could detect similar amounts of H3 and H4 by immunoblotting in sperm and spermatids and reduced amounts of histones H2A and H2B (Fig. 44). I next hypothesised that perhaps at least some of the chromatin structures that I observed in the sperm, but not in the spermatid chromatin after MNase digestion are due to the incorporation of sperm basic proteins into the sperm chromatin. Sperm basic proteins in *Xenopus* are more basic than histones, as evidenced by theoretical isoelectric point calculations (calculations performed using online Expasy tool http://web.expasy.org/cgi-bin/compute_pi/pi_tool): the least basic of the sperm basic proteins (Sp1) is still more basic than any canonical histone (Table 6). I have therefore reasoned that if any of the sperm chromatin structures observed after MNase digestion have sperm basic proteins as their components, they could be preferentially disrupted upon heparin treatment of chromatin. Heparin is a molecule of high negative charge and is therefore able to displace positively charged proteins from chromatin (Hildebrand *et al.* 1977). Treatment of permeabilised *Xenopus* sperm with heparin leads to chromatin dispersal and DNA release. Dispersal os sperm chromatin increased with the increasing concentration of heparin used (Fig. 45). Since the sperm basic proteins are more basic than histones, they should be displaced from sperm chromatin by the heparin treatment earlier than histones. Therefore, if some of the structures on sperm chromatin that protect the DNA from the MNase digestion are composed with the sperm basic proteins to a higher extent than the others, they could be displaced from the chromatin first. I have therefore decided to test the effect of sperm chromatin treatment with increasing doses of heparin on the pattern of MNase digestion. I have treated permeabilised sperm nuclei with increasing doses of heparin and subsequently performed MNase digestion.

Interestingly, ~110bp band (number 2) generated by MNase digestion is the first one to disappear upon increased heparin concentration (Fig. 46). This suggests that the band number 2 (~110bp long) could be composed of sperm basic proteins or be a mixture of sperm basic proteins and histones. Alternatively, the second band could also be composed of unstable histones. Regardless of the reason for the preferential displacement of the 2nd band, the important information is that these three chromatin structures in the sperm are not equally stable. The exact identity of these three bands and the nature of the DNA protected by these structures is a subject of follow-up studies currently conducted in the Gurdon laboratory.

Table 6. Theoretical isoelectric point of nuclear proteins in *Xenopus laevis.*

Name	H2A	H2B	H3	H4	Sp1	Sp4	Sp5
Isoelectric point	10.17	10.31	11.27	11.36	11.89	12.57	12.62

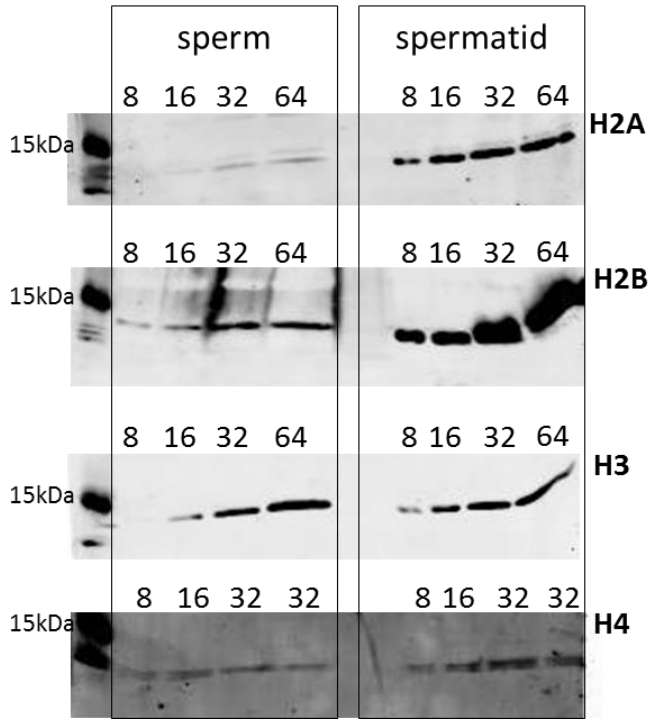


Fig. 44. Immunoblotting analysis for histone H2A, H2B, H3 and H4 on sperm and spermatids.

Immunoblotting analysis for core histones H2A, H2B, H3 and H4 on the protein lysates from sperm and spermatids. The number of cells from which the proteins were isolated is indicated above each lane (numbers represent thousands of cells, for example, '8' stands for proteins isolated from 8000 cells). Note that sperm and spermatids contain similar amount of histones H3 and H4, and that mature sperm has a reduced amount of histones H2A and H2B.

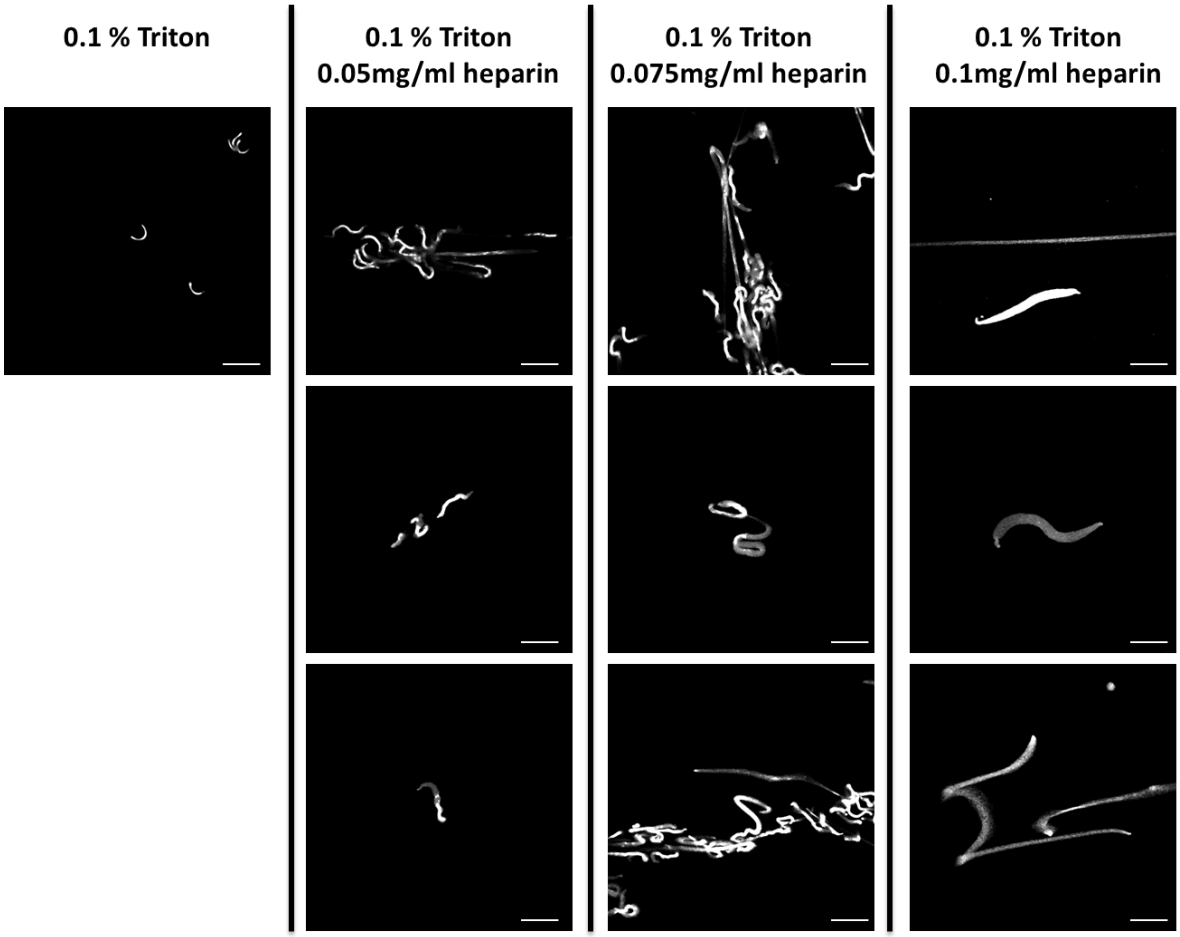


Fig. 45. Heparin treatment leads to chromatin dispersal in sperm.

Permeabilised sperm were treated with increasing doses of heparin and stained with DAPI. Three representative images are shown for each of the heparin treatment doses. Note that the higher the heparin dose, the more decondensed the sperm. Scale bar = 20um. All images were taken with the same magnification.

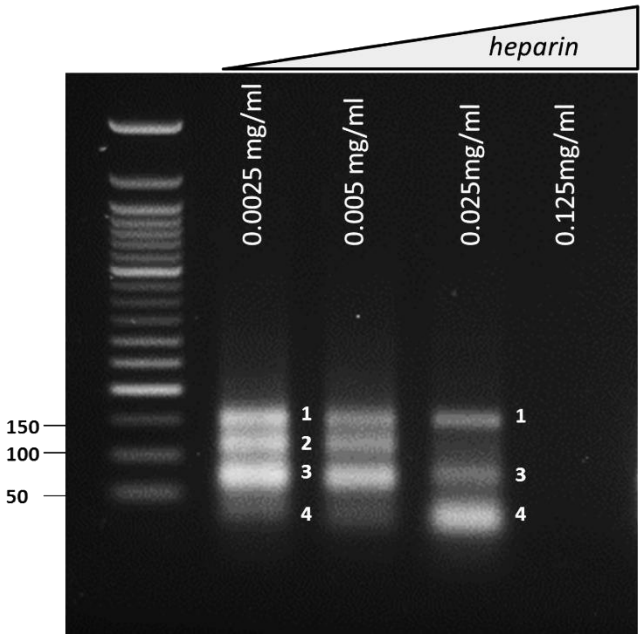


Fig. 46. Heparin treatment first releases the ~110bp structure from sperm chromatin.

Sperm chromatin was treated with increasing concentrations of heparin (0.0025 – 0.125 mg/ml) and subsequently digested with 2.5U MNase for 30mins. DNA was isolated and run on the gel. Note that the chromatin structure protecting a ~110bp DNA fragments (number 2) disappears as the first one with the increasing dose of heparin treatment. This is concomitant with the appearance of digested DNA (band number 4). Further increase of the heparin concentration leads to a loss of all the protective chromatin structures and digestion of all DNA in sperm.

7.4. ChIP-seq analysis for H3K27me3 does not reveal differences between sperm and spermatids

ChIP experiments were performed according to the advice provided by Dr Serap Erkek and Dr Antoine Peters. All the bioinformatic analyses were performed by Dr Angela Simeone.

My results obtained so far demonstrated that parentally-derived H3K27me3 marks regulate embryonic gene expression. Furthermore, I have shown that sperm and spermatids differ in their chromatin structure. I therefore reasoned that perhaps the misregulation of gene expression in spermatid-embryos, as compared to sperm-embryos, might be a consequence of improper epigenetic marking in the spermatid. As the majority of the misregulated genes are upregulated in spermatid-embryos, I hypothesised that the spermatid may lack the repressive H3K27me3 marks as compared with sperm. To address this, I performed chromatin immunopurification followed by a genome-wide sequencing (ChIP-seq) analysis for H3K27me3 mark on mononucleosomal chromatin isolated from *Xenopus laevis* sperm and spermatids. I have chosen the mononucleosomal chromatin fraction for the ChIP analysis, since this fraction is a canonical chromatin structure observed in both cell types and the identity of the other chromatin structures in sperm is currently unknown. I performed the ChIP experiments in three independent replicates (three different cell preparations from different frogs and ChIP experiments performed on different days). To assess the variability between the experiments, Dr Angela Simeone performed a Pearson correlation coefficient analysis for the replicates. The three biological replicates for the H3K27me3 mark showed an average Pearson correlation coefficient of 0.9 between replicates in sperm and 0.72 in spermatids. The ChIP-seq data was analysed first by looking at overall methylation levels and

second, by looking at localised regions of enrichment for histone marks (peaks) (see Fig. 47 for a detailed explanation).

7.4.1. Overall methylation levels analysis

I first wanted to test whether the misregulated genes have different histone H3K27me3 methylation levels than the genomic average. Quantification of methylation levels was performed by Dr Angela Simeone. Methylation levels were calculated as the total number of reads obtained for H3K27me3 (normalised to the number of reads obtained in the input sample) at the region around the transcriptional start site and also in the region encompassing the gene bodies, as H3K27me3 was shown before to spread in broad domains across the genes (Fig. 47) (Barski *et al.* 2007). The analysis of the H3K27me3 methylation levels at misregulated genes as compared with the genomic averages revealed that in *Xenopus laevis* misregulated genes are significantly more methylated (have more reads) for H3K27me3 in both sperm and in spermatids (p-values<0.05) (Fig. 48A). This confirms the previous finding made by cross-species comparison to human sperm (Fig. 39) – the misregulated genes in *Xenopus* indeed have more of H3K27me3 mark. Unexpectedly however, it seems that H3K27me3 cannot on its own explain the difference in gene expression between sperm- and spermatid-derived embryos, since it is enriched at misregulated genes in both cell types.

To make sure that meaningful differences in H3K27me3 methylation levels are not overlooked in the bioinformatic analysis described, I also decided to compare H3K27me3 levels at misregulated genes between sperm and spermatids. Therefore, in addition to the comparison of the methylation levels between the misregulated genes and the genomic averages (see above), Dr Angela Simeone also directly compared the methylation levels at

the misregulated genes between sperm and spermatids. The reasoning behind performing this additional analysis is that the mark could be overrepresented in both samples, but still be more abundant in sperm than in spermatids. The results of the comparison between sperm and spermatids showed H3K27me3 levels at misregulated genes was not significantly different between sperm and spermatids (p-value = 0.7) (Fig. 48B and Fig. 49A). The same was true also when looking at individual gene tracks – no significant differences in H3K27me3 methylation patterns were observed between sperm and spermatids (Fig. 49B).

To conclude, the quantification of overall methylation levels at misregulated genes showed a similar level of enrichment over the genome-wide average for H3K27me3 in both cell types.

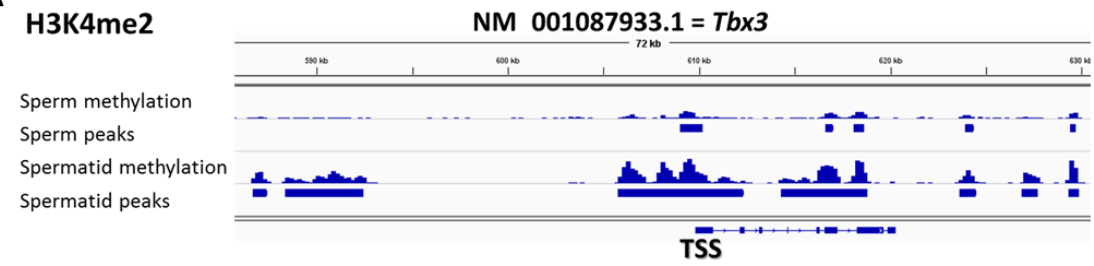
7.4.2. Peak analysis

To further characterise any potential differences in H3K27me3 level between sperm and spermatids, regions of enrichment (peaks) for histone marks were identified in the gene regulatory regions and in the gene bodies by Dr Angela Simeone (Fig. 47). In agreement with the findings for the overall methylation levels, genes with H3K27me3 peaks were significantly enriched within the misregulated gene list in both cell types, as compared to the genomic averages (p-values < 0.05) (Fig. 50A). It was also assessed (analysis by Dr Angela Simeone) how broad were the H3K27me3 peaks, as it could be that peaks for H3K27me3 differ in size between sperm and spermatids. Again, H3K27me3 peaks were of similar size in sperm and in spermatids (Fig. 50B), therefore the peak size of H3K27me3 could not explain the difference in gene expression between sperm- and spermatid-derived embryos.

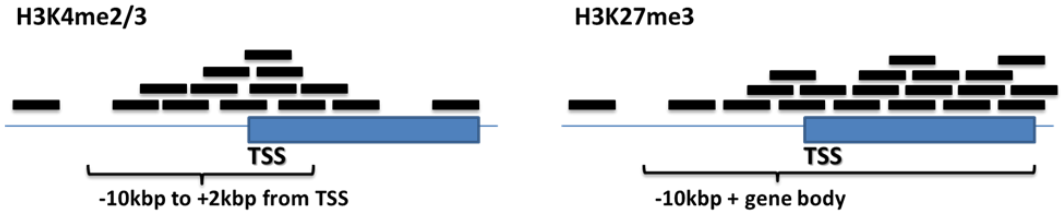
7.4.3. Summary – H3K27me3 marks are not different between sperm and spermatids

To summarise, both the quantitative assessment of methylation levels and the peak-oriented analysis led to the conclusion that H3K27me3 is enriched at misregulated genes above the genomic average both in sperm and spermatids. This suggests that, even though H3K27me3 is necessary for proper gene expression (Fig. 41), presence of this mark alone cannot explain differences in gene expression between sperm- and spermatid-derived embryos, as the mark is enriched in both cell types.

A



B



C

Quantification of overall methylation level

	Number of normalized reads
sperm	4
spermatid	23

D

Quantification of peaks

	Is the gene (*Tbx3*) positive for a peak?
sperm	Yes
spermatid	Yes

Fig. 47. Legend on the subsequent page.

Fig. 47. Schematic representation of ChIP-seq analysis of histone marks: overall methylation level versus localised enrichments (peaks).

(A) Quantification procedure is explained on the example of *tbx3* gene for H3K4me2 ChIP. Firstly, all the reads obtained from sequencing of ChIP samples were normalised to the reads obtained from sequencing of the corresponding input samples and to the total number of sequencing reads. Normalised methylation tracks are shown as 'Sperm methylation' and 'Spermatid methylation'. Regions of read enrichment for each mark ('peaks') were identified with the use of MACS2 (Zhang et al., 2008) and are depicted as horizontal bars spanning the regions of significant read enrichment and are visualised as tracks named 'Sperm peaks' and 'Spermatid peaks'. (B) Subsequently, overall methylation levels were quantified by summing up the normalised read number in the regions spanning -10kbp to +2kbp of gene transcriptional start sites (TSSs) for H3K4me2 and H3K4me3 and in the regions spanning -10kbp of TSS and the entire gene body for H3K27me3 (Bernstein *et al.* 2005, Barski *et al.* 2007, Akkers *et al.* 2009, van Heeringen *et al.* 2014). (C) Example of the results of quantification of normalised read number for *tbx3*. (D) The presence of peaks was evaluated in the same regions as above: -10kbp/+2kbp from TSS for H3K4me2/3 and -10kbp + gene body for H3K27me3. Table shows an example of the results of the peak presence assessment for *tbx3*.

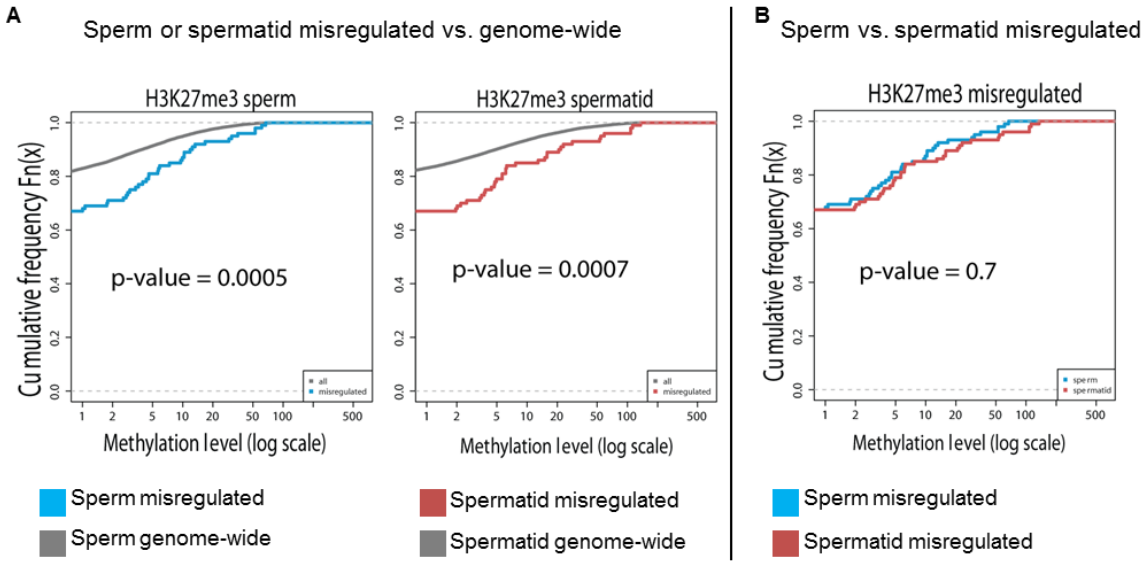


Fig. 48. Cumulative distribution curves for H3K27me3 in sperm and spermatids.

(A) Cumulative distribution curves of overall methylation levels for H3K27me3 compared between misregulated genes and the genome-wide averages in sperm or spermatids. Curves for the misregulated gene methylation level in sperm are shown in blue, for the spermatid – in red, and for the genome-wide average – in grey. P-values for the difference between the methylation level at misregulated genes and the genome-wide average are indicated in the graphs (Kolmogorov-Smirnov, KS-test). (B) Cumulative distribution curves of H3K27me3 methylation level at misregulated genes between sperm (blue) and spermatid (red). P-values for the difference between the methylation levels are indicated in the graphs (KS-test). Analysis for this figure was performed by Dr Angela Simeone. The figure was generated and kindly shared by Dr Angela Simeone.

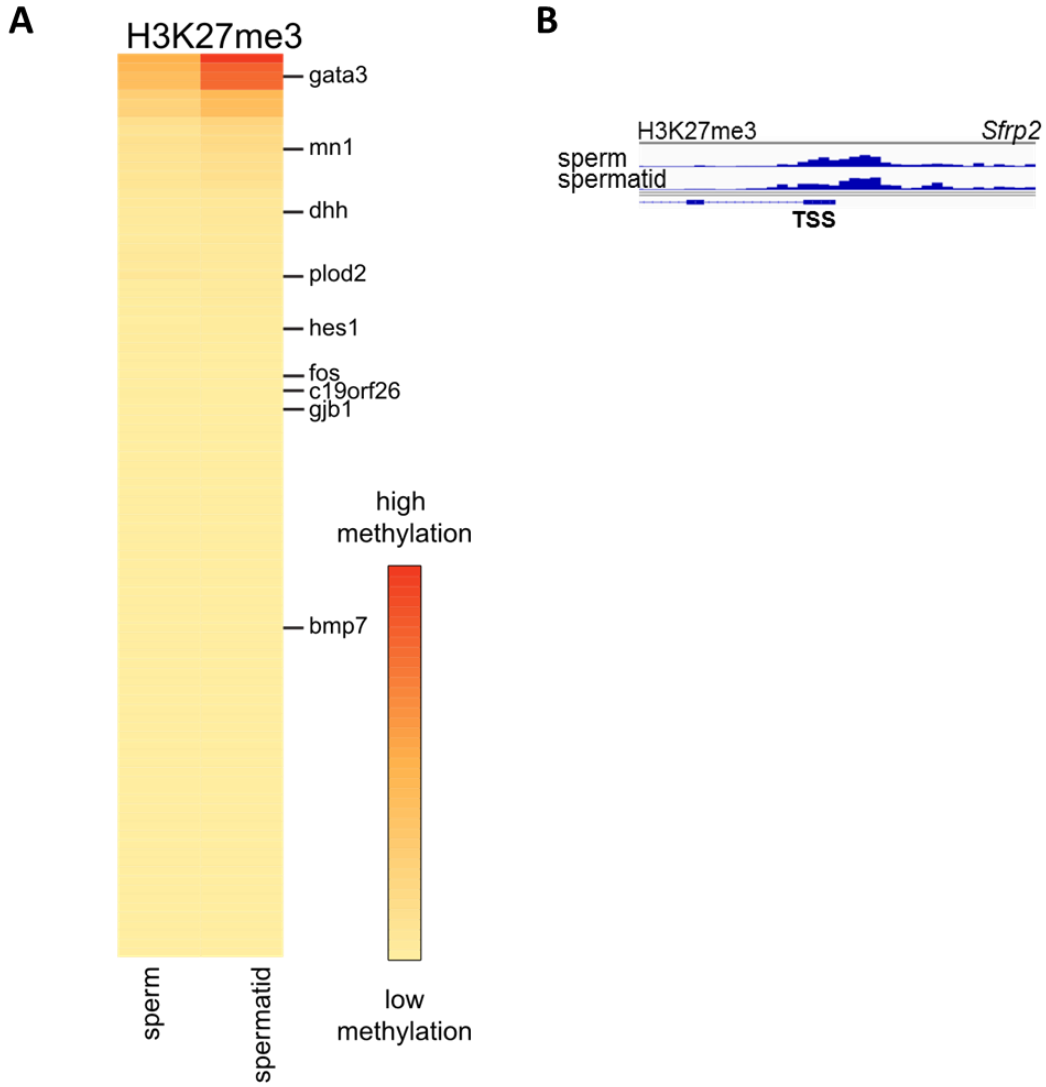


Fig. 49. H3K27me3 at misregulated genes is not different between sperm and spermatids.

(A) A heatmap showing the average normalised number of reads for misregulated genes in sperm and spermatids for H3K27me3. Genes (rows) are sorted from the most methylated to the least methylated in spermatids. (B) Methylation patterns of a representative misregulated gene in sperm and spermatids for H3K27me3. Track shows read numbers in the bound fraction, normalised to the input and to the total number of sequenced reads. Analysis for this figure was performed by Dr Angela Simeone. The panel 'A' of this figure was generated and kindly shared by Dr Angela Simeone.

A

Number of genes positive for H3K27me3 peaks in *Xenopus* sperm or spermatids

Mark	Cell type	Number of positive genes (% of all genes)	Number of positive genes within misregulated genes (% of misregulated genes)	Fold difference misregulated/ all	P-value (enrichment/depletion over genome-wide), proportion test
H3K27me3	Sperm	5366 (16%)	35 (35%)	2.2	0.0000001
H3K27me3	Spermatid	7770 (23%)	43 (43%)	1.9	0.000001

B

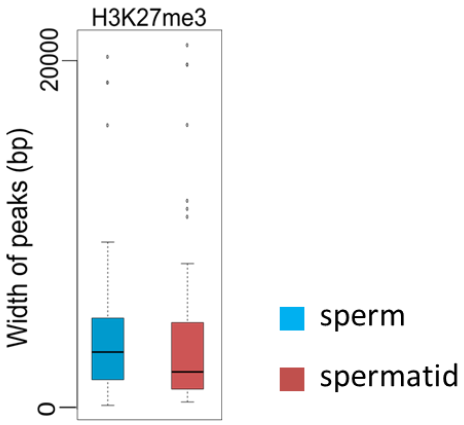


Fig. 50. Peak analysis for H3K27me3 reveals no differences between sperm and spermatids.

(A) Peaks were counted in the upstream regions and in the gene bodies for H3K27me3. Statistical analysis shows a significant enrichment for H3K27me3-positive genes amongst misregulated ones in sperm and spermatids (as compared to genome-wide average) (B) Box plot analysis of the H3K27me3 peak width in sperm and spermatids reveals that peaks are of similar size in both cell types. The analyses for this figure were performed by Dr Angela Simeone, who also created and kindly shared the panel 'B' of this figure.

7.5. Misregulated genes have more H3K4me2/3 activating marks in spermatids than in sperm

My ChIP-seq results for H3K27me3 revealed that this repressive mark is not different between sperm and spermatids; therefore, it cannot explain the difference in gene expression between sperm- and spermatid-derived embryos. I next hypothesised that since the majority of misregulated genes are upregulated in spermatid-embryos, perhaps the spermatid has more activating epigenetic marks than the sperm. To address this, I performed ChIP-seq analysis on mononucleosomal chromatin isolated from sperm and spermatids for two histone marks associated with gene activation: histone H3 lysine 4 dimethylation (H3K4me2) and histone H3 lysine 4 trimethylation (H3K4me3). Average Pearson correlation coefficient analysis performed by Dr Angela Simeone confirmed the reproducibility of the results (the exact average Pearson correlation coefficient values were: 0.9 for sperm H3K4me2; 0.55 for sperm H3K4me3; 0.95 for spermatid H3K4me2 and 0.8 for spermatid H3K4me3).

The ChIP-seq data were again analysed by two different approaches: first, looking at the overall methylation levels and second, by looking at the localised regions of enrichment for histone marks (peaks) (Fig. 47).

7.5.1. Overall methylation levels analysis

In the initial analysis histone methylation levels were compared at the misregulated genes with the genomic average. Methylation levels were again quantified (by Dr Angela Simeone) as the total number of reads obtained for H3K4me2 and H3K4me3 in the regions around the transcriptional start sites (TSS) (Fig. 47). Interestingly, there was a slight enrichment of H3K4me2 reads at misregulated genes over the genome-wide average in

spermatids (p-value = 0.1) (Fig. 51A), suggesting that the presence of high levels of H3K4me2 might be responsible for the gene upregulation in spermatid-embryos.

To further characterise this difference, histone methylation levels were compared at the misregulated genes between sperm and spermatids. Interestingly, misregulated genes had a significantly higher methylation in spermatids than in sperm for H3K4me2 and H3K4me3, (p-values < 0.05) (Fig. 51B and Fig. 52A). At individual gene tracks, H3K4me2 and H3K4me3 showed mainly quantitative differences in the methylation levels between sperm and spermatids (Fig. 52B). Interestingly, for some genes H3K4me2 additionally showed a different distribution of methylation between sperm and spermatids: in such genes H3K4me2-positive regions in spermatids were broader than in sperm and often contained several enriched regions close to the TSS (Fig. 52B).

To conclude, quantification of H3K4me2/3 levels at misregulated genes between spermatids and sperm revealed significantly higher levels of H3K4me2/3 in spermatids. Higher levels of H3K4me2/3 in spermatids correlate with the observed upregulation of genes in spermatid-derived embryos.

7.5.2. Peak analysis

Subsequently, peak analysis for H3K4me2 and for H3K4me3 was performed (by Dr Angela Simeone). Interestingly, in sperm, but not in spermatids, there was a two-fold decrease in the number of H3K4me3-positive genes among the misregulated ones as compared to the genomic average (Fig. 53A). Next, the size of peaks for H3K4me2 and H3K4me3 was analysed in sperm and spermatids. Interestingly, peaks for H3K4me2 and for H3K4me3 in spermatids were broader than in sperm (Fig. 53B).

7.5.3. Summary – H3K4me2/3 marks are more abundant in spermatids

To summarise, the activating marks H3K4me2 and H3K4me3 are more abundant in spermatids and depleted in sperm at misregulated genes. Enrichment of activating H3K4me2/3 in spermatids, as compared sperm, correlates well with the fact that the majority of misregulated genes (82/100) are overexpressed in spermatid-derived embryos, as compared to sperm-derived embryos. This suggests that sperm, as opposed to the spermatid, is epigenetically programmed for proper embryonic gene expression.

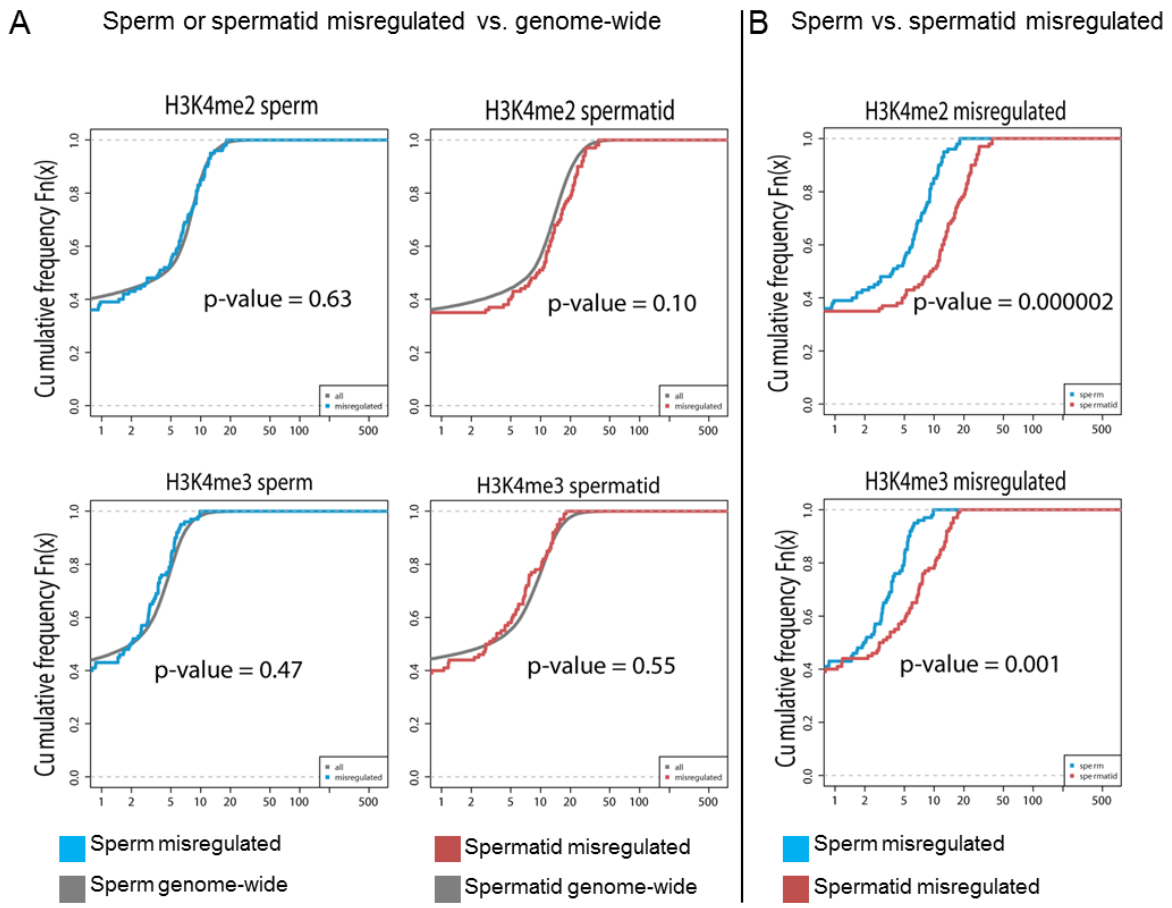


Fig. 51. Cumulative distribution curves for H3K4me2 and for H3K4me3 in sperm and spermatids.

(A) Cumulative distribution curves of overall methylation levels for H3K4me2 and H3K4me3 compared between misregulated genes and the genome-wide averages in sperm or spermatids. Curves for the misregulated genes methylation levels in sperm are shown in blue, for the spermatid – in red, and for the genome-wide averages – in grey. P-values for the difference between the methylation level at misregulated genes and the genome-wide averages are indicated in the graphs (Kolmogorov-Smirnov, KS-test). (B) Cumulative distribution curves of H3K4me2 and H3K4me3 methylation levels at misregulated genes between sperm (blue) and spermatids (red). P-values for the difference between the methylation levels are indicated in the graphs (KS-test). Analysis for this figure was performed by Dr Angela Simeone. The figure was generated and kindly shared by Dr Angela Simeone.

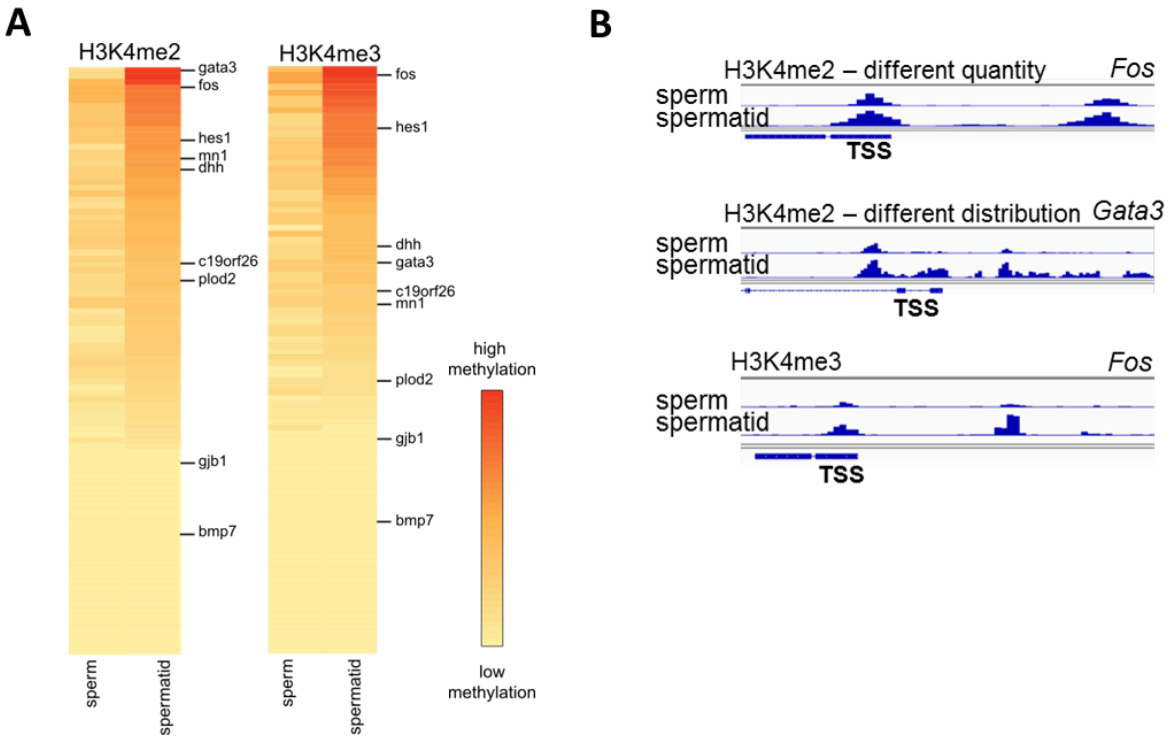


Fig. 52. H3K4me2 and H3K4me3 marks are more abundant in spermatids than in sperm at the misregulated genes.

(A) Heatmaps showing the average normalised number of reads for the misregulated genes in sperm and spermatids for H3K4me2 and for H3K4me3. Genes (rows) are sorted from the most methylated to the least methylated in spermatids, separately for each mark. (B) Methylation patterns of representative misregulated genes in sperm and spermatids for H3K4me2 and for H3K4me3. Tracks show read numbers in the bound fraction, normalised to the input and to the total number of sequenced reads. Analysis for this figure was performed by Dr Angela Simeone. The panel 'A' of this figure was generated and kindly shared by Dr Angela Simeone.

A

Number of genes positive for H3K27me3, H3K4me2 or H3K4me3 peaks in *Xenopus* sperm or spermatids

Mark	Cell type	Number of positive genes (% of all genes)	Number of positive genes within misregulated genes (% of misregulated genes)	Fold difference misregulated/ all	P-value (enrichment/depletion over genome-wide), proportion test
H3K4me2	Sperm	21109 (61%)	65 (65%)	1.1	0.3
H3K4me2	Spermatid	24660 (72%)	73 (73%)	1.0	0.4
H3K4me3	Sperm	2197 (6%)	3 (3%)	0.5	0.1
H3K4me3	Spermatid	20015 (58%)	61 (61%)	1.0	0.3

B

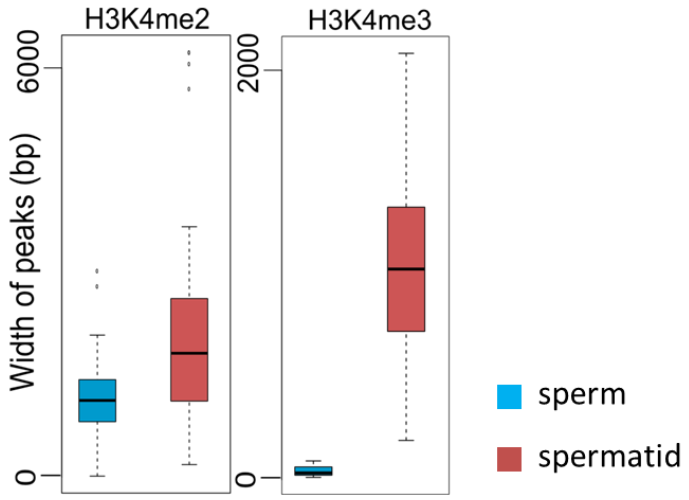


Fig. 53. Peak analysis for H3K4me2 and for H3K4me3 reveals depletion of H3K4me3 peaks in sperm and also shows that peaks for H3K4me2 and H3K4me3 are broader in spermatids.

(A) Peaks were counted in the upstream regions for H3K4me2 and for H3K4me3. Note a two-fold depletion of H3K4me3-positive genes amongst misregulated ones in sperm (as compared to the genome-wide average) (B) Box plot analysis of the H3K4me2 and H3K4me3 peaks width in sperm and spermatids reveals that peaks are broader in spermatids than in sperm. The analyses for this figure were performed by Dr Angela Simeone, who also created and kindly shared the panel 'B' of this figure.

7.6. Summary

In this chapter I investigated the roles of epigenetic signatures in the parental chromatin for the embryonic development. I first showed that the presence of H3K27me3 on the parental chromatin at fertilisation (but not afterwards) is needed for a proper regulation of embryonic gene expression. Then I looked into the chromatin structure of sperm and spermatids. It turned out that the sperm of *Xenopus laevis* has a uniquely structured chromatin, which is evidenced by unusual patterns of DNA bands released after the MNase digestion (DNA bands pattern are different from the canonical 'nucleosome ladder' and they are also different from the patterns observed in mouse or human sperm). Subsequent ChIP-seq analysis of mononucleosomal chromatin from sperm and spermatids for the repressive H3K27me3 and activating H3K4me2 and H3K4me3 revealed that sperm and spermatids do not differ in the abundance of H3K27me3, but that H3K4me2/3 was higher in spermatids than in sperm at misregulated genes. Interestingly, the average width of peaks for H3K4me2/3 was also higher in spermatids than in sperm. These results support the hypothesis that the upregulation of developmentally-important genes in spermatid-derived embryos is a consequence of the presence of H3K4me2 and H3K4me3 at higher levels in spermatids than in sperm. These also suggest that during spermiogenesis the spermatid has to epigenetically mature (lose H3K4me2/3 marks) in order to be correctly programmed to support embryonic development.

Chapter 8
Discussion

In this thesis I described my results investigating the nature of sperm programming for embryonic development. To address this question, I compared the sperm with its precursor cell, a spermatid. The results obtained support the hypothesis that programming of sperm is related to the acquisition of chromatin signatures that support a correct gene expression in the embryo. Specifically, I showed that spermatids, as opposed to sperm, retain on the chromatin activating H3K4me2/3 marks. Presence of these activating marks at misregulated genes in spermatids correlates with upregulation of these genes in spermatid-derived embryos. Therefore, my results suggest that sperm, as opposed to spermatids, are epigenetically mature to support normal embryonic development. There are however certain limitations to the conclusions that can be drawn from the results presented. Also, there are experiments, which are now carried in the Gurdon laboratory, that can potentially extend the findings presented in this thesis. Therefore, below I thematically group these matters and discuss them. I finish this section by explaining how the results presented in this thesis broaden our current understanding of sperm programming and how they relate to the phenomena of transgenerational inheritance of the phenotypic changes through paternal epigenetic marks.

8.1. Are the defects of spermatid-derived embryos a consequence of gene misexpression?

In order to understand the nature of sperm programming, I investigated the developmental defects of spermatid-derived embryos. Identifying the defects of the spermatid-derived embryos, not observed in sperm-derived embryos, allowed me to narrow down the aspects of development in which the sperm-derived embryos are better. I showed that replication problems, carry-over RNA or inefficient activation of rRNA transcription are unlikely to explain the developmental defects of spermatid-derived embryos. Interestingly, I identified 100 developmentally-important mRNAs as misregulated in spermatid-derived embryos (compared to sperm-derived embryos). Is this however the real cause of the developmental defects of spermatid-derived embryos?

The experiments presented in this thesis do not directly answer this question. Results shown allow correlating the developmental effects with the misexpression of these mRNAs. However, the presented results do not provide the functional evidence that could prove the hypothesis that misexpression of these mRNAs is indeed causative for the defects of spermatid-derived embryos. Experiments that could functionally test the relationship between the misexpression of these mRNAs and the occurrence of developmental defects in spermatid-derived embryos could be potentially designed in two different ways. First, one could try to overexpress the misregulated mRNAs in sperm-derived embryos to test whether developmental defects similar to those observed in spermatid-derived embryos would be induced. Alternatively, one could also downregulate proteins encoded by these mRNAs in spermatid-derived embryos to rescue their defects. Unfortunately however, none of these strategies are technically possible for multiple reasons. First, transcription factors need to be present at the correct concentrations and at very precise developmental time windows and

their expression should be limited only to specific tissues. Currently available protocols aiming at overexpressing or downregulating the proteins utilise mRNA or morpholino oligonucleotides injection, respectively. Injections are performed into 1-cell stage embryos; therefore the effects would happen in all the embryonic tissues and would not be limited to the desired tissues only. Secondly, the majority of these misregulated mRNAs are embryonically expressed with a precisely controlled time of expression. Injecting mRNAs or morpholinos into 1-cell stage embryos does not allow one to have control over the developmental timing of the expression. Last, but not least, it is difficult to imagine how one would express all the factors of interest at the precisely desired concentration. This would be even more problematic in the morpholino knockdown strategy, as for this strategy one would also need to precisely measure how the mRNA concentration encoding a given factor is reflected in the final concentration of the translated protein.

Interestingly, it has been demonstrated that about a 2-fold difference in the protein concentration of the transcription factor Xbra prevented nucleocytoplasmic hybrid embryos between *Xenopus laevis* and *Xenopus tropicalis* from the correct convergence/extension movements during gastrulation, which resulted in their failure to gastrulate successfully. These defects were partially rescued by correcting for the Xbra protein concentration (Narbonne *et al.* 2011). Therefore, if inappropriate concentration of a single transcription factor can be responsible for at least some gastrulation defects of the nucleocytoplasmic hybrid embryos (Narbonne *et al.* 2011), it is very likely that misregulation of 100 different mRNAs, which are developmentally-important and amongst which many are transcription factors, is indeed responsible for the developmental failure of spermatid-derived embryos. However, as mentioned before, the definite experimental evidence to support this statement is lacking and it is highly unlikely it would be technically possible to provide such evidence, at least with the currently available protocols.

8.2. Model for epigenetic programming of the sperm nucleus

My results show that expression of developmentally-important mRNAs is altered in spermatid-derived embryos as compared to sperm-derived embryos and that the majority of these mRNAs are upregulated in spermatid-derived embryos. With the help of Dr Kei Miyamoto and Dr Jerome Jullien I showed that the presence of H3K27me3 marks at fertilisation is necessary to prevent gene overexpression in embryos. This led me to the hypothesis that perhaps sperm has on its chromatin repressive H3K27me3 marks, which are absent on the spermatid chromatin, hence gene overexpression is observed in spermatid-derived, but not in sperm-derived embryos. However, ChIP-seq results showed that upregulation of genes in spermatid-derived embryos could not be explained by the absence of repressive H3K27me3 marks in spermatids, since these marks were enriched at misregulated genes in both sperm and spermatids. Instead, upregulation of misregulated genes correlated with the presence of higher levels of activating H3K4me2/3 marks in spermatid than in sperm.

These findings suggest that proper regulation of embryonic gene expression is ensured in the sperm chromatin by two separate epigenetic layers: the presence of repressive H3K27me3 marks and a simultaneous depletion of activating H3K4me2/3 marks (Fig. 54). The spermatid already has the repressive H3K27me3, but does not yet lose activating marks from the misregulated genes (Fig. 54). Therefore, spermatids, as opposed to sperm, are not epigenetically mature to support correct gene expression in the embryo due to the retention of activating marks on their chromatin. These findings are in agreement with reports in mouse: in contrast to H3K27me3, which is similar between sperm and spermatids, H3K4me3 levels are reduced in sperm (Erkek *et al.* 2013).

Curiously, the results described in this thesis show a similar low efficiency of normal development between spermatid-derived embryos and nuclear transfer-derived embryos (as compared with sperm-derived embryos). In others words, spermatid had a similarly low developmental potential as an early embryonic cell. Furthermore, it was shown in nuclear transfer experiments conducted in *Xenopus laevis* that ectopic expression of donor-cell genes in nuclear transfer-derived embryos was associated with the presence of H3K4me3 at these gene promoters in such embryos (Ng & Gurdon 2005, Ng & Gurdon 2008), as compared to control fertilised embryos. This is again similar to our observations that spermatids, as compared to sperm, had more of activating H3K4me2/3 marks on chromatin which correlated with the overexpression of embryonic genes in spermatid-derived embryos. All these results support the hypothesis that sperm is epigenetically-programmed to regulate the expression of embryonic genes.

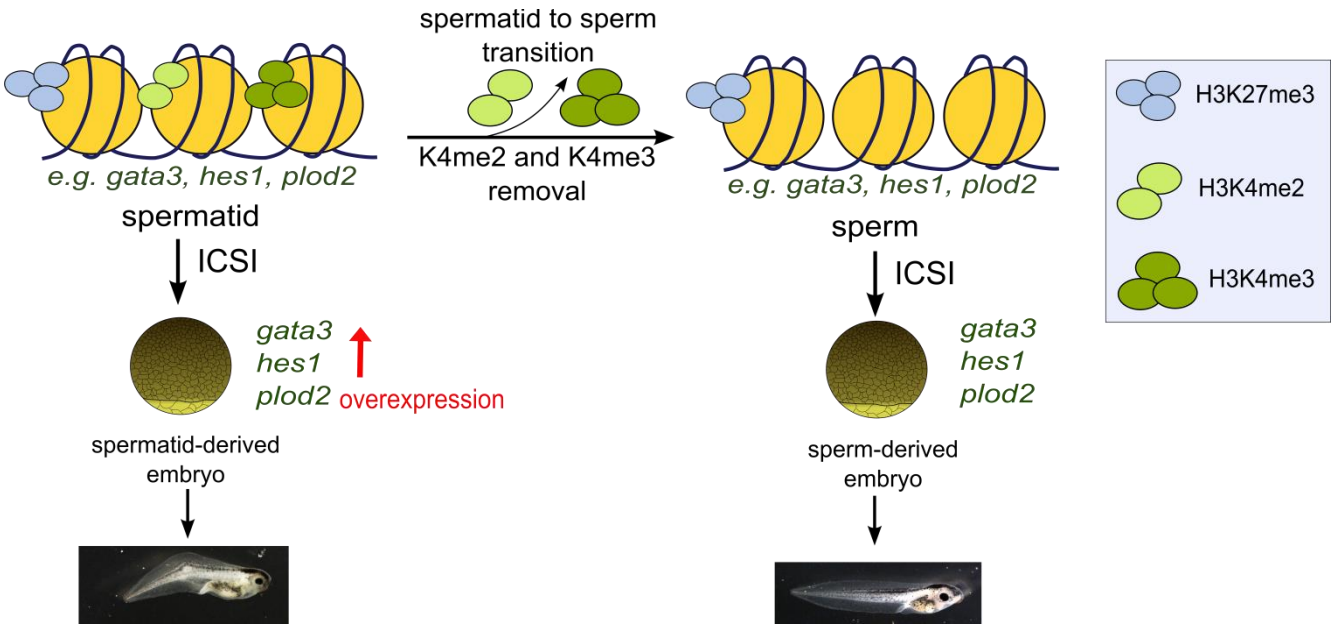


Fig. 54. Model for epigenetic programming of sperm nucleus for embryonic development.

The presence of H3K27me3 and the absence of H3K4me2/3 in sperm is important for a proper regulation of embryonic gene expression. Model is explained on examples of several misregulated genes *(gata3*, *hes1* and *plod2)*, which are upregulated in spermatid-derived embryos. In spermatids, these genes have H3K27me3 and H3K4me2/3 marks. Repressive H3K27me3 marks are retained in the chromatin of mature sperm, but activating marks H3K4me2/3 are lost. Aberrant retention of activating marks at the misregulated genes in spermatids correlates with overexpression of these genes in spermatid-derived embryos and with abnormal development of spermatid-derived embryos.

The remaining question is why the activating H3K4me2/3 marks were present at misregulated genes in the spermatid chromatin. As discussed earlier, misregulated genes were enriched for developmentally-important ones and many of them were transcription factors important for embryogenesis. Therefore, it is unlikely that those genes would be expressed in spermiogenesis. However, to exclude this possibility, I have performed a qRT-PCR analysis for selected misregulated genes (*mn1*, *hes1*, *gata3* and *sfrp2*) to test whether they were expressed in spermatids. These genes were identified by ChIP-seq as having H3K4me3 peaks exclusively in spermatids and also as having lower quantity of H3K4me2 methylation level in sperm than in spermatids. As expected, the results of the qRT-PCR analysis confirmed that

these genes were not expressed in spermatids (data not shown). Therefore, active expression of these genes in spermatids cannot explain the presence of H3K4me2/3 in spermatids.

Interestingly, it was shown that in spermatids, a global acetylation of histones is observed (Hazzouri *et al.* 2000), which was suggested to be involved in chromatin rearrangements that allow its ultimate compaction in several of ways. First, it was shown that acetylated histones can be recognised and bound by Brdt – bromodomain, testis-specific protein, and it was shown that such binding by Brdt leads to the chromatin compaction (Pivot-Pajot *et al.* 2003, Govin *et al.* 2006, Moriniere *et al.* 2009). Secondly, acetylation (but not their polyubiquitination) triggered degradation of core histones during spermiogenesis by testis-specific proteasomes (Qian *et al.* 2013), presumably facilitating protamine deposition on chromatin, which are the main chromatin component of the mature mammalian sperm. Last, it was also suggested that global acetylation increases the instability of the canonical histone-structured chromatin during spermiogenesis, therefore facilitating the incorporation of protamines (Gaucher *et al.* 2010). Therefore, it is plausible to think that the presence of activating H3K4me2/3 marks at genes in spermatids, similarly to what was reported for histone acetylation, does not necessarily reflect their transcriptional activity, but that it is important for structural changes occurring during spermatid to sperm transition. For example, analogously to what has been suggested for histone acetylation, methylation of H3K4 could create a more open chromatin structure in the spermatid, which in turn could facilitate the access of sperm basic proteins to chromatin and as a result, lead to a chromatin remodelling which allows its compaction in sperm. Then, if such H3K4me2/3-modified chromatin of spermatid is delivered to an egg, the egg, which does not have sperm basic proteins in the cytoplasm and other machinery required for remodelling of the chromatin to the sperm-like state, could interpret the H3K4me2/3 as transcriptionally-activating marks. This in turn

would lead to overexpression of genes having such marks, as indeed observed in spermatid-derived embryos.

8.3. How are the epigenetic marks transmitted to the embryo?

Another interesting aspect emerging from this study, which was only partially addressed in this thesis, is how the epigenetic marks present on sperm are transmitted to the embryo. In collaboration with other members of the Gurdon laboratory, I showed that the experimental removal of H3K27me3 marks at fertilisation had a pronounced effect on gene expression, whereas removal of the very same mark later during embryonic development did not have such effect (Fig. 41). In early *Xenopus* embryos the first twelve cell cycles are rapid, with no gene transcription. Only after reaching the mid-blastula transition stage, is the zygotic genome activated (Newport & Kirschner 1982). It was shown that in species in which such rapid cell cycle phases precede zygotic genome activation, histone post-translational modifications, such as H3K4me3 or H3K27me3 were not detected on chromatin during these rapid cell cycles (Akkers *et al.* 2009, Vastenhouw *et al.* 2010, Lindeman *et al.* 2011). Post-translational epigenetic marks were established again on chromatin only around the time of zygotic genome activation (Lindeman *et al.* 2011). It is possible that the presence of post-translational histone marks and the need for their re-establishment after each cell cycle would impede rapid DNA replication phases. During S-phase in *Drosophila* embryos H3K4me3 and H3K27me3 marks disappear completely from the chromatin; to re-appear again only in G2 phase. Interestingly, in *Drosophila* the enzymes that carry out these modifications: Trithorax and Enhancer-of-Zeste, respectively, remain associated with chromatin during S-phase. This suggests that the enzymes that modify histone tails (chromatin ‘writers’) also act as

placeholders for the actual histone marks (Petruk *et al.* 2012). Nascent chromatin capture experiments (performed by a pull down of chromatin at different phases of replication) showed that the enzymatic machinery associated with deposition of H3K27me3 mark on chromatin is indeed stably bound to chromatin in all cell cycle phases (Alabert *et al.* 2014). Interestingly, it was reported that Ezh2 (enhancer of zeste homolog 2), the enzyme that catalyses H3K27me3 deposition, was associated with chromatin in *Xenopus tropicalis* embryos and that this preceded H3K27me3 deposition (van Heeringen *et al.* 2014). This could suggest that writers of epigenetic marks could be, at the same time, their placeholders during the rapid cell cycle phases in the early embryo. A placeholder model fits well with the experimental data reported in this thesis – removal of epigenetic marks at fertilisation, but not after the mark has been already recognised by the placeholder, leads to the misregulation of gene expression (Fig. 41). It is also important to note that the removal of H3K27me3 marks late in embryogenesis, does not affect gene expression (Fig. 41). It is likely that in those conditions egg-derived Ezh2 is correctly targeted to genes, via binding to parentally-inherited H3K27me3, and can later on exert repression of these genes through an H3K27me3-independent mechanism.

An alternative possibility to the one described above is that the marks are present in the embryo all the time, faithfully recapitulated from the parental chromatin even in the rapid cell cycle phases, and the only time when they are absent from the chromatin is the S-phase itself. Then, due to the fact that the cell cycle phases in the early embryos are rapid, with almost no G1 and G2 phases (almost exclusively S-phases followed by mitoses) (Newport & Kirschner 1982), by probing the chromatin isolated from the early embryo, one would almost always look into the chromatin undergoing active DNA replication, and therefore devoid of epigenetic marks. Furthermore, it has been also shown that during the rapid cell cycle phases in zebrafish, the nucleosomes are not well positioned. Canonical nucleosome organisation

was only achieved around the time of zygotic genome activation (Zhang *et al.* 2014). Therefore, a globally low amount of well-positioned nucleosomes, plus the fact that early embryos are most of the time in S-phase, would prevent post-translational epigenetic marks from being detected at the earliest stages of embryogenesis, even if they were indeed present on the chromatin.

How to distinguish between the two hypotheses? One would ideally probe for the presence of the marks at different phases of the cell cycle. *In vitro* egg extracts in *Xenopus laevis* make such experiments possible. With the use of such egg extracts it was demonstrated that DNA methylation from the sperm chromatin is faithfully recapitulated after DNA replication and the molecular mechanism of DNA methylase targeting to newly replicated, hemimethylated DNA was revealed (Nishiyama *et al.* 2013). It would be important to perform such experiments to examine what is happening with histone marks. The use of egg extracts would allow one to test for the first time what happens to sperm-derived epigenetically marked histones after fertilisation. So far the only experimental data providing hints that sperm-derived modified histones may be retained after fertilisation come from immunostainings in mammalian systems (van der Heijden *et al.* 2006, van der Heijden *et al.* 2008), which however do not provide any information about the localisation of the retained marks or how global is this retention. A substantial advantage of using the *Xenopus laevis* egg extracts is the possibility of precisely controlling and monitoring the progression of the cell cycle – one could collect the samples at various time points from the start of the egg extract treatment (various phases of the cell cycle). Also, with the use of egg extracts it is possible to perform ChIP-qPCR and ChIP-seq analyses, to identify the localisation of the retained marks on the paternal chromatin. Currently it is not known whether at the time of fertilisation the egg replaces all sperm-derived nuclear proteins (including post-translationally histones) with its own, maternally-stored histones, or whether sperm-derived, epigenetically

marked histones are retained after this global chromatin remodelling. Also, even if such sperm-derived modified histones would survive the chromatin remodelling after the fertilisation, it is not known whether their marks would be reproduced after the DNA replication. Therefore, a precise timing of egg extract treatment would not only allow one to provide answers to these questions but also would enable discrimination between the two processes. Furthermore, the use of egg extracts creates the opportunity of a relatively easy examination of mechanisms allowing the possible retention of histone marks. One could perform immunodepletion of certain candidate factors from the extracts and examine the effects of their removal on the retention of histone marks. Experiments with the use of egg extracts aiming to determine what happens to H3K4me3 and H3K27me3 marks after fertilisation and during the first cell cycle, together with uncovering the underlying mechanisms of possible histone mark retention, are currently being designed in the Gurdon laboratory as follow-up studies on the results described in this thesis.

It would be also extremely important to track the epigenetic marks during the early phases of embryonic development *in vivo*. Even though egg extracts were shown to be able to recapitulate the early cell cycle events, a demonstration that the same events occur in embryos *in vivo* would be still required. Using *Xenopus* as a model system provides the advantage of almost unlimited material available for ChIP (or ChIP-seq) experiments. My preliminary experiments with ChIP for H3K4me3 followed by qPCR using *Xenopus laevis* gastrula stage embryos revealed that as little at 10 embryos give one a sufficient amount of material to reliably detect this histone mark (data not shown). Therefore, it would be interesting to test whether the marks observed on the chromatin of sperm and spermatids are indeed retained in the developing embryos (and whether the differences observed between sperm and spermatids between them are maintained between sperm- and spermatid-derived embryos).

Last, in the experiments that were designed to functionally test the importance of H3K27me3 by a K6B-mediated removal of these marks at fertilisation, the enzyme removed the marks from both the maternal and the paternal chromatin sets. Furthermore, the effects of H3K27me3 removal had apparent effects on gene expression only when it was present in the egg already at the time of fertilisation (Fig. 41). The mRNA was injected into fully-grown oocytes (isolated from PMSG-treated frogs), which are transcriptionally silent (Gilbert 2010), therefore it is unlikely that any major transcriptional events in the oocyte were disturbed by the H3K27me3 removal from the maternal chromatin before the sperm injection. However, this possibility cannot be excluded. Therefore, in order to be able to state precisely that the observed changes in transcription are due to the epigenetic mark removal at fertilisation, one would ideally examine the effects of delivering to the egg a paternal chromatin devoid of such marks. Also, in an ideal situation, such experiment would be performed in haploid embryos. As mentioned in the results section (Chapter 7.2), with the currently available protocols involving *in vitro* maturation of mRNA-injected oocytes, followed by ICSI, it is not possible to generate haploid, paternally-derived embryos. Therefore, the use of IVM/ICSI procedure does not allow the assessment of the effects of the histone mark removal at fertilisation solely on the paternal chromatin. However, a new experimental strategy is currently being designed in the Gurdon laboratory as a follow up of these experiments, in which recombinant histone demethylases are used. Such enzymes are able to remove histone marks from synthetic methylated peptides *in vitro,* therefore in theory they should be also able to remove such marks from a chromatin template. It is still technically impossible to remove the marks from the sperm chromatin, as it is highly condensed and any attempts of chromatin loosening lead to DNA dispersal (data not shown). However, Dr Jerome Jullien from the Gurdon laboratory is currently optimising the treatment of spermatids with histone demethylases *in vitro.* If such approach gives promising results, it should be possible to test

in the future whether the removal of histone marks solely on the paternal (spermatid) chromatin affects gene expression in the embryo. It would be especially interesting to test whether removal of H3K4me2/3 activating marks from the spermatid chromatin would be able to rescue (downregulate) the upregulation of misregulated genes in spermatid-derived embryos.

In the future (when a better annotation of the *Xenopus laevis* genome is available) it would be worth trying to target the enzymes that modify the chromatin (for example demethylases that remove H3K4me2/3 marks) to specific loci in the genome with the use of CRISPR-Cas9 system (Mali *et al.* 2013). This system is naturally used by bacteria as a defence mechanism against foreign nucleic acids. It is based on guide RNAs that bring the Cas9 nuclease to the foreign nucleotides for their degradation (Mali *et al.* 2013). Such guide RNAs can be custom designed to target any desired sequence in the genome and therefore this strategy has been widely applied to obtain gene knock-outs in many different systems (Cho *et al.* 2013, Cong *et al.* 2013, Friedland *et al.* 2013). Interestingly, it has been shown that using custom-designed targeting RNAs and nuclease-dead Cas9 fused to the proteins of interest it was possible to modify the levels of gene expression of the desired gene. For example, it was shown that by creating a fusion between a nuclease-dead Cas9 and a transcriptional transactivator protein and by simultaneously expressing in human cells appropriate guide RNAs, it was possible to selectively activate desired endogenous target genes (Perez-Pinera *et al.* 2013). Conversely, it was also shown that a another type of Cas9 modification can result in gene repression in bacteria, human and yeast by blocking transcriptional initiation and/or elongation (Gilbert *et al.* 2013, Qi *et al.* 2013), confirming that Cas9 system can be successfully applied not only for genome editing, but also for modulating gene expression. It would be therefore interesting to fuse epigenetic modifiers (for example H3K4me2/3 demethylase) to nuclease-dead Cas9 to try rescuing

(downregulating) the gene expression in spermatid-derived embryos. Doing such experiment in a targeted way (targeting all the 100 misregulated genes by injecting appropriate guide mRNAs into 1-cell stage embryos), would eliminate side effects which would occur if the enzyme would act on the whole chromatin.

8.4. What are the epigenetic changes occurring during spermiogenesis in *Xenopus laevis*?

ChIP-seq experiments described in this thesis probed for repressive H3K27me3 and for activating H3K4me2/3 marks in sperm and spermatids. Even though this experiments provided interesting and unexpected findings (H3K27me3 did not change between spermatids and sperm, whereas the activating H3K4me2/3 marks were more abundant in spermatids, which correlated well with the gene upregulation observed in spermatid-derived embryos), one has to realise that the picture obtained is far from complete. As mentioned in the introduction chapter, there are multiple epigenetic modifications known on DNA: DNA methylation, DNA hydroxymethylation, DNA formylation, DNA carboxylation; and even more histone modifications: methylation, acetylation, phosphorylation, ubiquitylation, sumoylation, ADP ribosylation, crotonylation and many others. Moreover, histone modifications occur on many different residues and different histones, for example both serine and threonine can become phosphorylated, or histones H3 and H4 can both be acetylated (Dawson & Kouzarides 2012). Plus, there are many histone variants existing and also, the positioning of histones itself can affect the chromatin structure. All these epigenetic changes affect each other and they can also attract different binding partners, which can further alter the transcriptional outcome of the target gene (Kouzarides 2007). On top of that,

it has been shown that protamines in mammalian sperm are also subjected to post-translational modifications (Brunner *et al.* 2014). All these suggest that the possible combination of various epigenetic marks that could affect the chromatin and the transcriptional state of a given gene is enormous. It is for sure not possible to probe for all the known DNA modifications, histone marks and histone variants. However, the more epigenetic profiles of spermatids and sperm are obtained, the more data can be utilised as an input into the model aiming at predicting gene expression status in the embryo based on the epigenetic states of the paternal chromatin.

Experiments currently conducted in the Gurdon laboratory aim to obtain epigenetic profiles of DNA methylation and also trimethylation of lysine 9 of histone H3 (H3K9me3) in sperm and spermatids. Investigations into these marks are of a particular interest, as the proteomic analysis of egg factors binding specifically to sperm, and not to the spermatid chromatin, identified HP1γ and Mbd3 proteins (Fig. 15). These proteins recognise and bind to H3K9me2/3 modification and to methylated cytosine in DNA, respectively. This therefore suggests that those two epigenetic marks: H3K9me3 and DNA methylation are likely to be different between sperm and spermatids. Furthermore, finding the functional connection between the epigenetic marks in sperm and spermatids and their readers from the egg side will provide a valuable addition to the current understanding of the inheritance of the epigenetic marks via the gametes.

The results described in this thesis also identify a unique chromatin structure in *Xenopus laevis* sperm. I showed that MNase digestion of chromatin leads to a release of DNA of three different sizes: ~75bp, ~110bp and ~150bp (Fig. 43). I also showed that ~110bp structure is the most unstable one, as it disappears first upon a treatment with an increasing concentration of a heparin (Fig. 46). What is the identity of these three different structures? What are they composed of? Are the main components of these structures histones, sperm

basic proteins or perhaps a mixture of both? Do they contain any specific sequence features? For example, the canonical nucleosome size, ~150bp structure could contain gene promoters or developmentally-important genes; or, the ~75bp structure could have mostly intronic sequences or housekeeping genes. The experiments currently ongoing in the Gurdon laboratory (in collaboration with Dr Angela Simeone and Dr Jerome Jullien) aim to address these questions by performing sequence analysis of DNA contained in each of these fragments and by chromatin immunopurifications against histones.

8.5. How the results obtained add to the current knowledge on sperm programming?

It is currently known that sperm of different species can bear epigenetic marks on DNA and also on histones (Hammoud *et al.* 2009, Brykczynska *et al.* 2010, Wu *et al.* 2011). It was correlated that genes bearing the repressive marks tend to be repressed during the earliest developmental stages (Brykczynska *et al.* 2010), whereas those with activating marks tend to be expressed early in embryogenesis (Wu *et al.* 2011). What is the novelty of the findings described in this thesis? First, it is not possible to manipulate the sperm nucleus to change the histone marks on chromatin. Therefore, comparing sperm and spermatids in the same type of assay allows bypassing the problem of inaccessibility of the sperm chromatin. By using these two different cell types, coming from the same lineage, having the same DNA and chromosome content, but differing in their developmental potential I was able to assess for the first time what is the source of developmental advantage of the sperm. I showed that the developmental advantage of sperm over the spermatids is likely related to the ability to correctly regulate embryonic gene expression. I also managed to link this ability to underlying differences in histone marks between sperm and spermatids: spermatids, which

overexpress the misregulated genes, have more of activating H3K4me2/3 marks. These results largely extend the currently available knowledge in the field, since even though I was not able to directly manipulate the epigenetic status of the sperm chromatin, I was able to compare the sperm to its direct precursor, that has a different epigenetic state. Furthermore, enzymatic removal of H3K27me3 marks from the parental chromatin at fertilisation proved for the first time that the presence of an epigenetic mark on the parental chromatin is indeed required for a correct regulation of embryonic gene expression. As such, the experiments presented in this thesis significantly advance our current understanding of the sperm programming. Furthermore, presented experiments became a basis for many follow-up projects currently carried in the Gurdon laboratory (as discussed above), therefore I strongly hope that the obtained results will also have interesting future implications.

8.6. Results described in this thesis in the context of the current knowledge on transgenerational inheritance of epigenetic information.

Concluding, the results described in this thesis provide experimental evidence supporting the hypothesis that the sperm is epigenetically programmed to regulate embryonic gene expression. This hypothesis is further strengthened by the fact that an incorrect pattern of histone modifications or DNA methylation was associated with cases of idiopathic infertility in humans (Hammoud *et al.* 2010, Hammoud *et al.* 2011). Furthermore, it was reported in mammals that epigenetic traits can be transgenerationally inherited from the father to the offspring (Braunschweig *et al.* 2012, Daxinger & Whitelaw 2012, Lambrot *et al.* 2013, Padmanabhan *et al.* 2013, Vassoler *et al.* 2013, Dias & Ressler 2014). For example, it was shown in mice that offspring of males fed on a low-protein diet had elevated expression of many hepatic genes, which was linked to changes in DNA methylation at the promoters of

these genes (Carone *et al.* 2010). In another example it was shown that paternal obesity in mice is transgenerationally inherited up to F2 generation and that this phenomenon was correlated with global changes in DNA methylation patterns and with altered expression profiles of mRNAs and microRNAs in the testes of the F0 obese male mouse (Fullston *et al.* 2013). There is therefore growing evidence for the existence and importance of transgenerational inheritance of epigenetic traits via the gametes. Not surprisingly, this subject receives more and more of the media attention, as the possibility that the environmental cues acting on the sperm/sperm progenitors can be transmitted and affect the phenotype of the offspring is revolutionising the current view based on genetic mutations as the main source of evolutionary adaptation (Grossniklaus *et al.* 2013, Hughes 2014, Kaiser 2014, Szyf 2014). Further research into the subject is needed not only because of its impact on the basic biology and our understanding of the non-genetic inheritance, but also because it is necessary to get insight into the potential effects of various environmental cues on the future fitness of the offspring from the perspective of human health. This is especially important as many of the reports on the transgenerational inheritance of epigenetic traits in animals point to the inheritance of metabolic adaptations. For example, it is crucial to understand the consequences of the diet on the health of the future offspring, especially in light of the growing numbers of obesity cases in humans. Furthermore, it was also reported that exposure of pregnant female rats to commonly used environmental toxins: vinclozin (fungicide) or methoxychlor (insecticide) induced transgenerational inheritance of reduced male fertility (complete infertility in 8% of the cases) in male mice in all subsequent generations tested, which was correlated with changes in global DNA methylation patterns transmitted through the male germline (Anway *et al.* 2005). Therefore, a better understanding of agents causing the heritable changes in the germline epigenome, as well as the mechanisms of the inheritance of such changes is clearly needed.

I hope that the results presented in this thesis will help our understanding of the mechanisms of transgenerational inheritance of epigenetic traits in general. I believe that *Xenopus*, in which it is possible to generate haploid sperm- and spermatid-derived embryos, to assess their developmental potential, transcriptional capacity, and manipulate the epigenetic state of chromatin, is a well suited model organism for further investigations of the nature of the sperm programming.

Chapter 9

References

Abe S 1988 Cell-Culture of Spermatogenic Cells from Amphibians. *Development Growth & Differentiation* **30** 209-218.

Abe S & Hiyoshi H 1991 Synthesis of sperm-specific basic nuclear proteins (SPs) in cultured spermatids from Xenopus laevis. *Exp Cell Res* **194** 90-94.

Agger K, Cloos PA, Christensen J, Pasini D, Rose S, Rappsilber J, Issaeva I, Canaani E, Salcini AE & Helin K 2007 UTX and JMJD3 are histone H3K27 demethylases involved in HOX gene regulation and development. *Nature* **449** 731-734.

Akkers RC, van Heeringen SJ, Jacobi UG, Janssen-Megens EM, Francoijs KJ, Stunnenberg HG & Veenstra GJ 2009 A hierarchy of H3K4me3 and H3K27me3 acquisition in spatial gene regulation in Xenopus embryos. *Dev Cell* **17** 425-434.

Alabert C, Bukowski-Wills JC, Lee SB, Kustatscher G, Nakamura K, de Lima Alves F, Menard P, Mejlvang J, Rappsilber J & Groth A 2014 Nascent chromatin capture proteomics determines chromatin dynamics during DNA replication and identifies unknown fork components. *Nat Cell Biol* **16** 281-293.

Alexa A, Rahnenfuhrer J & Lengauer T 2006 Improved scoring of functional groups from gene expression data by decorrelating GO graph structure. *Bioinformatics* **22** 1600-1607.

Alexeyenko A, Tamas I, Liu G & Sonnhammer EL 2006 Automatic clustering of orthologs and inparalogs shared by multiple proteomes. *Bioinformatics* **22** e9-15.

Ang YS, Tsai SY, Lee DF, Monk J, Su J, Ratnakumar K, Ding J, Ge Y, Darr H, Chang B, Wang J, Rendl M, Bernstein E, Schaniel C & Lemischka IR 2011 Wdr5 mediates self-renewal and reprogramming via the embryonic stem cell core transcriptional network. *Cell* **145** 183-197.

Anway MD, Cupp AS, Uzumcu M & Skinner MK 2005 Epigenetic transgenerational actions of endocrine disruptors and male fertility. *Science* **308** 1466-1469.

Aoki F, Worrad DM & Schultz RM 1997 Regulation of transcriptional activity during the first and second cell cycles in the preimplantation mouse embryo. *Dev Biol* **181** 296-307.

Arpanahi A, Brinkworth M, Iles D, Krawetz SA, Paradowska A, Platts AE, Saida M, Steger K, Tedder P & Miller D 2009 Endonuclease-sensitive regions of human spermatozoal chromatin are highly enriched in promoter and CTCF binding sequences. *Genome Res* **19** 1338-1349.

Bai L & Morozov AV 2010 Gene regulation by nucleosome positioning. *Trends Genet* **26** 476-483.

Balhorn R 2007 The protamine family of sperm nuclear proteins. *Genome Biol* **8** 227.

Banaszynski LA, Wen D, Dewell S, Whitcomb SJ, Lin M, Diaz N, Elsasser SJ, Chapgier A, Goldberg AD, Canaani E, Rafii S, Zheng D & Allis CD 2013 Hira-dependent histone H3.3 deposition facilitates PRC2 recruitment at developmental loci in ES cells. *Cell* **155** 107-120.

Bannister AJ, Zegerman P, Partridge JF, Miska EA, Thomas JO, Allshire RC & Kouzarides T 2001 Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. *Nature* **410** 120-124.

Barski A, Cuddapah S, Cui K, Roh TY, Schones DE, Wang Z, Wei G, Chepelev I & Zhao K 2007 High-resolution profiling of histone methylations in the human genome. *Cell* **129** 823-837.

Becker M, Becker A, Miyara F, Han Z, Kihara M, Brown DT, Hager GL, Latham K, Adashi EY & Misteli T 2005 Differential in vivo binding dynamics of somatic and oocyte-specific linker histones in oocytes and during ES cell nuclear transfer. *Mol Biol Cell* **16** 3887-3895.

Benvenuto G, Carpentieri ML, Salvatore P, Cindolo L, Bruni CB & Chiariotti L 1996 Cell-specific transcriptional regulation and reactivation of galectin-1 gene expression are controlled by DNA methylation of the promoter region. *Mol Cell Biol* **16** 2736-2743.

Bernstein BE, Humphrey EL, Erlich RL, Schneider R, Bouman P, Liu JS, Kouzarides T & Schreiber SL 2002 Methylation of histone H3 Lys 4 in coding regions of active genes. *Proc Natl Acad Sci U S A* **99** 8695-8700.

Bernstein BE, Kamal M, Lindblad-Toh K, Bekiranov S, Bailey DK, Huebert DJ, McMahon S, Karlsson EK, Kulbokas EJ, 3rd, Gingeras TR, Schreiber SL & Lander ES 2005 Genomic maps and comparative analysis of histone modifications in human and mouse. *Cell* **120** 169-181.

Bhaskara S, Jacques V, Rusche JR, Olson EN, Cairns BR & Chandrasekharan MB 2013 Histone deacetylases 1 and 2 maintain S-phase chromatin and DNA replication fork progression. *Epigenetics Chromatin* **6** 27.

Bitgood MJ & McMahon AP 1995 Hedgehog and Bmp genes are coexpressed at many diverse sites of cell-cell interaction in the mouse embryo. *Dev Biol* **172** 126-138.

Blow JJ & Laskey RA 1986 Initiation of DNA replication in nuclei and purified DNA by a cell-free extract of Xenopus eggs. *Cell* **47** 577-587.

Bochkarev A, Pfuetzner RA, Edwards AM & Frappier L 1997 Structure of the single-stranded-DNA-binding domain of replication protein A bound to DNA. *Nature* **385** 176-181.

Bogdanovic O & Veenstra GJ 2009 DNA methylation and methyl-CpG binding proteins: developmental requirements and function. *Chromosoma* **118** 549-565.

Boiani M, Eckardt S, Scholer HR & McLaughlin KJ 2002 Oct4 distribution and level in mouse clones: consequences for pluripotency. *Genes Dev* **16** 1209-1219.

Bonasio R, Tu S & Reinberg D 2010 Molecular signals of epigenetic states. *Science* **330** 612-616.

Bortvin A, Eggan K, Skaletsky H, Akutsu H, Berry DL, Yanagimachi R, Page DC & Jaenisch R 2003 Incomplete reactivation of Oct4-related genes in mouse embryos cloned from somatic nuclei. *Development* **130** 1673-1680.

Bouniol-Baly C, Nguyen E, Besombes D & Debey P 1997 Dynamic organization of DNA replication in one-cell mouse embryos: relationship to transcriptional activation. *Exp Cell Res* **236** 201-211.

Bourc'his D, Le Bourhis D, Patin D, Niveleau A, Comizzoli P, Renard JP & Viegas-Pequignot E 2001 Delayed and incomplete reprogramming of chromosome methylation patterns in bovine cloned embryos. *Curr Biol* **11** 1542-1546.

Boyes J & Bird A 1991 DNA methylation inhibits transcription indirectly via a methyl-CpG binding protein. *Cell* **64** 1123-1134.

Boyes J & Bird A 1992 Repression of genes by DNA methylation depends on CpG density and promoter strength: evidence for involvement of a methyl-CpG binding protein. *EMBO J* **11** 327-333.

Braunschweig M, Jagannathan V, Gutzwiller A & Bee G 2012 Investigations on transgenerational epigenetic response down the male line in F2 pigs. *PLoS One* **7** e30583.

Brunner AM, Nanni P & Mansuy IM 2014 Epigenetic marking of sperm by post-translational modification of histones and protamines. *Epigenetics Chromatin* **7** 2.

Brykczynska U, Hisano M, Erkek S, Ramos L, Oakeley EJ, Roloff TC, Beisel C, Schubeler D, Stadler MB & Peters AH 2010 Repressive and active histone methylation mark distinct promoters in human and mouse spermatozoa. *Nat Struct Mol Biol* **17** 679-687.

Bui HT, Seo HJ, Park MR, Park JY, Thuan NV, Wakayama T & Kim JH 2011 Histone deacetylase inhibition improves activation of ribosomal RNA genes and embryonic nucleolar reprogramming in cloned mouse embryos. *Biol Reprod* **85** 1048-1056.

Cao R, Wang L, Wang H, Xia L, Erdjument-Bromage H, Tempst P, Jones RS & Zhang Y 2002 Role of histone H3 lysine 27 methylation in Polycomb-group silencing. *Science* **298** 1039-1043.

Carone BR, Fauquier L, Habib N, Shea JM, Hart CE, Li R, Bock C, Li C, Gu H, Zamore PD, Meissner A, Weng Z, Hofmann HA, Friedman N & Rando OJ 2010 Paternally induced transgenerational

environmental reprogramming of metabolic gene expression in mammals. *Cell* **143** 1084-1096.

Cary PD, Crane-Robinson C, Bradbury EM & Dixon GH 1982 Effect of acetylation on the binding of N-terminal peptides of histone H4 to DNA. *Eur J Biochem* **127** 137-143.

Chahal SS, Matthews HR & Bradbury EM 1980 Acetylation of histone H4 and its role in chromatin structure and function. *Nature* **287** 76-79.

Chan MM, Smith ZD, Egli D, Regev A & Meissner A 2012 Mouse ooplasm confers context-specific reprogramming capacity. *Nat Genet* **44** 978-980.

Chen C, Yang MC & Yang TP 2001a Evidence that silencing of the HPRT promoter by DNA methylation is mediated by critical CpG sites. *J Biol Chem* **276** 320-328.

Chen GC, Guan LS, Yu JH, Li GC, Choi Kim HR & Wang ZY 2001b Rb-associated protein 46 (RbAp46) inhibits transcriptional transactivation mediated by BRCA1. *Biochem Biophys Res Commun* **284** 507-514.

Chen X, Hu T, Liang G, Yang M, Zong S, Miao S, Koide SS & Wang L 2008 A novel testis protein, RSB-66, interacting with INCA1 (inhibitor of Cdk interacting with cyclin A1). *Biochem Cell Biol* **86** 345-351.

Cho C, Jung-Ha H, Willis WD, Goulding EH, Stein P, Xu Z, Schultz RM, Hecht NB & Eddy EM 2003 Protamine 2 deficiency leads to sperm DNA damage and embryo death in mice. *Biol Reprod* **69** 211-217.

Cho SW, Kim S, Kim JM & Kim JS 2013 Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease. *Nature biotechnology* **31** 230-232.

Chowdhury I, Thompson WE, Welch C, Thomas K & Matthews R 2013 Prohibitin (PHB) inhibits apoptosis in rat granulosa cells (GCs) through the extracellular signal-regulated kinase 1/2 (ERK1/2) and the Bcl family of proteins. *Apoptosis* **18** 1513-1525.

Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, Hsu PD, Wu X, Jiang W, Marraffini LA & Zhang F 2013 Multiplex genome engineering using CRISPR/Cas systems. *Science* **339** 819-823.

Core LJ, Waterfall JJ & Lis JT 2008 Nascent RNA sequencing reveals widespread pausing and divergent initiation at human promoters. *Science* **322** 1845-1848.

Creyghton MP, Markoulaki S, Levine SS, Hanna J, Lodato MA, Sha K, Young RA, Jaenisch R & Boyer LA 2008 H2AZ is enriched at polycomb complex target genes in ES cells and is necessary for lineage commitment. *Cell* **135** 649-661.

Czermin B, Melfi R, McCabe D, Seitz V, Imhof A & Pirrotta V 2002 Drosophila enhancer of Zeste/ESC complexes have a histone H3 methyltransferase activity that marks chromosomal Polycomb sites. *Cell* **111** 185-196.

Dawson MA & Kouzarides T 2012 Cancer epigenetics: from mechanism to therapy. *Cell* **150** 12-27.

Daxinger L & Whitelaw E 2012 Understanding transgenerational epigenetic inheritance via the gametes in mammals. *Nat Rev Genet* **13** 153-162.

de Mateo S, Castillo J, Estanyol JM, Ballesca JL & Oliva R 2011 Proteomic characterization of the human sperm nucleus. *Proteomics* **11** 2714-2726.

de Rooij DG 2001 Proliferation and differentiation of spermatogonial stem cells. *Reproduction* **121** 347-354.

Dey A, Chitsaz F, Abbasi A, Misteli T & Ozato K 2003 The double bromodomain protein Brd4 binds to acetylated chromatin during interphase and mitosis. *Proc Natl Acad Sci U S A* **100** 8758-8763.

Dey A, Nishiyama A, Karpova T, McNally J & Ozato K 2009 Brd4 marks select genes on mitotic chromatin and directs postmitotic transcription. *Mol Biol Cell* **20** 4899-4909.

Dias BG & Ressler KJ 2014 Parental olfactory experience influences behavior and neural structure in subsequent generations. *Nat Neurosci* **17** 89-96.

Doerfler W 2008 In pursuit of the first recognized epigenetic signal--DNA methylation: a 1976 to 2008 synopsis. *Epigenetics* **3** 125-133.

Draker R, Ng MK, Sarcinella E, Ignatchenko V, Kislinger T & Cheung P 2012 A combination of H2A.Z and H4 acetylation recruits Brd2 to chromatin during transcriptional activation. *PLoS Genet* **8** e1003047.

du Chene I, Basyuk E, Lin YL, Triboulet R, Knezevich A, Chable-Bessia C, Mettling C, Baillat V, Reynes J, Corbeau P, Bertrand E, Marcello A, Emiliani S, Kiernan R & Benkirane M 2007 Suv39H1 and HP1gamma are responsible for chromatin-mediated HIV-1 transcriptional silencing and post-integration latency. *EMBO J* **26** 424-435.

Eggan K, Baldwin K, Tackett M, Osborne J, Gogos J, Chess A, Axel R & Jaenisch R 2004 Mice cloned from olfactory sensory neurons. *Nature* **428** 44-49.

Elling U, Klasen C, Eisenberger T, Anlag K & Treier M 2006 Murine inner cell mass-derived lineages depend on Sall4 function. *Proc Natl Acad Sci U S A* **103** 16319-16324.

Elsdale TR, Gurdon JB & Fischberg M 1960 A description of the technique for nuclear transplantation in Xenopus laevis. *J Embryol Exp Morphol* **8** 437-444.

Enninga J, Levay A & Fontoura BM 2003 Sec13 shuttles between the nucleus and the cytoplasm and stably interacts with Nup96 at the nuclear pore complex. *Mol Cell Biol* **23** 7271-7284.

Erdel F & Rippe K 2011 Chromatin remodelling in mammalian cells by ISWI-type complexes--where, when and why? *FEBS J* **278** 3608-3618.

Erkek S, Hisano M, Liang CY, Gill M, Murr R, Dieker J, Schubeler D, Vlag J, Stadler MB & Peters AH 2013 Molecular determinants of nucleosome retention at CpG-rich sequences in mouse spermatozoa. *Nat Struct Mol Biol* **20** 868-875.

Farcas AM, Blackledge NP, Sudbery I, Long HK, McGouran JF, Rose NR, Lee S, Sims D, Cerase A, Sheahan TW, Koseki H, Brockdorff N, Ponting CP, Kessler BM & Klose RJ 2012 KDM2B links the Polycomb Repressive Complex 1 (PRC1) to recognition of CpG islands. *Elife* **1** e00205.

Farthing CR, Ficz G, Ng RK, Chan CF, Andrews S, Dean W, Hemberger M & Reik W 2008 Global mapping of DNA methylation in mouse promoters reveals epigenetic reprogramming of pluripotency genes. *PLoS Genet* **4** e1000116.

Fontoura BM, Blobel G & Matunis MJ 1999 A conserved biogenesis pathway for nucleoporins: proteolytic processing of a 186-kilodalton precursor generates Nup98 and the novel nucleoporin, Nup96. *J Cell Biol* **144** 1097-1112.

Fontoura BM, Dales S, Blobel G & Zhong H 2001 The nucleoporin Nup98 associates with the intranuclear filamentous protein network of TPR. *Proc Natl Acad Sci U S A* **98** 3208-3213.

Frehlick LJ, Prado A, Calestagne-Morelli A & Ausio J 2007 Characterization of the PL-I-related SP2 protein from Xenopus. *Biochemistry* **46** 12700-12708.

Friedland AE, Tzur YB, Esvelt KM, Colaiacovo MP, Church GM & Calarco JA 2013 Heritable genome editing in C. elegans via a CRISPR-Cas9 system. *Nature methods* **10** 741-743.

Friedle H & Knochel W 2002 Cooperative interaction of Xvent-2 and GATA-2 in the activation of the ventral homeobox gene Xvent-1B. *J Biol Chem* **277** 23872-23881.

Fukuda A, Cao F, Morita S, Yamada K, Jincho Y, Tane S, Sotomaru Y & Kono T 2010 Identification of inappropriately reprogrammed genes by large-scale transcriptome analysis of individual cloned mouse blastocysts. *PLoS One* **5** e11274.

Fullston T, Ohlsson Teague EM, Palmer NO, DeBlasio MJ, Mitchell M, Corbett M, Print CG, Owens JA & Lane M 2013 Paternal obesity initiates metabolic disturbances in two generations of mice with incomplete penetrance to the F2 generation and alters the transcriptional profile of testis and sperm microRNA content. *FASEB J* **27** 4226-4243.

Gaggioli V, Le Viet B, Germe T & Hyrien O 2013 DNA topoisomerase IIalpha controls replication origin cluster licensing and firing time in Xenopus egg extracts. *Nucleic Acids Res*.

Ganier O, Bocquet S, Peiffer I, Brochard V, Arnaud P, Puy A, Jouneau A, Feil R, Renard JP & Mechali M 2011 Synergic reprogramming of mammalian cells by combined exposure to mitotic Xenopus egg extracts and transcription factors. *Proc Natl Acad Sci U S A* **108** 17331-17336.

Gao S, Chung YG, Parseghian MH, King GJ, Adashi EY & Latham KE 2004 Rapid H1 linker histone transitions following fertilization or somatic cell nuclear transfer: evidence for a uniform developmental program in mice. *Dev Biol* **266** 62-75.

Gaucher J, Boussouar F, Montellier E, Curtet S, Buchou T, Bertrand S, Hery P, Jounier S, Depaux A, Vitte AL, Guardiola P, Pernet K, Debernardi A, Lopez F, Holota H, Imbert J, Wolgemuth DJ, Gerard M, Rousseaux S & Khochbin S 2012 Bromodomain-dependent stage-specific male genome programming by Brdt. *EMBO J* **31** 3809-3820.

Gaucher J, Reynoird N, Montellier E, Boussouar F, Rousseaux S & Khochbin S 2010 From meiosis to postmeiotic events: the secrets of histone disappearance. *FEBS J* **277** 599-604.

Gaudet F, Hodgson JG, Eden A, Jackson-Grusby L, Dausman J, Gray JW, Leonhardt H & Jaenisch R 2003 Induction of tumors in mice by genomic hypomethylation. *Science* **300** 489-492.

Gilbert LA, Larson MH, Morsut L, Liu Z, Brar GA, Torres SE, Stern-Ginossar N, Brandman O, Whitehead EH, Doudna JA, Lim WA, Weissman JS & Qi LS 2013 CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. *Cell* **154** 442-451.

Gilbert SF 2010 *Developmental biology*. Sunderland, Mass.: Sinauer Associates.

Gillespie PJ & Blow JJ 2000 Nucleoplasmin-mediated chromatin remodelling is required for Xenopus sperm nuclei to become licensed for DNA replication. *Nucleic Acids Res* **28** 472-480.

Gillespie PJ, Gambus A & Blow JJ 2012 Preparation and use of Xenopus egg extracts to study DNA replication and chromatin associated proteins. *Methods* **57** 203-213.

Godde JS & Ura K 2008 Cracking the enigmatic linker histone code. *J Biochem* **143** 287-293.

Goldberg AD, Banaszynski LA, Noh KM, Lewis PW, Elsaesser SJ, Stadler S, Dewell S, Law M, Guo X, Li X, Wen D, Chapgier A, DeKelver RC, Miller JC, Lee YL, Boydston EA, Holmes MC, Gregory PD, Greally JM, Rafii S, Yang C, Scambler PJ, Garrick D, Gibbons RJ, Higgs DR, Cristea IM, Urnov FD, Zheng D & Allis CD 2010 Distinct factors control histone variant H3.3 localization at specific genomic regions. *Cell* **140** 678-691.

Govin J, Caron C, Lestrat C, Rousseaux S & Khochbin S 2004 The role of histones in chromatin remodelling during mammalian spermiogenesis. *Eur J Biochem* **271** 3459-3469.

Govin J, Lestrat C, Caron C, Pivot-Pajot C, Rousseaux S & Khochbin S 2006 Histone acetylation-mediated chromatin compaction during mouse spermatogenesis. *Ernst Schering Res Found Workshop*155-172.

Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q, Chen Z, Mauceli E, Hacohen N, Gnirke A, Rhind N, di Palma F, Birren BW, Nusbaum C, Lindblad-Toh K, Friedman N & Regev A 2011 Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat Biotechnol* **29** 644-652.

Green MR & Sambrook J 2012 *Molecular cloning : a laboratory manual*. Cold Spring Harbor, N.Y.: Cold Spring Harbor Laboratory Press.

Grossniklaus U, Kelly WG, Ferguson-Smith AC, Pembrey M & Lindquist S 2013 Transgenerational epigenetic inheritance: how important is it? *Nat Rev Genet* **14** 228-235.

Gurdon JB 1960 The Effects of Ultraviolet Irradiation on Uncleaved Eggs of *Xenopus laevis*. *Quarterly Journal of Microscopical Science* **101** 299-311.

Gurdon JB 1962 The developmental capacity of nuclei taken from intestinal epithelium cells of feeding tadpoles. *J Embryol Exp Morphol* **10** 622-640.

Gurdon JB 1976 Injected nuclei in frog oocytes: fate, enlargement, and chromatin dispersal. *J Embryol Exp Morphol* **36** 523-540.

Gurdon JB, Elsdale TR & Fischberg M 1958 Sexually mature individuals of Xenopus laevis from the transplantation of single somatic nuclei. *Nature* **182** 64-65.

Gurdon JB & Laskey RA 1970 The transplantation of nuclei from single cultured cells into enucleate frogs' eggs. *J Embryol Exp Morphol* **24** 227-248.

Gurdon JB & Uehlinger V 1966 "Fertile" intestine nuclei. *Nature* **210** 1240-1241.

Hagarman JA, Motley MP, Kristjansdottir K & Soloway PD 2013 Coordinate regulation of DNA methylation and H3K27me3 in mouse embryonic stem cells. *PLoS One* **8** e53880.

Hah N, Kolkman A, Ruhl DD, Pijnappel WW, Heck AJ, Timmers HT & Kraus WL 2010 A role for BAF57 in cell cycle-dependent transcriptional regulation by the SWI/SNF chromatin remodeling complex. *Cancer Res* **70** 4402-4411.

Hajkova P, Jeffries SJ, Lee C, Miller N, Jackson SP & Surani MA 2010 Genome-wide reprogramming in the mouse germ line entails the base excision repair pathway. *Science* **329** 78-82.

Halley-Stott RP, Pasque V, Astrand C, Miyamoto K, Simeoni I, Jullien J & Gurdon JB 2010 Mammalian nuclear transplantation to Germinal Vesicle stage Xenopus oocytes - a method for quantitative transcriptional reprogramming. *Methods* **51** 56-65.

Hamilton L 1963 An experimental analysis of the development of the haploid syndrome in embryos of Xenopus laevis. *J Embryol Exp Morphol* **11** 267-278.

Hammoud SS, Nix DA, Hammoud AO, Gibson M, Cairns BR & Carrell DT 2011 Genome-wide analysis identifies changes in histone retention and epigenetic modifications at developmental and imprinted gene loci in the sperm of infertile men. *Hum Reprod* **26** 2558-2569.

Hammoud SS, Nix DA, Zhang H, Purwar J, Carrell DT & Cairns BR 2009 Distinctive chromatin in human sperm packages genes for embryo development. *Nature* **460** 473-478.

Hammoud SS, Purwar J, Pflueger C, Cairns BR & Carrell DT 2010 Alterations in sperm DNA methylation patterns at imprinted loci in two classes of infertility. *Fertil Steril* **94** 1728-1733.

Han J, Li Q, McCullough L, Kettelkamp C, Formosa T & Zhang Z 2010 Ubiquitylation of FACT by the cullin-E3 ligase Rtt101 connects FACT to DNA replication. *Genes Dev* **24** 1485-1490.

Hargreaves DC, Horng T & Medzhitov R 2009 Control of inducible gene expression by signal-dependent transcriptional elongation. *Cell* **138** 129-145.

Hassig CA, Tong JK, Fleischer TC, Owa T, Grable PG, Ayer DE & Schreiber SL 1998 A role for histone deacetylase activity in HDAC1-mediated transcriptional repression. *Proc Natl Acad Sci U S A* **95** 3519-3524.

Hazzouri M, Pivot-Pajot C, Faure AK, Usson Y, Pelletier R, Sele B, Khochbin S & Rousseaux S 2000 Regulated hyperacetylation of core histones during mouse spermatogenesis: involvement of histone deacetylases. *Eur J Cell Biol* **79** 950-960.

He J, Shen L, Wan M, Taranova O, Wu H & Zhang Y 2013 Kdm2b maintains murine embryonic stem cell status by recruiting PRC1 complex to CpG islands of developmental genes. *Nat Cell Biol* **15** 373-384.

Henricksen LA, Carter T, Dutta A & Wold MS 1996 Phosphorylation of human replication protein A by the DNA-dependent protein kinase is involved in the modulation of DNA replication. *Nucleic Acids Res* **24** 3107-3112.

Hildebrand CE, Gurley LR, Tobey RA & Walters RA 1977 Action of heparin on mammalian nuclei. I. Differential extraction of histone H1 and cooperative removal of histones from chromatin. *Biochim Biophys Acta* **477** 295-311.

Hirasawa R, Matoba S, Inoue K & Ogura A 2013 Somatic donor cell type correlates with embryonic, but not extra-embryonic, gene expression in postimplantation cloned embryos. *PLoS One* **8** e76422.

Hisano M, Erkek S, Dessus-Babus S, Ramos L, Stadler MB & Peters AH 2013 Genome-wide chromatin analysis in mature mouse and human spermatozoa. *Nat Protoc* **8** 2449-2470.

Hiyoshi H, Uno S, Yokota T, Katagiri C, Nishida H, Takai M, Agata K, Eguchi G & Abe S 1991 Isolation of cDNA for a Xenopus sperm-specific basic nuclear protein (SP4) and evidence for expression of SP4 mRNA in primary spermatocytes. *Exp Cell Res* **194** 95-99.

Hochedlinger K & Jaenisch R 2002 Monoclonal mice generated by nuclear transfer from mature B and T donor cells. *Nature* **415** 1035-1038.

Hon GC, Rajagopal N, Shen Y, McCleary DF, Yue F, Dang MD & Ren B 2013 Epigenetic memory at embryonic enhancers identified in DNA methylation maps from adult mouse tissues. *Nat Genet* **45** 1198-1206.

Hsieh CL 1994 Dependence of transcriptional repression on CpG methylation density. *Mol Cell Biol* **14** 5487-5494.

Hughes V 2014 Epigenetics: The sins of the father. *Nature* **507** 22-24.

Hulstrand AM, Schneider PN & Houston DW 2010 The use of antisense oligonucleotides in Xenopus oocytes. *Methods* **51** 75-81.

Inoue A, Ogushi S, Saitou M, Suzuki MG & Aoki F 2011 Involvement of mouse nucleoplasmin 2 in the decondensation of sperm chromatin after fertilization. *Biol Reprod* **85** 70-77.

Iwano H, Nakamura M & Tajima S 2004 Xenopus MBD3 plays a crucial role in an early stage of development. *Dev Biol* **268** 416-428.

Izzo A, Kamieniarz K & Schneider R 2008 The histone H1 family: specific members, specific functions? *Biol Chem* **389** 333-343.

Jang MK, Mochizuki K, Zhou M, Jeong HS, Brady JN & Ozato K 2005 The bromodomain protein Brd4 is a positive regulatory component of P-TEFb and stimulates RNA polymerase II-dependent transcription. *Mol Cell* **19** 523-534.

Jenuwein T & Allis CD 2001 Translating the histone code. *Science* **293** 1074-1080.

Jeppesen P & Turner BM 1993 The inactive X chromosome in female mammals is distinguished by a lack of histone H4 acetylation, a cytogenetic marker for gene expression. *Cell* **74** 281-289.

Jiang CL, Jin SG & Pfeifer GP 2004 MBD3L1 is a transcriptional repressor that interacts with methyl-CpG-binding protein 2 (MBD2) and components of the NuRD complex. *J Biol Chem* **279** 52456-52464.

Jiang L, Zhang J, Wang JJ, Wang L, Zhang L, Li G, Yang X, Ma X, Sun X, Cai J, Huang X, Yu M, Wang X, Liu F, Wu CI, He C, Zhang B, Ci W & Liu J 2013 Sperm, but not oocyte, DNA methylome is inherited by zebrafish early embryos. *Cell* **153** 773-784.

Jin C & Felsenfeld G 2007 Nucleosome stability mediated by histone variants H3.3 and H2A.Z. *Genes Dev* **21** 1519-1529.

Jones PL, Veenstra GJ, Wade PA, Vermaak D, Kass SU, Landsberger N, Strouboulis J & Wolffe AP 1998 Methylated DNA and MeCP2 recruit histone deacetylase to repress transcription. *Nat Genet* **19** 187-191.

Jouve C, Palmeirim I, Henrique D, Beckers J, Gossler A, Ish-Horowicz D & Pourquie O 2000 Notch signalling is required for cyclic expression of the hairy-like gene HES1 in the presomitic mesoderm. *Development* **127** 1421-1429.

Jullien J, Astrand C, Halley-Stott RP, Garrett N & Gurdon JB 2010 Characterization of somatic cell nuclear reprogramming by oocytes in which a linker histone is required for pluripotency gene reactivation. *Proc Natl Acad Sci U S A* **107** 5483-5488.

Jullien J, Astrand C, Szenker E, Garrett N, Almouzni G & Gurdon JB 2012 HIRA dependent H3.3 deposition is required for transcriptional reprogramming following nuclear transfer to Xenopus oocytes. *Epigenetics Chromatin* **5** 17.

Kaiser J 2014 The epigenetics heretic. *Science* **343** 361-363.

Kang YK, Koo DB, Park JS, Choi YH, Chung AS, Lee KK & Han YM 2001 Aberrant methylation of donor genome in cloned bovine embryos. *Nat Genet* **28** 173-177.

Kanno T, Kanno Y, Siegel RM, Jang MK, Lenardo MJ & Ozato K 2004 Selective recognition of acetylated histones by bromodomain proteins visualized in living cells. *Mol Cell* **13** 33-43.

Kelley C, Yee K, Harland R & Zon LI 1994 Ventral expression of GATA-1 and GATA-2 in the Xenopus embryo defines induction of hematopoietic mesoderm. *Dev Biol* **165** 193-205.

Kent WJ 2002 BLAT--the BLAST-like alignment tool. *Genome Res* **12** 656-664.

Kikyo N, Wade PA, Guschin D, Ge H & Wolffe AP 2000 Active remodeling of somatic nuclei in egg cytoplasm by the nucleosomal ATPase ISWI. *Science* **289** 2360-2362.

Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R & Salzberg SL 2013 TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol* **14** R36.

Kim J, Lin JJ, Xu RH & Kung HF 1998 Mesoderm induction by heterodimeric AP-1 (c-Jun and c-Fos) and its involvement in mesoderm formation through the embryonic fibroblast growth factor/Xbra autocatalytic loop during the early development of Xenopus embryos. *J Biol Chem* **273** 1542-1550.

Kimelman D, Kirschner M & Scherson T 1987 The events of the midblastula transition in Xenopus are regulated by changes in the cell cycle. *Cell* **48** 399-407.

Kinoshita M, Hatada S, Asashima M & Noda M 1994 HMG-X, a Xenopus gene encoding an HMG1 homolog, is abundantly expressed in the developing nervous system. *FEBS Lett* **352** 191-196.

Kishigami S, Van Thuan N, Hikichi T, Ohta H, Wakayama S, Mizutani E & Wakayama T 2006 Epigenetic abnormalities of the mouse paternal zygotic genome associated with microinsemination of round spermatids. *Dev Biol* **289** 195-205.

Kishigami S, Wakayama S, Nguyen VT & Wakayama T 2004 Similar time restriction for intracytoplasmic sperm injection and round spermatid injection into activated oocytes for efficient offspring production. *Biol Reprod* **70** 1863-1869.

Kouzarides T 2007 Chromatin modifications and their function. *Cell* **128** 693-705.

Krawetz SA 2005 Paternal contribution: new insights and future challenges. *Nat Rev Genet* **6** 633-642.

Kuzmichev A, Nishioka K, Erdjument-Bromage H, Tempst P & Reinberg D 2002 Histone methyltransferase activity associated with a human multiprotein complex containing the Enhancer of Zeste protein. *Genes Dev* **16** 2893-2905.

Kwon H, Imbalzano AN, Khavari PA, Kingston RE & Green MR 1994 Nucleosome disruption and enhancement of activator binding by a human SW1/SNF complex. *Nature* **370** 477-481.

Kwon SH & Workman JL 2011 The changing faces of HP1: From heterochromatin formation and gene silencing to euchromatic gene expression: HP1 acts as a positive regulator of transcription. *Bioessays* **33** 280-289.

Labit H, Goldar A, Guilbaud G, Douarche C, Hyrien O & Marheineke K 2008 A simple and optimized method of producing silanized surfaces for FISH and replication mapping on combed DNA fibers. *Biotechniques* **45** 649-652, 654, 656-648.

Lachner M, O'Carroll D, Rea S, Mechtler K & Jenuwein T 2001 Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. *Nature* **410** 116-120.

Laherty CD, Yang WM, Sun JM, Davie JR, Seto E & Eisenman RN 1997 Histone deacetylases associated with the mSin3 corepressor mediate mad transcriptional repression. *Cell* **89** 349-356.

Lambrot R, Xu C, Saint-Phar S, Chountalos G, Cohen T, Paquet M, Suderman M, Hallett M & Kimmins S 2013 Low paternal dietary folate alters the mouse sperm epigenome and is associated with negative pregnancy outcomes. *Nat Commun* **4** 2889.

Lamond AI & Earnshaw WC 1998 Structure and function in the nucleus. *Science* **280** 547-553.

Laskey R 2005 Solving mysteries of DNA replication and frog cloning. *Cell* **123** 760-762.

Lassalle B & Testart J 1991 Sequential transformations of human sperm nucleus in human egg. *J Reprod Fertil* **91** 393-402.

Le Guezennec X, Vermeulen M, Brinkman AB, Hoeijmakers WA, Cohen A, Lasonder E & Stunnenberg HG 2006 MBD2/NuRD and MBD3/NuRD, two distinct complexes with different biochemical and functional properties. *Mol Cell Biol* **26** 843-851.

Lee CK, Shibata Y, Rao B, Strahl BD & Lieb JD 2004 Evidence for nucleosome depletion at active regulatory regions genome-wide. *Nat Genet* **36** 900-905.

Lee DY, Hayes JJ, Pruss D & Wolffe AP 1993 A positive role for histone acetylation in transcription factor access to nucleosomal DNA. *Cell* **72** 73-84.

Lee SY, Yoon J, Lee HS, Hwang YS, Cha SW, Jeong CH, Kim JI, Park JB, Lee JY, Kim S, Park MJ, Dong Z & Kim J 2011 The function of heterodimeric AP-1 comprised of c-Jun and c-Fos in activin mediated Spemann organizer gene expression. *PLoS One* **6** e21796.

Lehnertz B, Ueda Y, Derijck AA, Braunschweig U, Perez-Burgos L, Kubicek S, Chen T, Li E, Jenuwein T & Peters AH 2003 Suv39h-mediated histone H3 lysine 9 methylation directs DNA methylation to major satellite repeats at pericentric heterochromatin. *Curr Biol* **13** 1192-1200.

Lemaitre JM, Danis E, Pasero P, Vassetzky Y & Mechali M 2005 Mitotic remodeling of the replicon and chromosome structure. *Cell* **123** 787-801.

Lemaitre JM, Vassetzky Y & Mechali M 2001 Analysis of Chromatin Assembly, Chromatin Domains, and DNA Replication Using Xenopus Systems.

LeRoy G, Rickards B & Flint SJ 2008 The double bromodomain proteins Brd2 and Brd3 couple histone acetylation to transcription. *Mol Cell* **30** 51-60.

Li A, Maffey AH, Abbott WD, Conde e Silva N, Prunell A, Siino J, Churikov D, Zalensky AO & Ausio J 2005 Characterization of nucleosomes consisting of the human testis/sperm-specific histone H2B variant (hTSH2B). *Biochemistry* **44** 2529-2535.

Li H & Durbin R 2009 Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25** 1754-1760.

Li W & Godzik A 2006 Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* **22** 1658-1659.

Li Z, Gadue P, Chen K, Jiao Y, Tuteja G, Schug J, Li W & Kaestner KH 2012 Foxa2 and H2A.Z mediate nucleosome depletion during embryonic stem cell differentiation. *Cell* **151** 1608-1616.

Lindeman LC, Andersen IS, Reiner AH, Li N, Aanes H, Ostrup O, Winata C, Mathavan S, Muller F, Alestrom P & Collas P 2011 Prepatterning of developmental gene expression by modified histones before zygotic genome activation. *Dev Cell* **21** 993-1004.

Lindroth AM, Park YJ, McLean CM, Dokshin GA, Persson JM, Herman H, Pasini D, Miro X, Donohoe ME, Lee JT, Helin K & Soloway PD 2008 Antagonism between DNA and H3K27 methylation at the imprinted Rasgrf1 locus. *PLoS Genet* **4** e1000145.

Lister R, Pelizzola M, Dowen RH, Hawkins RD, Hon G, Tonti-Filippini J, Nery JR, Lee L, Ye Z, Ngo QM, Edsall L, Antosiewicz-Bourget J, Stewart R, Ruotti V, Millar AH, Thomson JA, Ren B & Ecker JR 2009 Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature* **462** 315-322.

Liu W, Ma Q, Wong K, Li W, Ohgi K, Zhang J, Aggarwal AK & Rosenfeld MG 2013 Brd4 and JMJD6-associated anti-pause enhancers in regulation of transcriptional pause release. *Cell* **155** 1581-1595.

Liu WM, Pang RT, Chiu PC, Wong BP, Lao K, Lee KF & Yeung WS 2012 Sperm-borne microRNA-34c is required for the first cleavage division in mouse. *Proc Natl Acad Sci U S A* **109** 490-494.

Loiodice I, Alves A, Rabut G, Van Overbeek M, Ellenberg J, Sibarita JB & Doye V 2004 The entire Nup107-160 complex, including three new members, is targeted as one entity to kinetochores in mitosis. *Mol Biol Cell* **15** 3333-3344.

Loppin B, Bonnefoy E, Anselme C, Laurencon A, Karr TL & Couble P 2005 The histone H3.3 chaperone HIRA is essential for chromatin assembly in the male pronucleus. *Nature* **437** 1386-1390.

Maatouk DM, Kellam LD, Mann MR, Lei H, Li E, Bartolomei MS & Resnick JL 2006 DNA methylation is a primary mechanism for silencing postmigratory primordial germ cell genes in both germ cell and somatic cell lineages. *Development* **133** 3411-3418.

Maeno M, Mead PE, Kelley C, Xu RH, Kung HF, Suzuki A, Ueno N & Zon LI 1996 The role of BMP-4 and GATA-2 in the induction and differentiation of hematopoietic mesoderm in Xenopus laevis. *Blood* **88** 1965-1972.

Mali P, Esvelt KM & Church GM 2013 Cas9 as a versatile tool for engineering biology. *Nature methods* **10** 957-963.

Martianov I, Brancorsini S, Catena R, Gansmuller A, Kotaja N, Parvinen M, Sassone-Corsi P & Davidson I 2005 Polar nuclear localization of H1T2, a histone H1 variant, required for

spermatid elongation and DNA condensation during spermiogenesis. *Proc Natl Acad Sci U S A* **102** 2808-2813.

Martianov I, Fimia GM, Dierich A, Parvinen M, Sassone-Corsi P & Davidson I 2001 Late arrest of spermiogenesis and germ cell apoptosis in mice lacking the TBP-like TLF/TRF2 gene. *Mol Cell* **7** 509-515.

Martin M 2011 *Cutadapt removes adapter sequences from high-throughput sequencing reads*.

Matzuk MM, McKeown MR, Filippakopoulos P, Li Q, Ma L, Agno JE, Lemieux ME, Picaud S, Yu RN, Qi J, Knapp S & Bradner JE 2012 Small-molecule inhibition of BRDT for male contraception. *Cell* **150** 673-684.

McKittrick E, Gafken PR, Ahmad K & Henikoff S 2004 Histone H3.3 is enriched in covalent modifications associated with active chromatin. *Proc Natl Acad Sci U S A* **101** 1525-1530.

McLay DW & Clarke HJ 2003 Remodelling the paternal chromatin at fertilization in mammals. *Reproduction* **125** 625-633.

Meistrich ML, Mohapatra B, Shirley CR & Zhao M 2003 Roles of transition nuclear proteins in spermiogenesis. *Chromosoma* **111** 483-488.

Miller D 2007 Spermatozoal RNA as reservoir, marker and carrier of epigenetic information: implications for cloning. *Reprod Domest Anim* **42 Suppl 2** 2-9.

Miyamoto K, Teperek M, Yusa K, Allen GE, Bradshaw CR & Gurdon JB 2013 Nuclear Wave1 is required for reprogramming transcription in oocytes and for normal development. *Science* **341** 1002-1005.

Mizutani E, Yamagata K, Ono T, Akagi S, Geshi M & Wakayama T 2012 Abnormal chromosome segregation at early cleavage is a major cause of the full-term developmental failure of mouse clones. *Dev Biol* **364** 56-65.

Monesi V, Geremia R, D'Agostino A & Boitani C 1978 Biochemistry of male germ cell differentiation in mammals: RNA synthesis in meiotic and postmeiotic cells. *Curr Top Dev Biol* **12** 11-36.

Moriniere J, Rousseaux S, Steuerwald U, Soler-Lopez M, Curtet S, Vitte AL, Govin J, Gaucher J, Sadoul K, Hart DJ, Krijgsveld J, Khochbin S, Muller CW & Petosa C 2009 Cooperative binding of two acetylation marks on a histone tail by a single bromodomain. *Nature* **461** 664-668.

Motamedi MR, Hong EJ, Li X, Gerber S, Denison C, Gygi S & Moazed D 2008 HP1 proteins form distinct complexes and mediate heterochromatic gene silencing by nonoverlapping mechanisms. *Mol Cell* **32** 778-790.

Muller J, Hart CM, Francis NJ, Vargas ML, Sengupta A, Wild B, Miller EL, O'Connor MB, Kingston RE & Simon JA 2002 Histone methyltransferase activity of a Drosophila Polycomb group repressor complex. *Cell* **111** 197-208.

Nakayama J, Rice JC, Strahl BD, Allis CD & Grewal SI 2001 Role of histone H3 lysine 9 methylation in epigenetic control of heterochromatin assembly. *Science* **292** 110-113.

Nakazaki H, Reddy AC, Mania-Farnell BL, Shen YW, Ichi S, McCabe C, George D, McLone DG, Tomita T & Mayanil CS 2008 Key basic helix-loop-helix transcription factor genes Hes1 and Ngn2 are regulated by Pax3 during mouse embryonic development. *Dev Biol* **316** 510-523.

Narbonne P, Simpson DE & Gurdon JB 2011 Deficient induction response in a Xenopus nucleocytoplasmic hybrid. *PLoS Biol* **9** e1001197.

Nardelli J, Thiesson D, Fujiwara Y, Tsai FY & Orkin SH 1999 Expression and genetic interaction of transcription factors GATA-2 and GATA-3 during development of the mouse central nervous system. *Dev Biol* **210** 305-321.

Newport J & Kirschner M 1982 A major developmental transition in early Xenopus embryos: I. characterization and timing of cellular changes at the midblastula stage. *Cell* **30** 675-686.

Ng RK & Gurdon JB 2005 Epigenetic memory of active gene transcription is inherited through somatic cell nuclear transfer. *Proc Natl Acad Sci U S A* **102** 1957-1962.

Ng RK & Gurdon JB 2008 Epigenetic memory of an active gene state depends on histone H3.3 incorporation into chromatin in the absence of transcription. *Nat Cell Biol* **10** 102-109.

Nicklay JJ, Shechter D, Chitta RK, Garcia BA, Shabanowitz J, Allis CD & Hunt DF 2009 Analysis of histones in Xenopus laevis. II. mass spectrometry reveals an index of cell type-specific modifications on H3 and H4. *J Biol Chem* **284** 1075-1085.

Nicolas E, Morales V, Magnaghi-Jaulin L, Harel-Bellan A, Richard-Foy H & Trouche D 2000 RbAp48 belongs to the histone deacetylase complex that associates with the retinoblastoma protein. *J Biol Chem* **275** 9797-9804.

Nieuwkoop PD & Faber J 1994 *Normal table of Xenopus laevis (Daudin) : a systematical and chronological survey of the development from the fertilized egg till the end of metamorphosis*. New York: Garland Pub.

Nishiyama A, Yamaguchi L, Sharif J, Johmura Y, Kawamura T, Nakanishi K, Shimamura S, Arita K, Kodama T, Ishikawa F, Koseki H & Nakanishi M 2013 Uhrf1-dependent H3K23 ubiquitylation couples maintenance DNA methylation and replication. *Nature* **502** 249-253.

Noma K, Allis CD & Grewal SI 2001 Transitions in distinct histone H3 methylation patterns at the heterochromatin domain boundaries. *Science* **293** 1150-1155.

Ogonuki N, Inoue K, Miki H, Mochida K, Hatori M, Okada H, Takeiri S, Shimozawa N, Nagashima H, Sankai T & Ogura A 2005 Differential development of rabbit embryos following microinsemination with sperm and spermatids. *Mol Reprod Dev* **72** 411-417.

Ogryzko VV, Schiltz RL, Russanova V, Howard BH & Nakatani Y 1996 The transcriptional coactivators p300 and CBP are histone acetyltransferases. *Cell* **87** 953-959.

Ohta H, Sakaide Y & Wakayama T 2009 Functional analysis of male mouse haploid germ cells of various differentiation stages: early and late round spermatids are functionally equivalent in producing progeny. *Biol Reprod* **80** 511-517.

Okuwaki M, Sumi A, Hisaoka M, Saotome-Nakamura A, Akashi S, Nishimura Y & Nagata K 2012 Function of homo- and hetero-oligomers of human nucleoplasmin/nucleophosmin family proteins NPM1, NPM2 and NPM3 during sperm chromatin remodeling. *Nucleic Acids Res* **40** 4861-4878.

Ostermeier GC, Miller D, Huntriss JD, Diamond MP & Krawetz SA 2004 Reproductive biology: delivering spermatozoan RNA to the oocyte. *Nature* **429** 154.

Ozsolak F, Song JS, Liu XS & Fisher DE 2007 High-throughput mapping of the chromatin structure of human promoters. *Nat Biotechnol* **25** 244-248.

Padmanabhan N, Jia D, Geary-Joo C, Wu X, Ferguson-Smith AC, Fung E, Bieda MC, Snyder FF, Gravel RA, Cross JC & Watson ED 2013 Mutation in folate metabolism causes epigenetic instability and transgenerational effects on development. *Cell* **155** 81-93.

Papageorgis P, Lambert AW, Ozturk S, Gao F, Pan H, Manne U, Alekseyev YO, Thiagalingam A, Abdolmaleky HM, Lenburg M & Thiagalingam S 2010 Smad signaling is required to maintain epigenetic silencing during breast cancer progression. *Cancer Res* **70** 968-978.

Paradowska AS, Miller D, Spiess AN, Vieweg M, Cerna M, Dvorakova-Hortova K, Bartkuhn M, Schuppe HC, Weidner W & Steger K 2012 Genome wide identification of promoter binding sites for H4K12ac in human sperm and its relevance for early embryonic development. *Epigenetics* **7** 1057-1070.

Pasque V, Gillich A, Garrett N & Gurdon JB 2011a Histone variant macroH2A confers resistance to nuclear reprogramming. *EMBO J* **30** 2373-2387.

Pasque V, Jullien J, Miyamoto K, Halley-Stott RP & Gurdon JB 2011b Epigenetic factors influencing resistance to nuclear reprogramming. *Trends Genet* **27** 516-525.

Pasque V, Miyamoto K & Gurdon JB 2010 Efficiencies and mechanisms of nuclear reprogramming. *Cold Spring Harb Symp Quant Biol* **75** 189-200.

Perez-Montero S, Carbonell A, Moran T, Vaquero A & Azorin F 2013 The embryonic linker histone H1 variant of Drosophila, dBigH1, regulates zygotic genome activation. *Dev Cell* **26** 578-590.

Perez-Pinera P, Kocak DD, Vockley CM, Adler AF, Kabadi AM, Polstein LR, Thakore PI, Glass KA, Ousterout DG, Leong KW, Guilak F, Crawford GE, Reddy TE & Gersbach CA 2013 RNA-

guided gene activation by CRISPR-Cas9-based transcription factors. *Nature methods* **10** 973-976.

Pessot CA, Brito M, Figueroa J, Concha, II, Yanez A & Burzio LO 1989 Presence of RNA in the sperm nucleus. *Biochem Biophys Res Commun* **158** 272-278.

Petruk S, Sedkov Y, Johnston DM, Hodgson JW, Black KL, Kovermann SK, Beck S, Canaani E, Brock HW & Mazo A 2012 TrxG and PcG proteins but not methylated histones remain associated with DNA through replication. *Cell* **150** 922-933.

Philpott A & Leno GH 1992 Nucleoplasmin remodels sperm chromatin in Xenopus egg extracts. *Cell* **69** 759-767.

Philpott A, Leno GH & Laskey RA 1991 Sperm decondensation in Xenopus egg cytoplasm is mediated by nucleoplasmin. *Cell* **65** 569-578.

Pittoggi C, Magnano AR, Sciamanna I, Giordano R, Lorenzini R & Spadafora C 2001 Specific localization of transcription factors in the chromatin of mouse mature spermatozoa. *Mol Reprod Dev* **60** 97-106.

Pivot-Pajot C, Caron C, Govin J, Vion A, Rousseaux S & Khochbin S 2003 Acetylation-dependent chromatin reorganization by BRDT, a testis-specific bromodomain-containing protein. *Mol Cell Biol* **23** 5354-5365.

Plath K, Fang J, Mlynarczyk-Evans SK, Cao R, Worringer KA, Wang H, de la Cruz CC, Otte AP, Panning B & Zhang Y 2003 Role of histone H3 lysine 27 methylation in X inactivation. *Science* **300** 131-135.

Poot RA, Bozhenok L, van den Berg DL, Hawkes N & Varga-Weisz PD 2005 Chromatin remodeling by WSTF-ISWI at the replication site: opening a window of opportunity for epigenetic inheritance? *Cell Cycle* **4** 543-546.

Popp C, Dean W, Feng S, Cokus SJ, Andrews S, Pellegrini M, Jacobsen SE & Reik W 2010 Genome-wide erasure of DNA methylation in mouse primordial germ cells is affected by AID deficiency. *Nature* **463** 1101-1105.

Potok ME, Nix DA, Parnell TJ & Cairns BR 2013 Reprogramming the maternal zebrafish genome after fertilization to match the paternal methylation pattern. *Cell* **153** 759-772.

Qi LS, Larson MH, Gilbert LA, Doudna JA, Weissman JS, Arkin AP & Lim WA 2013 Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell* **152** 1173-1183.

Qian MX, Pang Y, Liu CH, Haratake K, Du BY, Ji DY, Wang GF, Zhu QQ, Song W, Yu Y, Zhang XX, Huang HT, Miao S, Chen LB, Zhang ZH, Liang YN, Liu S, Cha H, Yang D, Zhai Y, Komatsu T, Tsuruta F, Li H, Cao C, Li W, Li GH, Cheng Y, Chiba T, Wang L, Goldberg AL, Shen Y & Qiu XB 2013 Acetylation-mediated proteasomal degradation of core histones during DNA repair and spermatogenesis. *Cell* **153** 1012-1024.

Rais Y, Zviran A, Geula S, Gafni O, Chomsky E, Viukov S, Mansour AA, Caspi I, Krupalnik V, Zerbib M, Maza I, Mor N, Baran D, Weinberger L, Jaitin DA, Lara-Astiaso D, Blecher-Gonen R, Shipony Z, Mukamel Z, Hagai T, Gilad S, Amann-Zalcenstein D, Tanay A, Amit I, Novershtern N & Hanna JH 2013 Deterministic direct reprogramming of somatic cells to pluripotency. *Nature* **502** 65-70.

Ramos I, Martin-Benito J, Finn R, Bretana L, Aloria K, Arizmendi JM, Ausio J, Muga A, Valpuesta JM & Prado A 2010 Nucleoplasmin binds histone H2A-H2B dimers through its distal face. *J Biol Chem* **285** 33771-33778.

Ravel C, Chantot-Bastaraud S, El Houate B, Berthaut I, Verstraete L, De Larouziere V, Lourenco D, Dumaine A, Antoine JM, Mandelbaum J, Siffroi JP & McElreavey K 2007 Mutations in the protamine 1 gene associated with male infertility. *Mol Hum Reprod* **13** 461-464.

Ray-Gallet D, Quivy JP, Scamps C, Martini EM, Lipinski M & Almouzni G 2002 HIRA is critical for a nucleosome assembly pathway independent of DNA synthesis. *Mol Cell* **9** 1091-1100.

Read EM, Rodaway AR, Neave B, Brandon N, Holder N, Patient RK & Walmsley ME 1998 Evidence for non-axial A/P patterning in the nonneural ectoderm of Xenopus and zebrafish pregastrula embryos. *Int J Dev Biol* **42** 763-774.

Reddington JP, Perricone SM, Nestor CE, Reichmann J, Youngson NA, Suzuki M, Reinhardt D, Dunican DS, Prendergast JG, Mjoseng H, Ramsahoye BH, Whitelaw E, Greally JM, Adams IR, Bickmore WA & Meehan RR 2013 Redistribution of H3K27me3 upon DNA hypomethylation results in de-repression of Polycomb target genes. *Genome Biol* **14** R25.

Reversade B & De Robertis EM 2005 Regulation of ADMP and BMP2/4/7 at opposite embryonic poles generates a self-regulating morphogenetic field. *Cell* **123** 1147-1160.

Reversade B, Kuroda H, Lee H, Mays A & De Robertis EM 2005 Depletion of Bmp2, Bmp4, Bmp7 and Spemann organizer signals induces massive brain formation in Xenopus embryos. *Development* **132** 3381-3392.

Richly H, Rocha-Viegas L, Ribeiro JD, Demajo S, Gundem G, Lopez-Bigas N, Nakagawa T, Rospert S, Ito T & Di Croce L 2010 Transcriptional activation of polycomb-repressed genes by ZRF1. *Nature* **468** 1124-1128.

Risley MS & Eckhardt RA 1979 Dissociation and separation of *Xenopus laevis* spermatogenic cells. *Journal of Experimental Zoology* **207** 93-106.

Risley MS & Eckhardt RA 1981 H1 histone variants in Xenopus laevis. *Dev Biol* **84** 79-87.

Risley MS, Eckhardt RA, Mann M & Kasinsky HE 1982 Determinants of sperm nuclear shaping in the genus Xenopus. *Chromosoma* **84** 557-569.

Robinson MD, McCarthy DJ & Smyth GK 2010 edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26** 139-140.

Roger B, Moisand A, Amalric F & Bouvet P 2002 rDNA transcription during Xenopus laevis oogenesis. *Biochem Biophys Res Commun* **290** 1151-1160.

Roosen-Runge EC 1969 Comparative aspects of spermatogenesis. *Biol Reprod* **1** Suppl 1:24-31.

Rowles A & Blow JJ 1997 Chromatin proteins involved in the initiation of DNA replication. *Curr Opin Genet Dev* **7** 152-157.

Santenard A, Ziegler-Birling C, Koch M, Tora L, Bannister AJ & Torres-Padilla ME 2010 Heterochromatin formation in the mouse embryo requires critical residues of the histone variant H3.3. *Nature cell biology* **12** 853-862.

Santos-Rosa H, Schneider R, Bannister AJ, Sherriff J, Bernstein BE, Emre NC, Schreiber SL, Mellor J & Kouzarides T 2002 Active genes are tri-methylated at K4 of histone H3. *Nature* **419** 407-411.

Saxena UH, Powell CM, Fecko JK, Cacioppo R, Chou HS, Cooper GM & Hansen U 2009 Phosphorylation by cyclin C/cyclin-dependent kinase 2 following mitogenic stimulation of murine fibroblasts inhibits transcriptional activity of LSF during G1 progression. *Mol Cell Biol* **29** 2335-2345.

Schones DE, Cui K, Cuddapah S, Roh TY, Barski A, Wang Z, Wei G & Zhao K 2008 Dynamic regulation of nucleosome positioning in the human genome. *Cell* **132** 887-898.

Schubeler D, Lorincz MC, Cimbora DM, Telling A, Feng YQ, Bouhassira EE & Groudine M 2000 Genomic targeting of methylated DNA: influence of methylation on transcription, replication, chromatin structure, and histone acetylation. *Mol Cell Biol* **20** 9103-9112.

Sendler E, Johnson GD, Mao S, Goodrich RJ, Diamond MP, Hauser R & Krawetz SA 2013 Stability, delivery and functions of human sperm RNAs at fertilization. *Nucleic Acids Res* **41** 4104-4117.

Seyedin SM & Kistler WS 1980 Isolation and characterization of rat testis H1t. An H1 histone variant associated with spermatogenesis. *J Biol Chem* **255** 5949-5954.

Shechter D, Nicklay JJ, Chitta RK, Shabanowitz J, Hunt DF & Allis CD 2009 Analysis of histones in Xenopus laevis. I. A distinct index of enriched variants and modifications exists in each cell type and is remodeled during developmental transitions. *J Biol Chem* **284** 1064-1074.

Sible JC, Erikson E, Hendrickson M, Maller JL & Gautier J 1998 Developmental regulation of MCM replication factors in Xenopus laevis. *Curr Biol* **8** 347-350.

Simpson RT 1978 Structure of chromatin containing extensively acetylated H3 and H4. *Cell* **13** 691-699.

Smallwood A, Esteve PO, Pradhan S & Carey M 2007 Functional cooperation between HP1 and DNMT1 mediates gene silencing. *Genes Dev* **21** 1169-1178.

Smallwood A, Hon GC, Jin F, Henry RE, Espinosa JM & Ren B 2012 CBX3 regulates efficient RNA processing genome-wide. *Genome Res* **22** 1426-1436.

Smith RC, Dworkin-Rastl E & Dworkin MB 1988 Expression of a histone H1-like protein is restricted to early Xenopus development. *Genes Dev* **2** 1284-1295.

Smith SJ, Fairclough L, Latinkic BV, Sparrow DB & Mohun TJ 2006 Xenopus laevis transgenesis by sperm nuclear injection. *Nat Protoc* **1** 2195-2203.

Snowden AW, Gregory PD, Case CC & Pabo CO 2002 Gene-specific targeting of H3K9 methylation is sufficient for initiating repression in vivo. *Curr Biol* **12** 2159-2166.

Song F, Mahmood S, Ghosh S, Liang P, Smiraglia DJ, Nagase H & Held WA 2009 Tissue specific differentially methylated regions (TDMR): Changes in DNA methylation during development. *Genomics* **93** 130-139.

Song F, Smith JF, Kimura MT, Morrow AD, Matsuyama T, Nagase H & Held WA 2005 Association of tissue-specific differentially methylated regions (TDMs) with differential gene expression. *Proc Natl Acad Sci U S A* **102** 3336-3341.

Sridharan R, Gonzales-Cope M, Chronis C, Bonora G, McKee R, Huang C, Patel S, Lopez D, Mishra N, Pellegrini M, Carey M, Garcia BA & Plath K 2013 Proteomic and genomic approaches reveal critical functions of H3K9 methylation and heterochromatin protein-1gamma in reprogramming to pluripotency. *Nat Cell Biol* **15** 872-882.

Steger K, Wilhelm J, Konrad L, Stalf T, Greb R, Diemer T, Kliesch S, Bergmann M & Weidner W 2008 Both protamine-1 to protamine-2 mRNA ratio and Bcl2 mRNA content in testicular spermatids and ejaculated spermatozoa discriminate between fertile and infertile men. *Hum Reprod* **23** 11-16.

Stengel S, Fiebig U, Kurth R & Denner J 2010 Regulation of human endogenous retrovirus-K expression in melanomas by CpG methylation. *Genes Chromosomes Cancer* **49** 401-411.

Strahl BD & Allis CD 2000 The language of covalent histone modifications. *Nature* **403** 41-45.

Sung LY, Gao S, Shen H, Yu H, Song Y, Smith SL, Chang CC, Inoue K, Kuo L, Lian J, Li A, Tian XC, Tuck DP, Weissman SM, Yang X & Cheng T 2006 Differentiated cells are more efficient than adult stem cells for cloning by somatic cell nuclear transfer. *Nat Genet* **38** 1323-1328.

Surani MA, Hayashi K & Hajkova P 2007 Genetic and epigenetic regulators of pluripotency. *Cell* **128** 747-762.

Suzuki T, Minami N, Kono T & Imai H 2007 Comparison of the RNA polymerase I-, II- and III-dependent transcript levels between nuclear transfer and in vitro fertilized embryos at the blastocyst stage. *J Reprod Dev* **53** 663-671.

Syed SH, Boulard M, Shukla MS, Gautier T, Travers A, Bednar J, Faivre-Moskalenko C, Dimitrov S & Angelov D 2009 The incorporation of the novel histone variant H2AL2 confers unusual structural and functional properties of the nucleosome. *Nucleic Acids Res* **37** 4684-4695.

Szyf M 2014 Lamarck revisited: epigenetic inheritance of ancestral odor fear conditioning. *Nat Neurosci* **17** 2-4.

Tachiwana H, Kagawa W, Osakabe A, Kawaguchi K, Shiga T, Hayashi-Takanaka Y, Kimura H & Kurumizaka H 2010 Structural basis of instability of the nucleosome containing a testis-specific histone variant, human H3T. *Proc Natl Acad Sci U S A* **107** 10454-10459.

Talbert PB, Ahmad K, Almouzni G, Ausio J, Berger F, Bhalla PL, Bonner WM, Cande WZ, Chadwick BP, Chan SW, Cross GA, Cui L, Dimitrov SI, Doenecke D, Eirin-Lopez JM, Gorovsky MA, Hake SB, Hamkalo BA, Holec S, Jacobsen SE, Kamieniarz K, Khochbin S, Ladurner AG, Landsman D, Latham JA, Loppin B, Malik HS, Marzluff WF, Pehrson JR, Postberg J, Schneider R, Singh MB, Smith MM, Thompson E, Torres-Padilla ME, Tremethick DJ, Turner BM, Waterborg JH,

Wollmann H, Yelagandula R, Zhu B & Henikoff S 2012 A unified phylogeny-based nomenclature for histone variants. *Epigenetics Chromatin* **5** 7.

Tamada H, Van Thuan N, Reed P, Nelson D, Katoku-Kikyo N, Wudel J, Wakayama T & Kikyo N 2006 Chromatin decondensation and nuclear reprogramming by nucleoplasmin. *Mol Cell Biol* **26** 1259-1271.

Tan M, Luo H, Lee S, Jin F, Yang JS, Montellier E, Buchou T, Cheng Z, Rousseaux S, Rajagopal N, Lu Z, Ye Z, Zhu Q, Wysocka J, Ye Y, Khochbin S, Ren B & Zhao Y 2011 Identification of 67 histone marks and histone lysine crotonylation as a new type of histone modification. *Cell* **146** 1016-1028.

Tan MH, Au KF, Leong DE, Foygel K, Wong WH & Yao MW 2013 An Oct4-Sall4-Nanog network controls developmental progression in the pre-implantation mouse embryo. *Mol Syst Biol* **9** 632.

Tanaka H, Iguchi N, Isotani A, Kitamura K, Toyama Y, Matsuoka Y, Onishi M, Masai K, Maekawa M, Toshimori K, Okabe M & Nishimune Y 2005 HANP1/H1T2, a novel histone H1-like protein involved in nuclear formation and sperm fertility. *Mol Cell Biol* **25** 7107-7119.

Tanaka M, Hennebold JD, Macfarlane J & Adashi EY 2001 A mammalian oocyte-specific linker histone gene H1oo: homology with the genes for the oocyte-specific cleavage stage histone (cs-H1) of sea urchin and the B4/H1M histone of the frog. *Development* **128** 655-664.

Teperek M & Miyamoto K 2013 Nuclear reprogramming of sperm and somatic nuclei in eggs and oocytes. *Reprod Med Biol* **12** 133-149.

Teranishi T, Tanaka M, Kimoto S, Ono Y, Miyakoshi K, Kono T & Yoshimura Y 2004 Rapid replacement of somatic linker histones with the oocyte-specific linker histone H1foo in nuclear transfer. *Dev Biol* **266** 76-86.

Thomas PD, Campbell MJ, Kejariwal A, Mi H, Karlak B, Daverman R, Diemer K, Muruganujan A & Narechania A 2003 PANTHER: a library of protein families and subfamilies indexed by function. *Genome Res* **13** 2129-2141.

Tong JK, Hassig CA, Schnitzler GR, Kingston RE & Schreiber SL 1998 Chromatin deacetylation by an ATP-dependent nucleosome remodelling complex. *Nature* **395** 917-921.

Torres-Padilla ME, Bannister AJ, Hurd PJ, Kouzarides T & Zernicka-Goetz M 2006 Dynamic distribution of the replacement histone variant H3.3 in the mouse oocyte and preimplantation embryos. *The International journal of developmental biology* **50** 455-461.

Torres-Padilla ME, Parfitt DE, Kouzarides T & Zernicka-Goetz M 2007 Histone arginine methylation regulates pluripotency in the early mouse embryo. *Nature* **445** 214-218.

Toshimori K 2003 Biology of spermatozoa maturation: an overview with an introduction to this issue. *Microsc Res Tech* **61** 1-6.

Trapnell C, Pachter L & Salzberg SL 2009 TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* **25** 1105-1111.

Tsai SC, Valkov N, Yang WM, Gump J, Sullivan D & Seto E 2000 Histone deacetylase interacts directly with DNA topoisomerase II. *Nat Genet* **26** 349-353.

Turner BM, Birley AJ & Lavender J 1992 Histone H4 isoforms acetylated at specific lysine residues define individual chromosomes and chromatin domains in Drosophila polytene nuclei. *Cell* **69** 375-384.

Umehara T, Nakamura Y, Jang MK, Nakano K, Tanaka A, Ozato K, Padmanabhan B & Yokoyama S 2010a Structural basis for acetylated histone H4 recognition by the human BRD2 bromodomain. *J Biol Chem* **285** 7610-7618.

Umehara T, Nakamura Y, Wakamori M, Ozato K, Yokoyama S & Padmanabhan B 2010b Structural implications for K5/K12-di-acetylated histone H4 recognition by the second bromodomain of BRD2. *FEBS Lett* **584** 3901-3908.

Ura K, Nightingale K & Wolffe AP 1996 Differential association of HMG1 and linker histones B4 and H1 with dinucleosomal DNA: structural transitions and transcriptional repression. *EMBO J* **15** 4959-4969.

Vallee M, Gravel C, Palin MF, Reghenas H, Stothard P, Wishart DS & Sirard MA 2005 Identification of novel and known oocyte-specific genes using complementary DNA subtraction and microarray analysis in three different species. *Biol Reprod* **73** 63-71.

van der Heijden GW, Derijck AA, Ramos L, Giele M, van der Vlag J & de Boer P 2006 Transmission of modified nucleosomes from the mouse male germline to the zygote and subsequent remodeling of paternal chromatin. *Dev Biol* **298** 458-469.

van der Heijden GW, Dieker JW, Derijck AA, Muller S, Berden JH, Braat DD, van der Vlag J & de Boer P 2005 Asymmetry in histone H3 variants and lysine methylation between paternal and maternal chromatin of the early mouse zygote. *Mech Dev* **122** 1008-1022.

van der Heijden GW, Ramos L, Baart EB, van den Berg IM, Derijck AA, van der Vlag J, Martini E & de Boer P 2008 Sperm-derived histones contribute to zygotic chromatin in humans. *BMC Dev Biol* **8** 34.

van Heeringen SJ, Akkers RC, van Kruijsbergen I, Arif MA, Hanssen LL, Sharifi N & Veenstra GJ 2014 Principles of nucleation of H3K27 methylation during embryonic development. *Genome Res* **24** 401-410.

Van Thuan N, Wakayama S, Kishigami S, Ohta H, Hikichi T, Mizutani E, Bui HT & Wakayama T 2006 Injection of somatic cell cytoplasm into oocytes before intracytoplasmic sperm injection impairs full-term development and increases placental weight in mice. *Biol Reprod* **74** 865-873.

VanDemark AP, Blanksma M, Ferris E, Heroux A, Hill CP & Formosa T 2006 The structure of the yFACT Pob3-M domain, its interaction with the DNA replication factor RPA, and a potential role in nucleosome deposition. *Mol Cell* **22** 363-374.

Vassena R, Han Z, Gao S, Baldwin DA, Schultz RM & Latham KE 2007 Tough beginnings: alterations in the transcriptome of cloned embryos during the first two cell cycles. *Dev Biol* **304** 75-89.

Vassoler FM, White SL, Schmidt HD, Sadri-Vakili G & Pierce RC 2013 Epigenetic inheritance of a cocaine-resistance phenotype. *Nat Neurosci* **16** 42-47.

Vastenhouw NL, Zhang Y, Woods IG, Imam F, Regev A, Liu XS, Rinn J & Schier AF 2010 Chromatin signature of embryonic pluripotency is established during genome activation. *Nature* **464** 922-926.

Vermaak D, Wade PA, Jones PL, Shi YB & Wolffe AP 1999 Functional analysis of the SIN3-histone deacetylase RPD3-RbAp48-histone H4 connection in the Xenopus oocyte. *Mol Cell Biol* **19** 5847-5860.

Wade PA, Gegonne A, Jones PL, Ballestar E, Aubry F & Wolffe AP 1999 Mi-2 complex couples DNA methylation to chromatin remodelling and histone deacetylation. *Nat Genet* **23** 62-66.

Wang G, Guo Y, Zhou T, Shi X, Yu J, Yang Y, Wu Y, Wang J, Liu M, Chen X, Tu W, Zeng Y, Jiang M, Li S, Zhang P, Zhou Q, Zheng B, Yu C, Zhou Z, Guo X & Sha J 2013 In-depth proteomic analysis of the human sperm reveals complex protein compositions. *J Proteomics* **79** 114-122.

Wang H, Wang L, Erdjument-Bromage H, Vidal M, Tempst P, Jones RS & Zhang Y 2004 Role of histone H2A ubiquitination in Polycomb silencing. *Nature* **431** 873-878.

Wang S, Zhang B & Faller DV 2002 Prohibitin requires Brg-1 and Brm for the repression of E2F and cell growth. *EMBO J* **21** 3019-3028.

Ward WS & Coffey DS 1991 DNA packaging and organization in mammalian spermatozoa: comparison with somatic cells. *Biol Reprod* **44** 569-574.

Wee G, Koo DB, Song BS, Kim JS, Kang MJ, Moon SJ, Kang YK, Lee KK & Han YM 2006 Inheritable histone H4 acetylation of somatic chromatins in cloned embryos. *J Biol Chem* **281** 6048-6057.

Wills A, Dickinson K, Khokha M & Baker JC 2008 Bmp signaling is necessary and sufficient for ventrolateral endoderm specification in Xenopus. *Dev Dyn* **237** 2177-2186.

Winteringham LN, Kobelke S, Williams JH, Ingley E & Klinken SP 2004 Myeloid Leukemia Factor 1 inhibits erythropoietin-induced differentiation, cell cycle exit and p27Kip1 accumulation. *Oncogene* **23** 5105-5109.

Wu Q, Bruce AW, Jedrusik A, Ellis PD, Andrews RM, Langford CF, Glover DM & Zernicka-Goetz M 2009 CARM1 is required in embryonic stem cells to maintain pluripotency and resist differentiation. *Stem cells* **27** 2637-2645.

Wu SF, Zhang H & Cairns BR 2011 Genes for embryo development are packaged in blocks of multivalent chromatin in zebrafish sperm. *Genome Res* **21** 578-589.

Wysocka J, Swigut T, Milne TA, Dou Y, Zhang X, Burlingame AL, Roeder RG, Brivanlou AH & Allis CD 2005 WDR5 associates with histone H3 methylated at K4 and is essential for H3 K4 methylation and vertebrate development. *Cell* **121** 859-872.

Wysocka J, Swigut T, Xiao H, Milne TA, Kwon SY, Landry J, Kauer M, Tackett AJ, Chait BT, Badenhorst P, Wu C & Allis CD 2006 A PHD finger of NURF couples histone H3 lysine 4 trimethylation with chromatin remodelling. *Nature* **442** 86-90.

Yan W, Ma L, Burns KH & Matzuk MM 2003 HILS1 is a spermatid-specific linker histone H1-like protein implicated in chromatin remodeling during mammalian spermiogenesis. *Proc Natl Acad Sci U S A* **100** 10546-10551.

Yang M, Li S, Liang G, Sun L, Zhang X, Liu N, Pang H, Miao S, Wang L & Rao Z 2003 Crystallization and preliminary crystallographic analysis of RSB-66, a novel round spermatid-specific protein. *Acta Crystallogr D Biol Crystallogr* **59** 1853-1855.

Yang Z, Yik JH, Chen R, He N, Jang MK, Ozato K & Zhou Q 2005 Recruitment of P-TEFb for stimulation of transcriptional elongation by the bromodomain protein Brd4. *Mol Cell* **19** 535-545.

Yao YL & Yang WM 2003 The metastasis-associated proteins 1 and 2 form distinct protein complexes with histone deacetylase activity. *J Biol Chem* **278** 42560-42568.

Yokota T 1991 Nuclear Basic Proteins of Xenopus laevis Sperm: Their Characterization and Synthesis during Spermatogenesis. *Development, Growth & Differentiation* **33** 9-17.

Yoneda-Kato N & Kato JY 2008 Shuttling imbalance of MLF1 results in p53 instability and increases susceptibility to oncogenic transformation. *Mol Cell Biol* **28** 422-434.

Yoneda-Kato N, Tomoda K, Umehara M, Arata Y & Kato JY 2005 Myeloid leukemia factor 1 regulates p53 by suppressing COP1 via COP9 signalosome subunit 3. *EMBO J* **24** 1739-1749.

Yoo BK, Emdad L, Gredler R, Fuller C, Dumur CI, Jones KH, Jackson-Cook C, Su ZZ, Chen D, Saxena UH, Hansen U, Fisher PB & Sarkar D 2010 Transcription factor Late SV40 Factor (LSF) functions as an oncogene in hepatocellular carcinoma. *Proc Natl Acad Sci U S A* **107** 8357-8362.

Yoshida M, Kijima M, Akita M & Beppu T 1990 Potent and specific inhibition of mammalian histone deacetylase both in vivo and in vitro by trichostatin A. *J Biol Chem* **265** 17174-17179.

Yu YE, Zhang Y, Unni E, Shirley CR, Deng JM, Russell LD, Weil MM, Behringer RR & Meistrich ML 2000 Abnormal spermatogenesis and reduced fertility in transition nuclear protein 1-deficient mice. *Proc Natl Acad Sci U S A* **97** 4683-4688.

Zdobnov EM & Apweiler R 2001 InterProScan--an integration platform for the signature-recognition methods in InterPro. *Bioinformatics* **17** 847-848.

Zhang J, Tam WL, Tong GQ, Wu Q, Chan HY, Soh BS, Lou Y, Yang J, Ma Y, Chai L, Ng HH, Lufkin T, Robson P & Lim B 2006 Sall4 modulates embryonic stem cell pluripotency and early embryonic development by the transcriptional regulation of Pou5f1. *Nat Cell Biol* **8** 1114-1123.

Zhang P, Lee H, Brunzelle JS & Couture JF 2012 The plasticity of WDR5 peptide-binding cleft enables the binding of the SET1 family of histone methyltransferases. *Nucleic Acids Res* **40** 4237-4246.

Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM, Brown M, Li W & Liu XS 2008 Model-based analysis of ChIP-Seq (MACS). *Genome Biol* **9** R137.

Zhang Y, Ng HH, Erdjument-Bromage H, Tempst P, Bird A & Reinberg D 1999 Analysis of the NuRD subunits reveals a histone deacetylase core complex and a connection with DNA methylation. *Genes Dev* **13** 1924-1935.

Zhang Y, Vastenhouw NL, Feng J, Fu K, Wang C, Ge Y, Pauli A, van Hummelen P, Schier AF & Liu XS 2014 Canonical nucleosome organization at promoters forms during genome activation. *Genome Res* **24** 260-266.

Zhao M, Shirley CR, Hayashi S, Marcon L, Mohapatra B, Suganuma R, Behringer RR, Boissonneault G, Yanagimachi R & Meistrich ML 2004 Transition nuclear proteins are required for normal chromatin condensation and functional sperm development. *Genesis* **38** 200-213.

Zhao R, Nakamura T, Fu Y, Lazar Z & Spector DL 2011 Gene bookmarking accelerates the kinetics of post-mitotic transcriptional re-activation. *Nat Cell Biol* **13** 1295-1304.

Zheng J, Xia X, Ding H, Yan A, Hu S, Gong X, Zong S, Zhang Y & Sheng HZ 2008 Erasure of the paternal transcription program during spermiogenesis: the first step in the reprogramming of sperm chromatin for zygotic development. *Dev Dyn* **237** 1463-1476.

Zheng Z, Jia JL, Bou G, Hu LL, Wang ZD, Shen XH, Shan ZY, Shen JL, Liu ZH & Lei L 2012 rRNA genes are not fully activated in mouse somatic cell nuclear transfer embryos. *J Biol Chem* **287** 19949-19960.

Zhu D, Hunter SB, Vertino PM & Van Meir EG 2011 Overexpression of MBD2 in glioblastoma maintains epigenetic silencing and inhibits the antiangiogenic function of the tumor suppressor gene BAI1. *Cancer Res* **71** 5859-5870.

Zippo A, Serafini R, Rocchigiani M, Pennacchini S, Krepelova A & Oliviero S 2009 Histone crosstalk between H3S10ph and H4K16ac generates a histone code that mediates transcription elongation. *Cell* **138** 1122-1136.

Zon LI, Mather C, Burgess S, Bolce ME, Harland RM & Orkin SH 1991 Expression of GATA-binding proteins during embryonic development in Xenopus laevis. *Proc Natl Acad Sci U S A* **88** 10642-10646.

Chapter 10

Appendices

List of Appendices:

- Appendix 1: review article: 'Epigenetic reprogramming: is deamination key to active DNA demethylation?', *Teperek-Tkacz M., Pasque V., Gentsch G., Ferguson-Smith AC.*, Reproduction, 2011

- Appendix 2: review article: 'Nuclear reprogramming of sperm and somatic nuclei in eggs and oocytes.', *Teperek M., Miyamoto K.*, Reprod Med Biol., 2013

- Appendix 3: book chapter: 'Cloning of Amphibia' (part of 'Principles of Cloning'), *Teperek-Tkacz M., Byrne JA., Gurdon JB.*, Elsevier, 2014

- Appendix 4: research article: 'Nuclear Wave1 is required for reprogramming transcription in oocytes and for normal development.', *Miyamoto K., Teperek M., Yusa K., Allen GE., Bradshaw CR., Gurdon JB.*, Science, 2013

Appendix 1

Review article: ‘Epigenetic reprogramming: is deamination key to active DNA demethylation?’

Teperek-Tkacz M., Pasque V., Gentsch G., Ferguson-Smith AC., Reproduction, 2011

REPRODUCTION
REVIEW

Epigenetic reprogramming: is deamination key to active DNA demethylation?

Marta Teperek-Tkacz[1,2], Vincent Pasque[1,2], George Gentsch[3] and Anne C Ferguson-Smith[4]

[1]Wellcome Trust/Cancer Research UK Gurdon Institute, Cambridge, CB2 1QN, UK, [2]Department of Zoology, University of Cambridge, Cambridge, CB2 2EJ, UK, [3]Division of Systems Biology, MRC National Institute for Medical Research, London, NW7 1AA, UK and [4]Department of Physiology, Development and Neuroscience, University of Cambridge, Cambridge, CB2 3EG, UK

Correspondence should be addressed to M Teperek-Tkacz at Wellcome Trust/Cancer Research UK Gurdon Institute; Email: mt446@cam.ac.uk

Abstract

DNA demethylation processes are important for reproduction, being central in epigenetic reprogramming during embryonic and germ cell development. While the enzymes methylating DNA have been known for many years, identification of factors capable of mediating active DNA demethylation has been challenging. Recent findings suggest that cytidine deaminases may be key players in active DNA demethylation. One of the most investigated candidates is activation-induced cytidine deaminase (AID), best known for its role in generating secondary antibody diversity in B cells. We evaluate evidence for cytidine deaminases in DNA demethylation pathways in vertebrates and discuss possible models for their targeting and activity regulation. These findings are also considered along with alternative demethylation pathways involving hydroxymethylation.

Reproduction (2011) **142** 621–632

Introduction

The fifth carbon of cytosine in DNA can be either unmethylated or methylated to form 5-methylcytosine (5meC) and 5-hydroxymethylcytosine (5hmeC) on subsequent hydroxylation. This occurs on cytosines flanked by various nucleotides (Lister *et al.* 2009), but 5meC in a CpG dinucleotide context is the best characterised (Doerfler 2008). CpG dinucleotides are often enriched in promoters of genes and their methylation is associated with gene silencing (Chen *et al.* 2001, Kroft *et al.* 2001, Chan *et al.* 2004, Song *et al.* 2009). The opposite process, demethylation or replacement of methylated cytosines with unmethylated cytosines, can restore gene expression (Benvenuto *et al.* 1996, Papageorgis *et al.* 2010, Stengel *et al.* 2010). DNA methylation can therefore be used to establish correct gene expression patterns during development and differentiation (Maatouk *et al.* 2006, Song *et al.* 2009). Demethylation can be achieved by both passive and active mechanisms. Passive demethylation relies on DNA replication in the absence of DNA methyltransferase (DNMT) maintenance activity so that unmethylated cytosines are incorporated into new DNA strands. During active demethylation, methylated cytosines are replaced with unmethylated cytosines by an enzymatic process independent of DNA replication.

Active DNA demethylation is believed to occur on a global scale twice during mouse embryogenesis. First, the paternal genome is actively demethylated before the first cell division of the mouse zygote (Mayer *et al.* 2000, Oswald *et al.* 2000, Santos *et al.* 2002, Okada *et al.* 2010, Wossidlo *et al.* 2010). A second wave of global DNA demethylation occurs in primordial germ cells (PGCs) between embryonic days 11.25 and 13.5 (Hajkova *et al.* 2002, Feng *et al.* 2010, Surani & Hajkova 2010). Both events are likely to be involved in re-setting the genome for early development. Indeed, locus-specific DNA demethylation is required for reactivation of pluripotency genes during cell reprogramming (Simonsson & Gurdon 2004, Bhutani *et al.* 2010). It is also known that cloned embryos have defects in DNA methylation (Dean *et al.* 2001, Kang *et al.* 2001), which may result from aberrant gene reprogramming and lead to developmental abnormalities. Furthermore, demethylation of oncogenes is often associated with cancers (Nishigaki *et al.* 2005). Hence, DNA methylation dynamics are at the core of many developmentally regulated processes and their misregulation can lead to developmental defects and disease.

Potential DNA demethylases in vertebrates

The establishment of DNA methylation is achieved by DNMT enzymes that are well characterised in plants and animals. The mammalian DNMTs and their activity, specificity and regulation have been extensively

ISSN 1470–1626 (paper) 1741–7899 (online)
DOI: 10.1530/REP-11-0148
Online version via www.reproduction-online.org

reviewed (Bestor 2000, Hermann *et al.* 2004, Turek-Plewa & Jagodzinski 2005, Cheng & Blumenthal 2008). In contrast, enzymes involved in DNA demethylation have been identified in plants but their mammalian equivalents have been the subject of controversy. The bifunctional DNA glycosylases repressor of silencing 1 (ROS1) and Demeter (DME) are known to be the first enzymes in the demethylation pathway in plants (Fig. 1). They first recognise and bind methylated cytosine and then excise it from DNA through hydrolytic cleavage. This creates an abasic site that can be filled with an unmethylated cytosine by the DNA repair machinery (Zhu 2009). The mammalian glycosylases thymine DNA glycosylase (TDG) and methyl-CpG-binding domain protein 4 (MBD4) can both efficiently hydrolyse the N-glycosidic bond between thymine and deoxyribose, eventually leading to the base removal (Neddermann & Jiricny 1994, Hendrich *et al.* 1999), and they are capable of cleaving the bond between 5meC and deoxyribose *in vitro*; however, their activity in 5meCs is 30–40 times lower than in thymines (Zhu *et al.* 2000*a*, 2000*b*, Kim *et al.* 2009). *Tdg* deficiency in mice leads to embryonic lethality and aberrant *de novo* DNA methylation of developmentally regulated genes; however, DNA methylation levels at fertilisation and in PGCs have not been assessed in these mutant animals (Cortazar *et al.* 2011, Cortellino *et al.* 2011) and 5meC levels are not altered in *Mbd4*-deficient mice (Millar *et al.* 2002, Wong *et al.* 2002). Other proteins have been suggested to possess DNA demethylation activity, including DNMTs (Metivier *et al.* 2008), MBD2 (Bhattacharya *et al.* 1999, Detich *et al.* 2002), MBD3 (Brown *et al.* 2008) and growth arrest and DNA damage-inducible protein alpha (GADD45A; Barreto *et al.* 2007). GADD45A has been shown to contribute to active demethylation of plasmid DNA injected into frog oocytes (Barreto *et al.* 2007, Schafer *et al.* 2010) and locus-specific promoter demethylation in cultured mammalian cells (Schmitz *et al.* 2009, Schafer *et al.* 2010). However, *Gadd45a* mutant mice do not show defects in DNA methylation (Engel *et al.* 2009); hence, its role in DNA demethylation has also been questioned (Jin *et al.* 2008).

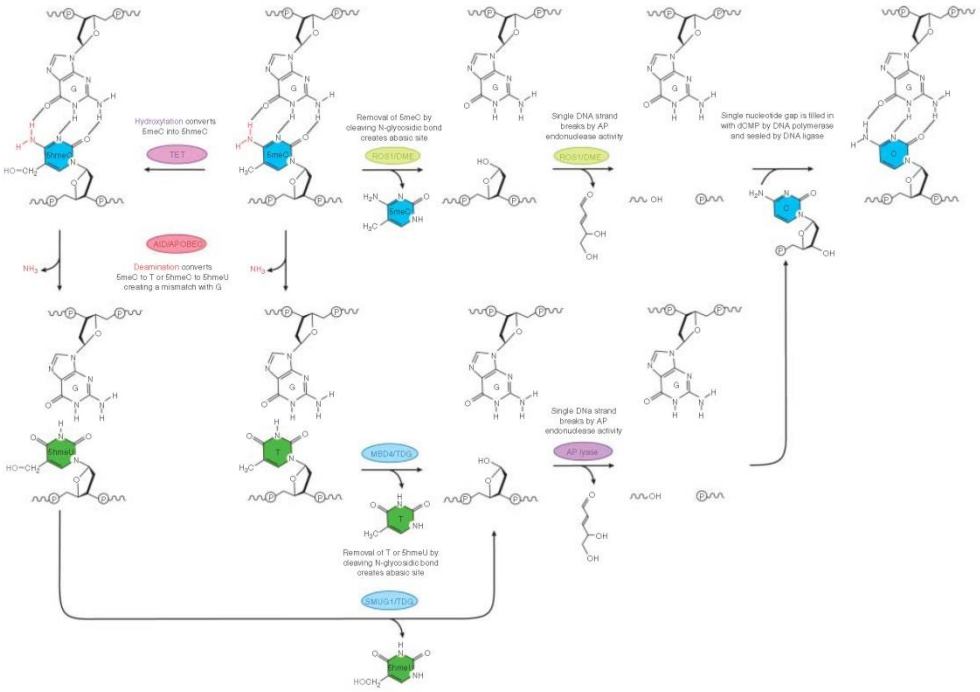


Figure 1 Possible mechanisms of active DNA demethylation. In plants, 5meC can be directly converted to unmethylated C by bifunctional DNA glycosylases ROS1 and DME, whereas in animals 5meC can be first hydroxylated to 5hmeC and subsequently deaminated by AID/APOBEC or directly deaminated by AID/APOBEC. These deamination products (5hmeU or T respectively) are further processed by the BER pathway.

Hydrolytic deamination catalysed by AID/APOBEC family members

Recent reports suggest that some members of the vertebrate-specific activation-induced cytidine deaminase (AID)/apolipoprotein B mRNA editing enzyme, catalytic polypeptide (APOBEC) family such as AID (Morgan *et al.* 2004, Rai *et al.* 2008, Bhutani *et al.* 2010, Popp *et al.* 2010), APOBEC1 (Morgan *et al.* 2004, Guo *et al.* 2011), APOBEC2 (Rai *et al.* 2008, Guo *et al.* 2011) and some proteins of the APOBEC3 branch (Guo *et al.* 2011) could play a key role in active DNA demethylation. AID and APOBEC proteins are zinc-dependent cytidine deaminases acting on single-stranded polynucleotides and deaminating cytosines in different contexts (Navaratnam *et al.* 1993, Teng *et al.* 1993, Bransteitter *et al.* 2003, Chelico *et al.* 2006). The zinc ion acting with the deaminase is coordinated by three amino acids: one histidine and two cysteines or three cysteines (Conticello *et al.* 2005). These residues are conserved within the motif signatures [H/C]xE and $PCx_{2-4}C$ (Gerber & Keller 1999). The latter motif is a zinc finger-like feature, a key to several proteins regulating chromatin modifications including DNA and histone methylation (Blackledge *et al.* 2010). This catalytic zinc site receives its activity by a fourth ligand, a water molecule, which is coordinated by the carboxylate ion of glutamate in [H/C]xE. The carboxylate group facilitates proton shuttling, which converts a water molecule – once it is trapped within the zinc coordination sphere – into a reactive hydroxide ion (Betts *et al.* 1994). Hydrolytic deamination of cytosines occurs through a nucleophilic attack of the zinc hydroxide on the pyrimidine carbon 4 carrying the amine group (Betts *et al.* 1994).

Among the AID/APOBEC family, *Aid* and *Apobec1* are expressed in mammalian oocytes and embryos at stages when global DNA demethylation occurs (Morgan *et al.* 2004). They can deaminate 5meC to thymine *in vitro* which, followed by G-T mismatch repair, could lead to DNA demethylation (Morgan *et al.* 2004). APOBEC1 was originally found to convert cytosine into uracil in the apolipoprotein B transcript (Navaratnam *et al.* 1993, Teng *et al.* 1993), whereas AID catalyses the same base conversion repetitively and preferentially on single-stranded DNA along immunoglobulin loci (Muramatsu *et al.* 1999, 2000, Pham *et al.* 2003). Interestingly, experiments by Rai *et al.* (2008) suggest that coupling of Aid and Apobec2 along with the glycosylase Mbd4 can lead to active DNA demethylation in zebrafish embryos.

Aid/Apobec-driven DNA demethylation in zebrafish

Demethylation of *in vitro* methylated circular or linearised DNA occurs when it is injected into zebrafish embryos at the one-cell stage. This is followed by its remethylation several hours later (Collas 1998, Rai *et al.* 2008). Recent overexpression and knockdown studies suggest that the presence of Aid/Apobec2 together with the Mbd4 DNA glycosylase and Gadd45 is responsible for the demethylation of plasmid DNA and concomitant demethylation of the embryonic genome (Rai *et al.* 2008). Knockdown of Aid or Mbd4 results in locus-specific hypermethylation and aberrant expression of genes important for neurogenesis. The repair of G-T mismatches is meant to be immediately initiated by the glycosylase Mbd4. Gadd45 may increase the DNA demethylation efficiency by promoting the Aid/Apobec2 physical interaction with Mbd4 (Rai *et al.* 2008). These results suggest that DNA glycosylases, known for their roles in DNA demethylation in plants, can also contribute to this process in vertebrates. However, according to this model, DNA glycosylases do not initiate the removal of 5meC as in plants, but only cut the N-glycosidic bond leading to the removal of thymines from G-T mismatches (Fig. 1). This may explain the lack of change in 5meC levels in *Mbd4* knockout mice, which instead have a higher frequency of G-T mutations (Millar *et al.* 2002, Wong *et al.* 2002).

While the model proposed by Rai *et al.* is appealing, caveats remain in the mechanistic understanding of Aid/Apobec2-induced demethylation. First, using cytidine deaminases in DNA demethylation seems risky. Cytidine deamination is mutagenic: besides creating G-T mismatches, the activity of Aid is primarily directed towards unmethylated cytosines, which converts them to uracils producing G-U mismatches, for example as occurs in B cells allowing secondary antibody diversification. Mismatches, if not repaired before DNA replication, will create permanent mutations. Secondly, these experiments have been mostly based on the artificial introduction of methylated plasmid DNA into zebrafish embryos. It is not known why this also leads to partial demethylation of the zebrafish genome itself. It is possible that the up-regulation of Aid/Apobec2 and Gadd45 expression and the stimulation of DNA demethylation is a type of immunological response to the presence of methylated exogenous DNA. Whether such a mechanism is common to other organisms is not yet known. Thirdly, it is not clear how symmetric demethylation of CpG in double-stranded DNA occurs since simultaneous 5meC deamination and subsequent nucleotide removal on both DNA strands would create mutagenic double-stranded breaks (Jiricny & Menigatti 2008). However, it is known that the genomes of vertebrate species such as fugu and zebrafish undergo CpG to TpG transitions over evolutionary time (Bird 1980, Glass *et al.* 2007). It cannot be excluded that these transitions occur through deamination of 5meC in CpG dinucleotides and subsequent erroneous repair. This is substantiated indirectly by the fact that CpG islands, generally protected from DNA methylation, do not show accumulation of C-to-T transitions.

Additionally, the role of Apobec2 remains controversial. The expression of APOBEC2 is mainly confined to

muscle and heart tissue (Liao *et al.* 1999). *Apobec2* knockout mice show a significant loss of body weight, a myofibre-type shift from fast to slow and centronuclear-like myopathy with age (Mikl *et al.* 2005, Sato *et al.* 2010). Recently, Apobec2 has been further implicated in left–right axis determination during early embryogenesis in *Xenopus* through an inhibition of the transforming growth factor β signalling pathway (Vonica *et al.* 2011). It is not clear how Apobec2 contributes to such a variety of developmental processes. Furthermore, its role as a potential DNA deaminase has been questioned due to lack of crucial residues (such as the positively charged amino acids found in AID) at the N-terminus facilitating binding to single-stranded DNA, and a tryptophan, which is located in the vicinity of the PCxxC motif and is required for APOBEC3G-catalysed DNA deamination (Pham *et al.* 2003, Chen *et al.* 2007, Sato *et al.* 2010).

Evidence for AID as a mammalian DNA demethylase

Two recent studies support a role for AID in DNA demethylation in mammalian systems (Bhutani *et al.* 2010, Popp *et al.* 2010). Bhutani *et al.* have shown that the use of siRNA against *Aid* interferes with reprogramming, demethylation and reactivation of the pluripotency genes *POU5F1* (*OCT4*) and *NANOG* in heterokaryons (fused cells containing multiple, genetically different nuclei) of human fibroblasts and mouse embryonic stem cells. Chromatin immunoprecipitation experiments implied that AID is bound to silent promoters of human somatic and mouse ES cells, but not to active, already unmethylated promoters in ES cell nuclei (Bhutani *et al.* 2010). These findings suggest a role for AID in promoter DNA demethylation, but leave the question open how AID associates with chromatin without causing immediate deamination in silent promoters. Moreover, the ES cells used for the fusions were actively dividing. Active divisions imply DNA replication that can therefore lead to passive DNA demethylation. It would be interesting to test whether AID is also important for active DNA demethylation in heterokaryons generated with cell cycle synchronised ES cells. In order to further confirm an involvement of AID in active DNA demethylation in heterokaryons, it would be worth performing the experiment using *Aid* mutant cells.

Genetic evidence for the involvement of AID in DNA demethylation has been obtained from examining the DNA methylation level in PGCs from *Aid* knockout mice (Popp *et al.* 2010). Although the data reporting *Aid* expression in PGCs at the time of global DNA demethylation has been challenged (Morgan *et al.* 2004, Hajkova *et al.* 2010), in $Aid^{-/-}$ mice, the erasure of DNA methylation marks in PGCs is up to three times lower compared with wild-type controls. Intriguingly, considerable DNA demethylation still occurs in $Aid^{-/-}$ PGCs (Popp *et al.* 2010). This suggests that residual DNA demethylation results from the activity of other deaminases such as the co-expressed *Apobec1* (Morgan *et al.* 2004, Hajkova *et al.* 2010), other mechanisms that do not require deamination, or a combination of both. Similarly, reduced levels of DNA demethylation in zebrafish have only been observed after simultaneous knockdown of Aid and Apobec2 (Rai *et al.* 2008), suggesting redundancy among members of the AID/APOBEC family.

A consensus mechanism of deaminase-mediated DNA demethylation

The results of most of the experiments described above provide an emerging consensus as to how cytidine deaminases can initiate the removal of 5meC from DNA (Fig. 1). The hydrolytic deamination of 5meC converts the base to thymine. It is important to note, however, that cytidine deaminases preferentially act on unmethylated cytosines in DNA. Both activities are considered mutagenic. If cytidine deaminases act on cytosine, the resultant uracil is recognised and repaired by uracil DNA glycosylases (Talpaert-Borle *et al.* 1982, Olsen *et al.* 1989). Thymine (T), resulting from deamination of 5meC, is a true base in genomic DNA, so mismatch repair proteins need to distinguish Ts in a mismatch with guanosines from correctly paired Ts in DNA. Thymine DNA glycosylases (TDG/MBD4) are able to selectively recognise such mismatches as they interact not only with the T but also with the opposing base pair (Barrett *et al.* 1998, Yoon *et al.* 2003, Maiti *et al.* 2008). Additionally, MBD4 not only recognises G-T mismatches but also has a methyl-binding domain (Hendrich *et al.* 1999, Wu *et al.* 2003). This domain could target MBD4 to 5meC and mark it as a potential site for deamination. Moreover, GADD45 may couple the action of MBD4 with AID/APOBEC2 (Rai *et al.* 2008). Physical interaction of deaminases with glycosylases could be critical for immediate recognition of mismatches generated by deamination of 5meC, therefore preventing the mutagenic activity of deaminases. Thymine DNA glycosylases (TDG/MBD4) are the first enzymes of the base excision repair (BER) pathway, which triggers the removal of the mispaired base T from the DNA (Fig. 1). This is further processed by endonucleases, such as AP-endonuclease 1 (APE1). In a recent screen for factors promoting DNA demethylation, the RING finger protein 4 (RNF4) has been identified. RNF4 has been shown to enhance DNA demethylation by coupling TDG and APE1 (Hu *et al.* 2010). After APE1-mediated hydrolytic cleavage of the phosphodiester DNA backbone, the DNA polymerase β (POLB) removes the remaining deoxyribose moiety and fills in the nascent single nucleotide gap with an unmethylated cytosine. Eventually, the break is sealed by a DNA ligase (Dalhus *et al.* 2009, Kunz *et al.* 2009). Additionally, non-enzymatic proteins such

as X-ray repair cross-complementing proteins 1 and 2 (XRCC1 and XRCC2) create a scaffold for the mismatch repair machinery (Tebbs *et al.* 1999, Adam *et al.* 2007).

Incorrectly paired nucleotides can also be excised by the nucleotide excision repair (NER) pathway. In contrast to the BER pathway, the NER pathway removes an ~29 bp long single-stranded DNA fragment including the incorrectly paired nucleotide(s). Subsequently, new DNA is synthesised using the undamaged strand as a template and DNA ligase creates covalent phosphodiester bonds (reviewed in Niehrs (2009)). It cannot be excluded that at least some mispaired nucleotides resulting from 5meC deamination can be repaired using NER enzymes. Indeed, knockdown of components of the NER pathway (Gadd45a, XPA, XPG and XPF) inhibits DNA demethylation in *Xenopus* oocytes (Barreto *et al.* 2007, Schmitz *et al.* 2009). DNA demethylation is also inhibited by treatment with chemicals specifically blocking the NER pathway, but not the BER pathway, both in *Xenopus* oocytes and cultured HEK293 cells (Schafer *et al.* 2010). Furthermore, it has been shown that DNA demethylation in mouse zygotes leads to the creation of DNA breaks. Interestingly, aphidicolin treatment blocks repair of the breaks (Wossidlo *et al.* 2010). Aphidicolin blocks DNA synthesis but has no inhibitory effects on the BER-specific POLB (the major polymerase of the BER pathway). This could suggest that repair pathways other than BER may be involved in DNA demethylation in mouse zygotes.

Importance of BER enzymes in mouse embryogenesis

If DNA demethylation was to be initiated by deaminases and resultant mismatches processed by DNA repair pathways, then the enzymes from these pathways should also be expressed at times when DNA demethylation occurs. Indeed, it has been reported that MBD4 together with other components of the BER pathway, including *Ape1*, *Polb* and DNA ligase III, are expressed at all stages of mouse preimplantation development (Ruddock-D'Cruz *et al.* 2008, May *et al.* 2009). Moreover, BER enzymes are present in the paternal pronucleus in the zygote and in PGCs at the time of global DNA demethylation (Hajkova *et al.* 2010). *Tdg*, *Polb* as well as *Xrcc1* and *Xrcc2* knockout mice are lethal either at the embryonic (Gu *et al.* 1994, Tebbs *et al.* 1999, Deans *et al.* 2000, Adam *et al.* 2007, Cortazar *et al.* 2007, 2011, Cortellino *et al.* 2011) or at the neonatal stage (Sugo *et al.* 2000), suggesting that the BER pathway is important for embryonic development. The spatial and temporal co-expression of these components and the reported interactions among them (Bennett *et al.* 1997, Vidal *et al.* 2001, Dianova *et al.* 2004, Parsons *et al.* 2005, Fitzgerald & Drohat 2008, Hu *et al.* 2010) suggest that they may form functional deaminase–BER complexes. The existence of such a complex would ensure that any deaminated 5meC is immediately recognised, repaired and replaced with an unmethylated cytosine, hence diminishing the risk of mutating the genome through deamination.

How is site-specific activity of deaminases achieved?

Since the activity of AID is mutagenic, cells maintain a tight control on its nuclear localisation. AID is kept away from DNA by using a strong cytoplasmic retention signal, as well as a strong nuclear export signal (Patenaude *et al.* 2009). AID is also actively imported to the nucleus, and its concentration is thought to be regulated by the proteasome; on proteasome inhibition, a ubiquitinated nuclear form is found (Aoufouchi *et al.* 2008). The mechanisms used to target cytidine deamination to particular sites are not clear. AID/APOBECs could interact with other targeting proteins, such as MBD4 that has a 5meC-binding domain. However, this would result in binding of deaminases to all methylated cytosines in DNA. Intriguingly, Bhutani *et al.* (2010) have detected AID bound to silent promoters both in human fibroblasts and in ES cells, which are not demethylated. It may be that AID binding is not sufficient to trigger its activity; hence, other cues are required. It has been reported that AID can only induce deamination in the context of single-stranded DNA (Larijani & Martin 2007, Larijani *et al.* 2007, Brar *et al.* 2008). It is not known whether melting of double-stranded DNA *in vivo* is sufficient to induce AID activity; however, the presence of single-stranded DNA could represent such a signal. But how could it become activated, if, as in most cases, the DNA in eukaryotic cells is double-stranded? We present three possible models for AID activation and targeting.

Model 1: activation of deaminases by active transcription

It has been suggested that DNA demethylation of silenced genes cannot occur without histone acetylation-induced transcription (D'Alessio *et al.* 2007). Transcription leads to a transient formation of single-stranded DNA (Leibovitch & Harel 1978, Leibovitch *et al.* 1979), and it has been hypothesised that transcription may be needed for AID targeting (Chaudhuri *et al.* 2003, Shen *et al.* 2009). When overexpressed in NIH 3T3 cells, AID more efficiently edits a *GFP* reporter gene transcribed at higher levels, suggesting that transcription may stimulate its activity (Yoshikawa *et al.* 2002). Recently, elongator complex protein 3 (ELP3) and three other proteins from the ELP family comprising the elongator complex have been reported to be necessary for paternal DNA demethylation in mouse zygotes (Okada *et al.* 2010). The elongator complex has been previously found to be associated with RNA polymerase II and involved in

transcriptional elongation (Otero *et al.* 1999). Interestingly, it has been shown that interactions with the elongator complex and transcriptional elongation factor SPT5 may direct AID to transcribed targets (Besmer *et al.* 2006, Pavri *et al.* 2010). SPT5 facilitates transcriptional targeting by delivering AID to stalled RNA polymerase II, which was suggested to occur more frequently in the presence of R loop secondary structures (Pavri *et al.* 2010). The importance of transcription for AID targeting is further substantiated by recent findings that the RNA exosome complex recruits AID to both strands of transcribed DNA to ensure simultaneous deamination of template and non-template DNA (Basu *et al.* 2011). The RNA exosome is meant to remove nascent RNA from template DNA to expose it to AID for binding. Subsequently, protein kinase A may stabilise single-stranded DNA and support the recruitment of the repair machinery (Vuong *et al.* 2009). The notion that transcription may be a prerequisite for active DNA demethylation is challenged by findings in zebrafish and mouse embryos. Active demethylation of plasmid DNA in zebrafish occurs in the absence of transcription (Collas 1998). Additionally, the onset of transcription in mouse zygotes at the one-cell stage occurs after the erasure of methyl marks from the paternal genome (Bouniol *et al.* 1995, Aoki *et al.* 1997). To conclude, even though many experiments point towards a central role for transcription in the activity of deaminases making this model an attractive one, it is possible that in some cases deaminases may be regulated using alternative pathways, as discussed below.

Model 2: deaminase targeting and chromatin modification

Changes in chromatin state could be sufficient to target AID/APOBECs. Some histone modifications, like di- and tri-methylation of lysine 9 of histone H3 (H3K9me2/3), are associated with chromatin compaction, whereas others, like acetylated lysine 9 of histone H3 (H3K9ac) or trimethylated lysine 4 of histone H3 (H3K4me3), can make the chromatin more accessible (Jenuwein & Allis 2001). It has been suggested that the presence of H3K9ac, H3K14ac and H3K4me3 may be important for AID targeting (Wang *et al.* 2009). Interestingly, at the time of global DNA demethylation in the mouse zygote, the paternal genome is devoid of repressive H3K9me2/3 marks (Liu *et al.* 2004, Santos *et al.* 2005). Similarly, active DNA demethylation in PGCs occurs after a loss of repressive H3K9me2 (Hajkova *et al.* 2008). Furthermore, the presence of H3K4me2/me3 is associated with pluripotency gene reactivation during cell reprogramming (Murata *et al.* 2010). Thus, it is conceivable that accessibility and state of the chromatin may be important to recruit AID to the sites of deamination.

Model 3: RNA-mediated deaminase targeting

Another appealing mechanism for directing deaminases to specific sites in the genome could involve RNA-mediated targeting. Recently, it has been shown that non-coding RNA can bind to a complementary rDNA promoter region to form a triple helix. Formation of DNA:RNA triplexes facilitates recruitment of the DNMT3B methyltransferase to the rDNA promoter (Schmitz *et al.* 2010). Enzyme targeting by interaction with RNAs is attractive, as it ensures a high degree of site specificity. It could be that some of the reported requirements of transcription for AID activity might reflect a need for the generation of guiding non-coding RNA (Chaudhuri *et al.* 2003, Shen *et al.* 2009). However, so far, there is no evidence supporting this hypothesis.

With the current state of knowledge, it is difficult to decide which of the proposed models (if any) is true. It could be that all, or some combination of them, are utilised at different developmental stages or in different model systems. It should also be noted that most of our current mechanistic understanding about AID is based on its immunological role and may not be relevant for deaminase-mediated DNA demethylation. More work has to be done in order to shed light on 5meC deaminase targeting, regulation and the reciprocal relations between active DNA demethylation, chromatin changes, transcription and the requirement of single-stranded DNA for deaminase activity.

One universal mechanism or several independent ones?

Despite growing evidence for the involvement of AID and other cytidine deaminases in DNA demethylation, many questions still remain (Box 1). First, there is a need for more genetic evidence for the role of cytidine deaminases in DNA demethylation. Although *Aid* knockout mice have significantly decreased global DNA demethylation levels in PGCs, they are viable and residual demethylation in PGCs is still observed (Popp *et al.* 2010). Because of possible redundancy of AID with other cytidine deaminases, it is important to simultaneously knock out other cytidine deaminases to see whether residual DNA demethylation is still observed and whether paternal DNA demethylation after fertilisation in the zygote still occurs. Furthermore, to examine the involvement of other components of the proposed DNA demethylation pathways, they too will need to be systematically depleted from the examined cells/animals. This will be challenging as the list of potential candidates is large (Table 1) and possible redundancy has to be taken into account when analysing the roles of particular factors. As we cannot exclude that distinct active DNA demethylation mechanisms exist, the interpretation of results between different systems should also be conducted with caution. It is important to

Box 1 Outstanding questions

- How are deaminases targeted to specific promoters?
- Is transcription necessary for active DNA demethylation?
- Is the presence of ssDNA sufficient to trigger deaminase activity?
- How are symmetric 5meCs removed in deaminase-driven pathway?
- Are deaminases involved in the global DNA demethylation, e.g. of the paternal genome in fertilised mouse zygotes? What other components of the deaminase-mediated DNA demethylation pathway are necessary for its efficiency?
- How many pathways are used (and integrated) to mediate active DNA demethylation in mammals?

consider that mechanisms may distinguish between genome-wide and locus-specific DNA demethylation. Moreover, it is very likely that various other enzymes from distinct pathways can act synergistically (or independently) to achieve DNA demethylation.

Hydroxylation and deamination

The involvement of other pathways has received considerable recent attention. In particular, recognition of a contribution of TET proteins to DNA demethylation comes from studies on zygotic reprogramming (Hajkova *et al.* 2010, Wossidlo *et al.* 2011). TET proteins have the ability to convert 5meC to 5hmeC (Tahiliani *et al.* 2009, Ito *et al.* 2010). Interestingly, *Tet1* and *Tet2* are highly expressed in embryonic stem cells and in PGCs at the time of global DNA demethylation and are induced during reprogramming of fibroblasts to induced pluripotent stem cells (Hajkova *et al.* 2010, Koh *et al.* 2011). However, recent findings in stem cells with depletion of TET1 or TET2 demonstrate that the correlation of TET activity and DNA methylation pattern is complex (Ito *et al.* 2010, Ko *et al.* 2010, Koh *et al.* 2011) and TET1 protein, as well as being involved in DNA demethylation and sustaining transcriptional activity of several genes, may also be responsible for the silencing of others (Ficz *et al.* 2011, Williams *et al.* 2011, Wu *et al.* 2011). It is not known whether DNA demethylation in PGCs is impaired in *Tet*-deficient mice. In contrast, *Tet3* is expressed in mouse oocytes and is present in the early zygote. Absence of TET3 during the time of active demethylation results in failure to demethylate the paternal genome, suggesting that the conversion of 5meC to 5hmeC may constitute an intermediate step in the active demethylation process (Iqbal *et al.* 2011, Wossidlo *et al.* 2011).

It is possible that both enzymatic pathways involving deaminases and TET proteins could cooperatively lead to DNA demethylation. Indeed, it has been shown recently that co-expression of TET1 and AID or other cytidine deaminases can increase the efficiency of DNA demethylation of reporter plasmid DNA transfected into cultured cells (Guo *et al.* 2011). AID has been proposed to preferentially deaminate 5hmeC generated by TET1 to produce 5-hydroxymethyl uracil (5hmeU), which can then be processed by glycosylases from the BER pathway (Guo *et al.* 2011). Furthermore, it has been shown recently that TDG can initiate the removal of 5hmeU generated by AID-mediated deamination of 5hmeC (Cortellino *et al.* 2011; Fig. 1). This is an attractive model, because TET proteins, in contrast to cytidine deaminases, can efficiently act on 5meC in double-stranded DNA (Tahiliani *et al.* 2009, Ito *et al.* 2010), and because the product of 5hmeC deamination – 5hmeU – can be recognised by BER pathway uracil glycosylases. However, under at least some developmental circumstances, this orchestrated process of oxidation and deamination may not be required for active DNA demethylation. It has been shown that during DNA demethylation at fertilisation in mouse zygotes, some 5mC is converted to unmethylated cytosine without a 5hmeC intermediate,

Table 1 Example of proteins potentially involved in active DNA demethylation driven by cytidine deaminases.

Cytidine deaminases	Protein coupling deaminases, glycosylases and enzymes from the BER pathway	Proteins from the BER pathway
AID, APOBEC1, APOBEC2A, APOBEC2B, APOBEC3A, APOBEC3B, APOBEC3C, APOBEC3D (known as APOBEC3E), APOBEC3F, APOBEC3G, APOBEC3H, APOBEC4	GADD45A, GADD45B, GADD45G, RNF4	*Thymine DNA glycosylases:* MBD4, TDG *DNA endonucleases:* APE1, APE2 *DNA polymerases:* POLB *DNA ligases:* DNA ligase I, DNA ligase III *Protein coupling BER enzymes:* XRCC1

which would point towards the involvement of DNA demethylation mechanisms independent of TET proteins and possibly using cytidine deaminases (Wossidlo *et al.* 2011). One cannot rule out that various mechanisms and distinct enzymes might act at different developmental stages. For example, active demethylation pathways in the germline may be different from those operating in the zygote or at other developmental stages. Consistent with this, it is known that imprinting marks are not erased in the first wave of global demethylation after fertilisation in mouse zygotes (Mayer *et al.* 2000, Oswald *et al.* 2000), but that they are removed in the second wave occurring in mouse PGCs (Hajkova *et al.* 2002, Sato *et al.* 2003, Surani & Hajkova 2010). In addition, several distinct active DNA demethylation mechanisms may be active at the same time to target different regions of the genome. Clearly, our understanding of active DNA demethylation mechanisms is still limited.

Conclusions

In summary, despite several lines of evidence supporting the hypothesis that DNA demethylation can start with 5meC deamination, it is likely that, *in vivo*, an AID-driven pathway is not the only one leading to active DNA demethylation. It is crucial to get further insights into key mechanisms of DNA demethylation control. Determining the mechanisms of DNA demethylation at a molecular level will be important for understanding the erasure and establishment of the normal epigenetic programme in the germline and in the zygote and will provide insights into the epigenetic perturbations implicated in assisted reproductive technologies (Maher 2005, Laprise 2009). Ultimately, it may contribute to the design of treatments for diseases associated with aberrant DNA methylation, such as imprinting disorders or cancer. For example, even though direct links to genome-wide levels of 5mC have not been established, it has been shown that up-regulation of AID-mediated DNA demethylation pathway is associated with human colon cancers (Rai *et al.* 2010). In addition, deciphering mechanisms leading to DNA demethylation could help circumvent problems with inefficient DNA demethylation accompanying induced pluripotent stem cell derivation (Takahashi & Yamanaka 2006, Lister *et al.* 2011). We are getting closer to an understanding of how active DNA demethylation is achieved on a molecular level, and how cytidine deaminases contribute to this process where repair is desired rather than rejected as observed at immunoglobulin loci to cause antibody diversity in B cells.

Declaration of interest

The authors declare that there is no conflict of interest that could be perceived as prejudicing the impartiality of the research reported.

Funding

This work was supported by the grants from the MRC, Wellcome Trust [089613], [081277], [081562] and a Wallonia-Brussels International Excellence Grant.

Acknowledgements

We would like to thank R Shields, K Miyamoto, J Jullien and J Green for critically reading and commenting on the manuscript.

References

Adam J, Deans B & Thacker J 2007 A role for Xrcc2 in the early stages of mouse development. *DNA Repair* **6** 224–234. (doi:10.1016/j.dnarep.2006.10.024)

Aoki F, Worrad DM & Schultz RM 1997 Regulation of transcriptional activity during the first and second cell cycles in the preimplantation mouse embryo. *Developmental Biology* **181** 296–307. (doi:10.1006/dbio.1996.8466)

Aoufouchi S, Faili A, Zober C, D'Orlando O, Weller S, Weill JC & Reynaud CA 2008 Proteasomal degradation restricts the nuclear lifespan of AID. *Journal of Experimental Medicine* **205** 1357–1368. (doi:10.1084/jem.20070950)

Barreto G, Schafer A, Marhold J, Stach D, Swaminathan SK, Handa V, Doderlein G, Maltry N, Wu W, Lyko F *et al.* 2007 Gadd45a promotes epigenetic gene activation by repair-mediated DNA demethylation. *Nature* **445** 671–675. (doi:10.1038/nature05515)

Barrett TE, Savva R, Panayotou G, Barlow T, Brown T, Jiricny J & Pearl LH 1998 Crystal structure of a G:T/U mismatch-specific DNA glycosylase: mismatch recognition by complementary-strand interactions. *Cell* **92** 117–129. (doi:10.1016/S0092-8674(00)80904-6)

Basu U, Meng FL, Keim C, Grinstein V, Pefanis E, Eccleston J, Zhang T, Myers D, Wasserman CR, Wesemann DR *et al.* 2011 The RNA exosome targets the AID cytidine deaminase to both strands of transcribed duplex DNA substrates. *Cell* **144** 353–363. (doi:10.1016/j.cell.2011.01.001)

Bennett RA, Wilson DM III, Wong D & Demple B 1997 Interaction of human apurinic endonuclease and DNA polymerase beta in the base excision repair pathway. *PNAS* **94** 7166–7169. (doi:10.1073/pnas.94.14.7166)

Benvenuto G, Carpentieri ML, Salvatore P, Cindolo L, Bruni CB & Chiariotti L 1996 Cell-specific transcriptional regulation and reactivation of galectin-1 gene expression are controlled by DNA methylation of the promoter region. *Molecular and Cellular Biology* **16** 2736–2743.

Besmer E, Market E & Papavasiliou FN 2006 The transcription elongation complex directs activation-induced cytidine deaminase-mediated DNA deamination. *Molecular and Cellular Biology* **26** 4378–4385. (doi:10.1128/MCB.02375-05)

Bestor TH 2000 The DNA methyltransferases of mammals. *Human Molecular Genetics* **9** 2395–2402. (doi:10.1093/hmg/9.16.2395)

Betts L, Xiang S, Short SA, Wolfenden R & Carter CW Jr 1994 Cytidine deaminase. The 2.3 A crystal structure of an enzyme: transition-state analog complex. *Journal of Molecular Biology* **235** 635–656. (doi:10.1006/jmbi.1994.1018)

Bhattacharya SK, Ramchandani S, Cervoni N & Szyf M 1999 A mammalian protein with specific demethylase activity for mCpG DNA. *Nature* **397** 579–583. (doi:10.1038/17533)

Bhutani N, Brady JJ, Damian M, Sacco A, Corbel SY & Blau HM 2010 Reprogramming towards pluripotency requires AID-dependent DNA demethylation. *Nature* **463** 1042–1047. (doi:10.1038/nature08752)

Bird AP 1980 DNA methylation and the frequency of CpG in animal DNA. *Nucleic Acids Research* **8** 1499–1504. (doi:10.1093/nar/8.7.1499)

Blackledge NP, Zhou JC, Tolstorukov MY, Farcas AM, Park PJ & Klose RJ 2010 CpG islands recruit a histone H3 lysine 36 demethylase. *Molecular Cell* **2** 179–190. (doi:10.1016/j.molcel.2010.04.009)

Bouniol C, Nguyen E & Debey P 1995 Endogenous transcription occurs at the 1-cell stage in the mouse embryo. *Experimental Cell Research* **218** 57–62. (doi:10.1006/excr.1995.1130)

Bransteitter R, Pham P, Scharff MD & Goodman MF 2003 Activation-induced cytidine deaminase deaminates deoxycytidine on single-stranded DNA but requires the action of RNase. *PNAS* **100** 4102–4107. (doi:10.1073/pnas.0730835100)

Brar SS, Sacho EJ, Tessmer I, Croteau DL, Erie DA & Diaz M 2008 Activation-induced deaminase, AID, is catalytically active as a monomer on single-stranded DNA. *DNA Repair* **7** 77–87. (doi:10.1016/j.dnarep.2007.08.002)

Brown SE, Suderman MJ, Hallett M & Szyf M 2008 DNA demethylation induced by the methyl-CpG-binding domain protein MBD3. *Gene* **420** 99–106. (doi:10.1016/j.gene.2008.05.009)

Chan Y, Fish JE, D'Abreo C, Lin S, Robb GB, Teichert AM, Karantzoulis-Fegaras F, Keightley A, Steer BM & Marsden PA 2004 The cell-specific expression of endothelial nitric-oxide synthase: a role for DNA methylation. *Journal of Biological Chemistry* **279** 35087–35100. (doi:10.1074/jbc.M405063200)

Chaudhuri J, Tian M, Khuong C, Chua K, Pinaud E & Alt FW 2003 Transcription-targeted DNA deamination by the AID antibody diversification enzyme. *Nature* **422** 726–730. (doi:10.1038/nature01574)

Chelico L, Pham P, Calabrese P & Goodman MF 2006 APOBEC3G DNA deaminase acts processively $3' \rightarrow 5'$ on single-stranded DNA. *Nature Structural & Molecular Biology* **13** 392–399. (doi:10.1038/nsmb1086)

Chen C, Yang MC & Yang TP 2001 Evidence that silencing of the HPRT promoter by DNA methylation is mediated by critical CpG sites. *Journal of Biological Chemistry* **276** 320–328. (doi:10.1074/jbc.M007096200)

Chen KM, Martemyanova N, Lu Y, Shindo K, Matsuo H & Harris RS 2007 Extensive mutagenesis experiments corroborate a structural model for the DNA deaminase domain of APOBEC3G. *FEBS Letters* **581** 4761–4766. (doi:10.1016/j.febslet.2007.08.076)

Cheng X & Blumenthal RM 2008 Mammalian DNA methyltransferases: a structural perspective. *Structure* **16** 341–350. (doi:10.1016/j.str.2008.01.004)

Collas P 1998 Modulation of plasmid DNA methylation and expression in zebrafish embryos. *Nucleic Acids Research* **26** 4454–4461. (doi:10.1093/nar/26.19.4454)

Conticello SG, Thomas CJ, Petersen-Mahrt SK & Neuberger MS 2005 Evolution of the AID/APOBEC family of polynucleotide (deoxy)cytidine deaminases. *Molecular Biology and Evolution* **22** 367–377. (doi:10.1093/molbev/msi026)

Cortazar D, Kunz C, Saito Y, Steinacher R & Schar P 2007 The enigmatic thymine DNA glycosylase. *DNA Repair* **6** 489–504. (doi:10.1016/j.dnarep.2006.10.013)

Cortazar D, Kunz C, Selfridge J, Lettieri T, Saito Y, MacDougall E, Wirz A, Schuermann D, Jacobs AL, Siegrist F *et al.* 2011 Embryonic lethal phenotype reveals a function of TDG in maintaining epigenetic stability. *Nature* **470** 419–423. (doi:10.1038/nature09672)

Cortellino S, Xu J, Sannai M, Moore R, Caretti E, Cigliano A, Le Coz M, Devarajan K, Wessels A, Soprano D *et al.* 2011 Thymine DNA glycosylase is essential for active DNA demethylation by linked deamination-base excision repair. *Cell* **146** 67–79. (doi:10.1016/j.cell.2011.06.020)

D'Alessio AC, Weaver IC & Szyf M 2007 Acetylation-induced transcription is required for active DNA demethylation in methylation-silenced genes. *Molecular and Cellular Biology* **27** 7462–7474. (doi:10.1128/MCB.01120-07)

Dalhus B, Laerdahl JK, Backe PH & Bjoras M 2009 DNA base repair – recognition and initiation of catalysis. *FEMS Microbiology Reviews* **33** 1044–1078. (doi:10.1111/j.1574-6976.2009.00188.x)

Dean W, Santos F, Stojkovic M, Zakhartchenko V, Walter J, Wolf E & Reik W 2001 Conservation of methylation reprogramming in mammalian development: aberrant reprogramming in cloned embryos. *PNAS* **98** 13734–13738. (doi:10.1073/pnas.241522698)

Deans B, Griffin CS, Maconochie M & Thacker J 2000 Xrcc2 is required for genetic stability, embryonic neurogenesis and viability in mice. *EMBO Journal* **19** 6675–6685. (doi:10.1093/emboj/19.24.6675)

Detich N, Theberge J & Szyf M 2002 Promoter-specific activation and demethylation by MBD2/demethylase. *Journal of Biological Chemistry* **277** 35791–35794. (doi:10.1074/jbc.C200408200)

Dianova II, Sleeth KM, Allinson SL, Parsons JL, Breslin C, Caldecott KW & Dianov GL 2004 XRCC1-DNA polymerase beta interaction is required for efficient base excision repair. *Nucleic Acids Research* **32** 2550–2555. (doi:10.1093/nar/gkh567)

Doerfler W 2008 In pursuit of the first recognized epigenetic signal – DNA methylation: a 1976 to 2008 synopsis. *Epigenetics* **3** 125–133. (doi:10.4161/epi.3.3.6249)

Engel N, Tront JS, Erinle T, Nguyen N, Latham KE, Sapienza C, Hoffman B & Liebermann DA 2009 Conserved DNA methylation in Gadd45a(−/−) mice. *Epigenetics* **4** 98–99. (doi:10.4161/epi.4.2.7858)

Feng S, Jacobsen SE & Reik W 2010 Epigenetic reprogramming in plant and animal development. *Science* **330** 622–627. (doi:10.1126/science.1190614)

Ficz G, Branco MR, Seisenberger S, Santos F, Krueger F, Hore TA, Marques CJ, Andrews S & Reik W 2011 Dynamic regulation of 5-hydroxymethylcytosine in mouse ES cells and during differentiation. *Nature* **473** 398–402. (doi:10.1038/nature10008)

Fitzgerald ME & Drohat AC 2008 Coordinating the initial steps of base excision repair. Apurinic/apyrimidinic endonuclease 1 actively stimulates thymine DNA glycosylase by disrupting the product complex. *Journal of Biological Chemistry* **283** 32680–32690. (doi:10.1074/jbc.M805504200)

Gerber AP & Keller W 1999 An adenosine deaminase that generates inosine at the wobble position of tRNAs. *Science* **286** 1146–1149. (doi:10.1126/science.286.5442.1146)

Glass JL, Thompson RF, Khulan B, Figueroa ME, Olivier EN, Oakley EJ, Van Zant G, Bouhassira EE, Melnick A, Golden A *et al.* 2007 CG dinucleotide clustering is a species-specific property of the genome. *Nucleic Acids Research* **35** 6798–6807. (doi:10.1093/nar/gkm489)

Gu H, Marth JD, Orban PC, Mossmann H & Rajewsky K 1994 Deletion of a DNA polymerase beta gene segment in T cells using cell type-specific gene targeting. *Science* **265** 103–106. (doi:10.1126/science.8016642)

Guo JU, Su Y, Zhong C, Ming GL & Song H 2011 Hydroxylation of 5-methylcytosine by TET1 promotes active DNA demethylation in the adult brain. *Cell* **145** 423–434. (doi:10.1016/j.cell.2011.03.022)

Hajkova P, Erhardt S, Lane N, Haaf T, El-Maarri O, Reik W, Walter J & Surani MA 2002 Epigenetic reprogramming in mouse primordial germ cells. *Mechanisms of Development* **117** 15–23. (doi:10.1016/S0925-4773(02)00181-8)

Hajkova P, Ancelin K, Waldmann T, Lacoste N, Lange UC, Cesari F, Lee C, Almouzni G, Schneider R & Surani MA 2008 Chromatin dynamics during epigenetic reprogramming in the mouse germ line. *Nature* **452** 877–881. (doi:10.1038/nature06714)

Hajkova P, Jeffries SJ, Lee C, Miller N, Jackson SP & Surani MA 2010 Genome-wide reprogramming in the mouse germ line entails the base excision repair pathway. *Science* **329** 78–82. (doi:10.1126/science.1187945)

Hendrich B, Hardeland U, Ng HH, Jiricny J & Bird A 1999 The thymine glycosylase MBD4 can bind to the product of deamination at methylated CpG sites. *Nature* **401** 301–304. (doi:10.1038/45843)

Hermann A, Gowher H & Jeltsch A 2004 Biochemistry and biology of mammalian DNA methyltransferases. *Cellular and Molecular Life Sciences* **61** 2571–2587. (doi:10.1007/s00018-004-4201-1)

Hu XV, Rodrigues TM, Tao H, Baker RK, Miraglia L, Orth AP, Lyons GE, Schultz PG & Wu X 2010 Identification of RING finger protein 4 (RNF4) as a modulator of DNA demethylation through a functional genomics screen. *PNAS* **107** 15087–15092. (doi:10.1073/pnas.1009025107)

Iqbal K, Jin SG, Pfeifer GP & Szabo PE 2011 Reprogramming of the paternal genome upon fertilization involves genome-wide oxidation of 5-methylcytosine. *PNAS* **108** 3642–3647. (doi:10.1073/pnas.1014033108)

Ito S, D'Alessio AC, Taranova OV, Hong K, Sowers LC & Zhang Y 2010 Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. *Nature* **466** 1129–1133. (doi:10.1038/nature09303)

Jenuwein T & Allis CD 2001 Translating the histone code. *Science* **293** 1074–1080. (doi:10.1126/science.1063127)

Jin SG, Guo C & Pfeifer GP 2008 GADD45A does not promote DNA demethylation. *PLoS Genetics* **4** e1000013. (doi:10.1371/journal.pgen.1000013)

Jiricny J & Menigatti M 2008 DNA cytosine demethylation: are we getting close? *Cell* **135** 1167–1169. (doi:10.1016/j.cell.2008.12.008)

Kang YK, Koo DB, Park JS, Choi YH, Chung AS, Lee KK & Han YM 2001 Aberrant methylation of donor genome in cloned bovine embryos. *Nature Genetics* **28** 173–177. (doi:10.1038/88903)

Kim MS, Kondo T, Takada I, Youn MY, Yamamoto Y, Takahashi S, Matsumoto T, Fujiyama S, Shirode Y, Yamaoka I *et al.* 2009 DNA demethylation in hormone-induced transcriptional derepression. *Nature* **461** 1007–1012. (doi:10.1038/nature08456)

Ko M, Huang Y, Jankowska AM, Pape UJ, Tahiliani M, Bandukwala HS, An J, Lamperti ED, Koh KP, Ganetzky R *et al.* 2010 Impaired hydroxylation of 5-methylcytosine in myeloid cancers with mutant TET2. *Nature* **468** 839–843. (doi:10.1038/nature09586)

Koh KP, Yabuuchi A, Rao S, Huang Y, Cunniff K, Nardone J, Laiho A, Tahiliani M, Sommer CA, Mostoslavsky G *et al.* 2011 Tet1 and Tet2 regulate 5-hydroxymethylcytosine production and cell lineage specification in mouse embryonic stem cells. *Cell Stem Cell* **8** 200–213. (doi:10.1016/j.stem.2011.01.008)

Kroft TL, Jethanandani P, McLean DJ & Goldberg E 2001 Methylation of CpG dinucleotides alters binding and silences testis-specific transcription directed by the mouse lactate dehydrogenase C promoter. *Biology of Reproduction* **65** 1522–1527. (doi:10.1095/biolreprod65.5.1522)

Kunz C, Saito Y & Schar P 2009 DNA repair in mammalian cells: mismatched repair: variations on a theme. *Cellular and Molecular Life Sciences* **66** 1021–1038. (doi:10.1007/s00018-009-8739-9)

Laprise SL 2009 Implications of epigenetics and genomic imprinting in assisted reproductive technologies. *Molecular Reproduction and Development* **76** 1006–1018. (doi:10.1002/mrd.21058)

Larijani M & Martin A 2007 Single-stranded DNA structure and positional context of the target cytidine determine the enzymatic efficiency of AID. *Molecular and Cellular Biology* **27** 8038–8048. (doi:10.1128/MCB.01046-07)

Larijani M, Petrov AP, Kolenchenko O, Berru M, Krylov SN & Martin A 2007 AID associates with single-stranded DNA with high affinity and a long complex half-life in a sequence-independent manner. *Molecular and Cellular Biology* **27** 20–30. (doi:10.1128/MCB.00824-06)

Leibovitch SA & Harel J 1978 Active DNA transcription sites released from the genome of normal embryonic chicken cells. *Nucleic Acids Research* **5** 777–787. (doi:10.1093/nar/5.3.777)

Leibovitch SA, Leibovitch MP, Kruh J & Harel J 1979 Relationship between single-stranded DNA isolated from cultured muscular cells during differentiation and the transcription of messenger RNA. *European Journal of Biochemistry* **97** 327–333. (doi:10.1111/j.1432-1033.1979.tb13118.x)

Liao W, Hong SH, Chan BH, Rudolph FB, Clark SC & Chan L 1999 APOBEC-2, a cardiac- and skeletal muscle-specific member of the cytidine deaminase supergene family. *Biochemical and Biophysical Research Communications* **2** 398–404. (doi:10.1006/bbrc.1999.0925)

Lister R, Pelizzola M, Dowen RH, Hawkins RD, Hon G, Tonti-Filippini J, Nery JR, Lee L, Ye Z, Ngo QM *et al.* 2009 Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature* **462** 315–322. (doi:10.1038/nature08514)

Lister R, Pelizzola M, Kida YS, Hawkins RD, Nery JR, Hon G, Antosiewicz-Bourget J, O'Malley R, Castanon R, Klugman S *et al.* 2011 Hotspots of aberrant epigenomic reprogramming in human induced pluripotent stem cells. *Nature* **471** 68–73. (doi:10.1038/nature09798)

Liu H, Kim JM & Aoki F 2004 Regulation of histone H3 lysine 9 methylation in oocytes and early pre-implantation embryos. *Development* **131** 2269–2280. (doi:10.1242/dev.01116)

Maatouk DM, Kellam LD, Mann MR, Lei H, Li E, Bartolomei MS & Resnick JL 2006 DNA methylation is a primary mechanism for silencing postmigratory primordial germ cell genes in both germ cell and somatic cell lineages. *Development* **133** 3411–3418. (doi:10.1242/dev.02500)

Maher ER 2005 Imprinting and assisted reproductive technology. *Human Molecular Genetics* **14** R133–R138. (doi:10.1093/hmg/ddi107)

Maiti A, Morgan MT, Pozharski E & Drohat AC 2008 Crystal structure of human thymine DNA glycosylase bound to DNA elucidates sequence-specific mismatch recognition. *PNAS* **105** 8890–8895. (doi:10.1073/pnas.0711061105)

May A, Kirchner R, Muller H, Hartmann P, El Hajj N, Tresch A, Zechner U, Mann W & Haaf T 2009 Multiplex rt-PCR expression analysis of developmentally important genes in individual mouse preimplantation embryos and blastomeres. *Biology of Reproduction* **80** 194–202. (doi:10.1095/biolreprod.107.064691)

Mayer W, Niveleau A, Walter J, Fundele R & Haaf T 2000 Demethylation of the zygotic paternal genome. *Nature* **403** 501–502. (doi:10.1038/35000656)

Metivier R, Gallais R, Tiffoche C, Le Peron C, Jurkowska RZ, Carmouche RP, Ibberson D, Barath P, Demay F, Reid G *et al.* 2008 Cyclical DNA methylation of a transcriptionally active promoter. *Nature* **452** 45–50. (doi:10.1038/nature06544)

Mikl MC, Watt IN, Lu M, Reik W, Davies SL, Neuberger MS & Rada C 2005 Mice deficient in APOBEC2 and APOBEC3. *Molecular and Cellular Biology* **16** 7270–7277. (doi:10.1128/MCB.25.16.7270-7277.2005)

Millar CB, Guy J, Sansom OJ, Selfridge J, MacDougall E, Hendrich B, Keightley PD, Bishop SM, Clarke AR & Bird A 2002 Enhanced CpG mutability and tumorigenesis in MBD4-deficient mice. *Science* **297** 403–405. (doi:10.1126/science.1073354)

Morgan HD, Dean W, Coker HA, Reik W & Petersen-Mahrt SK 2004 Activation-induced cytidine deaminase deaminates 5-methylcytosine in DNA and is expressed in pluripotent tissues: implications for epigenetic reprogramming. *Journal of Biological Chemistry* **279** 52353–52360. (doi:10.1074/jbc.M407695200)

Muramatsu M, Sankaranand VS, Anant S, Sugai M, Kinoshita K, Davidson NO & Honjo T 1999 Specific expression of activation-induced cytidine deaminase (AID), a novel member of the RNA-editing deaminase family in germinal center B cells. *Journal of Biological Chemistry* **274** 18470–18476. (doi:10.1074/jbc.274.26.18470)

Muramatsu M, Kinoshita K, Fagarasan S, Yamada S, Shinkai Y & Honjo T 2000 Class switch recombination and hypermutation require activation-induced cytidine deaminase (AID), a potential RNA editing enzyme. *Cell* **102** 553–563. (doi:10.1016/S0092-8674(00)00078-7)

Murata K, Kouzarides T, Bannister AJ & Gurdon JB 2010 Histone H3 lysine 4 methylation is associated with the transcriptional reprogramming efficiency of somatic nuclei by oocytes. *Epigenetics & Chromatin* **3** 4. (doi:10.1186/1756-8935-3-4)

Navaratnam N, Morrison JR, Bhattacharya S, Patel D, Funahashi T, Giannoni F, Teng BB, Davidson NO & Scott J 1993 The p27 catalytic subunit of the apolipoprotein B mRNA editing enzyme is a cytidine deaminase. *Journal of Biological Chemistry* **28** 20709–20712.

Neddermann P & Jiricny J 1994 Efficient removal of uracil from G.U mispairs by the mismatch-specific thymine DNA glycosylase from HeLa cells. *PNAS* **91** 1642–1646. (doi:10.1073/pnas.91.5.1642)

Niehrs C 2009 Active DNA demethylation and DNA repair. *Differentiation* **77** 1–11. (doi:10.1016/j.diff.2008.09.004)

Nishigaki M, Aoyagi K, Danjoh I, Fukaya M, Yanagihara K, Sakamoto H, Yoshida T & Sasaki H 2005 Discovery of aberrant expression of R-RAS by cancer-linked DNA hypomethylation in gastric cancer using microarrays. *Cancer Research* **65** 2115–2124. (doi:10.1158/0008-5472.CAN-04-3340)

Okada Y, Yamagata K, Hong K, Wakayama T & Zhang Y 2010 A role for the elongator complex in zygotic paternal genome demethylation. *Nature* **463** 554–558. (doi:10.1038/nature08732)

Olsen LC, Aasland R, Wittwer CU, Krokan HE & Helland DE 1989 Molecular cloning of human uracil-DNA glycosylase, a highly conserved DNA repair enzyme. *EMBO Journal* **8** 3121–3125.

Oswald J, Engemann S, Lane N, Mayer W, Olek A, Fundele R, Dean W, Reik W & Walter J 2000 Active demethylation of the paternal genome in the mouse zygote. *Current Biology* **10** 475–478. (doi:10.1016/S0960-9822(00)00448-6)

Otero G, Fellows J, Li Y, de Bizemont T, Dirac AM, Gustafsson CM, Erdjument-Bromage H, Tempst P & Svejstrup JQ 1999 Elongator, a multisubunit component of a novel RNA polymerase II holoenzyme for transcriptional elongation. *Molecular Cell* **3** 109–118. (doi:10.1016/S1097-2765(00)80179-3)

Papageorgis P, Lambert AW, Ozturk S, Gao F, Pan H, Manne U, Alekseyev YO, Thiagalingam A, Abdolmaleky HM, Lenburg M *et al.* 2010 Smad signaling is required to maintain epigenetic silencing during breast cancer progression. *Cancer Research* **70** 968–978. (doi:10.1158/0008-5472.CAN-09-1872)

Parsons JL, Dianova II & Dianov GL 2005 APE1-dependent repair of DNA single-strand breaks containing 3′-end 8-oxoguanine. *Nucleic Acids Research* **33** 2204–2209. (doi:10.1093/nar/gki518)

Patenaude AM, Orthwein A, Hu Y, Campo VA, Kavli B, Buschiazzo A & Di Noia JM 2009 Active nuclear import and cytoplasmic retention of activation-induced deaminase. *Nature Structural & Molecular Biology* **16** 517–527. (doi:10.1038/nsmb.1598)

Pavri R, Gazumyan A, Jankovic M, Di Virgilio M, Klein I, Ansarah-Sobrinho C, Resch W, Yamane A, Reina San-Martin B,

Barreto V *et al.* 2010 Activation-induced cytidine deaminase targets DNA at sites of RNA polymerase II stalling by interaction with Spt5. *Cell* **143** 122–133. (doi:10.1016/j.cell.2010.09.017)

Pham P, Bransteitter R, Petruska J & Goodman MF 2003 Processive AID-catalysed cytosine deamination on single-stranded DNA simulates somatic hypermutation. *Nature* **424** 103–107. (doi:10.1038/nature01760)

Popp C, Dean W, Feng S, Cokus SJ, Andrews S, Pellegrini M, Jacobsen SE & Reik W 2010 Genome-wide erasure of DNA methylation in mouse primordial germ cells is affected by AID deficiency. *Nature* **463** 1101–1105. (doi:10.1038/nature08829)

Rai K, Huggins IJ, James SR, Karpf AR, Jones DA & Cairns BR 2008 DNA demethylation in zebrafish involves the coupling of a deaminase, a glycosylase, and gadd45. *Cell* **135** 1201–1212. (doi:10.1016/j.cell.2008.11.042)

Rai K, Sarkar S, Broadbent TJ, Voas M, Grossmann KF, Nadauld LD, Dehghanizadeh S, Hagos FT, Li Y, Toth RK *et al.* 2010 DNA demethylase activity maintains intestinal cells in an undifferentiated state following loss of APC. *Cell* **142** 930–942. (doi:10.1016/j.cell.2010.08.030)

Ruddock-D'Cruz NT, Xue J, Wilson KJ, Heffernan C, Prashadkumar S, Cooney MA, Sanchez-Partida LG, French AJ & Holland MK 2008 Dynamic changes in the localization of five members of the methyl binding domain (MBD) gene family during murine and bovine preimplantation embryo development. *Molecular Reproduction and Development* **75** 48–59. (doi:10.1002/mrd.20712)

Santos F, Hendrich B, Reik W & Dean W 2002 Dynamic reprogramming of DNA methylation in the early mouse embryo. *Developmental Biology* **241** 172–182. (doi:10.1006/dbio.2001.0501)

Santos F, Peters AH, Otte AP, Reik W & Dean W 2005 Dynamic chromatin modifications characterise the first cell cycle in mouse embryos. *Developmental Biology* **280** 225–236. (doi:10.1016/j.ydbio.2005.01.025)

Sato S, Yoshimizu T, Sato E & Matsui Y 2003 Erasure of methylation imprinting of Igf2r during mouse primordial germ-cell development. *Molecular Reproduction and Development* **65** 41–50. (doi:10.1002/mrd.10264)

Sato Y, Probst HC, Tatsumi R, Ikeuchi Y, Neuberger MS & Rada C 2010 Deficiency in APOBEC2 leads to a shift in muscle fiber-type, diminished body mass and myopathy. *Journal of Biological Chemistry* **285** 7111–7118. (doi:10.1074/jbc.M109.052977)

Schafer A, Schomacher L, Barreto G, Doderlein G & Niehrs C 2010 Gemcitabine functions epigenetically by inhibiting repair mediated DNA demethylation. *PLoS ONE* **5** e14060. (doi:10.1371/journal.pone.0014060)

Schmitz KM, Schmitt N, Hoffmann-Rohrer U, Schafer A, Grummt I & Mayer C 2009 TAF12 recruits Gadd45a and the nucleotide excision repair complex to the promoter of rRNA genes leading to active DNA demethylation. *Molecular Cell* **33** 344–353. (doi:10.1016/j.molcel.2009.01.015)

Schmitz KM, Mayer C, Postepska A & Grummt I 2010 Interaction of noncoding RNA with the rDNA promoter mediates recruitment of DNMT3b and silencing of rRNA genes. *Genes and Development* **24** 2264–2269. (doi:10.1101/gad.590910)

Shen HM, Poirier MG, Allen MJ, North J, Lal R, Widom J & Storb U 2009 The activation-induced cytidine deaminase (AID) efficiently targets DNA in nucleosomes but only during transcription. *Journal of Experimental Medicine* **206** 1057–1071. (doi:10.1084/jem.20082678)

Simonsson S & Gurdon J 2004 DNA demethylation is necessary for the epigenetic reprogramming of somatic cell nuclei. *Nature Cell Biology* **6** 984–990. (doi:10.1038/ncb1176)

Song F, Mahmood S, Ghosh S, Liang P, Smiraglia DJ, Nagase H & Held WA 2009 Tissue specific differentially methylated regions (TDMR): changes in DNA methylation during development. *Genomics* **93** 130–139. (doi:10.1016/j.ygeno.2008.09.003)

Stengel S, Fiebig U, Kurth R & Denner J 2010 Regulation of human endogenous retrovirus-K expression in melanomas by CpG methylation. *Genes, Chromosomes & Cancer* **49** 401–411. (doi:10.1002/gcc.20751)

Sugo N, Aratani Y, Nagashima Y, Kubota Y & Koyama H 2000 Neonatal lethality with abnormal neurogenesis in mice deficient in DNA polymerase beta. *EMBO Journal* **19** 1397–1404. (doi:10.1093/emboj/19.6.1397)

Surani MA & Hajkova P 2010 Epigenetic reprogramming of mouse germ cells toward totipotency. *Cold Spring Harbor Symposia on Quantitative Biology* **75** 211–218. (doi:10.1101/sqb.2010.75.010)

Tahiliani M, Koh KP, Shen Y, Pastor WA, Bandukwala H, Brudno Y, Agarwal S, Iyer LM, Liu DR, Aravind L *et al.* 2009 Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science* **324** 930–935. (doi:10.1126/science.1170116)

Takahashi K & Yamanaka S 2006 Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* **126** 663–676. (doi:10.1016/j.cell.2006.07.024)

Talpaert-Borle M, Campagnari F & Creissen DM 1982 Properties of purified uracil-DNA glycosylase from calf thymus. An *in vitro* study using synthetic DNA-like substrates. *Journal of Biological Chemistry* **257** 1208–1214.

Tebbs RS, Flannery ML, Meneses JJ, Hartmann A, Tucker JD, Thompson LH, Cleaver JE & Pedersen RA 1999 Requirement for the Xrcc1 DNA base excision repair gene during early mouse development. *Developmental Biology* **208** 513–529. (doi:10.1006/dbio.1999.9232)

Teng B, Burant CF & Davidson NO 1993 Molecular cloning of an apolipoprotein B messenger RNA editing protein. *Science* **260** 1816–1819. (doi:10.1126/science.8511591)

Turek-Plewa J & Jagodzinski PP 2005 The role of mammalian DNA methyltransferases in the regulation of gene expression. *Cellular & Molecular Biology Letters* **10** 631–647.

Vidal AE, Boiteux S, Hickson ID & Radicella JP 2001 XRCC1 coordinates the initial and late stages of DNA abasic site repair through protein–protein interactions. *EMBO Journal* **20** 6530–6539. (doi:10.1093/emboj/20.22.6530)

Vonica A, Rosa A, Arduini BL & Brivanlou AH 2011 APOBEC2, a selective inhibitor of TGFbeta signaling, regulates left–right axis specification during early embryogenesis. *Developmental Biology* **350** 13–23. (doi:10.1016/j.ydbio.2010.09.016)

Vuong BQ, Lee M, Kabir S, Irimia C, Macchiarulo S, McKnight GS & Chaudhuri J 2009 Specific recruitment of protein kinase A to the immunoglobulin locus regulates class-switch recombination. *Nature Immunology* **10** 420–426. (doi:10.1038/ni.1708)

Wang L, Wuerffel R, Feldman S, Khamlichi AA & Kenter AL 2009 S region sequence, RNA polymerase II, and histone modifications create chromatin accessibility during class switch recombination. *Journal of Experimental Medicine* **206** 1817–1830. (doi:10.1084/jem.20081678)

Williams K, Christensen J, Pedersen MT, Johansen JV, Cloos PA, Rappsilber J & Helin K 2011 TET1 and hydroxymethylcytosine in transcription and DNA methylation fidelity. *Nature* **473** 343–348. (doi:10.1038/nature10066)

Wong E, Yang K, Kuraguchi M, Werling U, Avdievich E, Fan K, Fazzari M, Jin B, Brown AM, Lipkin M *et al.* 2002 Mbd4 inactivation increases Cright-arrowT transition mutations and promotes gastrointestinal tumor formation. *PNAS* **99** 14937–14942. (doi:10.1073/pnas.232579299)

Wossidlo M, Arand J, Sebastiano V, Lepikhov K, Boiani M, Reinhardt R, Scholer H & Walter J 2010 Dynamic link of DNA demethylation, DNA strand breaks and repair in mouse zygotes. *EMBO Journal* **29** 1877–1888. (doi:10.1038/emboj.2010.80)

Wossidlo M, Nakamura T, Lepikhov K, Marques CJ, Zakhartchenko V, Boiani M, Arand J, Nakano T, Reik W & Walter J 2011 5-Hydroxymethylcytosine in the mammalian zygote is linked with epigenetic reprogramming. *Nature Communications* **2** 241. (doi:10.1038/ncomms1240)

Wu P, Qiu C, Sohail A, Zhang X, Bhagwat AS & Cheng X 2003 Mismatch repair in methylated DNA. Structure and activity of the mismatch-specific thymine glycosylase domain of methyl-CpG-binding protein MBD4. *Journal of Biological Chemistry* **278** 5285–5291. (doi:10.1074/jbc.M210884200)

Wu H, D'Alessio AC, Ito S, Xia K, Wang Z, Cui K, Zhao K, Eve Sun Y & Zhang Y 2011 Dual functions of Tet1 in transcriptional regulation in mouse embryonic stem cells. *Nature* **473** 389–393. (doi:10.1038/nature09934)

Yoon JH, Iwai S, O'Connor TR & Pfeifer GP 2003 Human thymine DNA glycosylase (TDG) and methyl-CpG-binding protein 4 (MBD4) excise thymine glycol (Tg) from a Tg:G mispair. *Nucleic Acids Research* **31** 5399–5404. (doi:10.1093/nar/gkg730)

Yoshikawa K, Okazaki IM, Eto T, Kinoshita K, Muramatsu M, Nagaoka H & Honjo T 2002 AID enzyme-induced hypermutation in an actively transcribed gene in fibroblasts. *Science* **296** 2033–2036. (doi:10.1126/science.1071556)

Zhu JK 2009 Active DNA demethylation mediated by DNA glycosylases. *Annual Review of Genetics* **43** 143–166. (doi:10.1146/annurev-genet-102108-134205)

Zhu B, Zheng Y, Angliker H, Schwarz S, Thiry S, Siegmann M & Jost JP 2000*a* 5-Methylcytosine DNA glycosylase activity is also present in the human MBD4 (G/T mismatch glycosylase) and in a related avian sequence. *Nucleic Acids Research* **28** 4157–4165. (doi:10.1093/nar/28.21.4157)

Zhu B, Zheng Y, Hess D, Angliker H, Schwarz S, Siegmann M, Thiry S & Jost JP 2000*b* 5-Methylcytosine-DNA glycosylase activity is present in a cloned G/T mismatch DNA glycosylase associated with the chicken embryo DNA demethylation complex. *PNAS* **97** 5135–5139. (doi:10.1073/pnas.100107597)

Received 3 May 2011
First decision 27 June 2011
Revised manuscript received 31 August 2011
Accepted 12 September 2011

Appendix 2

Review article: 'Nuclear reprogramming of sperm and somatic nuclei in eggs and oocytes.'

Teperek M., Miyamoto K., Reprod Med Biol., 2013

Reprod Med Biol
DOI 10.1007/s12522-013-0155-z

REVIEW ARTICLE

Nuclear reprogramming of sperm and somatic nuclei in eggs and oocytes

Marta Teperek · Kei Miyamoto

Received: 27 March 2013 / Accepted: 18 May 2013

Abstract Eggs and oocytes have a prominent ability to reprogram sperm nuclei for ensuring embryonic development. The reprogramming activity that eggs/oocytes intrinsically have towards sperm is utilised to reprogram somatic nuclei injected into eggs/oocytes in nuclear transfer (NT) embryos. NT embryos of various species can give rise to cloned animals, demonstrating that eggs/oocytes can confer totipotency even to somatic nuclei. However, many studies indicate that reprogramming of somatic nuclei is not as efficient as that of sperm nuclei. In this review, we explain how and why sperm and somatic nuclei are differentially reprogrammed in eggs/oocytes. Recent studies have shown that sperm chromatin is epigenetically modified to be adequate for early embryonic development, while somatic nuclei do not have such modifications. Moreover, epigenetic memories encoded in sperm chromatin are transgenerationally inherited, implying unique roles of sperm. We also discuss whether somatic nuclei can be artificially modified to acquire sperm-like chromatin states in order to increase the efficiency of nuclear reprogramming.

Keywords Epigenetic modification · Fertilisation · Nuclear reprogramming · Nuclear transfer · Spermatogenesis

M. Teperek · K. Miyamoto (✉)
The Wellcome Trust/Cancer Research UK Gurdon Institute, The Henry Wellcome Building of Cancer and Developmental Biology, University of Cambridge, Tennis Court Road, CB2 1QN Cambridge, United Kingdom
e-mail: k.miyamoto@gurdon.cam.ac.uk

M. Teperek · K. Miyamoto
Department of Zoology, University of Cambridge, Cambridge, United Kingdom

Published online: 04 June 2013

Introduction

During the course of development, totipotent embryonic cells differentiate and become committed to distinct lineages. Once a cell is differentiated, it cannot be dedifferentiated in normal development. However, it was proved that it is possible to dedifferentiate a cell experimentally in a process called cloning. Successful cloning experiments have been first performed in *Xenopus laevis*: when cells from embryos/tadpoles were transferred into enucleated eggs, arrested at the metaphase of the second meiotic division (MII; the term "MII egg" is used in frogs, while "MII oocyte" is used for mammals), healthy embryos with the genetic material of the donor cells were generated, reaching adulthood and sexual maturity [1]. Decades later, Dolly the sheep was born as the first mammalian somatic cell nuclear transfer (SCNT) animal [2], followed by reports of successfully cloned individuals across many mammalian species, including the mouse [3], cow [4], goat [5], pig [6, 7], cat [8], rabbit [9], horse [10], rat [11], dog [12], ferret [13], red deer [14], and camel [15]. However, regardless of the species used, embryos obtained by SCNT develop to the adulthood inefficiently compared to embryos obtained by fertilisation. Furthermore, it is generally regarded that the more differentiated somatic cell nuclei are when used as donors in SCNT, the less efficient the process is; for example, in frogs, when blastula or gastrula nuclei were used as donors for SCNT, 36 % of resulting embryos were able to develop to feeding tadpoles, whereas when the nuclei of intestinal epithelium were used, only 1.5 % embryos reached a feeding tadpole stage [16]. Although this rule may not be fully applied to mammalian cloning [17, 18], it is still valid to conclude that embryonic cells are a better source for SCNT [19].

Why is SCNT inefficient or, in other words, why are somatic cells resistant to reprogramming? The sperm which fertilises the egg is also a highly specialised cell. However, despite being specialised, its nucleus can almost invariably support normal embryonic development. Therefore, the specialisation of the sperm that occurs during spermatogenesis prepares it to undergo efficient reprogramming by the egg after fertilisation. Namely, the sperm is developmentally 'programmed' to be 'reprogrammed' and to sustain normal embryonic development. In this review we first describe sperm 'programming': we review the important nuclear changes occurring during spermatogenesis that allow the mature sperm to acquire its unique chromatin features. Secondly, we discuss which of the unique sperm features may be responsible for the high efficiency of embryonic development after fertilisation, with a focus on protein and RNA contents, as well as on epigenetic modifications present on the chromatin of the mature sperm. Thirdly, we describe reprogramming events occurring upon NT in order to compare them to those occurring at fertilization. Thereafter, we summarise the most commonly observed abnormalities in cloned embryos, which are largely attributed to features of somatic cell chromatin. Lastly, we focus on the main differences between sperm and somatic cells that are likely related to reprogramming efficiency and discuss possible ways of improving cloning efficiency by making somatic cells more sperm-like.

Unique features of sperm chromatin

Sperm specialisation occurs during spermatogenesis as a series of precisely controlled events changing a progenitor germ cell into a spermatozoon. The length of the process, as well as the precise mechanisms controlling each step of spermatogenesis, differs amongst vertebrates, though general concepts are similar [20]. Somatic precursors of sperm cells, spermatogonia (developed from male germ cells), can undergo either a proliferative or a differentiative cell division. The former produces more spermatogonial cells, the latter results in the formation of a spermatogonial cell and a primary spermatocyte [21]. Primary spermatocytes enter meiotic division, producing secondary spermatocytes, which, upon completion of the meiotic division, form spermatids [20]. Spermatids have the number of chromosomes and the DNA content already reduced and do not undergo any other divisions. However, they resemble somatic cells morphologically, as well as at the molecular level, and in order to transform into highly specialised mature sperm they have to complete a series of substantial structural and morphological changes, called spermiogenesis [22] (Fig. 1).

Changes in the nuclear composition

One of the most striking changes occurring during spermiogenesis is the compaction of the sperm nucleus. Interestingly, it has been calculated that the volume of DNA of mouse sperm is six times smaller than the DNA in mitotic chromosomes [23]. The high condensation of sperm DNA is possibly due to the presence of protamines. Protamines are small and highly basic proteins that become incorporated into the chromatin during spermiogenesis in place of core histones that are the major component of the chromatin in spermatids [24]. The process of protamine incorporation is complex and requires many intermediate steps. Firstly, histone variants are thought to be incorporated alongside canonical histones, which are subsequently modified post-translationally and replaced by transition proteins (explained later) [25, 26]. Eventually, transition proteins are replaced by protamines, which are the major chromatin component of the mature sperm (Fig. 1). Furthermore, there are a lot of other changes occurring: transcription ceases, and a lot of proteins disappear from the maturing sperm nucleus whereas others are specifically accumulated. Below, we briefly describe our current understanding of nuclear changes occurring during spermiogenesis.

Histone variants and histone modifications

As mentioned above, incorporation of histone variants is the first major event allowing the nuclear maturation of a spermatid. There are numerous histone variants expressed specifically in testis, for example H2A.lap and H2A.Bbd as variants of histone H2A [27, 28], TH2B as a variant of H2B [29, 30], H1t and Hils1 as variants of H1 [31, 32] or H3t as a variant of histone H3 [33]. It has been shown that H2A.lap1 marks transcriptional start sites of specific transcripts expressed during spermatogenesis [28]. Interestingly, other histone variants are thought to facilitate incorporation of protamines. In vitro studies, in which nucleosomes have been assembled with testis-specific histone variants (H3t, TH2B, H2A.L2) and with somatic-type histones, indicate that such combinatorial nucleosomes are unstable [34–36]. Therefore, it is hypothesized that when these histones start to be expressed during spermatogenesis they form nucleosomes with somatic-like histones present in the spermatids. This, in turn, confers the instability of such combinatorial nucleosomes and allows the incorporation of transition proteins [37]. Furthermore, the instability of histone-containing nucleosomes can also be due to the presence of post-translational modifications that can alter the chromatin-binding properties. Such modifications can be present not only on core histones, but also on testis-specific histone variants [38–40]. One of the

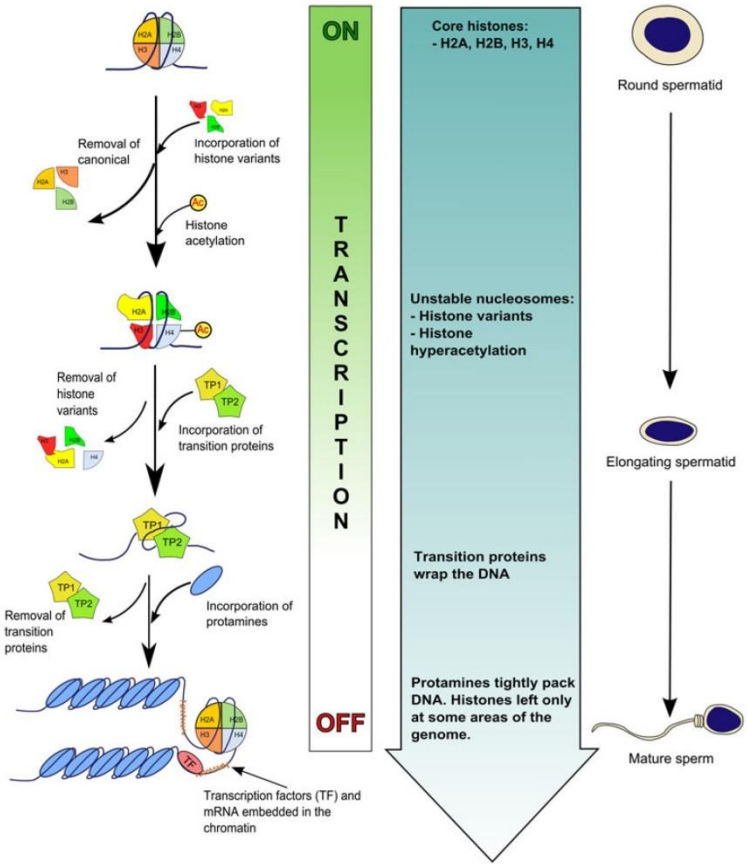


Fig. 1 Major chromatin changes during spermiogenesis. In order to achieve a sperm-like chromatin state, the round spermatid which enters the spermiogenesis process undergoes a series of chromatin remodelling events. First, canonical core histones packing the chromatin in the round spermatid can be replaced by histone variants, which together with global histone acetylation, leads to instability of the nucleosome structure. Subsequently, transition proteins are incorporated in place of unstable nucleosomes in the elongating spermatid. Finally, transition proteins are replaced with protamines. The mature sperm chromatin is mainly composed of protamines, with interspersed histones and with tightly associated mRNAs and transcription factors. All these processes are occurring in parallel with the cessation of transcription—round spermatids are transcriptionally active, whereas no transcription is detected in the mature sperm

most striking and the best described changes in histone post-translational modifications is a wave of core histone acetylation during spermiogenesis. Immunohistochemical studies revealed that acetylated forms of H2A, H2B, H3 and H4 are present specifically in elongating spermatids (no acetylation in round spermatids and no acetylation in mature sperm), suggesting that the transient presence of this mark may be implicated in exchange of histones to protamines and condensation of sperm DNA (Fig. 1) [41]. Further studies led to the identification of a protein, named Brdt (Bromodomain testis-specific), which specifically binds to acetylated lysine residues on histone H4 [42, 43]. Studies using mammalian culture cells ectopically expressing Brdt have shown that its binding to acetylated histones leads to chromatin compaction [42]. It has been also demonstrated that Brdt remodels chromatin in an ATP-independent way and that it interacts with Smarce1, which is a member of SWI/SNF remodelling complex [44]. Importantly, male mice lacking Brdt are infertile and have abnormal, misshapen sperm [45], suggesting that Brdt binding to acetylated histones may be indeed a crucial step allowing a proper condensation of sperm chromatin.

Transition proteins and protamines

Apart from histones and their modifications, other factors which are implicated in protamine deposition are transition proteins. Global histone acetylation precedes their appearance [37]. There are two transition proteins: transition protein 1 (TP1) and transition protein 2 (TP2). Their function is redundant, since mice with deletion of TP1 or TP2 alone are fertile (albeit with a reduced fertility) [46, 47]. However, knockout of both TP1 and TP2 results in infertile mice with abnormal sperm [26], suggesting that

transition proteins are important for a proper spermatogenesis. Finally, the protamines, which replace transition proteins towards the end of spermiogenesis, are also essential for male fertility. It has been demonstrated that even heterozygous male mice, which are mutant for protamine 1 or protamine 2, are infertile [48]. Detailed analyses of protamine 2 mutant mice revealed that haploinsufficiency of protamine 2 affected the DNA integrity of the mature sperm [49]. Interestingly, abnormal levels or mutations of protamines also correlate with infertility in humans [50, 51].

Cessation of transcription and disappearance of basal transcriptional machinery

Spermiogenesis is however not only a process leading to the acquisition of unique sperm-specific proteins. Numerous proteins typical for somatic cells disappear in the course of spermiogenesis. The most remarkable is the disappearance of RNA polymerase II and hence, cessation of transcription in maturing spermatids; round spermatids are actively transcribing, while there is no transcription in mature sperm (Fig. 1) [52]. The same happens to basal transcription factors: TAF1 and TLF (TBP-like factor) also disappear during spermatid condensation [53–55]. Therefore, it could be that cessation of transcription and removal of the majority of basal transcriptional machinery is important for the paternal nucleus to sustain the embryonic gene expression pattern after fertilization.

Which components of the sperm make it 'programmed' to be 'reprogrammed'?

Since there are numerous factors being acquired and lost during sperm maturation (Fig. 1), it is not easy to dissect the ones that are responsible for rendering the sperm easily reprogrammable by eggs/oocyte, as opposed to the somatic cell after nuclear transfer. We can however classify them into five categories: (1) protamines, (2) sperm-derived transcriptional regulators, (3) sperm-derived RNAs, (4) post-translational modification of histones in the mature sperm, and (5) DNA methylation profiles in the mature sperm. We will discuss below the potential involvement of each of these categories in making sperm easily reprogrammable.

Protamines

One of the biggest differences between sperm and somatic cells is the fact that DNA of a somatic cell is wrapped around histones, whereas DNA of sperm is tightly packed by protamines. Containing protamines instead of canonical histones can have at least two roles in facilitating the reprogramming at fertilisation. Firstly, after fertilisation, the protamines from the sperm are removed by an egg/oocyte with the aid of a maternal protein called nucleoplasmin [56–59]. Subsequently, oocyte-derived histones are incorporated into paternal DNA to allow the assembly of chromatin [60, 61]. On the other hand, a somatic cell does not have protamines, and hence it may be more difficult for the transplanted somatic cell to exchange its histones for the oocyte-derived ones. In addition, somatic cell histones often bear post-translational modifications associated with active gene states, according to the lineage from which the cell is originated. As a result, some genes characteristic for that lineage could continue to be inappropriately expressed in the embryo. Indeed, it has been shown in *Xenopus* that upon NT of somatic cells derived from somites (muscle precursors expressing a gene called *MyoD*), resulting embryos continued to aberrantly express *MyoD* (discussed in the later section) [62]. Therefore, it is tempting to speculate that having protamines instead of histones may be beneficial for sperm to erase the developmental program which is often encoded in histone marks: this is because at fertilisation protamines are efficiently replaced with oocyte-derived histones and because the newly incorporated histones are likely to be modified according to the embryonic developmental program. Indeed, it has been shown that embryos derived from round spermatid injection (round spermatids do not yet have protamines deposited on their chromatin) display epigenetic abnormalities as compared to embryos obtained by mature sperm injection (which do have protamines on their DNA): round spermatid-derived embryos have been shown to have elevated DNA methylation levels and histone H3K9me3 marks present on the paternal chromatin, which are not normally observed in sperm-derived embryos [63]. Apart from these potential roles of protamines in efficient histone exchange, the presence of protamines on sperm DNA and its tight packaging likely protects the DNA from any physical damage. It has been shown that in rabbits in which no offspring could be derived after the protamine-free round spermatid injection, the developmental arrest is likely due to abnormal ploidy of the resulting embryos [64]. Similarly, during the NT procedure somatic cells are exposed to numerous micromanipulations, and because they do not have protamines tightly protecting their DNA, these procedures could lead to DNA damage. In fact, it has been suggested that one of the major causes of developmental arrest of NT embryos is a result of DNA loss [65], which might be a consequence of DNA damage. With the current state of knowledge it is therefore difficult to discriminate whether the presence of protamines on the DNA helps to epigenetically program the paternal chromatin for embryonic development or whether it prevents the DNA

from physical damage (or both). In addition, it has been shown that even though the embryonic development after round spermatid injection is generally less efficient than the development after sperm injection, it can reproducibly give rise to normal offspring in several species [66–70], suggesting that the chromatin of round spermatid, even though it is not protaminated, can support full term development. However, since mice deficient with protamines are infertile [48], having protamines on the DNA is clearly beneficial for the development and hence the lack of protamines might be one of the explanations for the low efficiency of SCNT.

Sperm-derived transcriptional regulators

As mentioned before, it has been demonstrated that towards the end of spermiogenesis, transcription ceases and components of the basal transcriptional machinery disappear [55]. However, it has also been shown that the mature sperm contains transcription factors (Oct-1, Ets-1, C/EBP and TBP) associated with the hypersensitive regions of chromatin [71] (Fig. 1). It is therefore likely that such chromatin-associated factors could be delivered to the oocyte at fertilization. Furthermore, proteomic analysis of the mature sperm led to the identification of more proteins which are involved in transcriptional regulation, for example, Bromodomain-containing protein 7 (Brd7) or Polycomb protein Suz12 [72]. Upon delivery to the oocyte, these factors might help to pattern gene expression characteristic for the developing embryo. In contrast, a somatic cell during a nuclear transfer procedure may deliver transcriptional regulators responsible for the maintenance of its own differentiated cell state to the oocyte. This might interfere with normal embryonic gene expression since the embryo requires factors enabling the establishment of totipotency and not cell-type specific, differentiation factors. In fact, injection of the somatic cell cytoplasm into oocytes impairs normal development of fertilised embryos and indeed it has been demonstrated that such embryos have decreased expression levels of pluripotency factor Oct4 [73].

Sperm-derived RNAs

It has been recently discovered that mature sperm, despite having a very low amount of cytoplasm, carries numerous mRNAs [74], including mRNAs encoding transcription factors, for example cMyc [75]. These mRNAs have been shown to be delivered to the oocyte after fertilisation [76]. What is the role of these sperm-derived mRNAs? It is currently not well understood, but there are several hypotheses. The first one suggests that these mRNAs remain tightly associated with sperm chromatin and can somehow protect some regions of the sperm DNA from protamination. As a result, this would allow histone retention on these regions. Such protamine-free regions could be then selectively activated in the developing embryo (see section below) [74]. A second hypothesis is that RNA stored in the sperm can exert influence on the gene expression profile of the resulting embryo. It could be either by providing mRNA which serves as a template for synthesis of transcription factors important for embryogenesis, such as cMyc, or by inducing paramutations. Paramutation is a heritable epigenetic change resulting in a mutant phenotype in the absence of the actual genetic mutation. For example, it has been shown that wild type progeny of mice generated from heterozygous fathers harbouring a mutation in a *cKit* gene (demonstrated with a white tail phenotype) also have white tails. This paramutation was shown to be induced by antisense RNA carried over to the oocyte by the sperm at fertilisation [77]. It is also possible that sperm-derived RNA can be directly involved in regulating embryonic development. Indeed, it has been recently shown that sperm-borne microRNA-34c is required for the first cleavage division of the mouse embryo by directly regulating antiproliferative protein Bcl2 [78]. Therefore, there is evidence that carry-over sperm RNA can influence the pattern of embryonic gene expression and even phenotypes of the progeny. Such instructive information about embryonic development in a form of RNAs is likely to be absent or altered in somatic cells (presence of somatic-cell specific RNAs), therefore affecting the normality of embryonic development after NT.

Post-translational modifications of histones in the mature sperm

Another explanation for why the sperm so efficiently supports embryonic development can be due to the presence of particular post-translational modifications on histones retained in mature sperm. As explained before, towards the end of spermiogenesis, histones are replaced with protamines. However, it has been recently shown that both in mice and in humans some histones remain in the mature sperm [79, 80]. It has also been described that both in mouse and in human, sperm-derived histones (at least some of them) are transmitted to the oocyte at fertilization and retained in the paternal pronucleus [61, 81, 82]. Moreover, retention of histones in the mature sperm is nonrandom; namely, histones are retained at promoter regions of developmentally important genes. Furthermore, these retained histones bear epigenetic marks and their patterns correlate well with future embryonic gene expression: genes with an activating mark, H3K4me3, are switched on early in development, whereas genes with a

repressive mark, H3K27me3, are either repressed or activated late [79, 80]. Interestingly, this phenomenon has also been observed in species in which sperm is not protaminated (core histones are the major component of chromatin in the mature gamete). For example, in zebrafish, which does not have protamines, developmentally important promoters also have signatures of H3K4me3 and H3K27me3, which correlate with future embryonic gene expression [83, 84]. Interestingly, it has been described that infertile patients have abnormal histone retention profiles [85]. All these suggest that sperm may be able to provide the embryo with a 'patterning' message, encoded in its histone marks, as to which genes should be expressed and when during embryogenesis. Sperm of infertile patients, which do not have this message (or have incorrect messages), cannot support proper embryonic development. Such lack of the correct patterning message could also explain inefficient embryonic development after nuclear transfer. Somatic cells have their epigenetic marks on histones established according to their own developmental profile, and those epigenetic marks are likely different from those of embryos.

DNA methylation profiles in the mature sperm

Another epigenetic feature of the mature sperm that may render it suitable to support embryonic development is DNA methylation. Cytosine in DNA can be methylated, which leads to creation of 5-methylcytosine. Methylated cytosines are often associated with gene repression [86–90], whereas unmethylated cytosines usually mark transcriptionally active genes [91, 92]. Differential cytosine methylation is also observed at gene promoters in mature sperm. Interestingly, cytosine methylation states in gene promoters in sperm are very similar to those of embryonic stem cells [93] and correlate with gene expression patterns in the embryos. Genes whose promoters are unmethylated in the mature sperm are expressed early during embryogenesis, whereas methylated genes are expressed late or repressed in embryogenesis [79, 80, 84]. Moreover, there is a correlation between abnormal DNA methylation patterns in the mature sperm and infertility in humans [94]. Therefore, DNA methylation can be yet another way for the sperm to contribute a developmental message to the embryo, another way the sperm differs from somatic cells, and which may explain the low efficiency of SCNT. In fact, a recent genome-wide study clearly indicates that DNA methylation patterns of SCNT embryos resemble those of donor somatic genome rather than those of the paternal genome of fertilised embryos [95], suggesting that DNA methylation profiles are indeed different between sperm-derived and somatic cell-derived embryos.

Reprogramming of somatic nuclei in eggs and oocytes

We have so far discussed how sperm nuclei are efficiently reprogrammed in eggs/oocytes. However, cloned embryos are produced in a very different way from normal fertilised embryos and also undergo unique reprogramming processes. In this section we describe a series of reprogramming events in cloned embryos. To produce a cloned embryo, nuclear transplantation of a somatic nucleus to an enucleated egg/oocyte is performed (Fig. 2). Eggs/oocytes at the metaphase II (MII) stage are widely used as recipients for NT. Since the MII-arrested oocytes have a strong maturation promoting factor (MPF) activity, the transplanted somatic nucleus undergoes nuclear envelope breakdown and premature chromosome condensation (PCC) (Fig. 2) [96, 97]. It seems that exposure of somatic chromatin to the MII egg/oocyte cytoplasm is an important determinant for successful reprogramming by increasing the number of origins of DNA replication in order to facilitate robust DNA replication cycles in embryos [98, 99]. It has also been postulated that PCC allows oocyte reprogramming factors to gain access to somatic chromatin since the barrier of nuclear envelopes disappears. Epigenetic modifications on histone tails, such as deacetylation and phosphorylation [100, 101], are also induced soon after NT (within 2 h) (Fig. 2). Moreover, some core and linker histones are rapidly removed from somatic chromatin or exchanged to oocyte counterparts (Fig. 2). For example, macroH2A, associated with repressive chromatin, is readily removed after NT (Fig. 2) [102]. Subsequently, MII-arrested NT oocytes are activated and start early embryonic development. NT embryos then form pseudo-pronuclei, where somatic chromatin undergoes extensive decondensation (Fig. 2) [97]. In mouse, heterochromatin reorganization is induced in the pseudo-pronuclei of NT embryos [103]. Additionally, active DNA demethylation of the somatic genome is observed during the one-cell stage (Fig. 2) [95, 104]. Histone modifications of the pseudo-pronuclei are significantly different from those of original somatic chromatin and more resemble those of fertilised pronuclei, although quantitative differences of such histone modifications have been reported [100, 105–107] (see below). Towards the end of the first cell cycle in mouse, transcriptional activators are gradually accumulated in pronuclei (Fig. 2) [108] to allow major embryonic gene activation at the 2 cell stage (Fig. 2). Subsequently, genes required for pluripotency such as *Oct4* and *Nanog* are activated [109]. Once NT embryos have reached the blastocyst stage, embryonic stem cells can be derived. It is possible that reprogramming is continuing even at the blastocyst stage because abnormal histone modifications prominent at the early embryonic stages are corrected in the blastocyst stage embryos [107].

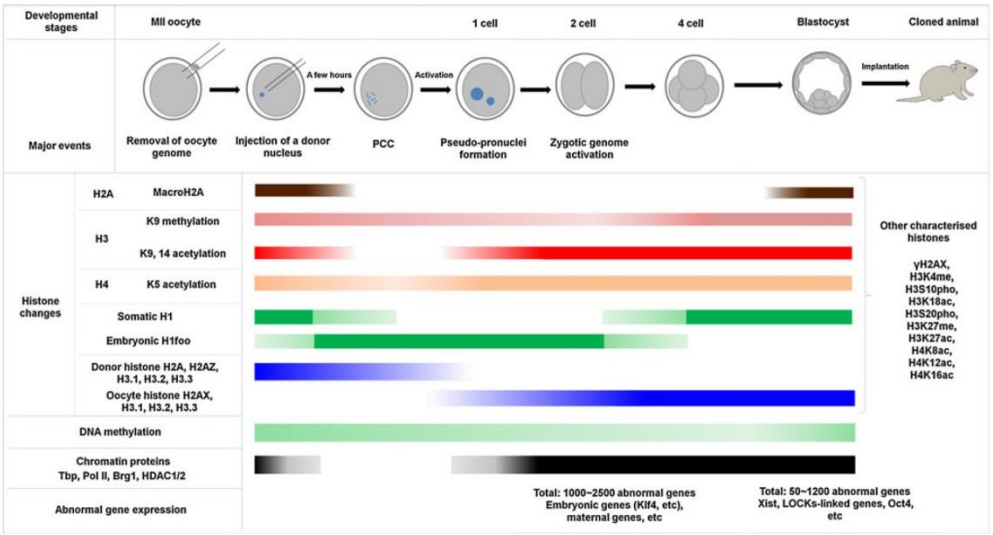


Fig. 2 Nuclear reprogramming in somatic cell nuclear transfer embryos. A series of reprogramming events occurring in NT embryos are depicted (mainly focusing on mouse NT embryos). A somatic nucleus injected into an enucleated oocyte at the metaphase II stage undergoes premature chromosome condensation (PCC). After activation, pseudo-pronuclei are formed. Transcription of embryonic genes starts from the 2-cell stage. Cloned embryos can be implanted to foster mothers and in most cases less than a few percentage of embryos can develop to term. Histone modifications, histone variants, DNA methylation, and chromatin proteins show dynamic changes during preimplantation development. Many of them are important for development of cloned embryos, but are abnormally regulated compared to fertilised counterparts. Abnormal gene expression in cloned embryos has also been reported. The numbers of abnormally expressed genes are based on transcriptome studies of 2-cell embryos [151, 152] and those for blastocyst embryos [150, 151, 156]

There are some alternative methods for NT. Oocytes at the telophase stage and enucleated embryos at the one-cell stage can also be used as recipients for NT as long as pronuclear factors are not removed during enucleation [110–112]. It is not clear whether reprogramming events induced in these NTs are the same as the ones in NT to metaphase II oocytes. Apart from these NTs, which can support full-term development, somatic nuclei injected into the giant nucleus of *Xenopus* oocytes at the meiotic prophase I stage also known as the germinal vesicle (GV) stage can be reprogrammed to express embryonic genes [113]. This type of NT is adequate for analysing mechanisms of transcriptional reprogramming [114, 115] since it does not require cell divisions and new protein synthesis. However, this type of NT does not support full term development.

There are a few maternal factors known to be important for development of NT embryos. For example, nucleoli in mouse and porcine NT embryos are derived from the nucleolus of the recipient oocyte [116]. When maternal nucleoli are removed, NT embryos are deprived of nucleoli and fail to develop. These results suggest that maternally inherited nucleoli are necessary for development. Tpt1 protein is an example of another maternal factor important for development of bovine NT embryos with the mechanisms still unknown [117]. Additionally, a screen for maternal factors required for reprogramming has identified DJ-1 as a maternal protein required for development of porcine NT embryos [118]. DJ-1 inhibits P53 activation in NT embryos; this also suggests that P53 is prone to be activated in NT embryos. Furthermore, it has been shown that maternal Tet3-mediated hydroxymethylation is induced in cloned mouse embryos [119] and is responsible for DNA demethylation on the Oct4 promoter [120]. More studies to identify maternal factors important for development of cloned embryos are needed for understanding the basis of egg/oocyte-mediated reprogramming.

Abnormalities in cloned embryos in relation to somatic cell characteristics

A number of reprogramming events, discussed in the previous section, are induced in NT embryos. NT embryos

that have successfully gone through these events can achieve full-term development. However, it is often the case that NT embryos fail to complete at least a part of reprogramming events. In this section, we discuss what kinds of errors are prone to occur in NT embryos during reprogramming. We then relate the defective reprogramming events in NT embryos to specific characteristics of somatic chromatin.

DNA demethylation and histone modification

In mouse, upon fertilisation, sperm DNA undergoes active and global demethylation during the one-cell stage. Low methylation levels are maintained until the blastocyst stage when global remethylation starts [121]. In cloned embryos, the extent of DNA demethylation in the transplanted somatic nuclei is less than that in paternal pronuclei [104]. Such abnormally high DNA methylation states in cloned embryos have been found across many mammalian species [104, 122, 123]. Bisulfite sequencing analyses have revealed that high DNA methylation states remain at the various genomic regions including the centric and pericentric repeats [124] and the *Oct4* regulatory regions [125, 126]. A recent genome-wide DNA methylation study has identified more than 20 genes whose DNA methylation remains abnormally high after NT compared to fertilised embryos [95]. This study also shows that some repetitive elements, such as long interspersed elements (LINEs) and long terminal repeats (LTRs), are resistant to DNA demethylation in cloned embryos [95]. Since these genomic regions are efficiently demethylated in paternal pronuclei, the refractory nature towards DNA demethylation seems largely attributed to somatic genomes.

Histone modifications of somatic cell chromatin are dramatically changed after NT to oocytes [100, 101]. Although histone tails of somatic chromatin are reprogrammed to resemble those of embryonic chromatin, a number of abnormal histone modifications have been reported in NT embryos [100, 105–107, 127]. Notably, abnormally high levels of histone H3K9 methylation are retained in NT embryos [100, 105, 107] and the high H3K9 methylation is likely one of the reasons why cloned embryos exhibit a low developmental capacity. Indeed, removal of H3K9 methylation prior to NT improves cloning efficiencies [128, 129] although causal relationship between loss of H3K9 methylation and reprogramming in NT embryos has to be further tested. H3K9 methylation also restricts reprogramming in iPS cells [130, 131], supporting the idea that somatic chromatin acquires H3K9 methylation to stabilize differentiation states and hence to resist nuclear reprogramming. A more targeted way of removing histone marks before NT, such as locus-specific H3K9 methylation removal, may further boost reprogramming efficiency. It has also been shown that modulating histone acetylation levels by histone deacetylase inhibitors can greatly improve cloning efficiencies [17, 132–136]. For example, histone H4K5 in NT embryos shows less acetylation than in fertilised embryos [127] and this abnormal acetylation state can be partially corrected by Trichostatin A treatment [137]. Furthermore, histone deacetylase inhibitors also improve many aspects of reprogramming including transcription and chromatin reorganization [100, 138, 139]. Histone modifications on H3K27 play an important role in gene regulation especially in relation with polycomb complexes. Many somatic genes are repressed through H3K27 methylation. After NT of somatic cells to oocytes, higher H3K27 methylation [140] and lower H3K27 acetylation [106] than in fertilised embryos are observed in 1- and 2-cell stage embryos. In addition, localisation of PRC2 components in cloned embryos is significantly different from fertilised embryos [140] and abnormal expression of polycomb-associated genes is observed [141]. Trichostatin A treatment was shown to increase a H3K27 acetylation level in cloned embryos to the one of fertilised embryos [106]. It would be interesting to see whether the elevated level of H3K27 acetylation is linked to the improvement of other chromatin signatures, such as H3K27 methylation and association of polycomb complexes.

In addition to the above mentioned histone modifications, linker histones and histone variants show dynamic changes upon NT of somatic cells to oocytes. The somatic type linker histone H1 is rapidly exchanged with the embryonic type linker histone B4/H1foo upon NT to oocytes [142–144]. Incorporation of histone B4 into the somatic chromatin creates accessible states of chromatin [144, 145]. A recent study indicates that histone variants H3.1, H3.2 and H3.3, as well as H2A and H2A.Z, are rapidly removed from transplanted nuclei in mouse NT embryos [146]. At the same time, oocyte-derived histone H3 variants and H2AX are incorporated into the transplanted nuclei. Interestingly, abnormally high amounts of H3.1 are incorporated into transplanted somatic nuclei compared to fertilised embryos (Fig. 2) [146]. This might be attributed to the features of donor chromatin that prefer H3.1 incorporation for heterochromatin formation or by the fact that chaperones for H3.1 can be carried over with the donor cell. In accordance with the incorporation of oocyte-derived H2AX into transplanted nuclei [146], phosphorylated H2AX (γH2AX) is found in pseudo-pronuclei of NT embryos although the number of γH2AX foci is smaller than that of fertilised embryos [147]. γH2AX in 1-cell embryos is proposed to be associated with DNA repair and DNA demethylation.

Embryonic gene activation

Transcriptionally silent NT embryos start to express embryonic genes from the time of zygotic genome activation (ZGA) onwards. Hundreds of embryonic genes that are silenced in somatic cells can be activated in NT embryos. However, many defects in properly turning on embryonic genes have been reported [148–152]. Abnormal gene expression in NT embryos is observed as early as ZGA [151, 152]. Microarray analysis of 2-cell mouse embryos has revealed that more than 2000 mRNAs are misregulated in NT embryos compared to fertilised embryos [152]. Transcription by RNA polymerase I is also disturbed in cloned embryos around this stage [139, 153]. This early abnormal gene expression can be rescued by treatment with histone deacetylase inhibitors [135, 139], suggesting that abnormal gene expression in early embryos is at least partially due to the somatic epigenome, which is not adequate to support the embryonic gene expression program. In fact, some somatic genes continue to express in NT embryos at the 2 cell stage [154]. Equally importantly, maternally stored transcripts are not properly degraded in NT embryos [152]. This is also true in rhesus-bovine interspecies NT embryos [155], implying that inappropriate degradation of maternal transcripts can cause a critical problem for subsequent embryonic development. Although NT embryos tend to display numerous abnormalities at the early developmental stages, the number of abnormally expressed genes in cloned embryos usually decreases at the blastocyst stage [151, 156]. This could be partially because only successfully reprogrammed embryos are able to reach the blastocyst stage. Nevertheless, these results support the idea that reprogramming is continued throughout early embryonic development. However, there are some embryonic genes that are resistant to reprogramming even at the blastocyst stage, such as *Oct4* [148] and *Sox2* [150]. Abnormal expression of such important pluripotency genes at this stage may severely affect subsequent development after implantation.

Recently, large-scale transcriptome analyses have identified gene loci that are often abnormally regulated in cloned blastocyst embryos [150, 157]. Genes in the "large organized chromatin K9-modifications" (LOCKs) regions are repressed in NT embryos. LOCKs are enriched with histone H3K9me2, a repressive mark [158]. Differentiated cells exhibit larger occupancy of the LOCKs in the genome than embryonic cells [158]. Since H3K9 methylation is difficult to remove in cloned embryos (see above), it is plausible that H3K9 methylation in LOCKs of somatic chromatin causes abnormal gene repression in NT embryos. It would be interesting to test functional significance of the abnormal gene repression associated with LOCKs in NT embryos. Another set of abnormally regulated genes is X chromosome-linked genes. This is due to the ectopic expression of *Xist* from the active X chromosome [157]. Suppressing excess *Xist* RNA by gene knockout [157] or siRNA-mediated knockdown [159] greatly improves the cloning efficiency. These results indicate that misregulation of a single important gene can cause detrimental effects on development of cloned embryos.

Epigenetic memory of somatic chromatin in cloned embryos

As mentioned above, somatic-like chromatin states can be transmitted to NT embryos and possibly result in inappropriate gene expression. It is especially clear that oocytes often fail to reverse silenced states of some embryonic genes in transplanted somatic chromatin because of layers of silencing mechanisms, including DNA methylation and H3K9 methylation. Apart from these types of reprogramming errors, there is evidence that active states of somatic genes are inherited in cloned embryos and the somatic type of gene expression continues in early embryos. In this section, we discuss examples of such abnormal somatic gene expression. For example, when myoblasts are used as donor cells for mouse NT, the myoblast cloned embryos start to express glucose transporter type 4 (*Glut4*), which is expressed in muscles but not in normal preimplantation embryos, at the late one-cell stage [154]. Similarly, the neuroectodermal marker gene is ectopically expressed in the endoderm cells in *Xenopus* NT embryos in which neuroectoderm nuclei are used as donors, and vice versa [160]. Interestingly, premature transcription of such somatic gene expression in NT embryos is also observed [160], suggesting that NT embryos are incapable of properly repressing some somatic genes of transplanted nuclei. Ng and Gurdon [62] provide a mechanistic insight into this epigenetic memory of somatic gene expression. Upon NT of somatic cells derived from somites (muscle precursors expressing a gene called *MyoD*), resulting embryos continue to aberrantly express *MyoD*. It has been shown that this memory of an active gene is dependent on trimethylation of lysine 4 on histone H3.3 in the promoter region of *MyoD* [62]. Another study proposes that a low level of histone H4K5 acetylation in somatic chromatin is inherited in cloned embryos and this may contribute to aberrant gene expression [127]. This idea needs to be further tested in individual genes that carry H4K5 acetylation. It is possible that more histone marks and chromatin proteins associated with active states of somatic gene expression would be discovered in future.

Other abnormalities in cloned embryos

Apart from the above mentioned chromatin defects, numerous other abnormalities have been reported in cloned embryos and animals, including defective placentas [161, 162], mitochondria heteroplasmy [17, 163], obesity [164], abnormal offspring syndrome [165], and short life spans [166]. In this section, we introduce other critical defects of cloned embryos, which are related to or derived from characteristics of somatic cells. In fertilised 1-cell mouse embryos, centromeres and pericentric heterochromatin regions of chromosomes associate with the periphery of the nucleolar precursor bodies [167, 168]. After NT of somatic nuclei into oocytes, the somatic type of heterochromatin often fails to be reprogrammed to the embryonic-like heterochromatin. Many cloned embryos show an abnormally high number of centromeres that do not associate with nucleolar precursor bodies [103]. Since centromeres are crucial for chromosome segregation, such abnormal positioning of centromeres may result in abnormal distribution of chromosomes to the daughter cells. In fact, cloned embryos often show improper distribution of chromosomes [169]. Recently, Mizutani et al. [65] demonstrated that abnormal chromosome segregation during early embryonic development is detrimental to full term development of cloned embryos. In relation to the abnormal chromosome segregation, somatic NT embryos exhibit abnormal spindle chromosome complex formation [170, 171] and spindle associating factors, such as calmodulin and NuMA, are found to be less abundant in somatic cell cloned embryos than in embryonic cell cloned embryos and fertilized embryos [170, 171]. Since defects in chromosome segregation seem to be one of the major causes of poor development of cloned animals, it would be interesting to determine to what extent abnormal heterochromatin and spindle formation affect early chromosome segregation defects.

Can we impose sperm-like states on somatic cells to increase the cloning efficiency?

As described above, there are a lot of differences between the sperm and somatic cells and each one of them, or a combination of these differences, may explain the low efficiency of cloning, as compared to efficient embryonic development after fertilisation. An ultimate goal in the reprogramming field is to increase the efficiency of animal cloning. It is reasonable to think that imposing sperm-like changes onto somatic cells will make them more prone to reprogramming. However, is such an approach feasible? In this section we discuss different possibilities of making somatic cells more like sperm.

Protamines

As mentioned above, the presence of protamines in the mature sperm is crucial for its ability to support correct embryonic development for at least a couple of reasons. Firstly, it may physically protect DNA from damage. Secondly, it may help to remove the majority of histones together with their marks and allow incorporation of oocyte-derived ones. Therefore, one could speculate that the easiest way to make a somatic cell more reprogrammable would be to wrap its DNA with protamines. However, it may be technically very challenging, because the deposition of protamines onto sperm chromatin is a complicated, multi-step process. It has been shown that impairments of single stages of this process (for example, removal of acetyl-lysine binding protein Brdt, or depletion of transition proteins) are detrimental for the sperm maturation and result in infertility [26, 45]. Therefore, in order to correctly deposit protamines on somatic cells, it is likely that one would not only need to introduce protamines into somatic cells, but also recapitulate other events occurring during spermiogenesis. Furthermore, this should be done in a sequential manner. Firstly, histone variants need to be expressed. The second step is to induce global histone acetylation, followed by histone removal and transition protein deposition. Introduction of protamines is only the last stage of the process. To summarise, if one could deposit protamines onto the chromatin of somatic cells, this might be a step forward in increasing the cloning efficiency. However, at the current state of methodology available, the correct deposition of protamines seems to be a technically challenging objective.

Carry-over proteins and RNAs

The mature sperm carries numerous proteins and RNAs which, as discussed above, could also be a part of sperm programming to support efficient embryonic development. Therefore, the next possible way to increase the efficiency of SCNT is to supply somatic cells with sperm proteins and RNAs. This would be challenging if one wants to isolate only the proteins and RNAs which are developmentally relevant, since numerous proteins and RNAs have been identified in the mature sperm and it is not well known which, if any, are important for embryonic development [72, 74]. A feasible approach is to isolate total protein and RNA from sperm and co-inject them into an oocyte together with a cell during the NT procedure. However, it cannot be excluded that the developmentally important proteins/RNAs are tightly associated with the sperm chromatin [74]. If this is the case, isolation of these factors could be technically demanding. Another approach is to directly deliver sperm into NT oocytes in which somatic

nuclei have been transplanted beforehand and subsequently remove the paternal pronucleus from the cloned embryos. This way, cloned oocytes are also activated by sperm injection in a similar way to normal fertilisation [97, 172]. However, this method did not improve cloning efficiency [172], which can be due to the fact that factors which are developmentally important may remain tightly associated with the sperm chromatin. Alternatively, enucleated early zygotes, in which sperm contents are delivered into recipient, can be used for NT [110, 111, 173].

Histone marks and DNA methylation

Another explanation for a sperm being superior to a somatic cell in its reprogramming capacity could be the unique pattern of its epigenetic modifications: histone and DNA methylation marks present on sperm chromatin can possibly pattern future embryonic gene expression. It is likely that imposing sperm-like epigenetic changes onto somatic cells would improve development of NT embryos. However, at the current state of knowledge this approach seems difficult. Firstly, so far the only genome-wide profiles in mouse and human sperm are available for H3K4me3, H3K27me3, H4K12ac, and DNA methylation [79, 80, 174]. It is known that epigenetic marks often work in combination and there are many more marks known than just these three; for example histone arginine methylation [175–177] or recently identified histone crotonylation [178]. It could be that the instructive information for embryonic development is a complex combination of various marks at various loci. If one can identify specific histone marks on a certain gene of sperm that are important for subsequent development, it would be worth trying to mimic such marks on somatic cell chromatin. Achieving gene locus-specific modification is currently difficult. However, with the development of zinc finger or TALE protein targeting approaches [179–184], it might be possible to induce sperm-like histone modifications in a gene-specific manner by targeting histone-modifying enzymes in the future.

Changing somatic cells into sperm?

The breakthrough study by Yamanaka and Takahashi [185] made it possible to derive pluripotent stem cells (iPS cells) from somatic cells by overexpression of transcription factors. iPS cells can differentiate into any type of cell in the body. Recently, Hayashi et al. [186] have succeeded in making spermatozoa from iPS and ES cells, and the produced spermatozoa supported full-term development after intracytoplasmic sperm injection. In theory it is therefore possible to obtain functional spermatozoa from somatic cells. However, this process requires multiple steps including dedifferentiation to an embryonic state, redifferentiation to primordial germ-like cells, and, ultimately, spermatogenesis. In addition, the resulting spermatozoa are haploid, carry paternal imprinting and require a maternal genome for successful development. Furthermore, the ultimate process of changing somatic cells into sperm—the spermatogenesis itself, is achieved by transferring primordial germ-like cells into testis. Therefore, this route may not be appropriate for the purpose of increasing cloning efficiency. However, this might be useful to produce a sperm deprived of a specific sperm factor important for subsequent development because gene knockdown or knockout can be achieved efficiently in pluripotent stem cells. An alternative route could be to directly change somatic cells to sperm-like or sperm progenitor-like cells by factor overexpression or extract treatment. This seems extremely challenging. Nevertheless, some progress has been made in inducing testis-specific gene expression in somatic cells treated with testis extracts. These extract-treated cells supported better development of cloned embryos than control non-treated cells [187]. It would be interesting to investigate why and how testis extract-treated cells are better reprogrammed in eggs/oocytes.

Conclusions and perspectives

Recent advances in genetic and epigenetic analytic approaches allowed the identification of some unique features of sperm chromatin that are absent in somatic cells. For example, histones are retained in sperm chromatin at promoter regions of developmentally important genes for subsequent gene expression in embryos. Such sperm chromatin modifications are likely to support embryonic development after fertilisation. Somatic chromatin does not have such "fine-tuning" for correct embryonic gene expression. Therefore, it is a likely explanation for why SCNT embryos often show abnormal reprogramming events compared to fertilised embryos. Somatic chromatin features, such as epigenetic memory, often remain in cloned embryos and can interfere with normal development. Moreover, other non-chromatin related factors, such as the presence of certain somatic-like transcription factors or presence of somatic-like spindle associating factors, can also impede the development of NT embryos. Therefore, it is remarkable that cloned embryos can sometimes develop to totally normal individuals. This is a good example illustrating that early embryonic development is characterised by both amazingly accurate programming and surprising plasticity. Nevertheless, there must be some essential requirements for the accomplishment of successful embryonic development. Identification of such roadblocks to development is a key challenge for

developmental and reproductive biology. Studying differences between the sperm and somatic cells in respect to their abilities to be reprogrammed by the eggs/oocytes would help to unravel the key requirements for successful development.

Acknowledgments We are grateful to Dr. J. B. Gurdon for useful discussions and Dr. M. J. Koziol and Mr. S. Wang for their critical reading of the manuscript. M. T. is supported by the Wellcome Trust [089613]. K. M. is a Research Fellow at Wolfson College in the University of Cambridge and is funded by the Herchel Smith Postdoctoral Fellowship and Great Britain Sasakawa Foundation.

Conflict of interest We have no conflict of interest.

References

1. Gurdon JB, Elsdale TR, Fischberg M. Sexually mature individuals of *Xenopus laevis* from the transplantation of single somatic nuclei. Nature. 1958;182:64–5.
2. Wilmut I, Schnieke AE, McWhir J, Kind AJ, Campbell KH. Viable offspring derived from fetal and adult mammalian cells. Nature. 1997;385:810–3.
3. Wakayama T, Perry AC, Zuccotti M, Johnson KR, Yanagimachi R. Full-term development of mice from enucleated oocytes injected with cumulus cell nuclei. Nature. 1998;394:369–74.
4. Kato Y, Tani T, Sotomaru Y, Kurokawa K, Kato J, Doguchi H, et al. Eight calves cloned from somatic cells of a single adult. Science. 1998;282:2095–8.
5. Baguisi A, Behboodi E, Melican DT, Pollock JS, Destrempes MM, Cammuso C, et al. Production of goats by somatic cell nuclear transfer. Nat Biotechnol. 1999;17:456–61.
6. Onishi A, Iwamoto M, Akita T, Mikawa S, Takeda K, Awata T, et al. Pig cloning by microinjection of fetal fibroblast nuclei. Science. 2000;289:1188–90.
7. Polejaeva IA, Chen SH, Vaught TD, Page RL, Mullins J, Ball S, et al. Cloned pigs produced by nuclear transfer from adult somatic cells. Nature. 2000;407:86–90.
8. Shin T, Kraemer D, Pryor J, Liu L, Rugila J, Howe L, et al. A cat cloned by nuclear transplantation. Nature. 2002;415:859.
9. Chesne P, Adenot PG, Viglietta C, Baratte M, Boulanger L, Renard JP. Cloned rabbits produced by nuclear transfer from adult somatic cells. Nat Biotechnol. 2002;20:366–9.
10. Galli C, Lagutina I, Crotti G, Colleoni S, Turini P, Ponderato N, et al. Pregnancy: a cloned horse born to its dam twin. Nature. 2003;424:635.
11. Zhou Q, Renard JP, Le Friec G, Brochard V, Beaujean N, Cherifi Y, et al. Generation of fertile cloned rats by regulating oocyte activation. Science. 2003;302:1179.
12. Lee BC, Kim MK, Jang G, Oh HJ, Yuda F, Kim HJ, et al. Dogs cloned from adult somatic cells. Nature. 2005;436:641.
13. Li Z, Sun X, Chen J, Liu X, Wisely SM, Zhou Q, et al. Cloned ferrets produced by somatic cell nuclear transfer. Dev Biol. 2006;293:439–48.
14. Berg DK, Li C, Asher G, Wells DN, Oback B. Red deer cloned from antler stem cells and their differentiated progeny. Biol Reprod. 2007;77:384–94.
15. Wani NA, Wernery U, Hassan FA, Wernery R, Skidmore JA. Production of the first cloned camel by somatic cell nuclear transfer. Biol Reprod. 2010;82:373–9.
16. Gurdon JB. The developmental capacity of nuclei taken from intestinal epithelium cells of feeding tadpoles. J Embryol Exp Morphol. 1962;10:622–40.
17. Ogura A, Inoue K, Wakayama T. Recent advancements in cloning by somatic cell nuclear transfer. Philos Trans R Soc Lond B Biol Sci. 2013;368:20110329.
18. Oback B, Wells DN. Donor cell differentiation, reprogramming, and cloning efficiency: elusive or illusive correlation? Mol Reprod Dev. 2007;74:646–54.
19. Hiiragi T, Solter D. Reprogramming is essential in nuclear transfer. Mol Reprod Dev. 2005;70:417–21.
20. Roosen-Runge EC. Comparative aspects of spermatogenesis. Biol Reprod. 1969;1(Suppl 1):24–31.
21. de Rooij DG. Proliferation and differentiation of spermatogonial stem cells. Reproduction. 2001;121:347–54.
22. Toshimori K. Biology of spermatozoa maturation: an overview with an introduction to this issue. Microsc Res Tech. 2003;61:1–6.
23. Ward WS, Coffey DS. DNA packaging and organization in mammalian spermatozoa: comparison with somatic cells. Biol Reprod. 1991;44:569–74.
24. Balhorn R. The protamine family of sperm nuclear proteins. Genome Biol. 2007;8:227.
25. Meistrich ML, Mohapatra B, Shirley CR, Zhao M. Roles of transition nuclear proteins in spermiogenesis. Chromosoma. 2003;111:483–8.
26. Zhao M, Shirley CR, Hayashi S, Marcon L, Mohapatra B, Suganuma R, et al. Transition nuclear proteins are required for normal chromatin condensation and functional sperm development. Genesis. 2004;38:200–13.
27. Ishibashi T, Li A, Eirin-Lopez JM, Zhao M, Missiaen K, Abbott DW, et al. H2A.Bbd: an X-chromosome-encoded histone involved in mammalian spermiogenesis. Nucleic Acids Res. 2010;38:1780–9.
28. Soboleva TA, Nekrasov M, Pahwa A, Williams R, Huttley GA, Tremethick DJ. A unique H2A histone variant occupies the transcriptional start site of active genes. Nat Struct Mol Biol. 2012;19:25–30.
29. Kim YJ, Hwang I, Tres LL, Kierszenbaum AL, Chae CB. Molecular cloning and differential expression of somatic and testis-specific H2B histone genes during rat spermatogenesis. Dev Biol. 1987;124:23–34.
30. Govin J, Caron C, Lestrat C, Rousseaux S, Khochbin S. The role of histones in chromatin remodelling during mammalian spermiogenesis. Eur J Biochem. 2004;271:3459–69.
31. Grimes SR Jr, van Wert J, Wolfe SA. Regulation of transcription of the testis-specific histone H1t gene by multiple promoter elements. Mol Biol Rep. 1997;24:175–84.
32. Yan W, Ma L, Burns KH, Matzuk MM. HILS1 is a spermatid-specific linker histone H1-like protein implicated in chromatin remodeling during mammalian spermiogenesis. Proc Natl Acad Sci USA. 2003;100:10546–51.
33. Tachiwana H, Osakabe A, Kimura H, Kurumizaka H. Nucleosome formation with the testis-specific histone H3 variant, H3t, by human nucleosome assembly proteins in vitro. Nucleic Acids Res. 2008;36:2208–18.
34. Li A, Maffey AH, Abbott WD, Conde e Silva N, Prunell A, Siino J, et al. Characterization of nucleosomes consisting of the human testis/sperm-specific histone H2B variant (hTSH2B). Biochemistry. 2005;44:2529–35.
35. Syed SH, Boulard M, Shukla MS, Gautier T, Travers A, Bednar J, et al. The incorporation of the novel histone variant H2AL2 confers unusual structural and functional properties of the nucleosome. Nucleic Acids Res. 2009;37:4684–95.

36. Tachiwana H, Kagawa W, Osakabe A, Kawaguchi K, Shiga T, Hayashi-Takanaka Y, et al. Structural basis of instability of the nucleosome containing a testis-specific histone variant, human H3T. Proc Natl Acad Sci USA. 2010;107:10454–9.
37. Gaucher J, Reynoird N, Montellier E, Boussouar F, Rousseaux S, Khochbin S. From meiosis to postmeiotic events: the secrets of histone disappearance. FEBS J. 2010;277:599–604.
38. Rose KL, Li A, Zalenskaya I, Zhang Y, Unni E, Hodgson KC, et al. C-terminal phosphorylation of murine testis-specific histone H1t in elongating spermatids. J Proteome Res. 2008;7:4070–8.
39. Lu S, Xie YM, Li X, Luo J, Shi XQ, Hong X, et al. Mass spectrometry analysis of dynamic post-translational modifications of TH2B during spermatogenesis. Mol Hum Reprod. 2009;15:373–8.
40. Sarg B, Chwatal S, Talasz H, Lindner HH. Testis-specific linker histone H1t is multiply phosphorylated during spermatogenesis. Identification of phosphorylation sites. J Biol Chem. 2009;284:3610–8.
41. Hazzouri M, Pivot-Pajot C, Faure AK, Usson Y, Pelletier R, Sele B, et al. Regulated hyperacetylation of core histones during mouse spermatogenesis: involvement of histone deacetylases. Eur J Cell Biol. 2000;79:950–60.
42. Pivot-Pajot C, Caron C, Govin J, Vion A, Rousseaux S, Khochbin S. Acetylation-dependent chromatin reorganization by BRDT, a testis-specific bromodomain-containing protein. Mol Cell Biol. 2003;23:5354–65.
43. Moriniere J, Rousseaux S, Steuerwald U, Soler-Lopez M, Curtet S, Vitte AL, et al. Cooperative binding of two acetylation marks on a histone tail by a single bromodomain. Nature. 2009;461:664–8.
44. Dhar S, Thota A, Rao MR. Insights into role of bromodomain, testis-specific (Brdt) in acetylated histone H4-dependent chromatin remodeling in mammalian spermiogenesis. J Biol Chem. 2012;287:6387–405.
45. Shang E, Nickerson HD, Wen D, Wang X, Wolgemuth DJ. The first bromodomain of Brdt, a testis-specific member of the BET sub-family of double-bromodomain-containing proteins, is essential for male germ cell differentiation. Development. 2007;134:3507–15.
46. Yu YE, Zhang Y, Unni E, Shirley CR, Deng JM, Russell LD, et al. Abnormal spermatogenesis and reduced fertility in transition nuclear protein 1-deficient mice. Proc Natl Acad Sci USA. 2000;97:4683–8.
47. Zhao M, Shirley CR, Yu YE, Mohapatra B, Zhang Y, Unni E, et al. Targeted disruption of the transition protein 2 gene affects sperm chromatin structure and reduces fertility in mice. Mol Cell Biol. 2001;21:7243–55.
48. Cho C, Willis WD, Goulding EH, Jung-Ha H, Choi YC, Hecht NB, et al. Haploinsufficiency of protamine-1 or -2 causes infertility in mice. Nat Genet. 2001;28:82–6.
49. Cho C, Jung-Ha H, Willis WD, Goulding EH, Stein P, Xu Z, et al. Protamine 2 deficiency leads to sperm DNA damage and embryo death in mice. Biol Reprod. 2003;69:211–7.
50. Ravel C, Chantot-Bastaraud S, El Houate B, Berthaut I, Verstraete L, De Larouziere V, et al. Mutations in the protamine 1 gene associated with male infertility. Mol Hum Reprod. 2007;13:461–4.
51. Steger K, Wilhelm J, Konrad L, Stalf T, Greb R, Diemer T, et al. Both protamine-1 to protamine-2 mRNA ratio and Bcl2 mRNA content in testicular spermatids and ejaculated spermatozoa discriminate between fertile and infertile men. Hum Reprod. 2008;23:11–6.
52. Monesi V, Geremia R, D'Agostino A, Boitani C. Biochemistry of male germ cell differentiation in mammals: RNA synthesis in meiotic and postmeiotic cells. Curr Top Dev Biol. 1978;12:11–36.
53. Martianov I, Fimia GM, Dierich A, Parvinen M, Sassone-Corsi P, Davidson I. Late arrest of spermiogenesis and germ cell apoptosis in mice lacking the TBP-like TLF/TRF2 gene. Mol Cell. 2001;7:509–15.
54. Martianov I, Brancorsini S, Gansmuller A, Parvinen M, Davidson I, Sassone-Corsi P. Distinct functions of TBP and TLF/TRF2 during spermatogenesis: requirement of TLF for heterochromatic chromocenter formation in haploid round spermatids. Development. 2002;129:945–55.
55. Zheng J, Xia X, Ding H, Yan A, Hu S, Gong X, et al. Erasure of the paternal transcription program during spermiogenesis: the first step in the reprogramming of sperm chromatin for zygotic development. Dev Dyn. 2008;237:1463–76.
56. Philpott A, Leno GH, Laskey RA. Sperm decondensation in Xenopus egg cytoplasm is mediated by nucleoplasmin. Cell. 1991;65:569–78.
57. Ramos I, Martin-Benito J, Finn R, Bretana L, Aloria K, Arizmendi JM, et al. Nucleoplasmin binds histone H2A–H2B dimers through its distal face. J Biol Chem. 2010;285:33771–8.
58. Inoue A, Ogushi S, Saitou M, Suzuki MG, Aoki F. Involvement of mouse nucleoplasmin 2 in the decondensation of sperm chromatin after fertilization. Biol Reprod. 2011;85:70–7.
59. Okuwaki M, Sumi A, Hisaoka M, Saotome-Nakamura A, Akashi S, Nishimura Y, et al. Function of homo- and hetero-oligomers of human nucleoplasmin/nucleophosmin family proteins NPM1, NPM2 and NPM3 during sperm chromatin remodeling. Nucleic Acids Res. 2012;40:4861–78.
60. McLay DW, Clarke HJ. Remodelling the paternal chromatin at fertilization in mammals. Reproduction. 2003;125:625–33.
61. van der Heijden GW, Dieker JW, Derijck AA, Muller S, Berden JH, Braat DD, et al. Asymmetry in histone H3 variants and lysine methylation between paternal and maternal chromatin of the early mouse zygote. Mech Dev. 2005;122:1008–22.
62. Ng RK, Gurdon JB. Epigenetic memory of an active gene state depends on histone H3.3 incorporation into chromatin in the absence of transcription. Nat Cell Biol. 2008;10:102–9.
63. Kishigami S, Van Thuan N, Hikichi T, Ohta H, Wakayama S, Mizutani E, et al. Epigenetic abnormalities of the mouse paternal zygotic genome associated with microinsemination of round spermatids. Dev Biol. 2006;289:195–205.
64. Ogonuki N, Inoue K, Miki H, Mochida K, Hatori M, Okada H, et al. Differential development of rabbit embryos following microinsemination with sperm and spermatids. Mol Reprod Dev. 2005;72:411–7.
65. Mizutani E, Yamagata K, Ono T, Akagi S, Geshi M, Wakayama T. Abnormal chromosome segregation at early cleavage is a major cause of the full-term developmental failure of mouse clones. Dev Biol. 2012;364:56–65.
66. Hirabayashi M, Kato M, Aoto T, Ueda M, Hochi S. Rescue of infertile transgenic rat lines by intracytoplasmic injection of cryopreserved round spermatids. Mol Reprod Dev. 2002;62:295–9.
67. Haigo K, Yamauchi Y, Yazama F, Yanagimachi R, Horiuchi T. Full-term development of hamster embryos produced by injection of round spermatids into oocytes. Biol Reprod. 2004;71:194–8.
68. Ogura A, Matsuda J, Yanagimachi R. Birth of normal young after electrofusion of mouse oocytes with round spermatids. Proc Natl Acad Sci USA. 1994;91:7460–2.
69. Kimura Y, Yanagimachi R. Mouse oocytes injected with testicular spermatozoa or round spermatids can develop into normal offspring. Development. 1995;121:2397–405.
70. Kishigami S, Wakayama S, Nguyen VT, Wakayama T. Similar time restriction for intracytoplasmic sperm injection and round spermatid injection into activated oocytes for efficient offspring production. Biol Reprod. 2004;70:1863–9.

71. Pittoggi C, Magnano AR, Sciamanna I, Giordano R, Lorenzini R, Spadafora C. Specific localization of transcription factors in the chromatin of mouse mature spermatozoa. Mol Reprod Dev. 2001;60:97–106.
72. de Mateo S, Castillo J, Estanyol JM, Ballesca JL, Oliva R. Proteomic characterization of the human sperm nucleus. Proteomics. 2011;11:2714–26.
73. Van Thuan N, Wakayama S, Kishigami S, Ohta H, Hikichi T, Mizutani E, et al. Injection of somatic cell cytoplasm into oocytes before intracytoplasmic sperm injection impairs full-term development and increases placental weight in mice. Biol Reprod. 2006;74:865–73.
74. Miller D, Ostermeier GC. Towards a better understanding of RNA carriage by ejaculate spermatozoa. Hum Reprod Update. 2006;12:757–67.
75. Kumar G, Patel D, Naz RK. c-MYC mRNA is present in human sperm cells. Cell Mol Biol Res. 1993;39:111–7.
76. Ostermeier GC, Miller D, Huntriss JD, Diamond MP, Krawetz SA. Reproductive biology: delivering spermatozoan RNA to the oocyte. Nature. 2004;429:154.
77. Rassoulzadegan M, Grandjean V, Gounon P, Vincent S, Gillot I, Cuzin F. RNA-mediated non-Mendelian inheritance of an epigenetic change in the mouse. Nature. 2006;441:469–74.
78. Liu WM, Pang RT, Chiu PC, Wong BP, Lao K, Lee KF, et al. Sperm-borne microRNA-34c is required for the first cleavage division in mouse. Proc Natl Acad Sci USA. 2012;109:490–4.
79. Hammoud SS, Nix DA, Zhang H, Purwar J, Carrell DT, Cairns BR. Distinctive chromatin in human sperm packages genes for embryo development. Nature. 2009;460:473–8.
80. Brykczynska U, Hisano M, Erkek S, Ramos L, Oakeley EJ, Roloff TC, et al. Repressive and active histone methylation mark distinct promoters in human and mouse spermatozoa. Nat Struct Mol Biol. 2010;17:679–87.
81. van der Heijden GW, Derijck AA, Ramos L, Giele M, van der Vlag J, de Boer P. Transmission of modified nucleosomes from the mouse male germline to the zygote and subsequent remodeling of paternal chromatin. Dev Biol. 2006;298:458–69.
82. van der Heijden GW, Ramos L, Baart EB, van den Berg IM, Derijck AA, van der Vlag J, et al. Sperm-derived histones contribute to zygotic chromatin in humans. BMC Dev Biol. 2008;8:34.
83. Lindeman LC, Andersen IS, Reiner AH, Li N, Aanes H, Ostrup O, et al. Prepatterning of developmental gene expression by modified histones before zygotic genome activation. Dev Cell. 2011;21:993–1004.
84. Wu SF, Zhang H, Cairns BR. Genes for embryo development are packaged in blocks of multivalent chromatin in zebrafish sperm. Genome Res. 2011;21:578–89.
85. Hammoud SS, Nix DA, Hammoud AO, Gibson M, Cairns BR, Carrell DT. Genome-wide analysis identifies changes in histone retention and epigenetic modifications at developmental and imprinted gene loci in the sperm of infertile men. Hum Reprod. 2011;26:2558–69.
86. Chen C, Yang MC, Yang TP. Evidence that silencing of the HPRT promoter by DNA methylation is mediated by critical CpG sites. J Biol Chem. 2001;276:320–8.
87. Kroft TL, Jethanandani P, McLean DJ, Goldberg E. Methylation of CpG dinucleotides alters binding and silences testis-specific transcription directed by the mouse lactate dehydrogenase C promoter. Biol Reprod. 2001;65:1522–7.
88. Chan Y, Fish JE, D'Abreo C, Lin S, Robb GB, Teichert AM, et al. The cell-specific expression of endothelial nitric-oxide synthase: a role for DNA methylation. J Biol Chem. 2004;279:35087–100.
89. Song F, Mahmood S, Ghosh S, Liang P, Smiraglia DJ, Nagase H, et al. Tissue specific differentially methylated regions (TDMR): changes in DNA methylation during development. Genomics. 2009;93:130–9.
90. Song F, Smith JF, Kimura MT, Morrow AD, Matsuyama T, Nagase H, et al. Association of tissue-specific differentially methylated regions (TDMs) with differential gene expression. Proc Natl Acad Sci USA. 2005;102:3336–41.
91. Benvenuto G, Carpentieri ML, Salvatore P, Cindolo L, Bruni CB, Chiariotti L. Cell-specific transcriptional regulation and reactivation of galectin-1 gene expression are controlled by DNA methylation of the promoter region. Mol Cell Biol. 1996;16:2736–43.
92. Stengel S, Fiebig U, Kurth R, Denner J. Regulation of human endogenous retrovirus-K expression in melanomas by CpG methylation. Genes Chromosomes Cancer. 2010;49:401–11.
93. Farthing CR, Ficz G, Ng RK, Chan CF, Andrews S, Dean W, et al. Global mapping of DNA methylation in mouse promoters reveals epigenetic reprogramming of pluripotency genes. PLoS Genet. 2008;4:e1000116.
94. Hammoud SS, Purwar J, Pflueger C, Cairns BR, Carrell DT. Alterations in sperm DNA methylation patterns at imprinted loci in two classes of infertility. Fertil Steril. 2010;94:1728–33.
95. Chan MM, Smith ZD, Egli D, Regev A, Meissner A. Mouse ooplasm confers context-specific reprogramming capacity. Nat Genet. 2012;44:978–80.
96. Campbell KH, Loi P, Otaegui PJ, Wilmut I. Cell cycle co-ordination in embryo cloning by nuclear transfer. Rev Reprod. 1996;1:40–6.
97. Wakayama T. Production of cloned mice and ES cells from adult somatic cells by nuclear transfer: how to improve cloning efficiency? J Reprod Dev. 2007;53:13–26.
98. Ganier O, Bocquet S, Peiffer I, Brochard V, Arnaud P, Puy A, et al. Synergic reprogramming of mammalian cells by combined exposure to mitotic Xenopus egg extracts and transcription factors. Proc Natl Acad Sci USA. 2011;108:17331–6.
99. Lemaitre JM, Danis E, Pasero P, Vassetzky Y, Mechali M. Mitotic remodeling of the replicon and chromosome structure. Cell. 2005;123:787–801.
100. Wang F, Kou Z, Zhang Y, Gao S. Dynamic reprogramming of histone acetylation and methylation in the first cell cycle of cloned mouse embryos. Biol Reprod. 2007;77:1007–16.
101. Bui HT, Van Thuan N, Wakayama T, Miyano T. Chromatin remodeling in somatic cells injected into mature pig oocytes. Reproduction. 2006;131:1037–49.
102. Chang CC, Gao S, Sung LY, Corry GN, Ma Y, Nagy ZP, et al. Rapid elimination of the histone variant MacroH2A from somatic cell heterochromatin after nuclear transfer. Cell Reprogram. 2010;12:43–53.
103. Martin C, Brochard V, Migne C, Zink D, Debey P, Beaujean N. Architectural reorganization of the nuclei upon transfer into oocytes accompanies genome reprogramming. Mol Reprod Dev. 2006;73:1102–11.
104. Dean W, Santos F, Stojkovic M, Zakhartchenko V, Walter J, Wolf E, et al. Conservation of methylation reprogramming in mammalian development: aberrant reprogramming in cloned embryos. Proc Natl Acad Sci USA. 2001;98:13734–8.
105. Santos F, Zakhartchenko V, Stojkovic M, Peters A, Jenuwein T, Wolf E, et al. Epigenetic marking correlates with developmental potential in cloned bovine preimplantation embryos. Curr Biol. 2003;13:1116–21.
106. Hayashi-Takanaka Y, Yamagata K, Wakayama T, Stasevich TJ, Kainuma T, Tsurimoto T, et al. Tracking epigenetic histone modifications in single cells using Fab-based live endogenous modification labeling. Nucleic Acids Res. 2011;39:6475–88.
107. Wu X, Li Y, Xue L, Wang L, Yue Y, Li K, et al. Multiple histone site epigenetic modifications in nuclear transfer and in vitro fertilized bovine embryos. Zygote. 2011;19:31–45.

108. Gao T, Zheng J, Xing F, Fang H, Sun F, Yan A, et al. Nuclear reprogramming: the strategy used in normal development is also used in somatic cell nuclear transfer and parthenogenesis. Cell Res. 2007;17:135–50.
109. Hamatani T, Carter MG, Sharov AA, Ko MS. Dynamics of global gene expression changes during mouse preimplantation development. Dev Cell. 2004;6:117–31.
110. Greda P, Karasiewicz J, Modlinski JA. Mouse zygotes as recipients in embryo cloning. Reproduction. 2006;132:741–8.
111. Egli D, Rosains J, Birkhoff G, Eggan K. Developmental reprogramming after chromosome transfer into mitotic mouse zygotes. Nature. 2007;447:679–85.
112. Bordignon V, Smith LC. Telophase-stage host ooplasts support complete reprogramming of roscovitine-treated somatic cell nuclei in cattle. Cloning Stem Cells. 2006;8:305–17.
113. Jullien J, Pasque V, Halley-Stott RP, Miyamoto K, Gurdon JB. Mechanisms of nuclear reprogramming by eggs and oocytes: a deterministic process? Nat Rev Mol Cell Biol. 2011;12:453–9.
114. Pasque V, Gillich A, Garrett N, Gurdon JB. Histone variant macroH2A confers resistance to nuclear reprogramming. EMBO J. 2011;30:2373–87.
115. Miyamoto K, Pasque V, Jullien J, Gurdon JB. Nuclear actin polymerization is required for transcriptional reprogramming of Oct4 by oocytes. Genes Dev. 2011;25:946–58.
116. Ogushi S, Palmieri C, Fulka H, Saitou M, Miyano T, Fulka J Jr. The maternal nucleolus is essential for early embryonic development in mammals. Science. 2008;319:613–6.
117. Tani T, Shimada H, Kato Y, Tsunoda Y. Bovine oocytes with the potential to reprogram somatic cell nuclei have a unique 23-kDa protein, phosphorylated transcriptionally controlled tumor protein (TCTP). Cloning Stem Cells. 2007;9:267–80.
118. Miyamoto K, Nagai K, Kitamura N, Nishikawa T, Ikegami H, Binh NT, et al. Identification and characterization of an oocyte factor required for development of porcine nuclear transfer embryos. Proc Natl Acad Sci USA. 2011;108:7040–5.
119. Wossidlo M, Nakamura T, Lepikhov K, Marques CJ, Zakhartchenko V, Boiani M, et al. 5-Hydroxymethylcytosine in the mammalian zygote is linked with epigenetic reprogramming. Nat Commun. 2011;2:241.
120. Gu TP, Guo F, Yang H, Wu HP, Xu GF, Liu W, et al. The role of Tet3 DNA dioxygenase in epigenetic reprogramming by oocytes. Nature. 2011;477:606–10.
121. Smith ZD, Chan MM, Mikkelsen TS, Gu H, Gnirke A, Regev A, et al. A unique regulatory phase of DNA methylation in the early mammalian embryo. Nature. 2012;484:339–44.
122. Beaujean N, Taylor J, Gardner J, Wilmut I, Meehan R, Young L. Effect of limited DNA methylation reprogramming in the normal sheep embryo on somatic cell nuclear transfer. Biol Reprod. 2004;71:185–93.
123. Yang J, Yang S, Beaujean N, Niu Y, He X, Xie Y, et al. Epigenetic marks in cloned rhesus monkey embryos: comparison with counterparts produced in vitro. Biol Reprod. 2007;76:36–42.
124. Yamagata K, Yamazaki T, Miki H, Ogonuki N, Inoue K, Ogura A, et al. Centromeric DNA hypomethylation as an epigenetic signature discriminates between germ and somatic cell lineages. Dev Biol. 2007;312:419–26.
125. Yamazaki Y, Fujita TC, Low EW, Alarcon VB, Yanagimachi R, Marikawa Y. Gradual DNA demethylation of the Oct4 promoter in cloned mouse embryos. Mol Reprod Dev. 2006;73:180–8.
126. Kawasumi M, Unno Y, Matsuoka T, Nishiwaki M, Anzai M, Amano T, et al. Abnormal DNA methylation of the Oct-4 enhancer region in cloned mouse embryos. Mol Reprod Dev. 2009;76:342–50.
127. Wee G, Koo DB, Song BS, Kim JS, Kang MJ, Moon SJ, et al. Inheritable histone H4 acetylation of somatic chromatins in cloned embryos. J Biol Chem. 2006;281:6048–57.
128. Epsztejn-Litman S, Feldman N, Abu-Remaileh M, Shufaro Y, Gerson A, Ueda J, et al. De novo DNA methylation promoted by G9a prevents reprogramming of embryonically silenced genes. Nat Struct Mol Biol. 2008;15:1176–83.
129. Antony J, Oback F, Chamley LW, Oback B, Laible G. Transient JMJD2B-Mediated Reduction of H3K9me3 Levels Improves Reprogramming of Embryonic Stem Cells into Cloned Embryos. Mol Cell Biol. 2013;33:974–83.
130. Soufi A, Donahue G, Zaret KS. Facilitators and impediments of the pluripotency reprogramming factors' initial engagement with the genome. Cell. 2012;151:994–1004.
131. Chen J, Liu H, Liu J, Qi J, Wei B, Yang J, et al. H3K9 methylation is a barrier during somatic cell reprogramming into iPSCs. Nat Genet. 2013;45:34–42.
132. Zhao J, Ross JW, Hao Y, Spate LD, Walters EM, Samuel MS, et al. Significant improvement in cloning efficiency of an inbred miniature pig by histone deacetylase inhibitor treatment after somatic cell nuclear transfer. Biol Reprod. 2009;81:525–30.
133. Dai X, Hao J, Hou XJ, Hai T, Fan Y, Yu Y, et al. Somatic nucleus reprogramming is significantly improved by m-carboxycinnamic acid bishydroxamide, a histone deacetylase inhibitor. J Biol Chem. 2010;285:31002–10.
134. Kishigami S, Mizutani E, Ohta H, Hikichi T, Thuan NV, Wakayama S, et al. Significant improvement of mouse cloning technique by treatment with trichostatin A after somatic nuclear transfer. Biochem Biophys Res Commun. 2006;340:183–9.
135. Van Thuan N, Bui HT, Kim JH, Hikichi T, Wakayama S, Kishigami S, et al. The histone deacetylase inhibitor scriptaid enhances nascent mRNA production and rescues full-term development in cloned inbred mice. Reproduction. 2009;138:309–17.
136. Ono T, Li C, Mizutani E, Terashita Y, Yamagata K, Wakayama T. Inhibition of class IIb histone deacetylase significantly improves cloning efficiency in mice. Biol Reprod. 2010;83:929–37.
137. Iager AE, Ragina NP, Ross PJ, Beyhan Z, Cunniff K, Rodriguez RM, et al. Trichostatin A improves histone acetylation in bovine somatic cell nuclear transfer early embryos. Cloning Stem Cells. 2008;10:371–9.
138. Bui HT, Wakayama S, Kishigami S, Park KK, Kim JH, Thuan NV, et al. Effect of trichostatin A on chromatin remodeling, histone modifications, DNA replication, and transcriptional activity in cloned mouse embryos. Biol Reprod. 2010;83:454–63.
139. Bui HT, Seo HJ, Park MR, Park JY, Thuan NV, Wakayama T, et al. Histone deacetylase inhibition improves activation of ribosomal RNA genes and embryonic nucleolar reprogramming in cloned mouse embryos. Biol Reprod. 2011;85:1048–56.
140. Bao S, Miyoshi N, Okamoto I, Jenuwein T, Heard E, Azim Surani M. Initiation of epigenetic reprogramming of the X chromosome in somatic nuclei transplanted to a mouse oocyte. EMBO Rep. 2005;6:748–54.
141. Zhang M, Wang F, Kou Z, Zhang Y, Gao S. Defective chromatin structure in somatic cell cloned mouse embryos. J Biol Chem. 2009;284:24981–7.
142. Teranishi T, Tanaka M, Kimoto S, Ono Y, Miyakoshi K, Kono T, et al. Rapid replacement of somatic linker histones with the oocyte-specific linker histone H1foo in nuclear transfer. Dev Biol. 2004;266:76–86.
143. Gao S, Chung YG, Parseghian MH, King GJ, Adashi EY, Latham KE. Rapid H1 linker histone transitions following fertilization or somatic cell nuclear transfer: evidence for a uniform developmental program in mice. Dev Biol. 2004;266:62–75.
144. Jullien J, Astrand C, Halley-Stott RP, Garrett N, Gurdon JB. Characterization of somatic cell nuclear reprogramming by oocytes in which a linker histone is required for pluripotency gene reactivation. Proc Natl Acad Sci USA. 2010;107:5483–8.

145. Saeki H, Ohsumi K, Aihara H, Ito T, Hirose S, Ura K, et al. Linker histone variants control chromatin dynamics during early embryogenesis. Proc Natl Acad Sci USA. 2005;102:5697–702.
146. Nashun B, Akiyama T, Suzuki MG, Aoki F. Dramatic replacement of histone variants during genome remodeling in nuclear-transferred embryos. Epigenetics. 2011;6:1489–97.
147. Wossidlo M, Arand J, Sebastiano V, Lepikhov K, Boiani M, Reinhardt R, et al. Dynamic link of DNA demethylation, DNA strand breaks and repair in mouse zygotes. EMBO J. 2010;29:1877–88.
148. Boiani M, Eckardt S, Scholer HR, McLaughlin KJ. Oct4 distribution and level in mouse clones: consequences for pluripotency. Genes Dev. 2002;16:1209–19.
149. Bortvin A, Eggan K, Skaletsky H, Akutsu H, Berry DL, Yanagimachi R, et al. Incomplete reactivation of Oct4-related genes in mouse embryos cloned from somatic nuclei. Development. 2003;130:1673–80.
150. Fukuda A, Cao F, Morita S, Yamada K, Jincho Y, Tane S, et al. Identification of inappropriately reprogrammed genes by large-scale transcriptome analysis of individual cloned mouse blastocysts. PLoS One. 2010;5:e11274.
151. Cao F, Fukuda A, Watanabe H, Kono T. The transcriptomic architecture of mouse Sertoli cell clone embryos reveals temporal-spatial-specific reprogramming. Reproduction. 2013; 145:277–88.
152. Vassena R, Han Z, Gao S, Baldwin DA, Schultz RM, Latham KE. Tough beginnings: alterations in the transcriptome of cloned embryos during the first two cell cycles. Dev Biol. 2007;304:75–89.
153. Zheng Z, Jia JL, Bou G, Hu LL, Wang ZD, Shen XH, et al. rRNA genes are not fully activated in mouse somatic cell nuclear transfer embryos. J Biol Chem. 2012;287:19949–60.
154. Gao S, Chung YG, Williams JW, Riley J, Moley K, Latham KE. Somatic cell-like features of cloned mouse embryos prepared with cultured myoblast nuclei. Biol Reprod. 2003;69:48–56.
155. Wang K, Otu HH, Chen Y, Lee Y, Latham K, Cibelli JB. Reprogrammed transcriptome in rhesus-bovine interspecies somatic cell nuclear transfer embryos. PLoS One. 2011;6: e22197.
156. Smith SL, Everts RE, Tian XC, Du F, Sung LY, Rodriguez-Zas SL, et al. Global gene expression profiles reveal significant nuclear reprogramming by the blastocyst stage after cloning. Proc Natl Acad Sci USA. 2005;102:17582–7.
157. Inoue K, Kohda T, Sugimoto M, Sado T, Ogonuki N, Matoba S, et al. Impeding Xist expression from the active X chromosome improves mouse somatic cell nuclear transfer. Science. 2010;330:496–9.
158. Wen B, Wu H, Shinkai Y, Irizarry RA, Feinberg AP. Large histone H3 lysine 9 dimethylated chromatin blocks distinguish differentiated from embryonic stem cells. Nat Genet. 2009;41:246–50.
159. Matoba S, Inoue K, Kohda T, Sugimoto M, Mizutani E, Ogonuki N, et al. RNAi-mediated knockdown of Xist can rescue the impaired postimplantation development of cloned mouse embryos. Proc Natl Acad Sci USA. 2011;108:20621–6.
160. Ng RK, Gurdon JB. Epigenetic memory of active gene transcription is inherited through somatic cell nuclear transfer. Proc Natl Acad Sci USA. 2005;102:1957–62.
161. Jouneau A, Zhou Q, Camus A, Brochard V, Maulny L, Collignon J, et al. Developmental abnormalities of NT mouse embryos appear early after implantation. Development. 2006;133:1597–607.
162. Lin J, Shi L, Zhang M, Yang H, Qin Y, Zhang J, et al. Defects in trophoblast cell lineage account for the impaired in vivo development of cloned embryos generated by somatic nuclear transfer. Cell Stem Cell. 2011;8:371–5.
163. Takeda K, Akagi S, Kaneyama K, Kojima T, Takahashi S, Imai H, et al. Proliferation of donor mitochondrial DNA in nuclear transfer calves (Bos taurus) derived from cumulus cells. Mol Reprod Dev. 2003;64:429–37.
164. Tamashiro KL, Wakayama T, Akutsu H, Yamazaki Y, Lachey JL, Wortman MD, et al. Cloned mice have an obese phenotype not transmitted to their offspring. Nat Med. 2002;8:262–7.
165. Niemann H, Lucas-Hahn A. Somatic cell nuclear transfer cloning: practical applications and current legislation. Reprod Domest Anim. 2012;47(Suppl 5):2–10.
166. Ogonuki N, Inoue K, Yamamoto Y, Noguchi Y, Tanemura K, Suzuki O, et al. Early death of mice cloned from somatic cells. Nat Genet. 2002;30:253–4.
167. Martin C, Beaujean N, Brochard V, Audouard C, Zink D, Debey P. Genome restructuring in mouse embryos during reprogramming and early development. Dev Biol. 2006;292:317–32.
168. Puschendorf M, Terranova R, Boutsma E, Mao X, Isono K, Brykczynska U, et al. PRC1 and Suv39 h specify parental asymmetry at constitutive heterochromatin in early mouse embryos. Nat Genet. 2008;40:411–20.
169. Kawasumi M, Anzai M, Takehara T, Mitani T, Kato H, Saeki K, et al. Abnormal distribution of chromosomes in the first division of nuclear transferred mouse embryos. J Reprod Dev. 2007;53:615–22.
170. Simerly C, Dominko T, Navara C, Payne C, Capuano S, Gosman G, et al. Molecular correlates of primate nuclear transfer failures. Science. 2003;300:297.
171. Miyara F, Han Z, Gao S, Vassena R, Latham KE. Non-equivalence of embryonic and somatic cell nuclei affecting spindle composition in clones. Dev Biol. 2006;289:206–17.
172. Kishikawa H, Wakayama T, Yanagimachi R. Comparison of oocyte-activating agents for mouse cloning. Cloning. 1999;1:153–9.
173. Schurmann A, Wells DN, Oback B. Early zygotes are suitable recipients for bovine somatic nuclear transfer and result in cloned offspring. Reproduction. 2006;132:839–48.
174. Paradowska AS, Miller D, Spiess AN, Vieweg M, Cerna M, Dvorakova-Hortova K, et al. Genome wide identification of promoter binding sites for H4K12ac in human sperm and its relevance for early embryonic development. Epigenetics. 2012;7:1057–70.
175. Jenuwein T, Allis CD. Translating the histone code. Science. 2001;293:1074–80.
176. Kouzarides T. Chromatin modifications and their function. Cell. 2007;128:693–705.
177. Bannister AJ, Kouzarides T. Regulation of chromatin by histone modifications. Cell Res. 2011;21:381–95.
178. Tan M, Luo H, Lee S, Jin F, Yang JS, Montellier E, et al. Identification of 67 histone marks and histone lysine crotonylation as a new type of histone modification. Cell. 2011;146: 1016–28.
179. Liu Q, Segal DJ, Ghiara JB, Barbas CF 3rd. Design of polydactyl zinc-finger proteins for unique addressing within complex genomes. Proc Natl Acad Sci USA. 1997;94:5525–30.
180. Bae KH, Kwon YD, Shin HC, Hwang MS, Ryu EH, Park KS, et al. Human zinc fingers as building blocks in the construction of artificial transcription factors. Nat Biotechnol. 2003;21: 275–80.
181. Boch J, Scholze H, Schornack S, Landgraf A, Hahn S, Kay S, et al. Breaking the code of DNA binding specificity of TAL-type III effectors. Science. 2009;326:1509–12.
182. Moscou MJ, Bogdanove AJ. A simple cipher governs DNA recognition by TAL effectors. Science. 2009;326:1501.
183. Sander JD, Maeder ML, Reyon D, Voytas DF, Joung JK, Dobbs D. ZiFiT (Zinc Finger Targeter): an updated zinc finger engineering tool. Nucleic Acids Res. 2010;38:W462–8.

184. Miller JC, Tan S, Qiao G, Barlow KA, Wang J, Xia DF, et al. A TALE nuclease architecture for efficient genome editing. Nat Biotechnol. 2011;29:143–8.
185. Takahashi K, Yamanaka S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell. 2006;126:663–76.
186. Hayashi K, Ohta H, Kurimoto K, Aramaki S, Saitou M. Reconstitution of the mouse germ cell specification pathway in culture by pluripotent stem cells. Cell. 2011;146:519–32.
187. Roh S, Choi HY, Park SK, Won C, Kim BW, Kim JH, et al. Porcine nuclear transfer using somatic donor cells altered to express male germ cell function. Reprod Fertil Dev. 2009;21:882–91.

Appendix 3

Book chapter: ‘Cloning of Amphibia’ (part of ‘Principles of Cloning’)

Teperek-Tkacz M., Byrne JA., Gurdon JB., Elsevier, 2014

Chapter 14

Cloning of Amphibia

Marta Teperek-Tkacz[1,2], James A. Byrne[3] and J. B. Gurdon[1,2]
[1]Wellcome Trust/Cancer Research UK Gurdon Institute, Cambridge, UK, [2]Department of Zoology, University of Cambridge, Cambridge, UK,
[3]Molecular and Medical Pharmacology, Broad Stem Cell Center, University of California, LA, USA

Chapter Outline

INTRODUCTION AND BACKGROUND

The original interest in developing a procedure for somatic cell nuclear transfer (cloning) in animals was to provide a test of whether all of the different kinds of cells in the vertebrate body have the same genetic constitution. The idea of a nuclear transplantation experiment to test this proposition dates back to Rauber (1886); see McKinnell (1978). In 1883, August Weismann proposed that chromosomes did not divide equally during development, but that different chromosomal determinants entered different cells (Weismann, 1893). The first direct empirical evidence against this hypothesis came from Hans Driesch, who discovered that separated sea urchin blastomeres could develop into complete larvae (Driesch, 1894), thus demonstrating that all of the genetic constituents necessary for complete development were retained in blastomeres and providing the first experimentally observable instance of regulative development. Spemann also had this question in mind when he devised an egg ligature experiment in 1928 (Spemann, 1938). Spemann succeeded in showing that the nucleus of an eight-cell newt embryo could, when combined with cytoplasm from such an embryo, generate normal development. This demonstrated the totipotency of a nucleus of an eight-cell embryo, but did not show whether totipotency of the genome continued beyond this stage. Briggs and King (1952) were the first to carry out a successful nuclear transplantation experiment in which the nucleus of a *Rana pipiens* blastula cell was transplanted to an enucleated egg and was able to give rise to normal swimming tadpoles. Later work by Briggs and King (1957) found that nuclei from the endoderm cells of late gastrula and later embryos were no longer able to support normal development. This led to the idea that there may be some stable restriction of developmental potentiality that takes place before or during gastrulation. In 1958 nuclear transplantation in *Xenopus laevis* was successful, such that nuclei of embryo cells could be transplanted to enucleated eggs and sustain normal development. The *Xenopus* work agreed with the results of Briggs and King in that the extent to which transplanted nuclei supported normal development decreased as more differentiated cells were used as donors (Gurdon, 1960). Surprisingly, at that time, it was then found that some nuclei from differentiated cells of the intestinal epithelium could also yield normal tadpole development (Gurdon *et al.*, 1962a). Finally, some of the *Xenopus* tadpoles derived from transplanted

Principles of Cloning. DOI: http://dx.doi.org/10.1016/B978-0-12-386541-0.00014-X

intestinal epithelium nuclei became sexually normal and mature adults (Gurdon and Uehlinger, 1966). Therefore, although the extent to which normal development was achieved decreased substantially as the nuclei of more differentiated cells were used, it was nevertheless clear that some differentiated intestinal cells contained entirely totipotent nuclei.

Here we review three aspects of amphibian cloning: (1) development of nuclear transfer embryos using second meiotic metaphase eggs; (2) immediate changes in transplanted nuclei; and (3) mechanisms of nuclear reprogramming by meiotic prophase oocytes.

DEVELOPMENT OF NUCLEAR TRANSFER EMBRYOS USING SECOND MEIOTIC METAPHASE EGGS

Techniques and Methods

Some amphibian eggs, like those of *Xenopus*, do not need more activation than is provided normally by sperm or experimentally by the injection needle. Activation by these means releases oscillatory waves of calcium concentration that initiate development. Unfortunately, it has not so far been found possible to introduce a nucleus into an egg without causing activation; ideally it would be possible to transplant a nucleus to an egg and then prevent activation for a few hours so that the transplanted nucleus can have time to initiate changes from its somatic state. As it is, DNA replication and cell division are initiated at a defined time, 90 minutes at 18°C in *Xenopus*, after activation. The inability to delay the onset of the first cell division of the egg that receives a transplanted nucleus seems often to give insufficient time for the nucleus of a slow dividing cell to complete its replication cycle before cell division. This can lead to incompletely replicated chromosomes being torn apart at the first cell division (see below).

The nuclear transfer procedure originally devised by Briggs and King (1952) involved sucking a donor cell into a small micropipette the diameter of which is such that the donor cell wall is broken but the donor cell nucleus is not damaged. The whole ruptured cell is then injected into the egg (Figure 14.1). This procedure works well for all somatic cells so long as they are large enough to be handled this way (i.e., about 15 µm in diameter). For much smaller donor cells it is convenient to use a chemical means of cell membrane rupture, and this can be done by use of streptolysin O. Chan and Gurdon (1996) have shown that results obtained by chemical cell membrane rupture are comparable with those achieved with physical rupture.

It has not been easy to synchronize donor cells in amphibian embryos. Such attempts as have been made have not shown a significant effect of cell cycle phase on nuclear transplant embryo development (reviewed by Gurdon, 1986).

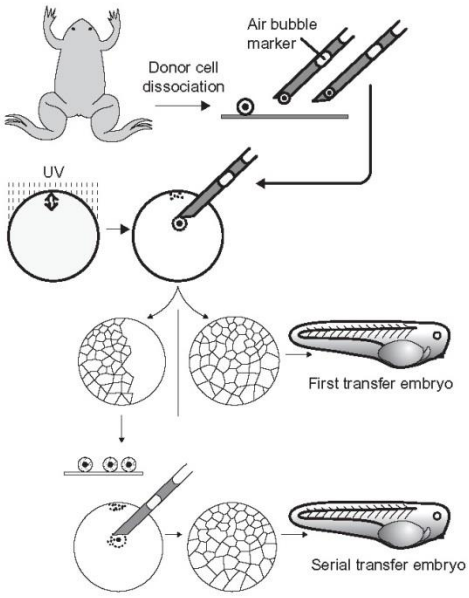


FIGURE 14.1 Nuclear transfer procedure for *Xenopus*, illustrated with transfers from adult foot skin cells. Nuclear transfers from specialist cells often yield a substantial number of partially cleaved blastulae, which can be used, as shown here, as a source of nuclei for serial transfers. The nuclear transfer procedure for other amphibian species differs from that shown mainly by the method of enucleation (ultraviolet radiation, UV, in this figure) and in some cases by the inclusion of an egg activation step.

Genome Conservation

Primary Nuclear Transfers

If the nucleus of a clearly differentiated cell can functionally replace the egg and sperm pronuclei in a zygote, this means that the genes required for development, as far as the nuclear transplant embryo differentiates, must have been conserved in that donor cell. For all somatic cell nuclear transfer experiments, from Briggs and King (1957) onwards, it has been clear that the more differentiated the donor cell is, the less far does the nuclear transplant embryo develop. The most normal nuclear transplant embryo development is obtained by nuclei of early embryonic cells, but even in these cases it is rare for more than about one-third of the nuclear transfers to yield entirely normal development, reaching, for example, feeding larvae. One should rather be impressed by the fact that somatic cell nuclear transfer works at all than that many resulting embryos develop abnormally.

The most informative nuclear transplantation results come from work on the endoderm/intestine cell lineage in *Xenopus*. Starting with intestinal epithelium cells of feeding larvae, 1.5% of total nuclear transfers became normal feeding tadpoles. Several of these intestinal nuclear transfer embryos were reared to sexual maturity, and both males and females were fertile.

The most informative nuclear transplantation results come from work on the endoderm/intestine cell lineage in *Xenopus* (Gurdon, 1962a). Starting with intestinal epithelium cells of feeding larvae, 1.5% of total nuclear transfers became normal feeding tadpoles. This figure becomes 7% if the results of serial nuclear transfers are included (Figure 14.1 and see below), or 24% if only normally cleaved embryos are included. The percentage of intestine first and serial transfers reaching muscular response stage tadpoles was 20%. Several of these intestinal nuclear transfer embryos were reared to sexual maturity, and both males and females were fertile to the extent that laboratory-reared animals usually are (Gurdon and Uehlinger, 1966). These frogs carried the one nucleolus nuclear marker. These results established the principle that a complete genome can be conserved during the process of cell differentiation and sexual maturation.

This principle was supported by later experiments using cultured cells grown from a range of adult tissue, including skin, lung, and kidney. From all these adult tissues, apparently normal, swimming tadpoles were obtained, containing a wide range of (if not all) cell types present at this stage (Laskey and Gurdon, 1970). A similar result was subsequently achieved for keratinized skin cells of adult frogs (Gurdon *et al.*, 1975). In all these cases, the swimming larvae did not survive to adulthood, but nevertheless showed that cell differentiation to adult tissues does not entail any loss or stable inactivation of genes required for all cell types that compose a swimming tadpole. The possibility was raised that the sexually mature, intestinal epithelium-derived frogs could have arisen from a failure of the enucleation of the recipient eggs. However, the use of the one nucleolus genetic marker (Elsdale *et al.*, 1960) proved beyond doubt that it was entirely the nucleus of a differentiated intestinal epithelium cell that could give rise to sexually mature animals. As with the intestinal epithelial cell nuclear transplants, the percentage of total nuclear transfers reaching the swimming tadpole stage was low, ranging from 1% to 10%. In the course of time, improved techniques for nuclear transplantation enabled the success rate of somatic cell nuclear transfer from differentiated cells to be much improved in *Rana pipiens*, the species originally used by Briggs and King (Hennen, 1970).

Among these early nuclear transfer experiments, clones of adult frogs were obtained (see, for example, Gurdon, 1962b). Cloning in amphibia attracted considerable public interest on the grounds that this procedure might eventually be applied to humans, and this interest was much accentuated by the production of fertile adult sheep by nuclear transfer from differentiated mammary gland cells (Wilmut *et al.*, 1997). This work, together with subsequent work with the mouse (Hochedlinger and Jaenisch, 2002), confirmed that the process of cell differentiation does not involve any stable genetic change; this conclusion in mammalian work also demonstrated that the nuclei of adult cells, apart from differentiated cells, shared genetic totipotency.

In the early days of nuclear transplantation in amphibia, a number of species were used apart from *Rana pipiens* and *Xenopus laevis*. These include *Pleurodeles* (Gallien *et al.*, 1973) and *Ambystoma* (Briggs *et al.*, 1964). The success rate of nuclear transfer in these amphibian species was less than in *Xenopus*, but supported the general principle that genetic totipotency is maintained during cell differentiation.

It is important to note that the nuclei of germline cells, which must in principle be totipotent, do not always yield normal nuclear transplant embryo development. Smith (1965) found that nuclei of tadpole germ cells yielded normal development, after nuclear transfer, but not as frequently as blastula nuclei; a substantial proportion of the germ-cell derived embryos developed abnormally.

An attempt to assess the developmental capacity of cancer cell nuclei has been undertaken using cells from the Herpes virus-associated *Rana pipens* adenocarcinoma (McKinnell *et al.*, 1969). Some of the resulting nuclear transplant embryos reached the muscular response stage of development. However, there is the lingering question of whether these embryos are derived from nuclei of cancerous, as opposed to normal somatic, cells.

Serial Nuclear Transfers

From an early stage in the field, serial nuclear transplantation has been carried out (King and Briggs, 1956). In this case an embryo resulting from the transfer of the somatic nucleus is itself used to provide nuclei for subsequent transfer to another set of recipient eggs (Figure 14.1). Early experiments found very large differences in the normality or abnormality of members of such serial nuclear transfer clones. In many cases (Di Berardino and King, 1967) chromosomal defects were observed in the clone of embryos, and these were not present in the donor cells themselves. Interestingly these chromosome defects continue to arise after the first division. Therefore these defects must have arisen as a result of nuclear transfer. This could well be because chromosome replication was imperfect in the initial nuclear transfer experiments, and/or that the rapid rate of cell division during *Xenopus* early embryogenesis does not usually permit sufficient time for complete chromosome replication.

In *Xenopus* serial nuclear transfer experiments, Gurdon (1962a) saw what would appear to be an improvement in the normality of development as serial nuclear transfer was carried out. This was true when the first nuclear transfer embryos were themselves defective, as in the case of partially cleaved blastulae. Taking the normal cells from such defective blastulae whose development would arrest at an early stage, it was found that some entirely normal tadpoles could be obtained among the serial transfer embryos (Figure 14.1). The best explanation for this phenomenon is that the first transfer embryos were a mosaic of cells, some of which contained normal chromosome sets and others of which did not. Such partial embryos commonly arise when a transplanted nucleus is forced into division at the first cell cycle but does not complete its replication and moves entirely to one of the first two blastomeres. It then has a second opportunity to complete its replication, thereby creating an embryo with one normal or nearly normal cell and with another non-nucleate cell. The nucleated cell, having now completed chromosome replication, is able to divide and produce a clone of more normal cells within the defective embryo. By serially transplanting such nuclei, the true developmental capacity of the original donor nucleus is revealed. Such a serial nuclear transfer procedure made it possible to demonstrate normal or nearly normal development from a higher proportion of the original donor cells than could be achieved by considering only the first nuclear transfer embryos.

It has also been possible to take normal cells from such defective first nuclear transfer embryos and graft them to a host. Using suitably marked nuclei, it is possible to demonstrate the derivation of normal cells, such as functional muscle and nerve, from original donor nuclei that yielded defective early embryos which would soon die (Byrne *et al.*, 2002). As a result of these serial nuclear transfer and graft procedures, it was possible to show that functional muscle and nerve cells, contained in spontaneously moving tadpoles, could be derived from as many as 30% of all the nuclei transplanted from intestinal epithelial cells (Gurdon, 1962a). Using the results of first nuclear transfers only, this percentage was less than 20%.

It might be pointed out that the proportion of differentiated cells whose nuclei give entirely normal development is very small, often amounting to 2% or less. This does not necessarily mean that 98% of the donor cell population was genetically defective or had stably repressed genes. Many of the defective embryos probably owe their defects to incomplete DNA replication after the initial nuclear transfer from somatic cells with a long DNA replication stage (see above).

Non-Enucleated Recipient Eggs

It is technically straightforward to transplant somatic nuclei into non-enucleated eggs. When this happens, the resulting embryos are often triploid or triploid with chromosome deficiencies. Such embryos develop less well than a perfect haploid embryo (androgenetic or gynogenetic), and no better than enucleated eggs with the same transplanted nuclei (Brun, 1973). The aneuploidy resulting from incomplete replication of donor somatic nuclei could well account for this. It has not so far been useful to pursue this line of investigation.

Cell-Type Specificity of Abnormal Development and Epigenetic Memory

Some genes, such as 5S oocyte active genes, are rapidly repressed after nuclear transfer to eggs, and then reactivated as nuclear transplant embryos progress through development (Wakefield and Gurdon, 1983). However, at an early stage in the amphibian nuclear transfer field, it was important to know whether the developmental abnormalities seen in nuclear transplant embryos represented the particular cell type from which nuclei were taken. It could be asked, for example, whether nuclear transplant embryos resulting from endoderm nuclei might develop with more normal endoderm-derived cells than with cells of ectodermal, nerve, or mesodermal lineages (Briggs and King, 1960). If such a result was obtained, this could reflect some developmental repression of ectoderm or mesoderm genes no longer required, in this case, for the endodermal lineage. An extensive test was performed by Di Berardino and King (1967) to recognize a so-called "endoderm syndrome," in which endodermal tissues might be more normal when embryos were derived from transplanted endoderm nuclei than from neural or mesodermal nuclei. Only 3 embryos out of 23 conformed to the "endoderm syndrome." These results did not therefore give significant support to the idea that nuclear transplant embryo deficiencies tend to correlate inversely with the origin of donor nuclei. In all nuclear transplant embryo experiments it has often been found that endodermal tissues are less defective than neural or mesodermal tissues in abnormal nuclear transplant embryos. This could well be because endodermal tissues differentiate later in development than other lineages, and therefore these cells might survive better. In *Xenopus*, Simnett (1964) tested this idea by transplanting nuclei from neurectoderm tissue. He found that the deficiencies in neurectoderm-derived nuclear transplant embryos were similar to those of endoderm-derived embryos. This again failed to reveal donor cell-type-specific deficiencies.

In more recent work benefiting from the greatly improved precision with which the expression of individual genes can be measured, evidence has come to light regarding the phenomenon of "epigenetic memory." In this case, somatic nuclei were transplanted from muscle and muscle progenitor tissue and the resulting embryos were assayed for the expression of muscle, nerve, or endoderm

genes. Many of such embryos were defective and included partially cleaved embryos of the kind referred to above. The more normal cells of such partially cleaved embryos revealed a remarkable tendency for genes to be overexpressed in the nuclear transplant embryos when donor cells expressing these genes were used. Thus, about half of all nuclear transplant embryos derived from muscle nuclei showed some degree (often very high) of overexpression of muscle genes. A similar effect was observed when nuclei were taken from neurectodermal tissues (Ng and Gurdon, 2008). The extent of inappropriate gene expression of, for example, muscle genes in neural tissue of nuclear transplant embryos derived from muscle donor nuclei was very variable. In some cases, the overexpression of muscle genes in neural tissue was almost as high as the expression of these genes in muscle tissue. This, at first sight surprising, result seems to reflect a remarkable memory of an active gene state. In *Xenopus*, there is little or no transcription after fertilization or nuclear transfer until the mid-blastula stage. Therefore, in these cases a muscle nucleus is transplanted, ceases transcription, and then re-initiates transcription of donor specific genes after the blastula stage.

When a muscle nucleus is originally transplanted, it tends to resurrect its transcriptional pattern in non-muscle tissue at the mid-blastula stage. This is an example of the epigenetic memory of an active gene state.

When a muscle nucleus is originally transplanted, it tends to resurrect its transcriptional pattern in non-muscle tissue at the mid-blastula stage. This is an example of the epigenetic memory of an active gene state. Such a phenomenon has subsequently been recognized in induced pluripotency experiments, although it has been reported that the incidence of epigenetic memory is significantly lower in egg (metaphase II oocyte)-reprogrammed cells compared to factor-reprogrammed cells (Kim *et al.*, 2010; Polo *et al.*, 2010), highlighting the possibility that there may be as-yet unidentified factors in the metaphase II oocyte that could help reduce the epigenetic memory reported following factor-based reprogramming in both mouse (Kim *et al.*, 2010) and human induced pluripotent stem cells (Kim *et al.*, 2011).

There may be as-yet unidentified factors in the metaphase II oocyte that could help reduce the epigenetic memory reported following factor-based reprogramming in both mouse and human induced pluripotent stem cells.

It is also important to note that efforts to further remove epigenetic memory observed following somatic cell nuclear transfer may also have implications for removing epigenetic memory observed following factor-based reprogramming, and that the amphibian provides a powerful system for investigating the mechanisms underlying these reprogramming questions. In the original work with *Xenopus* (Ng and Gurdon, 2008) there was evidence that this epigenetic memory depended upon a high content of the variant histone H3.3 in *Xenopus* eggs. It seems that histone H3.3 was incorporated into the transplanted nuclei and their daughters during the early cleavage divisions of the egg. It is known that histone H3.3 can be incorporated, via its special chaperone HIRA, throughout the cell cycle, giving it some potential advantage for incorporation compared to the core histones H3.1 or H3.2 (Ray-Gallet *et al.*, 2002; Tagami *et al.*, 2004). According to this idea, somatic nuclei deriving from a particular tissue, carrying active genes, would tend to have these genes maintained in a potentially active transcriptional state. This would seem, in some cases, to persist through multiple cell divisions until the mid-blastula stage. Another aspect of this *Xenopus* work on epigenetic memory showed that it was most often observed with lineage-determining genes and not with terminal cell-type specific genes. It was suggested that a tendency for epigenetic memory to exist would be obscured by the series of sequential gene activations leading to terminal cell differentiation. If, as was suggested, there is a balance between active genes in donor nuclei continuing their transcriptional activity and the opposite effect of egg components reprogramming such nuclei, then such memory would be most easily observed in the early lineage-determining genes.

Nuclear Transfers between Different Species

It is well known that combinations of nucleus and cytoplasm from different species do not develop normally (reviewed by Gurdon, 1986). However, combinations of nucleus and cytoplasm from sub-species, such as *Xenopus laevis* and *laevis victorianus*, survive entirely normally, as indeed happens with normal fertilization under wild conditions. In fact, the definition of a species is often considered to be one in which hybrids are not viable. In *Xenopus*, nuclear transfers between *Xenopus laevis* and *Xenopus tropicalis* were indeed developmentally defective; *tropicalis* nuclei in *laevis* eggs survived much better than *laevis* nuclei in *tropicalis* eggs, possibly because small *tropicalis* eggs are technically difficult to work with.

It is interesting to ask, which genes in these cross-species combinations cause the defective development? What are the cytoplasmic substances which fail to activate genes of another species? Using the *Xenopus laevis/tropicalis* combination, the surprising conclusion was that many of the earliest genes transcribed at the time of zygotic gene activation are transcribed fairly normally in this lethal combination. Further analysis showed that the developmental defects arose significantly because the signaling processes

associated with normal gastrulation operate quantitatively at about half the normal level. This can account for developmental deficiencies (Narbonne *et al.*, 2011).

Cross-species nuclear transfers of this kind were also carried out in a combination in which nuclei of the urodele *Pleurodeles* were transplanted to enucleated eggs of *Xenopus*, a combination of species separated by many millions of years in evolution. Surprisingly, some of the *Pleurodeles* genes were successfully activated by the *Xenopus* egg cytoplasm (Narbonne and Gurdon, 2012). There is, however, a limit to the ability of the nucleus of one species to initiate the earliest stages of development in eggs of another.

IMMEDIATE CHANGES IN TRANSPLANTED NUCLEI

DNA Replication

Somatic nuclei undergo dramatic changes in structure and activity very soon after transfer to oocytes in different stages of meiosis, and this has opened the way to understanding the molecular changes involved. After nuclear transfer to eggs, sperm and transplanted somatic nuclei undergo a substantial enlargement and decondensation of chromatin. As soon as 20 minutes after fertilization at 18°C, DNA synthesis is initiated in the sperm nuclei. This first round of DNA synthesis is completed in the next 20 minutes, and cell division starts 90 minutes after fertilization (Graham *et al.*, 1966). H3 thymidine labeling has shown that transplanted somatic nuclei continue to synthesize DNA for 40 minutes after nuclear injection, and often continue to do this up to the time of the first cell division (Figure 14.2). This prolonged period of DNA replication seems to reflect inadequately rapid and complete DNA replication, probably because the DNA synthesis period of somatic nuclei often lasts about 6 hours. Incomplete DNA replication is often followed by irregular cleavage and the production of partial blastulae (see above) (Graham *et al.*, 1966). The mechanism by which eggs induce DNA replication in somatic nuclei has been extensively analyzed by Laskey and colleagues (Laskey, 1989). This first synthetic step of transplanted nuclei is caused by a high concentration of MCM in the egg and the opening up of chromatin so that multiple origins of replication are initiated, thereby enabling DNA replication of the whole genome to be completed in a very short time (Laskey, 2005; Lemaitre *et al.*, 2005). The nuclei of somatic cells resist this accelerated DNA replication. It is possible that this resistance to induced rapid DNA replication has the same mechanistic basis as resistance to transcription (see below).

The mechanistic analysis of the initiation of DNA replication in eggs has depended to a very large extent on the successful production of extracts from eggs; this in turn owes its success to minimal dilution in making such extracts (Lohka and Masui, 1983), and to the large size and ready availability of *Xenopus* eggs. It is likely that, if we could understand the resistance to more rapid DNA replication and relieve this constraint, the efficiency and normality of development after nuclear transfer could be radically improved.

Oocytes Undergoing Meiotic Division

The fully-grown oocyte of *Xenopus* responds rapidly to a progesterone pulse and undergoes completion of first meiosis and the progression of the second meiotic division as far as second metaphase. During this time, the oocyte's lampbrush chromosomes become highly condensed and undergo the two meiotic reduction divisions to change the tetraploid chromosome content of an oocyte to the haploid dose of an unfertilized egg. If multiple somatic nuclei are injected into these oocytes as they undergo maturation, a remarkable condensation of chromosomes takes place, and these are arranged on multiple spindles distributed throughout the oocyte (Figure 14.3). It is remarkable that multiple somatic nuclei injected into growing oocytes (see below) become enlarged and active in RNA synthesis, whereas the same nuclei injected into meiotic maturation oocytes are converted into condensed chromosomes on spindles; the same nuclei injected into activated eggs undergo massive induction of DNA synthesis and inhibition of transcription.

Meiotic Prophase Oocytes

It was found, some time ago, that multiple somatic nuclei could be injected into the nucleus (germinal vesicle) of growing oocytes in first meiotic prophase. Such nuclei undergo a massive enlargement and decondensation of chromatin (Gurdon, 1976). To this extent, they come to resemble the oocyte's own germinal vesicle nucleus. There is no induction of DNA replication in such nuclei and they become intensely active in RNA synthesis, as are

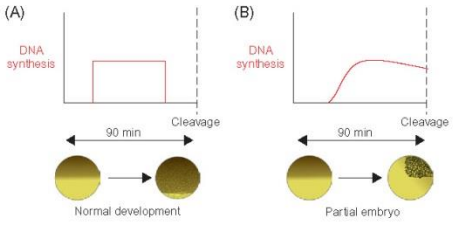


FIGURE 14.2 Effects of incomplete DNA synthesis on embryonic development after nuclear transfer. (A) shows a timely termination of DNA replication and proper embryonic development after fertilization. (B) A somatic cell nucleus after nuclear transfer is unable to complete DNA replication before the cleavage onset, resulting in formation of a partial embryo.

MECHANISMS OF NUCLEAR REPROGRAMMING BY MEIOTIC PROPHASE OOCYTES

Changes Induced by Nuclear Transfer to Oocytes

The massive swelling and chromatin decondensation referred to above is accompanied by a huge increase in accumulated RNA, as revealed by autoradiography of nuclei transplanted to H3-U containing oocytes. An important question is whether the induced transcription is global, i.e., applies to all genes in the genome, or is selective, with some genes being activated and others repressed. The first indication of a selective change in gene activity after nuclear transfer came from experiments in which somatic nuclei were transplanted between two amphibian species, namely *Xenopus* and *Pleurodeles*. This cross-species combination made it possible to distinguish proteins attributable to new activity of transplanted nuclei as opposed to those resulting from pre-existing recipient oocyte mRNA. 14C-labeled proteins were distinguished by 2D gel electrophoresis (De Robertis and Gurdon, 1977). Many genes that were expressed in the donor *Xenopus* cultured somatic cells were repressed after transfer to oocytes of *Pleurodeles*. Conversely, some of the genes characteristically expressed in *Xenopus* oocytes were induced in the *Xenopus* nuclei transplanted to *Pleurodeles* oocytes. These changes were seen on a background of genes typically expressed by maternal mRNA in *Pleurodeles* oocytes. These results demonstrated that the transfer of somatic nuclei to oocytes induces oocyte-specific genes, and that somatic genes are repressed. Therefore, at the level of protein synthesis, nuclear transfers to oocytes demonstrate selective gene activation and repression. It was not known at what level of gene expression these changes took place.

A valuable advance was to be able to monitor changes in gene expression induced by oocytes at the level of transcription. Byrne *et al.* (2003) transplanted the nuclei of mammalian cells to the *Xenopus* oocyte germinal vesicle. When adult human white blood cell nuclei or adult mouse thymocyte nuclei were injected into the GV of a *Xenopus* oocyte, the expression of an embryonic pluripotency marker gene, Oct4, was detected from transplanted nuclei. The silencing of a thymocyte-specific gene, Thy1, in transplanted thymocyte nuclei was also seen. Unexpectedly, some genes like MyoD are strongly activated in nuclei injected into oocytes (Biddle *et al.*, 2009). These results established that somatic cell nuclear transfers to meiotic prophase oocytes showed selective gene activation and repression at the transcriptional level.

Using mammalian nuclei, for which there is no background of mammalian transcripts in *Xenopus* oocytes, it was possible to determine many precise details of this system (Halley-Stott *et al.*, 2010). There can be several hundred re-initiations of transcription per gene per day. Nearly all transplanted nuclei undergo transcriptional activation and are colonized by B4 linker histone (Jullien *et al.*, 2010). The efficiency of transcriptional activation is seen in each injected nucleus, when up to 500 nuclei are injected. This experimental system can be used to observe epigenetic changes that take place in nuclei transplanted to oocytes (Murata *et al.*, 2010).

The high efficiency and long duration of transcription by nuclei transplanted to oocytes render this system "like a living test-tube." This is because any component of any size can be introduced into the oocyte and its effect on ongoing transcription seen, in a way that is not possible with nuclei introduced into extracts of cells *in vitro*.

Components Necessary for Reprogramming

For some genes, DNA methylation stabilizes a repressed state. In the case of Oct4, the oocyte is able to reverse this (Simonsson and Gurdon, 2004). The remarkable efficiency with which the oocyte reprograms somatic cell nuclei prompted investigators to identify the components of the reprogramming machinery. However, the oocyte contains numerous proteins accumulated during oogenesis, and selecting good candidate oocyte reprogramming factors (CORFs) was challenging. The idea was to find CORFs matching two criteria: the CORF should (1) be present in the oocyte in large amounts; and (2) be able to robustly access the transplanted nuclei. The oocyte contains a large storage of an oocyte-specific histone variant B4, which is robustly incorporated into the nuclei of somatic cells incubated in extracts prepared from eggs (Dimitrov and Wolffe, 1996). It has been also shown that mammalian homologue of B4, histone H1foo, becomes incorporated into somatic cells after nuclear transfer (Teranishi *et al.*, 2004). Therefore, it has been hypothesized that this linker histone variant may be involved in reprogramming of somatic cell nuclei after nuclear transfer to oocytes. Indeed, *Xenopus* oocytes which have been injected with an antibody specifically blocking histone B4, or which have been loaded with a dominant-negative version of histone B4, could no longer reprogram injected somatic cell nuclei (Jullien *et al.*, 2010). Additionally, H1foo has recently been demonstrated to act as a global epigenomic modulator that not only decondenses chromatin but can also, by itself, maintain cells in an undifferentiatied state (Hayakawa *et al.*, 2012). Interestingly, B4 linker histone turned out to be not the only histone variant necessary for reprogramming of somatic cells in the frog oocyte. As mentioned previously, *Xenopus* oocytes and, subsequently, developing embryos also contain a H3.3 histone variant, which was implied in the maintenance of epigenetic memory in cloned frog

embryos (Ng and Gurdon, 2008). A homologue of this histone variant has been shown to be essential for the remodeling of the sperm nucleus at fertilization in flies (Loppin *et al.*, 2005). Finally, it has been demonstrated that oocytes in which H3.3 histone incorporation was inhibited were unable to reprogram transplanted somatic cell nuclei (Jullien *et al.*, 2012). These experiments showed that oocyte-specific linker histones B4 and H3.3 are necessary components of the reprogramming machinery utilized by the frog oocyte.

Unexpectedly, it turned out that histone variants are not the only necessary components needed for reprogramming in the frog oocyte. Actin is a structural protein that is abundant in the oocyte, and was thought to be necessary to create a physical scaffold for the large frog oocyte GV (Bohnsack *et al.*, 2006). It has now been shown that filamentous actin becomes enriched in somatic nuclei upon transplantation to the GV (Miyamoto *et al.*, 2011), suggesting that nuclear actin may play a role in the reprogramming of transplanted somatic cells. Indeed, inhibition of nuclear actin polymerization prevented reactivation of pluripotency genes in transplanted mammalian nuclei (Miyamoto *et al.*, 2011), confirming the role of this structural protein as a necessary component of the reprogramming machinery.

It is interesting to speculate that the oocyte-based linker histone and other decondensation/remodeling factors in metaphase II oocytes may play a role in loosening chromatin following somatic cell nuclear transfer, thereby providing reprogramming factors' access to epigenetically repressed genes (Gurdon and Wilmut, 2011). Whether these chromatin-loosening factors can play a role in augmenting induced pluripotent stem cell-based reprogramming is yet to be determined.

Resistance to Reprogramming

Despite the remarkable efficiency with which the *Xenopus* oocyte can reprogram transplanted nuclei of somatic cells, even from distantly related species, it is known that the extent of transcriptional reprogramming varies depending on the cell types used as nuclear donors. Since the oocyte components are always the same, this suggests that some cells are more resistant than the others towards reprogramming. A good illustration of this phenomenon is an experiment in which four different mouse somatic cell types have been used as donors for nuclear transplantation into *Xenopus* GVs: C2C12 cells (muscle lineage precursors), MEFs (mouse embryonic fibroblasts), ES cells (embryonic stem cells) and RA-ES cells (retinoic acid-differentiated ES cells) (Halley-Stott *et al.*, 2010). Ninety-six hours after nuclear transfer to the GV, RNA from such nuclear transfer oocytes was isolated, used for the generation of a cDNA template and, subsequently, a quantitative PCR analysis was performed to assess the number of transcripts of the pluripotency genes *Oct4* and *Sox2* that were synthesized by each cell type in the oocyte. ES cells served as an internal reference for the extent of gene reactivation in this experiment, because they naturally express these pluripotency genes. Interestingly, it seems that each cell type has a different ability to reactivate these genes (Table 14.1). For example, C2C12 cells reactivated *Oct4* five times better than MEFs, whereas MEFs reactivated *Sox2* ten times better than C2C12 cells. However, both C2C12 and MEFs were far less efficient in reactivating these genes than RA-ES: C2C12 and MEFs reactivated *Oct4* and *Sox2* only to 0.3–3.33% of transcript levels synthesized in ES cells, whereas RA-ES reactivated *Oct4* to the full ES cell level and *Sox2* to 83.3% of ES cell level (Halley-Stott *et al.*, 2010).

TABLE 14.1 Different Cell Types have a Different Ability to Reprogram *Oct4* and *Sox2* after the Nuclear Transfer to *Xenopus* Oocytes

	Number of Transcripts per Nucleus (% of ES Cell Level) in Each Cell after 96 Hours			
	C2C12	MEFs	RA-ES	ES
Oct4	20 (2.6%)	4 (0.5%)	800 (106%)	750 (100%)
Sox2	30 (0.3%)	300 (3.33%)	7500 (83.3%)	9000 (100%)

What is the molecular basis of the resistance to reprogramming? In the example above, all the cells used were murine; therefore, the genetic incompatibility between various cell types and frog oocytes should be the same. To dissect the mechanisms of the resistance to reprogramming, two different mouse cell types with different capabilities of reactivating silenced X chromosomes have been used (Pasque *et al.*, 2011). It has been shown that the inactive X of mouse post-implantation-derived epiblast stem cells (EpiSCs) can be reactivated after nuclear transfer to the GV of the *Xenopus* oocyte, whereas the inactive X of MEFs cannot be reactivated by this procedure. Immunofluorescent labeling has revealed that MEFs, which cannot reactivate the inactive X, accumulate a histone variant macroH2A in their nuclei, in contrast to EpiSCs nuclei. The removal of macroH2A from MEFs significantly improved the reactivation of inactive X from their nuclei after nuclear transfer. These results show that the presence of macroH2A histone confers some resistance to X-chromosome reprogramming in MEFs. MacroH2A is just one example of epigenetic change that can be responsible for the resistance to reprogramming. However, there are many other epigenetic mechanisms that can restrict reprogramming, including DNA methylation or silencing by post-translational histone modifications. Furthermore, it is likely that cells displaying a high resistance to reprogramming may utilize a combination of various epigenetic mechanisms to confer the stability of the silent state.

It is likely that cells displaying a high resistance to reprogramming may utilize a combination of various epigenetic mechanisms to confer the stability of the silent state.

Further Reading

A number of reviews have appeared at intervals, where more extensive references can be found. These include Kikyo and Wolffe (2000) and Di Berardino (2006).

ACKNOWLEDGEMENTS

J. B. Gurdon and M. Teperek-Tkacz acknowledge the core funding provided by the Wellcome Trust (092096) and CRUK (C6946/A14492).

REFERENCES

Biddle, A., Simeoni, I., Gurdon, J.B., 2009. *Xenopus* oocytes reactivate muscle gene transcription in transplanted somatic nuclei independently of myogenic factors. Development 136, 2695–2703.

Bohnsack, M.T., Stuven, T., Kuhn, C., Cordes, V.C., Gorlich, D., 2006. A selective block of nuclear actin export stabilizes the giant nuclei of *Xenopus* oocytes. Nat. Cell Biol. 8, 257–263.

Briggs, R., King, T.J., 1952. Transplantation of living nuclei from blastula cells into enucleated frogs' eggs. Proc. Natl. Acad. Sci. USA 38, 455–463.

Briggs, R., King, T.J., 1957. Changes in the nuclei of differentiating endoderm cells as revealed by nuclear transplantation. J. Morphol. 100, 269–312.

Briggs, R., King, T.J., 1960. Nuclear transplantation studies on the early gastrula (*Rana pipiens*). I. Nuclei of presumptive endoderm. Dev. Biol. 2, 252–270.

Briggs, R., Signoret, J., Humphrey, R.R., 1964. Transplantation of nuclei of various cell types from neurulae of the Mexican axolotl (*Ambystoma mesicanum*). Dev. Biol., 233–246.

Brun, R., 1973. Mammalian cells in *Xenopus* eggs. Nat. New Biol. 243, 26–27.

Byrne, J.A., Simonsson, S., Gurdon, J.B., 2002. From intestine to muscle: nuclear reprogramming through defective cloned embryos. Proc. Natl. Acad. Sci. USA 99, 6059–6063.

Byrne, J.A., Simonsson, S., Western, P.S., Gurdon, J.B., 2003. Nuclei of adult mammalian somatic cells are directly reprogrammed to oct-4 stem cell gene expression by amphibian oocytes. Curr. Biol. 13, 1206–1213.

Chan, A.P., Gurdon, J.B., 1996. Nuclear transplantation from stably transfected cultured cells of *Xenopus*. Intl. J. Devel. Biol. 40, 441–451.

De Robertis, E.M., Gurdon, J.B., 1977. Gene activation in somatic nuclei after injection into amphibian oocytes. Proc. Natl. Acad. Sci. USA 74, 2470–2474.

Di Berardino, M.A., 2006. Origin and progress of nuclear transfer in non-mammalian animals. Methods Mol. Biol. 348, 3–32.

Di Berardino, M.A., King, T.J., 1967. Development and cellular differentiation of neural nuclear-transplants of known karyotype. Dev. Biol. 15, 102–128.

Dimitrov, S., Wolffe, A.P., 1996. Remodeling somatic nuclei in *Xenopus laevis* egg extracts: molecular mechanisms for the selective release of histones H1 and H1(0) from chromatin and the acquisition of transcriptional competence. EMBO J. 15, 5897–5906.

Driesch, H., 1894. Analytische Theorie de organischen Entwicklung. W. Engelmann, Leipzig.

Elsdale, T.R., Gurdon, J.B., Fischberg, M., 1960. A description of the technique for nuclear transplantation in *Xenopus laevis*. J. Embryol. Exp. Morphol. 8, 437–444.

Gallien, C.L., Aimar, C., Guillet, F., 1973. Nucleocytoplasmic interactions during ontogenesis in individuals obtained by intra- and interspecific nuclear transplantation in the genus *Pleurodeles* (urodele amphibian). Dev. Biol. 33 (1), 154–170.

Graham, C.F., Arms, K., Gurdon, J.B., 1966. The induction of DNA synthesis by frog egg cytoplasm. Dev. Biol. 14, 349–381.

Gurdon, J.B., 1960. The developmental capacity of nuclei taken from differentiating endoderm cells of *Xenopus laevis*. J. Embryol. Exp. Morphol. 8, 505–526.

Gurdon, J.B., 1962a. The developmental capacity of nuclei taken from intestinal epithelium cells of feeding tadpoles. J. Embryol. Exp. Morphol. 10, 622–640.

Gurdon, J.B., 1962b. Multiple genetically-identical frogs. J. Hered. 53, 4–9.

Gurdon, J.B., 1968. Changes in somatic cell nuclei inserted into growing and maturing amphibian oocytges. J. Embryol. Exp. Morphol. 20, 401–414.

Gurdon, J.B., 1976. Injected nuclei in frog oocytes: fate, enlargement, and chromatin dispersal. J. Embryol. Exp. Morphol. 36, 523–540.

Gurdon, J.B., 1977. Methods for nuclear transplantation in Amphibia In: Stein, G. Stein, J. (Eds.), Methods in Cell Biology, 16 Academic Press, New York, NY, pp. 125–139.

Gurdon, J.B., 1986. Nuclear transplantation in eggs and oocytes. J. Cell Sci. (Suppl.) 4, 287–318.

Gurdon, J.B., Uehlinger, V., 1966. "Fertile" intestine nuclei. Nature 210, 1240–1241.

Gurdon, J.B., Wilmut, I., 2011. Nuclear transfer to eggs and oocytes. In Germ Cells. Cold Spring Harbor Perspect. Bio. 3 (6), 185–198.

Gurdon, J.B., Laskey, R.A., Reeves, O.R., 1975. The developmental capacity of nuclei transplanted from keratinized skin cells of adult frogs. J. Embryol. Exp. Morphol. 34, 93–112.

Halley-Stott, R.P., Pasque, V., Astrand, C., Miyamoto, K., Simeoni, I., Jullien, J., et al., 2010. Mammalian nuclear transplantation to Germinal Vesicle stage *Xenopus* oocytes – a method for quantitative transcriptional reprogramming. Methods 51, 56–65.

Hayakawa, K., Ohgane, J., Tanaka, S., Yagi, S., Shiota, K., 2012. Oocyte-specific linker histone H1foo is an epigenomic modulator that decondenses chromatin and impairs pluripotency. Epigenetics 7 (9), 1029–1036.

Hennen, S., 1970. Influence of spermine and reduced temperature on the ability of transplanted nuclei to promote normal development in eggs of *Rana pipiens*. Proc. Natl. Acad. Sci. USA 66, 630–637.

Hochedlinger, K., Jaenisch, R., 2002. Monoclonal mice generated by nuclear transfer from mature B and T donor cells. Nature 415, 1035–1038.

Jullien, J., Astrand, C., Halley-Stott, R.P., Garrett, N., Gurdon, J.B., 2010. Characterization of somatic cell nuclear reprogramming by oocytes in which a linker histone is required for pluripotency gene reactivation. Proc. Natl. Acad. Sci. USA 107, 5483–5488.

Jullien, J., Astrand, C., Szenker, E., Garrett, N., Almouzni, G., Gurdon, J., 2012. HIRA dependent H3.3 deposition is required for transcriptional reprogramming following nuclear transfer to *Xenopus* oocytes. Epigenetics Chromatin 5, 17.

Kikyo, N., Wolffe, A.P., 2000. Reprogramming nuclei: insights from cloning, nuclear transfer and heterokaryons. J. Cell Sci. 113 (Pt 1), 11–20.
Kim, K., Doi, A., Wen, B., Ng, K., Zhao, R., Cahan, P., et al., 2010. Epigenetic memory in induced pluripotent stem cells. Nature 467 (7313), 285–290.
Kim, K., Zhao, R., Doi, A., Ng, K., Unternaehrer, J., Cahan, P., et al., 2011. Donor cell type can influence the epigenome and differential potential of human induced pluripotent stem cells. Nat. Biotechnol. 29 (12), 1117–1119.
King, T.J., Briggs, R., 1956. Serial transplantation of embryonic nuclei. Cold Spring Harb. Symp. Quant. Biol. 21, 271–290.
Laskey, R.A., 1989. S phase of the cell cycle. Science 246 (4930), 609–614.
Laskey, R.A., 2005. Solving mysteries of DNA replication and frog cloning. Cell 123 (5), 760–762.
Laskey, R.A., Gurdon, J.B., 1970. Genetic content of adult somatic cells tested by nuclear transplantation from cultured cells. Nature 228, 1332–1334.
Lemaitre, J.M., Danis, E., Pasero, P., Vassetzky, Y., Méchali, M., 2005. Mitotic remodelling of the replicon and chromosome structure. Cell 123 (5), 787–801.
Lohka, M.J., Masui, Y., 1983. Formation *in vitro* of sperm pronuclei and mitotic chromosomes induced by amphibian ooplasmic components. Science 220 (4598), 719–721.
Loppin, B., Bonnefoy, E., Anselme, C., Laurencon, A., Karr, T.L., Couble, P., 2005. The histone H3.3 chaperone HIRA is essential for chromatin assembly in the male pronucleus. Nature 437, 1386–1390.
McKinnell, R.G., 1978. Cloning: Nuclear Transplantation in Amphibia. University of Minnesota Press, Minneapolis, MN.
McKinnell, R.G., Deggins, B.A., Labat, D.D., 1969. Transplantation of pluripotential nuclei from triploid frog tumors. Science 169, 394–396.
Miyamoto, K., Pasque, V., Jullien, J., Gurdon, J.B., 2011. Nuclear actin polymerization is required for transcriptional reprogramming of Oct4 by oocytes. Genes Dev. 25, 946–958.
Murata, K., Kouzarides, T., Bannister, A.J., Gurdon, J.B., 2010. Histone H3 lysine 4 methylation is associated with the transcriptional reprogramming efficiency of somatic nuclei by oocytes. Epigenetics Chromatin 3, 4.
Narbonne et al., 2011. doi: 10.1371/journal.pbio.1001197.
Narbonne, P., Gurdon, J.B., 2012. Amphibian interorder nuclear transfer embryos reveal conserved embryonic gene transcription, but deficient DNA replication or chromosome segregation. International Journal of Developmental Biology, special issue on Female Germ Cells in Development and Cancer (in press).
Ng, R.K., Gurdon, J.B., 2008. Epigenetic memory of an active gene state depends on histone H3.3 incorporation into chromatin in the absence of transcription. Nat. Cell Biol. 10, 102–109.
Pasque, V., Gillich, A., Garrett, N., Gurdon, J.B., 2011. Histone variant macroH2A confers resistance to nuclear reprogramming. EMBO J. 30, 2373–2387.
Polo, J.M., Liu, S., Figueroa, M.E., Kulalert, W., Eminli, S., Tan, K.Y., et al., 2010. Cell type of origin influences the molecular and functional properties of mouse induced pluripotent stem cells. Nat. Biotech. 28, 848–855.
Rauber, A., 1886. Personaltheil und germinaltheil des individuum. Zool. Anz. 9, 166–171.
Ray-Gallet, D., Quivy, J.P., Scamps, C., Martini, E.M., Lipinski, M., Almouzni, G., 2002. HIRA is critical for a nucleosome assembly pathway independent of DNA synthesis. Mol. Cell 9 (5), 1091–1100.
Simnett, J.D., 1964. The development of embryos derived from the transplantation of neural ectoderm cell nuclei in *Xenopus laevis*. Dev. Biol. 10, 467–486.
Simonsson, S., Gurdon, J.B., 2004. doi: 10.1038/ncb1176.
Smith, L.D., 1965. Transplantation of the nuclei of primordial somatic cells into enucleated eggs of *Rana pipiens*. Proc. Natl. Acad. Sci. 54, 101–107.
Spemann, H., 1938. Embryonic Development and Induction. Yale University Press, New Haven, CT.
Tagami, H., Ray-Gallet, D., Almouzni, G., Nakatani, Y., 2004. Histone H3.1 and H3.3 complexes mediate nucleosome assembly pathways dependent or independent of DNA synthesis. Cell 116 (1), 51–61.
Teranishi, T., Tanaka, M., Kimoto, S., Ono, Y., Miyakoshi, K., Kono, T., et al., 2004. Rapid replacement of somatic linker histones with the oocyte-specific linker histone H1foo in nuclear transfer. Dev. Biol. 266, 76–86.
Wakefield, L., Gurdon, J.B., 1983. Cytoplasmic regulation of 5S RNA genes in nuclear-transplant embryos. EMBO J. 2, 1613–1619.
Weismann, A., 1893. The Germ-Plasm: A Theory of Heredity. Scribner's, New York, NY.
Wilmut, I., Schnieke, A.E., McWhir, J., Kind, A.J., Campbell, K.H.S., 1997. Viable offspring derived from fetal and adult mammalian cells. Nature 385, 810–813.

Appendix 4

Research article: 'Nuclear Wave1 is required for reprogramming transcription in oocytes and for normal development.'

Miyamoto K., Teperek M., Yusa K., Allen GE., Bradshaw CR., Gurdon JB., Science, 2013

References and Notes

1. T. Alerstam, A. Hedenstrom, S. Åkesson, *Oikos* **103**, 247–260 (2003).
2. T. Alerstam, *Science* **313**, 791–794 (2006).
3. S. Åkesson, A. Hedenstrom, *Bioscience* **57**, 123–133 (2007).
4. T. Fransson *et al.*, *Nature* **414**, 35–36 (2001).
5. M. Wikelski *et al.*, *Nature* **423**, 704 (2003).
6. M. Liedvogel, S. Åkesson, S. Bensch, *Trends Ecol. Evol.* **26**, 561–569 (2011).
7. C. Mettke-Hofmann, E. Gwinner, *Proc. Natl. Acad. Sci. U.S.A.* **100**, 5863–5866 (2003).
8. N. Chernetsov, P. Berthold, U. Querner, *J. Exp. Biol.* **207**, 937–943 (2004).
9. K. Thorup *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **104**, 18115–18119 (2007).
10. C. Palacín, J. C. Alonso, J. A. Alonso, M. Magana, C. A. Martin, *J. Avian Biol.* **42**, 301–308 (2011).
11. L. Conradt, T. J. Roper, *Trends Ecol. Evol.* **20**, 449–456 (2005).
12. I. D. Couzin, J. Krause, N. R. Franks, S. A. Levin, *Nature* **433**, 513–516 (2005).
13. V. Guttal, I. D. Couzin, *Proc. Natl. Acad. Sci. U.S.A.* **107**, 16172–16177 (2010).
14. K. L. Jones *et al.*, *Conserv. Biol.* **16**, 789–799 (2002).
15. R. P. Urbanek *et al.*, *Proc. North Am. Crane Workshop* **9**, 213–223 (2005).
16. R. P. Urbanek, L. E. A. Fondow, S. E. Zimorski, M. A. Wellington, M. A. Nipper, *Bird Conserv. Int.* **20**, 43–54 (2010).
17. E. Schüz, *Die Vogelwarte* **2**, 63–78 (1949).
18. Materials and methods are available as supplementary materials on *Science* Online.
19. A. J. Wilson *et al.*, *J. Anim. Ecol.* **79**, 13–26 (2010).
20. E. Schüz, *Die Vogelwarte* **1**, 239–253 (1950).
21. H. Mouritsen, O. N. Larsen, *J. Exp. Biol.* **201**, 2927–2934 (1998).
22. E. Kuyt, *Aerial radio-tracking of whooping cranes migrating between Wood Buffalo National Park and Aransas National Wildlife Refuge, 1981–84*. Canadian Wildlife Service Occasional Paper no. 74 (Canadian Wildlife Service, Ottawa, 1992), pp. 1–53.
23. T. Alerstam, *J. Exp. Biol.* **199**, 9–19 (1996).
24. K. Thorup, T. Alerstam, M. Hake, N. Kjellén, *Proc. R. Soc. London Ser. B Biol. Sci.* **270** (suppl. 1), S8–S11 (2003).
25. I. Newton, *J. Ornithol.* **147**, 146–166 (2006).
26. A. Flack, B. Pettit, R. Freeman, T. Guilford, D. Biro, *Anim. Behav.* **83**, 703–709 (2012).
27. Whooping Crane Eastern Partnership, Technical Database; www.bringbackthecranes.org/technicaldatabase/index.html.
28. S. J. Converse, J. A. Royle, R. P. Urbanek, *J. Ornithol.* **152** (suppl. 2), 561–572 (2012).
29. J. W. Chapman *et al.*, *Curr. Biol.* **21**, R861–R870 (2011).
30. R. B. O'Hara, J. M. Cano, O. Ovaskainen, C. Teplitsky, J. S. Alho, *J. Evol. Biol.* **21**, 949–957 (2008).
31. L. H. Damgaard, *J. Anim. Sci.* **85**, 1363–1368 (2007).
32. P. Waldmann, *Evolution* **63**, 1640–1643 (2009).
33. R. B. O'Hara, *Ann. Zool. Fenn.* **46**, 124–137 (2009).

Acknowledgments: T.M. and W.F.F. were supported by NSF Advances in Biological Informatics award 1062411, T.M. and R.B.O. were supported by the Landes-Offensive zur Entwicklung Wissenschaftlich-ökonomischer Exzellenz programme (Hesse, Germany), and T.M. was supported by the Robert Bosch Foundation. Data were provided by the Whooping Crane Eastern Partnership and K. Jones and are deposited in the Dryad Repository: http://doi.org/10.5061/dryad.1r0f7. We thank F. Hailer, S. Servanty, E. Grant, J. Calabrese, R. Reynolds, S. Via, E. L. Neuschulz, C. Rushing, D. Bennu, and three anonymous reviewers for helpful comments and discussions.

Supplementary Materials

www.sciencemag.org/cgi/content/full/341/6149/999/DC1
Materials and Methods
Fig. S1
References (*27–33*)

27 February 2013; accepted 1 August 2013
10.1126/science.1237139

Nuclear Wave1 Is Required for Reprogramming Transcription in Oocytes and for Normal Development

Kei Miyamoto,[1,2]* Marta Teperek,[1,2] Kosuke Yusa,[3] George E Allen,[1,2] Charles R Bradshaw,[1,2] J. B. Gurdon[1,2]*

Eggs and oocytes have a remarkable ability to induce transcription of sperm after normal fertilization and in somatic nuclei after somatic cell nuclear transfer. This ability of eggs and oocytes is essential for normal development. Nuclear actin and actin-binding proteins have been shown to contribute to transcription, although their mode of action is elusive. Here, we find that *Xenopus* Wave1, previously characterized as a protein involved in actin cytoskeleton organization, is present in the oocyte nucleus and is required for efficient transcriptional reprogramming. Moreover, Wave1 knockdown in embryos results in abnormal development and defective *hox* gene activation. Nuclear Wave1 binds by its WHD domain to active transcription components, and this binding contributes to the action of RNA polymerase II. We identify Wave1 as a maternal reprogramming factor that also has a necessary role in gene activation in development.

Eggs and oocytes efficiently reprogram transplanted somatic nuclei to an embryonic state (*1*, *2*). This reprogramming ability of eggs and oocytes toward somatic nuclei is believed to relate to their natural activity to activate sperm nuclei at fertilization. Reprogramming factors are synthesized and accumulated during egg formation and are especially enriched in the amphibian oocyte nucleus, named the germinal vesicle (GV) (*1*). GVs also contain necessary factors for embryonic development (*3*, *4*). It is unclear what kinds of GV factors are required for reprogramming and for normal development, and how they contribute to these fundamental processes. To identify such a maternal factor, we have developed a nuclear transfer assay; hundreds of mammalian somatic cell nuclei are injected into the GV of *Xenopus* oocytes, and these nuclei undergo not only continuous transcription of active genes but also transcriptional reactivation of somatically silenced embryonic genes within 2 days (*1*). This system thus provides a unique opportunity to identify maternal factors responsible for reprogramming the transcription of somatic nuclei.

Previously, we found an important role of oocyte nuclear actin in transcriptional reprogramming (*5*). Actin dynamics are regulated by actin-binding proteins (ABPs) (*6*). Increasing evidence suggests that nuclear ABPs play crucial roles in transcriptional activation (*7*–*9*). Therefore, we tested the roles of nuclear ABPs in reprogramming and development. The effect of overexpressing ABPs in recipient *Xenopus* oocytes on transcriptional reprogramming of *Pou5f1* (*Oct4*) was examined by reverse transcription quantitative polymerase chain reaction (RT-QPCR) (fig. S1A) (*10*). Overexpression of two ABPs Toca1 (*5*) and Rac1 significantly enhanced *Oct4* transcription from transplanted mouse C2C12 myoblast cell nuclei ($P < 0.01$) (fig.

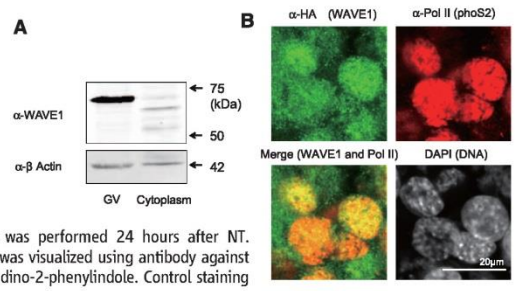


Fig. 1. Wave1 is present in the *Xenopus* oocyte nuclei and transplanted mouse nuclei. (A) Western blot analysis revealed that Wave1 is accumulated in the GV of the *Xenopus* oocyte. **(B)** Mouse C2C12 nuclei were injected into the GV overexpressing HA-NLS-Wave1 (fig. S1A). Immunofluorescence analysis was performed 24 hours after NT. HA-NLS-Wave1 localization was visualized using antibody against HA (αHA). DAPI, 4′,6′-diamidino-2-phenylindole. Control staining is shown in fig. S3.

[1]Wellcome Trust/Cancer Research UK Gurdon Institute, The Henry Wellcome Building of Cancer and Developmental Biology, Tennis Court Road, Cambridge CB2 1QN, UK. [2]Department of Zoology, University of Cambridge, Downing Street, Cambridge CB2 3EJ, UK. [3]Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, UK.

*Corresponding author. E-mail: k.miyamoto@gurdon.cam.ac.uk (K.M.); j.gurdon@gurdon.cam.ac.uk (J.B.G.)

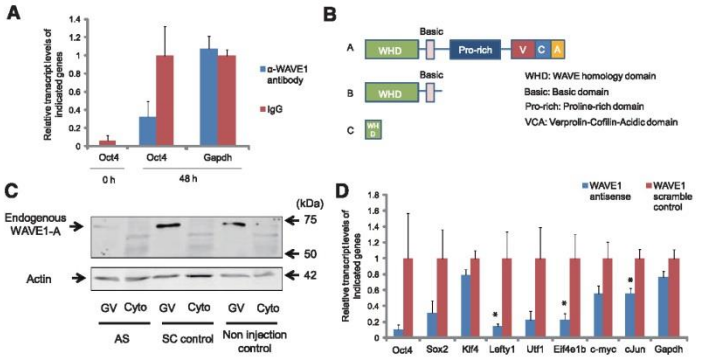


Fig. 2. Nuclear Wave1 is required for transcriptional reprogramming in oocytes. All error bars represent SEM. *$P < 0.05$. (**A**) Injection of antibody against Wave1 into GVs inhibits the transcriptional activation of *Oct4* in NT oocytes ($n = 12$). Immunoglobulin G was injected as control ($n = 13$). Relative fold increases of gene transcripts to control NT were measured by QPCR. (**B**) Three transcript variants expressed in *Xenopus* oocytes were identified (A, B, and C), and each protein domain is represented by different colors. (**C**) Knockdown of endogenous Wave1 proteins in the GV by antisense oligonucleotide injection (AS) as checked by Western blot. SC control means scrambled oligonucleotides–injected control. Cyto represents the oocyte cytoplasm. (**D**) Knockdown of Wave1 by AS inhibits transcriptional activation of many embryonic genes in NT oocytes (blue bars). $n = 6$.

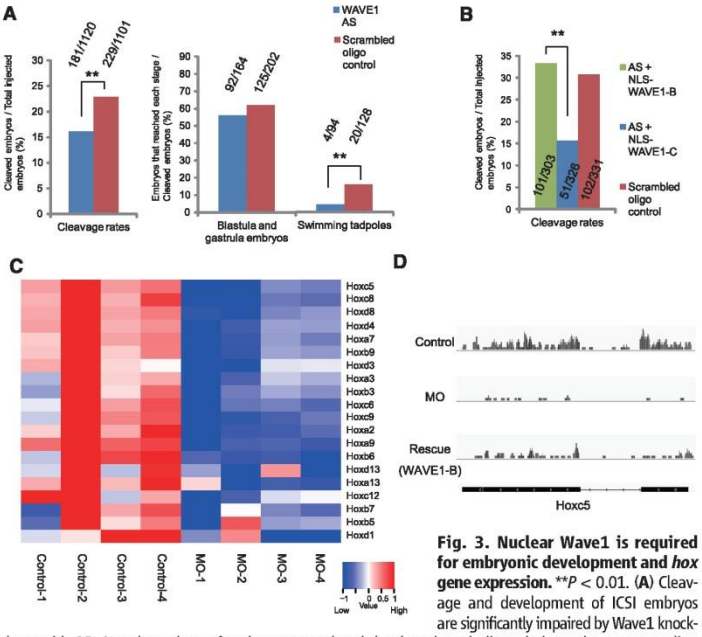


Fig. 3. Nuclear Wave1 is required for embryonic development and *hox* gene expression. **$P < 0.01$. (**A**) Cleavage and development of ICSI embryos are significantly impaired by Wave1 knockdown with AS. Actual numbers of embryos treated and developed are indicated above the corresponding graphs. (**B**) Abnormal cleavage caused by Wave1 knockdown is rescued by overexpression of NLS–Wave1-B, but not by NLS–Wave1-C. (**C**) Heat-map analysis shows lower expression of *hox* genes in Wave1-MO embryos (MO1-4) than control embryos. (High expression in red and low expression in blue.) (**D**) RNA-seq analysis shows that down-regulated *hoxc5* expression by Wave1 MO injection is partially rescued by overexpressing NLS–Wave1-B. Black boxes in *hoxc5* gene represent exons.

S1B). Downstream target proteins of Toca1 and RAC1 are Wiskott-Aldrich syndrome protein (WASP) and WASP family verprolin-homologous protein (WAVE) (*11*–*13*). WASP is involved in RNA polymerase II (Pol II)–mediated transcription and in transcriptional activation during T cell differentiation (*7*, *8*). However, N-wasp, the ubiquitous form of Wasp, is unlikely to be responsible for transcriptional reprogramming in *Xenopus* oocytes (supplementary text and fig. S1, C to E). Because the positive roles of N-WASP in transcription have been shown in cultured cells (*8*), oocytes may use different ABPs from somatic cells to regulate their productive transcription.

We tested the involvement of Wave in transcriptional reprogramming. WAVE plays a cytoplasmic role in actin reorganization as a downstream target of RAC (*13*). Rac1 protein is present in GVs, including in transplanted nuclei (fig. S2A). To our surprise, Wave1, one isoform of Wave enriched in brain, is also accumulated in the oocyte GV (Fig. 1A). We therefore focused on Wave1 as a candidate reprogramming factor. Little WAVE1 is detected in C2C12 nuclei before nuclear transfer (NT) (green color, fig. S2B). After NT, the accumulation of Wave1 was observed, especially in some somatic nuclei that showed extensive swelling (white arrow, fig. S2B), known to correlate with high transcriptional activity (*14*). Accordingly, localization of Wave1 and active RNA Pol II phosphorylated at serine 2 was examined. *Xenopus laevis* Wave1 tagged with a nuclear localization signal (NLS) and hemagglutinin (HA) (HA-NLS-Wave1) was expressed in GVs. A NLS was added to target Wave1 to the nucleus in order to focus on the nuclear role of Wave1. HA-NLS-Wave1 signals overlapped with active Pol II in transplanted nuclei (yellow color in the merged photo, Fig. 1B). Time course changes of HA-NLS-Wave1 and active Pol II in NT oocytes are shown in fig. S3. These results suggest that Wave1 is present in actively transcribing nuclei.

Subsequently, the importance of nuclear Wave1 in reprogramming embryonic genes in *Xenopus* NT oocytes was examined. We injected an antibody against Wave1 along with C2C12 murine myoblast nuclei to GVs. Wave1 antibodies inhibited *Oct4* activation (Fig. 2A). We then asked whether overexpressing Wave1 enhances transcriptional reprogramming in oocytes. For this experiment, we cloned three transcript variants of *wave1* (*wave1-A*, *-B* or *-C*) from *Xenopus* oocyte cDNAs (Fig. 2B and table S1), although the variant C was rarely detected. All three transcripts were expressed in GVs by mRNA injection (fig. S4A). Overexpression of Wave1-A and Wave1-B enhanced activation of embryonic genes (fig. S4B). Notably, Wave1-B significantly enhanced transcription from many of the genes examined, including housekeeping genes. In addition, we specifically knocked down oocyte Wave1 by antisense oligonucleotide (AS) injections (fig. S5A). ASs against *wave1* match sequences of both *wave1-A* and *wave1-B*. This method enabled us to stably decrease maternal *wave1* mRNAs (fig. S5B) and

to knock down Wave1 proteins (Fig. 2C). After injection of C2C12 nuclei into these Wave1-knocked down oocytes, activation of many embryonic genes was significantly impaired (Fig. 2D). We performed rescue experiments as depicted in fig. S5C using Wave1-B. Wave1-B expression successfully rescued activation of embryonic and oocyte-specific genes in oocytes ($P < 0.05$) (fig. S4C). Finally, we injected *Xenopus* sperm or 250 to 300 nuclei from XL177, a *Xenopus* cell line derived from tadpole epithelium, into oocytes having Wave1 knocked down to ask if Wave1 is important for transcriptional reprogramming in intra-species NT oocytes. Several embryonic genes that were identified as candidate target genes of Wave1 by RNA-sequencing (RNA-seq) (described below) were activated after NT (fig. S6A). Transcription from such genes was significantly inhibited by Wave1 knockdown (fig. S6B). These results indicate that nuclear Wave1 is necessary for efficient transcriptional reprogramming in *Xenopus* oocytes.

We next asked whether nuclear Wave1 enhances transcriptional reprogramming in the induced pluripotent stem (iPS) cell system. NLS-Wave was transiently expressed at various time points during mouse iPS cell production using the doxycycline-inducible system (*15*). Wave1 expression at day 6 or 9 enhanced activation of pluripotency genes (fig. S7). This result suggests that *Xenopus* nuclear Wave1 can have a positive effect on the mammalian reprogramming system.

To test the idea that a GV factor important for reprogramming is also required for development, the role of Wave1 in embryogenesis was examined. We developed a system in which the knockdown of maternal Wave1 is achieved during early embryonic development (*10*). We injected ASs into oocytes, incubated for 1 or 2 days, and in vitro matured to metaphase II (MII) eggs (fig. S8A) (*16*). These Wave1 knockdown eggs were used for intracytoplasmic sperm injection (ICSI) (fig. S8A). We obtained a healthy frog from ICSI embryos using in vitro matured eggs injected with control oligonucleotides (fig. S8B), which proved that this approach can support full-term development. Development of ICSI embryos was severely compromised by AS injection (Fig. 3A and fig. S8C). Moreover, the number of embryos that cleaved was decreased in AS-injected embryos (Fig. 3A), which suggested that Wave1 plays a role even before mid-blastula transition. These results agree with a previous report using bovine embryos (*17*). Abnormal cleavage was rescued by overexpressing NLS–Wave1-B (Fig. 3B). Wave1 inhibition in embryos was also accomplished by injection of morpholino oligonucleotides (MOs) into one-cell stage embryos (fig. S9A). This route toward knockdown is not effective at early cleavage stages, unlike AS injection to oocytes, and hence, defects in cleavage were not observed while development to the swimming tadpole is compromised (fig. S9B), in accord with the AS-knockdown results (Fig. 3A). These results indicate that nuclear Wave1 is required for normal embryonic development.

We then asked whether Wave1 is also involved in transcriptional activation in embryos. Transcriptional activation was analyzed by using Wave1 MO-injected embryos at the gastrula stage (stage12 to 12.5). RNA-seq analyses of Wave1 MO-injected embryos identified 964 misregulated transcripts among 23,560 total *Xenopus* transcripts (table S2) (420 down-regulated versus 544 up-regulated transcripts in Wave1 MO-injected embryos compared with control MO-injected embryos, false discovery rate < 0.05). Down-regulated genes identified by RNA-seq analyses were confirmed by RT-QPCR analyses (fig. S10). Pathway analysis of misregulated genes identified previously characterized functions of WAVE1, such as WASP family member and adenosine 3′,5′-monophosphate (cAMP) response element–binding protein (table S3) (*18*), which suggested the specific inhibition of Wave1 functions. A number of other pathways are newly identified (table S3). Remarkably, many *hox* genes were down-regulated in Wave1 MO-injected embryos (Fig. 3C). Down-regulation of *hox* genes transcription was partially rescued by expression of NLS–Wave1-B (Fig. 3D and fig. S11). In conclusion, nuclear Wave1 plays an important role in transcriptional activation in development. RNA-seq analyses also suggest that Wave1 seems to have specific downstream target genes rather than affecting global transcription in embryos.

We finally asked how nuclear Wave1 is involved in transcriptional activation. Results of our mass spectrometry experiments on Wave1-binding partners (supplementary text) prompted us to investigate whether nuclear Wave1 binds to the active transcription machinery. HA-NLS–Wave1-A, -B, and -C were expressed in embryos by mRNA injection, and extracts from these embryos were used for coimmunoprecipitation analysis. Active Pol II(phoS2) was bound to Wave1-A and -B, but not to Wave1-C (fig. S12A), in good agreement with transcriptional reprogramming experiments (fig. S4B). Wave1-A and -B were also coprecipitated with the C terminus of mixed-lineage leukemia (Mll) protein, which contains the SET domain and serves as a histone H3 lysine 4 (H3K4) methyltransferase. Although the protein band recognized by Mll antibody appeared lower than expected (*19*), we detected a specific band at the same size using human SET1 antibody (fig. S12B). We therefore concluded that the protein bound to Wave1 is a *Xenopus* SET domain–containing protein. These bindings were also observed when human WAVE1 was expressed (fig. S12C). This association was maintained even when the actin nucleation domain was removed from WAVE1 (*13*) (hWAVE1ΔVPH) (fig. S12C), which suggested that WAVE1 binding to the transcription apparatus is not through actin polymerization. We expressed different regions of the *Xenopus wave1* gene to identify a domain responsible for its binding to Pol II and the SET domain protein (Figs. 2B and 4A). The full Wave homology domain (WHD) was sufficient to mediate interaction with these

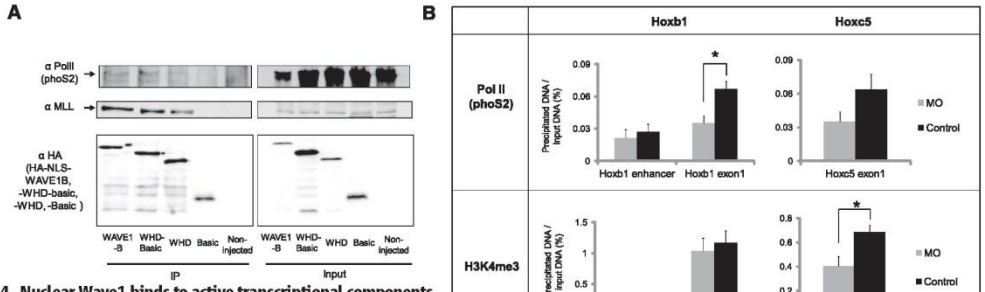


Fig. 4. Nuclear Wave1 binds to active transcriptional components through the WHD domain and is important for their activities in embryos. (**A**) Coimmunoprecipitation analysis in embryo extracts revealed that the WHD domain (Fig. 2B) is sufficient to bind to active Pol II and Mll. Five percent input was loaded. (**B**) Binding of elongating Pol II(phoS2) and H3K4me3 to *hoxb1* and *hoxc5* genes was examined by ChIP analyses. The levels of binding were compared in gastrula embryos with or without Wave1 MO injection (MO and Control, respectively). Relative fold increases of precipitated DNA over input DNA were measured by QPCR. $*P < 0.05$. $n = 4$ (Pol II) and 8 (H3K4me3). Error bars represent SEM.

transcription regulators (Fig. 4A). We next checked the binding of nuclear Wave1 to *hox* genes by chromatin immunoprecipitation (ChIP) analysis. Overexpressed NLS–Wave1-B was specifically enriched on the *hox* genes enhancer regions (fig. S13, A and B). Finally, we examined levels of active Pol II and H3K4 trimethylation (H3K4me3) on *hox* genes in Wave1 knockdown embryos. ChIP analysis revealed that binding of elongating Pol II was down-regulated on *hoxb* and *hoxc* genes, and H3K4me3 was also reduced on *hoxc* genes (Fig. 4B and fig. S14). In conclusion, nuclear Wave1 associates with active transcription machineries through the WHD domain and modulates their activities on *hox* genes.

Our results demonstrate a role of Wave1 in transcription in oocytes and embryos and provide in vivo evidence that an actin-binding protein plays an important nuclear role in embryonic development. This nuclear function of Wave1 is due to its N terminus WHD domain (Fig. 4). It is noteworthy that the WHD domain mediates protein complex formation with kinase proteins for actin reorganization in the cytoplasm (*18*), which suggested that Wave1 can have differential binding partners between the cytoplasm and nucleus. This explains why Wave1 can have a different function in nuclei from that in the cytoplasm.

A long-standing question in developmental biology is which oocyte factors stored in the GV can contribute to embryonic development and reprogramming. Our study identifies a maternal reprogramming protein that is required for normal early development and shows its mechanism of action.

References and Notes

1. J. Jullien, V. Pasque, R. P. Halley-Stott, K. Miyamoto, J. B. Gurdon, *Nat. Rev. Mol. Cell Biol.* **12**, 453–459 (2011).
2. K. Kim *et al.*, *Nature* **467**, 285–290 (2010).
3. S. Gao *et al.*, *Biol. Reprod.* **67**, 928–934 (2002).
4. S. Ogushi *et al.*, *Science* **319**, 613–616 (2008).
5. K. Miyamoto, V. Pasque, J. Jullien, J. B. Gurdon, *Genes Dev.* **25**, 946–958 (2011).
6. K. G. Campellone, M. D. Welch, *Nat. Rev. Mol. Cell Biol.* **11**, 237–251 (2010).
7. M. D. Taylor *et al.*, *Sci. Transl. Med.* **2**, 37ra44 (2010).
8. X. Wu *et al.*, *Nat. Cell Biol.* **8**, 756–763 (2006).
9. C. Ferrai *et al.*, *Mol. Biol. Cell* **20**, 3543–3551 (2009).
10. Materials and methods are available as supplementary materials on *Science* Online.
11. R. Fricke *et al.*, *Curr. Biol.* **19**, 1429–1437 (2009).
12. H. Y. Ho *et al.*, *Cell* **118**, 203–216 (2004).
13. H. Miki, S. Suetsugu, T. Takenawa, *EMBO J.* **17**, 6932–6941 (1998).
14. J. B. Gurdon, *J. Embryol. Exp. Morphol.* **20**, 401–414 (1968).
15. R. P. Koche *et al.*, *Cell Stem Cell* **8**, 96–105 (2011).
16. A. M. Hulstrand, P. N. Schneider, D. W. Houston, *Methods* **51**, 75–81 (2010).
17. V. Y. Rawe, C. Payne, C. Navara, G. Schatten, *Dev. Biol.* **276**, 253–267 (2004).
18. T. Takenawa, S. Suetsugu, *Nat. Rev. Mol. Cell Biol.* **8**, 37–48 (2007).
19. A. Yokoyama, I. Kitabayashi, P. M. Ayton, M. L. Cleary, M. Ohki, *Blood* **100**, 3710–3718 (2002).

Acknowledgments: We are grateful to D. Simpson for preparing eggs and oocytes; to A. Meissner for dox-inducible mouse embryo fibroblasts; to S. Suetsugu, T. Takenawa, and H. Miki for hWAVE1 constructs; and to Gurdon laboratory members for their critical reading of the manuscript. K.M. is a Research Fellow at Wolfson College and is supported by the Herchel Smith Postdoctoral Fellowship and the Great Britain Sasakawa Foundation. M.T. is funded by the Wellcome Trust (WT089613). K.Y. was supported by Wellcome Trust (WT077187). Gurdon laboratory is supported by grants from the Wellcome Trust (101050/Z/13/Z) and Medical Research Council (G1001690/1) to J.B.G. RNA-seq data were deposited in NCBI (SRA091722).

Supplementary Materials

www.sciencemag.org/cgi/content/full/341/6149/1002/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S14
Table S1 to S4
References (*20*–*36*)

10 May 2013; accepted 5 August 2013
10.1126/science.1240376

Pou5f1 Transcription Factor Controls Zygotic Gene Activation In Vertebrates

Manuel Leichsenring,[1]* Julia Maes,[1] Rebecca Mössner,[1] Wolfgang Driever,[1,2]† Daria Onichtchouk[1,2]†

The development of multicellular animals is initially controlled by maternal gene products deposited in the oocyte. During the maternal-to-zygotic transition, transcription of zygotic genes commences, and developmental control starts to be regulated by zygotic gene products. In *Drosophila*, the transcription factor Zelda specifically binds to promoters of the earliest zygotic genes and primes them for activation. It is unknown whether a similar regulation exists in other animals. We found that zebrafish Pou5f1, a homolog of the mammalian pluripotency transcription factor Oct4, occupies SOX-POU binding sites before the onset of zygotic transcription and activates the earliest zygotic genes. Our data position Pou5f1 and SOX-POU sites at the center of the zygotic gene activation network of vertebrates and provide a link between zygotic gene activation and pluripotency control.

In early metazoan development, the zygotic genome is not immediately transcribed; instead, factors expressed during oogenesis from the maternal genome control development. The controlled synchronous onset of expression of the earliest large wave of zygotic transcripts is termed zygotic genome activation (ZGA). In vertebrates, ZGA is triggered by the change of nuclear-to-cytoplasmic ratio during early cleavage stages, as well as by other genome-level mechanisms (*1*). ZGA is one component of the maternal-to-zygotic transition (MZT) period, which also includes the gradual degradation of maternal transcripts (*2*). In *Drosophila*, the transcription factor Zelda (Zld) selectively primes the earliest zygotic genes for activation at MZT (*3*–*8*) and controls ZGA. Because no homologs of Zld have been reported outside the insect clade, it is unclear whether similar specific control of transcription onset exists in early vertebrate development.

In zebrafish, zygotic transcription starts during the midblastula transition (MBT), after the 10th cell division (*9*). Pou5f1 (also named Pou2 or Pou5f3; www.zfin.org), Nanog, and the functionally redundant SoxB1 group of transcription factors (Sox2, Sox3, Sox19a, and Sox19b) are ubiquitously present in the zebrafish egg and early embryo (*10*–*13*). Their homologs are considered core transcription factors of the mammalian embryonic stem (ES) cell state. In mammalian ES cells, the Pou5f1-Sox2 complex cooperatively assembles to target gene promoters containing bipartite SOX-POU binding sites (*14*–*16*). In zebrafish, loss of Pou5f1 function leads to severe disturbance of developmental progress immediately after MBT (*17*–*20*). Time-resolved expression analysis of MZ*spg* mutants, devoid of both maternal and zygotic Pou5f1 function, revealed delays and desynchronization in expression of hundreds of genes (*21*), suggesting a role for Pou5f1 in temporal control of development. Here, we show that Pou5f1 in zebrafish selectively primes the earliest zygotic genes for activation, providing functionality similar to *Drosophila* Zelda (*3*–*8*).

To detect in vivo Pou5f1 and Sox2 chromatin binding events, we performed chromatin immunoprecipitation followed by parallel sequencing (ChIP-seq) and identified 7747 Pou5f1-bound regions at the pre-MBT 512-cell stage [2.75 hours postfertilization (hpf)], as well as 6670 Pou5f1-bound and 5924 Sox2-bound regions at the post-MBT stage (5 hpf; figs. S1 and S2 and table S1). Post- and pre-MBT Pou5f1 and Sox2 binding sites tended to colocalize (Fig. 1, A and D). To determine whether Pou5f1- or Sox2-bound genes correspond to targets of Pou5f1 and SoxB1 transcriptional regulation, we compared the list of genes bound by Pou5f1 or Sox2 within 20 kb upstream of their transcription start site (TSS) with lists of genes differentially expressed in embryos deficient for the respective transcription factors (*13*, *21*). More than 100 genes activated by Pou5f1 and SoxB1 were directly linked to Pou5f1 and

[1]Developmental Biology Unit, Institute Biology I, Faculty of Biology, Albert-Ludwigs-University, 79085 Freiburg, Germany. [2]Center for Biological Signaling Studies (BIOSS), Albert-Ludwigs-University, 79085 Freiburg, Germany.

*Present address: Institute of Molecular Biology, 55128 Mainz, Germany.

†Corresponding author. E-mail: daria.onichtchouk@biologie.uni-freiburg.de (D.O.); driever@biologie.uni-freiburg.de (W.D.)