CHARACTERIZATION OF THE GENETIC AND EPIGENETIC STATUS OF FERTILE AND INFERTILE MEN

by

Saher Sue Hammoud

A dissertation submitted to the faculty of
The University of Utah
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Physiology

The University of Utah

May 2011

The University of Utah Graduate School

STATEMENT OF DISSERTATION APPROVAL

The dissertation of **Saher Sue Hammoud**

has been approved by the following supervisory committee members:

Douglas T. Carrell	, Chair	**March 1, 2011** Date Approved
Bradley R. Cairns	, Member	**March 1, 2011** Date Approved
David Jones	, Member	**March 1, 2011** Date Approved
Mary T. Lucero	, Member	**March 1, 2011** Date Approved
Michael C. Sanguinetti	, Member	**March 1, 2011** Date Approved

and by **F. Edward Dudek**, Chair of the Department of **Physiology**

and by Charles A. Wight, Dean of The Graduate School.

ABSTRACT

The incidence of infertility is rising and estimated to affect one in six couples, yet a significant number of couples have an undefined cause for their infertility. The search for causes of male infertility has been the focus of many recent investigations, however, many questions remain. We hypothesized that genetic and epigenetic states of the mature sperm are important for establishing proper germ cell identity, function, and developmental capacity. We reason that perturbations to the genome or epigenome may underlie a subset of idiopathic cases of infertility. This dissertation addresses some potential genetic causes of male infertility; focusing on the genetic causes of an abnomal protamine ratio. Our gene re-sequencing studies examined whether single nucleotide polymorphisms (SNPs) in the untranslated regions of the protamine genes (Chapter 3) or in a transcriptional regulator of the protamine genes known as *YBX2* (Appendix A) may account for the altered P1/P2 ratio in infertile men. We concluded that gene mutations in these two regulatory elements were a rare cause of male infertility. As an alternative we turned to examine whether epigenetic changes in the germline may underlie some cases of male infertility. Initially we focused on characterizing the sperm chromatin of fertile men. A comprehensive analysis of the sperm epigenome revealed novel biological features that revised our understanding of the potential role of sperm chromatin in the creation of a totipotent gamete (Chapter 5). We then assessed whether changes in the chromatin landscape associate with infertility or poor reproductive outcome. Interesting,

we show significant changes in DNA methylation at a number of imprinted loci (Chapter 6 and 7) and subtle changes in the overall chromatin landscape (H3K4me and H3K27me) genome-wide that may have a detrimental effect on fecundity (Chapter 7). Taken together, these studies reinforce the dogma that infertility is a complex disease and that changes in the genetic and the epigenetic states of the mature sperm may account for some, but not all, cases of idiopathic infertility. Furthermore, these data suggest that germ cell chromatin may have a significant role in germ cell function and possibly contribute to early embryonic processes and infertility.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGMENTS

I was fortunate to have the opportunity to work with two very amazing mentors Douglas Carrell and Brad Cairns. Their training, stimulating discussions, intelligence, and enthusiasm brought out my best potential. Thank you very much for all your undue guidance and support, and for not firing me after all the expensive microarray bills I managed to generate. I would also like to send a special thanks to my thesis committee members Mary Lucero, Mike Sanguinetti, and David Jones for all the enlightening discussion and for the many things I have learned from you.

To my friends and colleagues in both the Carrell and Cairn's Lab: thank you for all help, reagents, and for making the lab a wonderful place to be. I would like specifically thank Benjamin Emery, Jahnvi Purwar, Andrew Oler all of whom I collaborated with on a few projects. I would also like to thank Lori Barnard for all her help throughout my graduate career and for trying to keep me on track and completing my goals.

Much of my graduate work was heavily dependent on next generation sequence and microarray experiments, which would have been inconceivable without the integral help of the microarray (Brian Dalley and Nicole Moss) and bioinformatics core facilities (David Nix and Brett Milash)

I would like to especially thank my very supportive and loving husband (Ahmad Hammoud) and my adorable twin girls (Maya and Lara) for their patience with my near total immersion in graduate school for the past 5 years. It was their support, love, and generosity that kept me moving forward in overwhelming times.

Last but not least, I would like to thank my mom, dad, and siblings for their support through difficult times. To my mom, a friend and a role model, there is not very much I can say to express my love and gratitude for all the sacrifices you made, but all I can say “I love you dearly”.

CHAPTER 1

INTRODUCTION

Infertility is a major health problem worldwide, affecting at least one in every eight couples. Many infertility problems are currently attributed to endocrinopathies, ejaculatory failure, gametogenesis dysfunction, tubal disease, uterine abnormalities, ovarian dysfunction and environmental exposure; however, a significant number of couples also have unexplained infertility. The high prevalence of unexplained infertility cases can be attributed, in part, to the poor understanding of the basic genetic and epigenetic mechanisms regulating fertility. We hypothesized that genetic and epigenetic states of the mature sperm are important for establishing proper germ cell identity, function, and developmental capacity. We reasoned that perturbations to the genome or epigenome may underlie a subset of idiopathic cases of male infertility and/or poor reproductive outcome. This dissertation aims to understand the contribution of sperm chromatin to the creation of a totipotent gamete, fertility, and development.

Human Sperm Chromatin Packaging: Protamines and Their Role in Fertility

During postmeiotic maturation of male haploid germ cells, chromatin undergoes dramatic reorganization including the exchange of canonical histones with intermediate highly basic transition nuclear proteins, which are subsequently replaced with protamines (discussed in Chapter 2 Review).[1] Protamines are arginine rich sperm specific nuclear proteins that are evolutionarily related to H1 linker histone.[2,3] The protamine family is comprised of protamines 1, 2, 3, and 4. Most mammals only express protamine 1 (P1), whereas mice and humans express P1 and P2, and at much lower abundance, P3 and P4. In mice, protamine haploinsufficiency or premature expression of the P1 results in male infertility,[4,5] suggesting that not only are protamine genes important for fertility, but the

temporal control of protamine gene transcription and translation are important to ensure proper sperm packaging and function.

In humans P1 and P2 are expressed in relatively equal quantities.[6] The P1/P2 ratio in human sperm donors of known normal fertility lies close to 1,[7] with a normal range from 0.8 to 1.2.[8,9] Patients with either low or high P1/P2 ratio have significantly reduced semen parameters (lower sperm concentration, motility, and morphology), increased sperm DNA damage, and reduced fertilization ability, as compared to patients with normal P1/P2 ratio (discussed in Chapter 2).[8,10-12] The reduced fertilization resulting from sperm abnormalities can be partially overcome by the use of intracytoplasmic sperm injection (ICSI); however, implantation and pregnancy rates remain significantly reduced.[8,12-15]

Interestingly, while a direct relationship between abnormal protamine expression and semen parameters (sperm count, motility, and morphology) or fertilization ability is apparent, the molecular mechanism for this association remains unclear. The reduction in P1 or P2 in infertile patients may be attributed to reduced protamine transcription, altered translation of the transcript or failed posttranslational modifications, but none of these scenarios would directly help explain the associated decline in sperm counts and function unless the regulation of protamines is linked to a broader control of spermatogenesis (discussed in Chapter 2).[16]

Protamine Gene Polymorphsims Are a Rare Cause of Infertility

To better understand the aetiology underlying a deregulated protamine ratio coupled with abnormal semen parameters, several mechanisms were proposed including

mutations in the protamine genes, untranslated regions (UTRs-Chapter 3), or in the transcriptional regulators (Contrin- *YBX2* – Appendix; discussed in Chapter 2).[17-19]

Candidate gene re-sequencing studies were performed to identify genetic variants that might alter protein structure or function and possibly explain the abnormality in protamine ratio and alteration in semen parameters in infertile men. The human protamine gene re-sequencing studies revealed many known and novel synonymous single nucleotide polymorphisms (SNP), however, these mutations were not significantly different between infertile patients with or without protamine alterations and fertile controls.[17,18,20,21] These data suggest that the presence of pathogenic mutations in the protamine genes are a rare cause of infertility.[17,19-24]

Since little genetic variability in the protamine genes was observed in a large population of infertile patients and controls, we and others subsequently extended the analysis to the untranslated regions (UTRs) of the protamine genes. UTRs (e.g., promoters and 3' UTRs) are predicted to be highly enriched in functional regulatory elements, which may be important in modulating gene expression or possibly accounting for changes in the transcriptional /translational regulation of the protamine genes.[25] Our study examined a total of 315 patients and we identified 14 single nucleotide polymorphisms (SNPs), of which 13 were novel SNPs in the UTRs of P1 and P2, and verified the existence of a variable length repeat (VLR), GAn, in the P2 5' region (discussed in Chapter 3).[26] However, SNP frequencies and VLR allelic frequencies were not significantly different between patients and fertile control populations.[26] In contrast, Gazquez et al. in a much smaller patient population (30 patients and 7 controls) identified a PRM1 –190 CA polymorphism that is significantly associated with abnormal sperm

head morphology and abnormal P1/P2 ratio in infertile patients.[24] The discrepancy between the two datasets may be attributed to differences in patient population or ethnicities. In conclusion, the literature general consensus is that alterations in the protamine genes or the untranslated regions are probably not singularly responsible for the protamine deficiency or abnormal spermatogenesis, but rather abnormal protamine ratio is a read-out of a general spermatogenic defect which affects a wide-range of spermatogenesis specific genes, including protamines. However, it remains possible that a combination of SNPs (haplotypes) together may predispose patients to abnormal protamine ratio and/or alteration in sperm parameters.

The application of candidate gene sequencing methods to elucidate genetic causes of male infertility have been inefficient and have yielded poor associations (with the exception of two gene mutations described in aryl hydrocarbon receptor repressor (*AHRR)* and methylenetetrahydrofolate reductase (MTHFR)).[27-31] Given the growing number of candidate 'fertility genes' important for proper sperm formation and function[32] and the complexity of the disease, genome-wide SNP studies acquired popularity due to the unprecedented amount of genetic information obtained from a single sample.[33-35] These preliminary genome wide studies are making significant contributions to the understanding of the complexity of the genetic element of male factor infertility. Furthermore, genomewide studies support the notion that infertility is a multifactorial disease, and like other complex diseases, numerous etiologies likely contribute to male infertility, with genetics or epigenetics being potential contributors.

Epigenetics and Its Cellular Function

Epigenetic modifications on the DNA sequence (DNA methylation) or on chromatin-associated proteins (i.e., histones) comprise the "cellular epigenome"; together these modifications play an important role in the regulation of gene expression.[36,37] Unlike the genome, the cellular epigenome is highly variable between cells and is dynamic and plastic in response to cellular stress and environmental cues.[38,39] The role of the cellular epigenome has been increasingly highlighted and has been implicated in many cellular and developmental processes such as embryonic reprogramming, cellular differentiation, imprinting, X chromosome inactivation, genomic stability, and complex diseases such as cancer.

DNA methylation in mammals is a post-replication modification that is predominantly found on cytosines of the dinucleotides sequence CpG. DNA methylation is a major regulator of gene expression and a primary chromatin attribute associated with the main mode of repression at transposons and retroviral elements.[40-44] DNA methylation is established by DNA methyltransferases (also known as DNMTs)[45,46] and recognized by methyl-binding proteins which then recruit chromatin silencing machinery to reinforce a repressive chromatin state.[47-49]

DNA methylation is just one component of a wider epigenetic program along with the histone code. Histone proteins are subject to over 100 known posttranslational modifications, including acetylation, methylation, ADP-ribosylation, ubiquitination, and phosphorylation.[50,51] These modifications occur on the histone tails and functionally impact transcription, replication, recombination, and repair. Many of the modifications and the chromatin regulators (example: MLL and EZH2) have been implicated in a range

of early embryonic developmental processes and physiologic pathways.[52-55]

Epigenetics and Its Role in Development

A central question in early development is how is totipotency and pluripotency established in early embryos? Studies in embryonic stem (ES) cells have produced many interesting concepts for pluripotency, such as the use of special chromatin attributes (histone modifications) and a transcriptional network to prevent cellular differentiation.[56-59] In 2006, Bernstien et al. first described a novel chromatin structure at the promoters of important developmental transcription factors in embryonic stem cells. This novel chromatin structure, termed bivalent domains, consists of two contrasting chromatin marks, a large region of Lys27 methylation (a repressive histone modification) overlapping with a smaller region of Lys4 methylation (an activating histone modification). It was subsequently shown that these bivalent promoters were also DNA hypomethylated and co-occupied by a network of transcription factors (Oct4, Sox2, Nanog, FOXD3, PRC2 etc). The combination of bivalent domains, DNA hypomethylation and co-occupancy of key transcription factors together maintains pluripotency and promotes self-renewal.[56-59] These findings in ES cells have led to the question of how and when pluripotency/totipotency is established. Specifically, can germ cell chromatin contribute to this poised state in ES cells or is this chromatin state established in the ES cell? To address these questions, we have focused on understanding the chromatin of human sperm.

The Emerging Role for Epigenetics in the Male Germline

During spermatogenesis male germ cells undergo unique and extensive chromatin and epigenetic remodeling soon after their specification (determination to become a spermatocyte) and during differentiation.[60] During mitosis and meiosis, male germ cell DNA is packaged in nucleosomes, comprised of histone 2A (H2A), histone 2B (H2B), histone 3 (H3) and histone 4 (H4), all of which are susceptible to covalent modifications, such as methylation, acetylation, ubiquitination and phosphorylation. Each of these chemical modifications to histones works alone or in concert to influence gene repression and/or activation and promote subsequent events in chromatin remodeling (discussed in Chapter 4).

Although sperm undergoes dramatic chromatin changes, the exchange process is incomplete; approximately 5-15% of the genome remains packaged in nucleosomes.[61,62] This chromatin composition led to the predominant notion that paternal contribution of epigenetic information to the embryo may be limited to the few paternally imprinted genes since the vast majority of the sperm genome is packaged by protamines and protamines are not known to propagate information via modifications. This low level of histone retention raised an intriguing question of whether the presence of nucleosomes may simply be due to inefficient protamine replacement, leading to a low random genomewide distribution with no function in the embryo, or the retained nucleosomes along with their attendant epigenetic modifications are of potential significance to the developing embryo. Before genomewide array and sequencing technologies were readily available, two groups examined the sperm chromatin composition at few loci required in embryogenesis, and both reported significant histone enrichment at these tested loci.[61,63]

These early studies suggested that histone retention might be enriched at developmental loci; however, a limited number of regions were evaluated.

To characterize the complete sperm epigenome, we applied genomewide approaches to examine the localization of retained histones. We found that nucleosomes were significantly enriched at developmental transcription factors, imprinted genes, miRNAs, and spermatogenesis gene promoters.[64] This work was subsequently validated by several other groups.[65,66] Furthermore, we found that spermatogenesis and cell cycle gene promoters retained H3K4me, a mark of gene activation, whereas developmental gene promoters retained H3K4me and H3K27me similar to embryonic stem cells.[64-66] This work suggests that the unique chromatin structure in ES cells may begin in the germ-line and is recapitulated in the embryo.

This differential poising at spermatogenesis and cell cycle vs. developmental transcription factors is not limited to human male germ cells, but was found to be conserved in both mouse and zebrafish sperm.[66,67] This conservation in histone marking raises the question of whether the retained nucleosomal regions are important for early embryonic development or are simply residual marks from the spermatogonial stem cell.

Epigenetics and Infertility

Retaining histones at developmental promoters and imprinted loci in fertile men expands the role of the paternal sperm epigenome and provides the paternal genome the opportunity to convey instructive epigenetic information to the offspring. Therefore, one may envision that infertile men with an altered P1/P2 ratio or histone retention may posit a risk for transmission of aberrant epigenetic marks to offspring. If an epigenetic role is

ascribed to the retained nucleosomes for humans, then there are obvious and profound implications for sperm with abnormal histone retention and protamine levels, and for the use of such sperm, or any immature sperm, to achieve a pregnancy with the use of assisted reproductive technologies (ART; Chapter 7).

Precedence for epigenetic abnormalities and infertility first stemmed from early work of in-vitro fertilization (IVF) in animal models. Offspring conceived by assisted reproductive technologies in animals had an increased incidence of large offspring syndrome.[68] In humans, a meta-analysis showed that children born from assisted reproductive technology (ART) have a four fold increased incidence of Beckwith-Weidemann syndrome compared with children conceived naturally.[69-72] This rise in imprinting errors was attributed to embryo or gamete manipulation, in-vitro culture conditions, hormonal stimulation, or ovulation induction.[70,72-75] However, an alternative hypothesis is that the increased incidence of imprinting disorders might be due to facilitation of conception using gametes of infertile couples that may have elevated risk of epigenomic errors. This view aligns with the limited number of reports, including our own data, which showed abnormal methylation patterns at a number of imprinted loci (SNRPN, MEST, LIT1) in the gametes of infertile men (Chapter 6).[76-82] These findings are concerning and require a better understanding of the effects of existent or acquired epimutations on embryo quality, IVF success rate, or the potential risk for transgenerational inheritance of epimutations.

Dissertation Overview

This dissertation describes the sperm genetic and epigenetic contributions to early development and infertility. We hypothesize that genetic and epigenetic states of the mature sperm are important for establishing proper germ cell identity, function, and developmental capacity. We reason that perturbations to the genome or epigenome may underlie a subset of idiopathic cases of male infertility and/or poor reproductive outcome. Initially, we investigated potential genetic causes of male infertility by evaluating protamine 1 and 2 UTRs in fertile and infertile men. Following the genetic analysis, the epigenetic status of fertile and infertile men was evaluated to determine whether epigenetic alteration in sperm chromatin contribute to the spectrum of abnormalities in spermatogenesis or poor reproductive outcome. Sperm DNA methylation and histone localization and characterization was performed using state of the art genome-wide technologies. The analysis of sperm chromatin in fertile men demonstrated that modified nucleosomes enriched specifically at imprinted loci, developmental and miRNA gene promoters, which raised the possibility that sperm chromatin packaging may be instructive to the zygote. Furthermore, when comparing the epigenetic profile of fertile and infertile men we find many subtle differences throughout the genome that may have cumulative detrimental effect on fertility. Together this body of work provides a better understanding of the sperm genetic and epigenetic states and their potential role in spermatogenesis, fertility, and early embryo development.

References

1. Kimmins S, Sassone-Corsi P. Chromatin remodelling and epigenetic features of germ cells. Nature 2005;434(7033):583-9.

2. Lewis JD, Saperas N, Song Y, Zamora MJ, Chiva M, Ausio J. Histone H1 and the origin of protamines. Proc Natl Acad Sci U S A 2004;101(12):4148-52.

3. Oliva R, Dixon GH. Vertebrate protamine genes and the histone-to-protamine replacement reaction. Prog Nucleic Acid Res Mol Biol 1991;40:25-94.

4. Lee K, Haugen HS, Clegg CH, Braun RE. Premature translation of protamine 1 mRNA causes precocious nuclear condensation and arrests spermatid differentiation in mice. Proc Natl Acad Sci U S A 1995;92(26):12451-5.

5. Cho C, Willis WD, Goulding EH, Jung-Ha H, Choi YC, Hecht NB, Eddy EM. Haploinsufficiency of protamine-1 or -2 causes infertility in mice. Nature Genetics 2001;28(1):82-6.

6. Corzett M, Mazrimas J, Balhorn R. Protamine 1: protamine 2 stoichiometry in the sperm of eutherian mammals. Mol Reprod Dev 2002;61(4):519-27.

7. Carrell DT, Liu L. Altered protamine 2 expression is uncommon in donors of known fertility, but common among men with poor fertilizing capacity, and may reflect other abnormalities of spermiogenesis. J Androl 2001;22(4):604-10.

8. Aoki VW, Liu L, Carrell DT. Identification and evaluation of a novel sperm protamine abnormality in a population of infertile males. Hum Reprod 2005;20(5):1298-306.

9. Oliva R. Protamines and male infertility. Hum Reprod Update 2006;12:417-435.

10. Torregrosa N, Dominguez-Fandos D, Camejo MI, Shirley CR, Meistrich ML, Ballesca JL, Oliva R. Protamine 2 precursors, protamine 1/protamine 2 ratio, DNA integrity and other sperm parameters in infertile patients. Hum Reprod 2006;21(8):2084-9.

11. Garcia-Peiro A, Martinez-Heredia J, Oliver-Bonet M, Abad C, Amengual MJ, Navarro J, Jones C, Coward K, Gosalvez J, Benet J. Protamine 1 to protamine 2 ratio correlates with dynamic aspects of DNA fragmentation in human sperm. Fertil Steril 2011;95(1):105-9.

12. Aoki VW, Moskovtsev SI, Willis J, Liu L, Mullen JB, Carrell DT. DNA integrity is compromised in protamine-deficient human sperm. J Androl 2005;26(6):741-8.

13. Chevaillier P, Mauro N, Feneux D, Jouannet P, David G. Anomalous protein complement of sperm nuclei in some infertile men. Lancet 1987;2(8562):806-7.

14. de Yebra L, Ballesca JL, Vanrell JA, Corzett M, Balhorn R, Oliva R. Detection of P2 precursors in the sperm cells of infertile patients who have reduced protamine P2 levels. Fertil Steril 1998;69(4):755-9.

15. Aoki VW, Liu L, Jones KP, Hatasaka HH, Gibson M, Peterson CM, Carrell DT. Sperm protamine 1/protamine 2 ratios are related to in vitro fertilization pregnancy rates and predictive of fertilization ability. Fertil Steril 2006;86(5):1408-15.

16. Carrell DT, Emery BR, Hammoud S. Altered protamine expression and diminished spermatogenesis: what is the link? Hum Reprod Update 2007;13(3):313-27.

17. Aoki VW, Christensen GL, Atkins JF, Carrell DT. Identification of novel polymorphisms in the nuclear protein genes and their relationship with human sperm protamine deficiency and severe male infertility. Fertil Steril 2006;86(5):1416-22.

18. Iguchi N, Yang S, Lamb DJ, Hecht NB. An SNP in protamine 1: a possible genetic cause of male infertility? J Med Genet 2006;43(4):382-4.

19. Schlicker M, Schnulle V, Schneppel L, Vorob'ev VI, Engel W. Disturbances of nuclear condensation in human spermatozoa: search for mutations in the genes for protamine 1, protamine 2 and transition protein 1. Hum Reprod 1994;9(12):2313-7.

20. Tanaka H, Miyagawa Y, Tsujimura A, Matsumiya K, Okuyama A, Nishimune Y. Single nucleotide polymorphisms in the protamine-1 and -2 genes of fertile and infertile human male populations. Mol Hum Reprod 2003;9(2):69-73.

21. Kichine E, Msaidie S, Bokilo AD, Ducourneau A, Navarro A, Levy N, Terriou P, Collignon P, Boetsch G, Chiaroni J and others. Low-frequency protamine 1 gene transversions c.102G->T and c.-107G->C do not correlate with male infertility. J Med Genet 2008;45(4):255-6.

22. Queralt R, Oliva R. Identification of conserved potential regulatory sequences of the protamine-encoding P1 genes from ten different mammals. Gene 1993;133(2):197-204.

23. de Yebra L, Ballesca JL, Vanrell JA, Bassas L, Oliva R. Complete selective absence of protamine P2 in humans. J Biol Chem 1993;268(14):10553-7.

24. Gazquez C, Oriola J, de Mateo S, Vidal-Taboada JM, Ballesca JL, Oliva R. A common protamine 1 promoter polymorphism (-190 C->A) correlates with abnormal sperm morphology and increased protamine P1/P2 ratio in infertile patients. J Androl 2008;29(5):540-8.

25. Xie X, Lu J, Kulbokas EJ, Golub TR, Mootha V, Lindblad-Toh K, Lander ES, Kellis M. Systematic discovery of regulatory motifs in human promoters and 3' UTRs by comparison of several mammals. Nature 2005;434(7031):338-45.

26. Hammoud S, Emery BR, Aoki VW, Carrell DT. Identification of genetic variation in the 5' and 3' non-coding regions of the protamine genes in patients with protamine deregulation. Arch Androl 2007;53(5):267-74.

27. Krausz C, Quintana-Murci L, McElreavey K. Prognostic value of Y deletion analysis: what is the clinical prognostic value of Y chromosome microdeletion analysis? Hum Reprod 2000;15(7):1431-4.

28. Lee HC, Jeong YM, Lee SH, Cha KY, Song SH, Kim NK, Lee KW, Lee S. Association study of four polymorphisms in three folate-related enzyme genes with non-obstructive male infertility. Hum Reprod 2006;21(12):3162-70.

29. Park JH, Lee HC, Jeong YM, Chung TG, Kim HJ, Kim NK, Lee SH, Lee S. MTHFR C677T polymorphism associates with unexplained infertile male factors. J Assist Reprod Genet 2005;22(9-10):361-8.

30. Masciarelli S, Horner K, Liu C, Park SH, Hinckley M, Hockman S, Nedachi T, Jin C, Conti M, Manganiello V. Cyclic nucleotide phosphodiesterase 3A-deficient mice as a model of female infertility. J Clin Invest 2004;114(2):196-205.

31. Watanabe M, Sueoka K, Sasagawa I, Nakabayashi A, Yoshimura Y, Ogata T. Association of male infertility with Pro185Ala polymorphism in the aryl hydrocarbon receptor repressor gene: implication for the susceptibility to dioxins. Fertil Steril 2004;82 Suppl 3:1067-71.

32. Matzuk MM, Lamb DJ. Genetic dissection of mammalian fertility pathways. Nat Cell Biol 2002;4 Suppl:s41-9.

33. Feugang JM, Kaya A, Page GP, Chen L, Mehta T, Hirani K, Nazareth L, Topper E, Gibbs R, Memili E. Two-stage genome-wide association study identifies integrin beta 5 as having potential role in bull fertility. BMC Genomics 2009;10:176.

34. Aston KI, Carrell DT. Genome-wide study of single-nucleotide polymorphisms associated with azoospermia and severe oligozoospermia. J Androl 2009;30(6):711-25.

35. Aston KI, Krausz C, Laface I, Ruiz-Castane E, Carrell DT. Evaluation of 172 candidate polymorphisms for association with oligozoospermia or azoospermia in a large cohort of men of European descent. Hum Reprod 2010;25(6):1383-97.

36. Suzuki MM, Bird A. DNA methylation landscapes: provocative insights from epigenomics. Nat Rev Genet 2008;9(6):465-76.

37. Bird A. Perceptions of epigenetics. Nature 2007;447(7143):396-8.

38. Margueron R, Trojer P, Reinberg D. The key to development: interpreting the histone code? Curr Opin Genet Dev 2005;15(2):163-76.

39. Jenuwein T, Allis CD. Translating the histone code. Science 2001;293(5532):1074-80.

40. Bird AP, Wolffe AP. Methylation-induced repression--belts, braces, and chromatin. Cell 1999;99(5):451-4.

41. Walsh CP, Chaillet JR, Bestor TH. Transcription of IAP endogenous retroviruses is constrained by cytosine methylation. Nature Genetics 1998;20(2):116-7.

42. Dodge JE, Okano M, Dick F, Tsujimoto N, Chen T, Wang S, Ueda Y, Dyson N, Li E. Inactivation of Dnmt3b in mouse embryonic fibroblasts results in DNA hypomethylation, chromosomal instability, and spontaneous immortalization. J Biol Chem 2005;280(18):17986-91.

43. Karpf AR, Matsui S. Genetic disruption of cytosine DNA methyltransferase enzymes induces chromosomal instability in human cancer cells. Cancer Res 2005;65(19):8635-9.

44. Chen T, Li E. Structure and function of eukaryotic DNA methyltransferases. Curr Top Dev Biol 2004;60:55-89.

45. Okano M, Bell DW, Haber DA, Li E. DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. Cell 1999;99(3):247-57.

46. Li E, Bestor TH, Jaenisch R. Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. Cell 1992;69(6):915-26.

47. Zhang Y, Ng HH, Erdjument-Bromage H, Tempst P, Bird A, Reinberg D. Analysis of the NuRD subunits reveals a histone deacetylase core complex and a connection with DNA methylation. Genes Dev 1999;13(15):1924-35.

48. Jones PL, Veenstra GJ, Wade PA, Vermaak D, Kass SU, Landsberger N, Strouboulis J, Wolffe AP. Methylated DNA and MeCP2 recruit histone deacetylase to repress transcription. Nature Genetics 1998;19(2):187-91.

49. Nan X, Ng HH, Johnson CA, Laherty CD, Turner BM, Eisenman RN, Bird A. Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. Nature 1998;393(6683):386-9.

50. Kouzarides T. Chromatin modifications and their function. Cell 2007;128(4):693-705.

51. Lee JS, Smith E, Shilatifard A. The language of histone crosstalk. Cell 2010;142(5):682-5.

52. Simon JA, Lange CA. Roles of the EZH2 histone methyltransferase in cancer epigenetics. Mutat Res 2008;647(1-2):21-9.

53. Mills AA. Throwing the cancer switch: reciprocal roles of polycomb and trithorax proteins. Nat Rev Cancer 2010;10(10):669-82.

54. Ringrose L, Paro R. Polycomb/Trithorax response elements and epigenetic memory of cell identity. Development 2007;134(2):223-32.

55. Schwartz YB, Pirrotta V. Polycomb silencing mechanisms and the management of genomic programmes. Nat Rev Genet 2007;8(1):9-22.

56. Wang J, Rao S, Chu J, Shen X, Levasseur DN, Theunissen TW, Orkin SH. A protein interaction network for pluripotency of embryonic stem cells. Nature 2006;444(7117):364-8.

57. Rao S, Orkin SH. Unraveling the transcriptional network controlling ES cell pluripotency. Genome Biol 2006;7(8):230.

58. Kim J, Chu J, Shen X, Wang J, Orkin SH. An extended transcriptional network for pluripotency of embryonic stem cells. Cell 2008;132(6):1049-61.

59. Jaenisch R, Young R. Stem cells, the molecular circuitry of pluripotency and nuclear reprogramming. Cell 2008;132(4):567-82.

60. Seki Y HK, Itoh K, Mizugaki M, Saitou M, Matsui Y. . Extensive and orderly reprogramming of genome-wide chromatin modifications associated with specification and early development of germ cells in mice. Dev Biol 2005;278:440–45.

61. Wykes SM, Krawetz SA. The structural organization of sperm chromatin. J Biol Chem 2003;278(32):29471-7.

62. Gatewood JM, Cook GR, Balhorn R, Bradbury EM, Schmid CW. Sequence-specific packaging of DNA in human sperm chromatin. Science 1987;236(4804):962-4.

63. Gardiner-Garden M, Ballesteros M, Gordon M, Tam PP. Histone- and protamine-DNA association: conservation of different patterns within the beta-globin domain in human sperm. Mol Cell Biol 1998;18(6):3350-6.

64. 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(7254):473-8.

65. Arpanahi A, Brinkworth M, Iles D, Krawetz SA, Paradowska A, Platts AE, Saida M, Steger K, Tedder P, Miller D. Endonuclease-sensitive regions of human spermatozoal chromatin are highly enriched in promoter and CTCF binding sequences. Genome Res 2009;19(8):1338-49.

66. Brykczynska U, Hisano M, Erkek S, Ramos L, Oakeley EJ, Roloff TC, Beisel C, Schubeler D, Stadler MB, Peters AH. Repressive and active histone methylation mark distinct promoters in human and mouse spermatozoa. Nat Struct Mol Biol 2010;17(6):679-87.

67. Vastenhouw NL, Zhang Y, Woods IG, Imam F, Regev A, Liu XS, Rinn J, Schier AF. Chromatin signature of embryonic pluripotency is established during genome activation. Nature 2010;464(7290):922-6.

68. Young LE, Sinclair KD, Wilmut I. Large offspring syndrome in cattle and sheep. Rev Reprod 1998;3(3):155-63.

69. Gosden R, Trasler J, Lucifero D, Faddy M. Rare congenital disorders, imprinted genes, and assisted reproductive technology. Lancet 2003;361(9373):1975-7.

70. DeBaun MR, Niemitz EL, Feinberg AP. Association of in vitro fertilization with Beckwith-Wiedemann syndrome and epigenetic alterations of LIT1 and H19. Am J Hum Genet 2003;72(1):156-60.

71. Niemitz EL, DeBaun MR, Fallon J, Murakami K, Kugoh H, Oshimura M, Feinberg AP. Microdeletion of LIT1 in familial Beckwith-Wiedemann syndrome. Am J Hum Genet 2004;75(5):844-9.

72. Cox GF, Burger J, Lip V, Mau UA, Sperling K, Wu BL, Horsthemke B. Intracytoplasmic sperm injection may increase the risk of imprinting defects. Am J Hum Genet 2002;71(1):162-4.

73. Ludwig M, Katalinic A, Gross S, Sutcliffe A, Varon R, Horsthemke B. Increased prevalence of imprinting defects in patients with Angelman syndrome born to subfertile couples. J Med Genet 2005;42(4):289-91.

74. Mann MR, Lee SS, Doherty AS, Verona RI, Nolen LD, Schultz RM, Bartolomei MS. Selective loss of imprinting in the placenta following preimplantation development in culture. Development 2004;131(15):3727-35.

75. Khosla S, Dean W, Brown D, Reik W, Feil R. Culture of preimplantation mouse embryos affects fetal development and the expression of imprinted genes. Biol Reprod 2001;64(3):918-26.

76. Kagami M, Nagai T, Fukami M, Yamazawa K, Ogata T. Silver-Russell syndrome in a girl born after in vitro fertilization: partial hypermethylation at the differentially methylated region of PEG1/MEST. J Assist Reprod Genet 2007;24(4):131-6.

77. 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 2009.

78. Marques CJ, Francisco T, Sousa S, Carvalho F, Barros A, Sousa M. Methylation defects of imprinted genes in human testicular spermatozoa. Fertil Steril 2010;94(2):585-94.

79. Marques CJ, Costa P, Vaz B, Carvalho F, Fernandes S, Barros A, Sousa M. Abnormal methylation of imprinted genes in human sperm is associated with oligozoospermia. Mol Hum Reprod 2008;14(2):67-74.

80. Marques CJ, Carvalho F, Sousa M, Barros A. Genomic imprinting in disruptive spermatogenesis. Lancet 2004;363(9422):1700-2.

81. Kobayashi H, Hiura H, John RM, Sato A, Otsu E, Kobayashi N, Suzuki R, Suzuki F, Hayashi C, Utsunomiya T and others. DNA methylation errors at imprinted loci after assisted conception originate in the parental sperm. Eur J Hum Genet 2009;17(12):1582-91.

82. Kobayashi H, Sato A, Otsu E, Hiura H, Tomatsu C, Utsunomiya T, Sasaki H, Yaegashi N, Arima T. Aberrant DNA methylation of imprinted loci in sperm from oligospermic patients. Human Molecular Genetics 2007;16(21):2542-51.

CHAPTER 2

ALTERED PROTAMINE EXPRESSION AND DIMINISHED SPERMATOGENESIS: WHAT IS THE LINK?

Published in *Human Reproduction Update*

Authors

Douglas T. Carrell

Benjamin R. Emery

Sue Hammoud

Human Reproduction Update, Vol.13, No.3 pp. 313–327, 2007 doi:10.1093/humupd/dml057
Advance Access publication January 5, 2007

Altered protamine expression and diminished spermatogenesis: what is the link?

Douglas T.Carrell[1,2,3,4], Benjamin R.Emery[1,2] and Sue Hammoud[1,2]

[1]Andrology and IVF Laboratories, Department of Surgery, [2]Department of Physiology and [3]Department of Obstetrics and Gynecology, University of Utah School of Medicine, Salt Lake City, UT, USA

[4]To whom correspondence should be addressed at: Andrology and IVF Laboratories, 675 S. Arapeen Dr, Suite #205, Salt Lake City, UT 84108, USA. E-mail: douglas.carrell@hsc.utah.edu

During the elongating spermatid stage of spermiogenesis, human sperm chromatin undergoes a complex transition in which histones are extensively replaced by protamines in a carefully regulated transition including histone modifications and intermediate and temporary replacement of the histones by sperm-specific transition proteins. The replacement of most histones by protamines 1 and 2 facilitates a high order of chromatin packaging necessary for normal sperm function and may also be necessary for DNA silencing and imprinting changes within the sperm cell. Protamines 1 and 2 are usually expressed in nearly equal quantities, but elevated or diminished protamine 1/protamine 2 ratios are observed in some infertile men and is often associated with severe spermatogenesis defects. Human and animal studies demonstrate that expression of the protamine proteins is uniquely regulated by transcription/translation factors, including storage of the mRNA in ribonucleoprotein (RNP) particles composed of the mRNA, transcription factors and a kinesin molecule necessary for transport of the RNP to the cytoplasm and removal of transcriptional activators from the nucleus. Recent studies indicate that most patients with abnormal protamine protein levels have elevated levels of protamine transcript in the mature sperm cell, indicating a possible defect in transcription or translation. The regulation of protamine expression is unique and includes several possible mechanisms which may be responsible for dysregulation of protamine expression and concurrent broad spectrum defects in spermatogenesis. We suggest two hypotheses: (i) that abnormal protamine expression is indicative of a generalized defect in mRNA storage and/or translation which affects other mRNA transcripts or (ii) that protamines may act as a checkpoint of spermatogenesis.

Key words: chromatin/gene expression/protamine/spermatogenesis/transition protein

Introduction

Sperm chromatin is a highly organized, compact structure consisting of DNA and heterogeneous nucleoproteins. The most abundant nucleoproteins in mature sperm are the protamines, positively charged molecules that replace histones during spermiogenesis. Protamines confer a higher order of DNA packaging in sperm than that found in somatic cells, and the condensed and insoluble nature of the highly condensed sperm chromatin protects the genetic integrity of the paternal genome during its transport through the male and female reproductive tracts (Gatewood *et al.*, 1987; Balhorn *et al.*, 1999; Brewer *et al.*, 2002). Protamine replacement may also be necessary for silencing of the paternal genome and reprogramming of the imprinting pattern of the gamete (Aoki and Carrell, 2003).

Humans express two protamines, protamine 1 (P1) and protamine 2 (P2), both of which are expressed in roughly equal quantities (Balhorn *et al.*, 1999; Corzett *et al.*, 2002). Protamines are highly basic sperm-specific nuclear proteins that are characterized by an arginine-rich core and cysteine residues (Dixon *et al.*, 1986; Krawetz and Dixon, 1988). The high level of arginine causes a net positive charge that facilitates strong DNA binding (Balhorn *et al.*, 2000). The cysteine residues facilitate the formation of multiple inter and intra-protamine disulphide bonds that are essential for the high order of chromatin packaging necessary for normal sperm function (Courtens and Loir, 1981; Loir and Lanneau, 1984; Singh and Rao, 1988; Le Lannic *et al.*, 1993; Szczygiel and Ward, 2002).

During spermiogenesis, protamines progressively replace somatic histones in a stepwise manner (Dixon *et al.*, 1986). First, somatic histones are replaced by testis-specific histone variants, which are then replaced by transition proteins (TP1 and TP2) in a process that involves extensive DNA rearrangement and remodeling (Ward *et al.*, 1989). During the elongating spermatid stage, the transition proteins are replaced in the condensing chromatin by protamines. In humans, ~ 85% of the histones are replaced by protamines. (Hecht, 1989, 1990; Oliva and Dixon, 1990; Dadoune, 1995; Steger, 1999). This sequential process

facilitates molecular remodelling of the male genome within the differentiating spermatid nucleus (Figure 1) (Sassone-Corsi, 2002).

Previous studies have shown that the mean P1/P2 ratio in human sperm is approximately 1.0 (Balhorn *et al.*, 1999; Carrell and Liu, 2001; Oliva, 2006). Sperm from some infertile men have been shown to have altered P1/P2 ratios and/or non-detectable P2 in mature sperm, whereas the occurrence of protamine abnormalities in sperm from fertile men is extremely rare (no known cases have been reported) (Balhorn *et al.*, 1988; Chevaillier *et al.*, 1987; Belokopytova *et al.*, 1993; de Yebra *et al.*, 1998; Aoki *et al.*, 2005a; Oliva, 2006). Additionally, transgenic mice with protamine haploinsufficiency have severely altered spermatogenesis and male infertility (Cho *et al.*, 2001). The link between abnormal protamine levels and infertility is intriguing because abnormal protamine expression has been associated with low sperm counts, decreased sperm motility and morphology, diminished fertilization ability and increased sperm chromatin damage, some of which are not intuitively linked to abnormal chromatin structure (Carrell and Liu, 2001; Mengual *et al.*, 2003; Aoki *et al.*, 2005a).

A direct relationship between abnormal protamine expression and sperm count, motility, morphology or fertilization ability is not readily apparent. The reduction in P1 or P2 in these patients may be explained by reduced protamine transcription, altered translation of the transcript or failed post-translational modifications, but none of these scenarios would directly explain the associated decline in sperm counts and function unless the regulation of protamine exchange is linked to a broader control of spermatogenesis.

This review will briefly summarize the current understanding of protamine replacement of histones, the link between altered protamine replacement and male infertility, the regulation of protamine

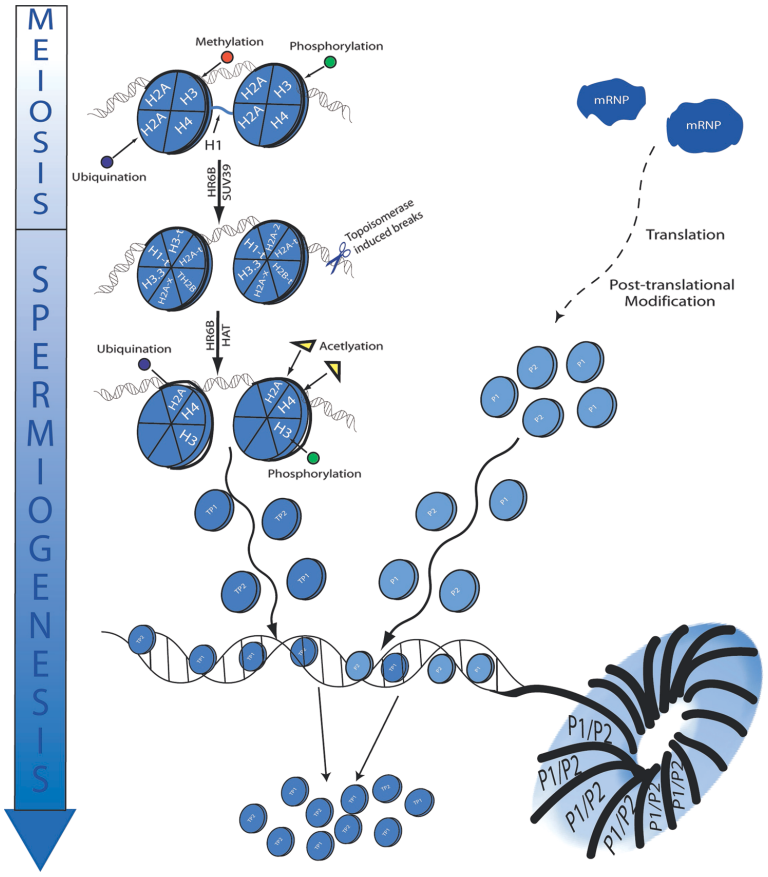


Figure 1. Diagram highlighting the key events in the transition of somatic histones to replacement by protamines. Somatic histones undergo site-specific methylation, phosphorylation and ubiquitination which facilitates their replacement by testis-specific histones (t) during meiosis. Hyperacetylation of H4-t is a key factor in relaxation of the DNA coil to facilitate replacement of the testis-specific histones by the transition proteins, whereas topoisomerase 1 relieves torsional stress by causing double-strand breaks which are subsequently re-ligated. Protamines 1 and 2, processed from a pool of RNP particles, undergo maturation before and during binding to the DNA and replacemnt of the transition proteins. HR6B, ubiquitin-conjugating enzyme E2B (UBE2B) (RAD6 homolog); HAT, histone acetyltransferase; Suv39, H3 Lys 9 histone methyltransferase.

expression during spermatogenesis and possible causes of altered protamine expression. Two possible models will be discussed regarding the link between abnormal protamine expression and aberrant spermatogenesis. The first hypothesis is that abnormal protamine expression is indicative of a general abnormality of spermatogenesis, possibly due to abnormal function of a transcriptional or translational regulator. Candidate regulatory factors will be discussed. The second hypothesis to be discussed is that the protamines may act as a checkpoint regulator of spermatogenesis and that abnormal protamine expression leads to induction of an apoptotic process that ends in severely diminished semen quality.

Clinical significance of abnormal protamine expression

Abnormal protamine expression is clearly associated with infertility, as recently thoroughly reviewed by Oliva (Oliva, 2006). Briefly, studies have identified males with undetectable P2, which has consistently been linked to severe male infertility (de Yebra *et al.*, 1993; Carrell and Liu, 2001). In mice, haploinsufficiency of the protamines has been shown to cause altered spermatogenesis, including lowered sperm counts and DNA damage (C. Cho *et al.*, 2001; Cho *et al.*, 2003). Numerous studies have demonstrated that an altered ratio of P1/P2, either increased or decreased, is associated with reduced fertility (de Yebra *et al.*, 1998; Carrell and Liu, 2001; Aoki *et al.*, 2006d; Oliva, 2006). Interestingly, no reports have been made of P1/P2 expression abnormalities in males of known fertility.

Initial studies suggested that the most common protamine abnormality in infertile men was an elevated P1/P2 ratio (Oliva, 2006). The elevated P1/P2 ratio is often the result of decreased P2 protein levels, concomitant with an increased level of P2 precursors (Carrell and Liu, 2001; de Yebra *et al.*, 1998; Aoki *et al.*, 2006d). Under expression of P2 accounts for the majority of the cases with high P1/P2 ratio, but subsequent studies have demonstrated that P1 dysregulation also accounts for some abnormalities (Aoki *et al.*, 2005a). However, P2 dysregulation is more common and this may be explained by the fact that the P2 gene is derived more recently than the P1 gene, which may suggest that the regulatory mechanisms governing P2 gene expression are not as stringent and more susceptible to variation than the P1 gene (Lewis *et al.*, 2003).

Human sperm protamine dysregulation is associated with diminished semen quality parameters, sperm functional ability and sperm DNA integrity (de Yebra *et al.*, 1993, 1998; Balhorn *et al.*, 1999; Carrell and Liu, 2001; Aoki *et al.*, 2005a). Aoki *et al.* (2005b) have shown that sperm concentration, motility and morphology are significantly reduced in patients with either a low or a high P1/P2 ratio when compared with patients with a normal P1/P2 ratio. In addition, an altered P1/P2 ratio is associated with decreased fertilization ability, although fertilization and pregnancy rates are not different when patients undergo intracytoplasmic sperm injection (ICSI) as opposed to standard in vitro fertilization (IVF) (Carrell and Liu, 2001; Nasr-Esfahani *et al.*, 2004; Aoki *et al.*, 2005b).

Protamines and DNA damage

One potential consequence of abnormal protamine expression is a susceptibility to DNA damage. Our laboratory has measured DNA integrity using an assay similar to the sperm chromatin structure assay and compared the DNA fragmentation index with protamine levels in human sperm. Patients with low P1/P2 ratio had significantly elevated DNA fragmentation when compared with patients with normal and high P1/P2 ratios (Aoki *et al.*, 2005b). Moreover, patients who under-expressed P1, P2 or both P1 and P2 had significantly elevated levels of DNA fragmentation compared with patients normally expressing P1and P2. Additionally, Torregrosa *et al.* (2006) have recently shown a positive correlation between TUNEL-positive sperm and the presence of P2 precursors. These studies emphasized the important role protamines play in protecting the genetic content of the mature sperm from nucleases.

A recent study evaluated the role of protamine abnormalities at an individual cell level by using fluorescence immunohistochemistry techniques to simultaneously evaluate protamine levels, cell viability and DNA damage as measured by the TUNEL assay (Aoki *et al.*, 2006c). Concurrently, global protamine levels were evaluated with a fraction of the semen sample that underwent standard nuclear protein extraction and electrophoresis. The data not only confirmed a close correlation between the mean protamine levels determined by fluorescence microscopy and the standard electrophoresis technique, but also showed that within a semen sample there is heterogeneity in protamine expression and a clear correlation between under-expression of protamines, DNA damage and lack of viability (Aoki *et al.*, 2006c). The intra-ejaculate protamine heterogeneity observed in this study is consistent with other reports using CMA3 and Aniline Blue staining to assess protamine quantity indirectly (Manicardi *et al.*, 1995; Hammadeh *et al.*, 2001), but novel in the direct link between protamine abnormalities in a given cell and DNA damage within the cell.

Possible mechanisms of DNA damage

DNA nicks may be induced through apoptotic processes (Cisternas and Moreno, 2006). Apoptosis regulates germ cell over proliferation and eliminates defective germ cells from the genetic pool (Hikim *et al.*, 1995; Rodriguez *et al.*, 1997; Pentikainen *et al.*, 1999). Apoptosis is characterized by DNA double-stranded breaks which occur as a result of activated endogenous DNA nucleases (Gorczyca *et al.*, 1993). In somatic cells, the apoptotic cascade involves the formation of apoptotic body; however, in highly differentiated spermatozoa, the sequence of events may differ as a result of the highly condensed sperm nucleus (Hikim *et al.*, 1995). Spermatozoa that are marked for apoptotic degradation may have normal mitochondrial activity, high or low motility (Barroso *et al.*, 2000) as well as normal morphology (Host *et al.*, 2000a,b). Oosterhuis *et al.* (2000) reported that 20% of ejaculated spermatozoa showed DNA strand breaks and the apoptotic marker annexin V, whereas, Sakkas *et al.* (1999) reported that DNA strand breaks and apoptotic markers did not co-exist together in the same mature spermatozoa. Ejaculated spermatozoa with apoptotic markers appeared to have escaped programmed cell death in a process called abortive apoptosis (Sakkas *et al.*, 1999). Therefore, it will be important to distinguish between cells that show high levels of DNA strand breaks and cells that are positive for apoptotic markers. It is inappropriate to assume that strand breaks are synonymous with apoptotic degeneration.

DNA damage may also be increased if the DNA nicking and ligating activities of topoisomerase II are defective. The presence of higher than usual levels of topoisomerase II found during the elongating spermatid stage is associated with high levels of DNA nicks (Roca and Mezquita, 1989; McPherson and Longo, 1993), possibly needed to relieve torsional stress caused by the negative supercoiling associated with histone to protamine transition (Balhorn, 1982; Risley *et al.*, 1986; McPherson and Longo, 1993). These nicks are not usually harmful, since they are usually re-ligated prior to completion of spermiogenesis and ejaculation (McPherson and Longo, 1993). However, if the activity of topoisomerase is blocked or disrupted, then DNA nicks remain in mature sperm or are not repaired properly (Morse-Gaudio and Risley, 1994).

Shaman *et al.* have recently demonstrated that topoisomerase II likely acts in two ways. First, it relieves torsional stress by causing double strand breaks which are re-ligated (termed sperm chromatin fragmentation). Second, it acts in conjunction with an extracellular nuclease to cause regulated double-strand breaks in protamine-bound DNA at $\sim$50 kb intervals, the DNA span of one loop bound to protamine (termed sperm DNA degradation) (Sotolongo *et al.*, 2005; Shaman *et al.*, 2006). In the absence of protamines, extensive degradation occurs. This topoisomerase/nuclease-induced DNA degradation may be a specialized apoptotic pathway in sperm, different from the normal function of topoisomerase in relieving torsional stress, followed by re-ligation of the DNA break (Shaman *et al.*, 2006).

Caron *et al.* (2001) suggested that the transient DNA nicks can be repaired by transition protein 1. Transition proteins have been found to have an undefined enzymatic activity that is responsible for repairing single-stranded breaks and UV-induced DNA lesions *in vivo*; therefore, the role of transition proteins extends beyond initiating DNA compaction to restoring transient DNA nicks. Evidence from the literature indicates that the disappearance of single-strand breaks during spermiogenesis is coincident with the presence of the transition proteins in elongating spermatid (Sakkas *et al.*, 1995; Kistler *et al.*, 1996; Smith and Haaf, 1998).

Reactive oxygen species and chromatin damage

In recent years, concern has been expressed about the generation of reactive oxygen species (ROS) in the male reproductive tract. High levels of ROS are toxic to sperm quality and function (Saleh and Agarwal, 2002). Elevated levels of ROS have been reported in 25–40% of the infertile patients (Padron *et al.*, 1997). Two factors that protect the DNA from oxidative stress are tight DNA packaging and antioxidants present in seminal plasma (Twigg *et al.*, 1998). Oxidative stress happens as a result of the imbalance between ROS generation and antioxidants scavenging activities (Sikka, 2001). Strong evidence suggests that the presence of single and double-stranded breaks observed in infertile patients is a result of ROS (Fraga *et al.*, 1996; Kodama *et al.*, 1997; Sun *et al.*, 1997; Aitken and Baker, 2004). The presence of 8-hydroxy-2-deoxyguanosine in seminal plasma has been used as a marker for oxidative DNA damage (Ames *et al.*, 1993). A significant positive correlation was established between DNA fragmentation and ROS (Barroso *et al.*, 2000). Furthermore, exposure of sperm to artificially produced ROS resulted in a significant increase in DNA damage in the form of deletions, frame shifts, DNA cross-links and chromosomal rearrangements (Twigg *et al.*, 1998a,b; Kemal Duru *et al.*, 2000). However, direct studies on ROS-induced damage in protamine deficient sperm have not been performed.

Protamine abnormalities and assisted reproduction techniques

The intrasample heterogeneity of protamine content is clinically significant for patients undergoing assisted reproductive technology (ART) (Aoki *et al.*, 2006c). Protamine-deficient patients undergoing human IVF/ICSI have been shown to have normal embryo quality, implantation and pregnancy rates (Carrell and Liu, 2001; Nasr-Esfahani *et al.*, 2004; Aoki *et al.*, 2005b). This heterogeneity in protamine concentration in sperm of a given semen sample may explain how ICSI appears to overcome poor semen quality and DNA damage (Aoki *et al.*, 2006e). It is possible that selection of the most morphologically normal motile sperm for ICSI injection inherently selects for sperm with normal protamine expression, although studies evaluating sperm morphology and protamines in individual sperm have not been performed. Interestingly, the above findings are inconsistent with the data presented from the mouse protamine-deficient haploinsufficiency model, in which haploinsufficient mice were found to have a higher rate of embryo death when ICSI was performed (Cho *et al.*, 2001, 2003). This difference in embryo lethality, reflecting a more severe effect of abnormal protamine expression in this model, may be the result of a homogeneous pathology throughout the seminiferous tubule rather than the variable expression seen in infertile men with protamine expression defects or may be the result of a lack of effect on sperm morphology in mice.

The long-term consequences of ICSI with DNA-damaged sperm is still not clear (Silber, 1995; Ludwig, 2005; Verpoest and Tournaye, 2006). Animal studies have suggested a sperm DNA damage threshold below which a normal embryo can develop (Ahmadi and Ng, 1999). Others have demonstrated that the oocyte may have the DNA repair system that aids in 'repairing' altered chromatin (Ashwood-Smith and Edwards, 1996; Perry *et al.*, 1999). Another concern regarding the use of sperm with abnormal chromatin is the potential for improper gene imprinting, since protamines have been suggested to be a possible regulator of normal genomic imprinting (Aoki and Carrell, 2003; Oliva, 2006) and since imprinting errors have been suggested to be elevated in patients undergoing ART (Allen and Reardon, 2005; Chang *et al.*, 2005; Ludwig *et al.*, 2005). Hartman *et al.* (2006) have recently noted no increase in imprinting errors in men with severe spermatogenesis defects. Our laboratory has recently reported that defects in global methylation are not observed in men with known protamine abnormalities (Aoki *et al.*, 2006b). Clearly, there is a need for further studies to evaluate specific gene imprinting in those patients and other potential defects in sperm with abnormal protamine replacement and/or DNA damage.

Protamine replacement of histones

Although protamine replacement is often termed 'a two-step process' (histones replaced by transition proteins which are

replaced by protamines), more steps are involved, including the expression and incorporation of testis-specific histone variants, histone hyperacetylation, replacement of histones with transition proteins and protamine incorporation and phosphorylation (Aoki and Carrell, 2003; Churikov *et al.*, 2004b). Each of these steps is critical to proper progression of chromatin maturation and spermiogenesis.

Histone modifications

In somatic cells, nucleosomes are composed of two molecules of histones 2A, 2B, 3 and 4 (H2A, H2B, H3 and H4). Histone 1 (H1 or linker histone) links inter-nucleosomal DNA. In addition to the somatic-type histone variants, spermatogenic cells express testis-specific histones that replace somatic histones (Dadoune, 2003). Although termed 'testis-specific' histones, at least one testis variant, an H3 variant, has been shown to be expressed in somatic cells (Govin *et al.*, 2005).

The characterization of testis-specific histone variants is in the early stages, but several testicular variants have been identified in the human for both the nucleosome and linker histones (Churikov *et al.*, 2004b; Govin *et al.*, 2005). The differences observed between testis-specific histones and somatic histones include structural differences in the N-terminal region, the core region and the C-terminal region. Interestingly, some testis-specific histone variants do not undergo 3′ polyadenylation and are translated early during spermatogenesis (Zalensky *et al.*, 2002; Churikov *et al.*, 2004a). Among those variants is an H2B variant which has been shown to localize in telomeres and may be important in meiosis (Gineitis *et al.*, 2000). Another key difference in the testis variant of H2B is the replacement of four prolines found in the N-terminal region of the somatic H2B with phosphorylatable amino acids, likely indicating that their function is regulated by phosphorylation (Churikov *et al.*, 2004a).

Recently, Zhang *et al.* (2006) have shown that increased levels of histone 2B in sperm is associated with lower levels of protamines. Although previous studies have demonstrated high levels of histones in the sperm of some infertility patients and an indirect link between histone retention and altered sperm protamine expression, this study is the first direct evidence of abnormal histone retention linked to altered protamine replacement. Future studies will likely focus on the further characterization of histone variants and whether abnormal expression of a testis-specific histone variant may be directly responsible for altered protamine replacement. In that regard, Tanaka *et al.* have recently evaluated the gene sequence of *HANP1* in infertile and fertile men. *HANP1* is the human orthologue of the mouse *Hanp1/Hit2* gene that encodes a testes variant of H1; homozygous disruptions of this gene in mice has previously been shown to cause male infertilty (Tanaka *et al.*, 2005, 2006). Although five single nucleotide polymorphisms (SNPs) were identified for *HANP1* in their study population, the SNPs did not appear to be linked to male infertility (Tanaka *et al.*, 2006). Further studies are warranted to evaluate both the protein and the gene in males with known protamine abnormalities.

Hyperacetylation of the histones is critical for normal progression of spermatogenesis and is regulated by an interplay of histone acetyl transferases and histone deacetylases (Candido and Dixon, 1972; Grimes and Henderson, 1984; Meistrich *et al.*, 1992; Hazzouri *et al.*, 2000; Marcon and Boissonneault, 2004). Histone hyperacetylation reduces the binding between nucleosomes and DNA, leading to chromatin relaxation (Hong *et al.*, 1993), and is also associated with the activation of topoisomerases in inducing strand breaks. Species that retain histones throughout spermiogenesis have relatively low levels of acetylated histones (Kennedy and Davies, 1980, 1981). It has been suggested that hyperacetylation of core histones may facilitate their displacement by protamines (Oliva and Mezquita, 1982, 1986; Oliva *et al.*, 1987), and a double bromodomain containing testis-specific factor (BRDT) has been identified in mice as a possible key factor in the transition process. BRDT has been shown to be capable of condensing acetylated chromatin (Pivot-Pajot *et al.*, 2003) by recruiting a highly expressed chaperone protein, CIA-II, to mediate histone removal (Umehara and Horikoshi, 2003) .

Sonnack *et al.* (2002) have demonstrated a relationship between decreased acetylation and abnormal spermatogenesis. They also observed increased acetylation in spermatocytes of testes exhibiting maturation arrest, indicating a possible relationship between premature hyperacetylation and maturation arrest. This same laboratory has also demonstrated that the administration of histone deacetylases results in severe infertility (Fenic *et al.*, 2004). Future studies will likely evaluate the degree of acetylation in protamine-deficient patients and experimental models with abnormal protamine expression.

Replacement of histones with transition proteins

DNA relaxation, as a result of hyperacetylation and topoisomerase activity, facilitates the exchange of histones with transition proteins which are proteins of intermediate basicity (Wilkins, 1956; Courtens and Loir, 1981; Luerssen *et al.*, 1989; Oliva and Dixon, 1991; Ward and Coffey, 1991; Kierszenbaum, 2001; Meistrich *et al.*, 2003). TP1 and TP2 mRNA are first seen in the post-meiotic, round spermatid stage in mice and are degraded around stages, 13–14. The proteins are observed in stages 12 and 13 and are removed by stage 14 when they are replaced with protamines (Figure 2) (Kistler *et al.*, 1996). There is significant overlap in

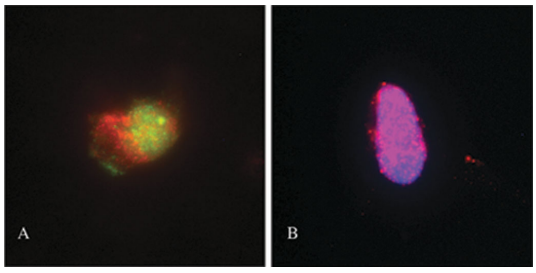


Figure 2. Immunofluorescence microscopy of the localization of DNA, transition protein 1 and protamine 1 in a round spermatid and elongating spermatid. (**A**) In the round spermatid, transition protein 1 is seen within the nucleus (green fluorescence) and protamine 1 is located in the cytoplasm (red fluorescence). (**B**) An elongating spermatid shows no staining for transition protein 1 within the nucleus, although mature protamine is observed throughout the nucleus (red fluorescence) (for methods, see Aoki *et al.*, 2006).

the expression of histones, transition proteins and protamines (Meistrich *et al.*, 2003). Expression of the proteins has been shown to have some overlap in human-elongating spermatids (Aoki *et al.*, 2006c).

TP1 is a 60 kDa protein with numerous basic amino acids distributed randomly throughout the molecule (Kistler *et al.*, 1975). TP1 has important DNA-destabilizing properties due to the presence of two tyrosine residues flanked by basic amino acids (Singh and Rao, 1988). TP2 is a 13 kDa protein that contains proline, serine, arginine and lysine residues (Grimes *et al.*, 1975). Two potential zinc finger domains exist in mouse and rat TP2 and may play an important role in the initiation of chromatin condensation and cessation of transcriptional activity during mammalian spermiogenesis (Baskaran and Rao, 1991). In transgenic mice, animals devoid of transition protein 2 had reduced amounts of processed P2 proteins and failed to complete chromatin compaction (Cho *et al.*, 2003; Meistrich *et al.*, 2003). In addition to incomplete chromatin condensation, sperm from TP2 null mice show an increase in DNA denaturation when compared with sperm from control mice (Zhao *et al.*, 2001). The increased denaturability of the DNA is believed to result from DNA strand breaks (Sailer *et al.*, 1995; Aravindan *et al.*, 1997).

Studies using double knock out mice for both TP1 and TP2 have shown that the absence of one transition protein does not affect the level of transcription or translation of the other transition protein or the protamines, but does affect the retention of the other transition protein through post-translational modifications (Shirley *et al.*, 2004; Zhao *et al.*, 2004). Although the redundancy is not complete, there is compensation for one transition protein by the other, as demonstrated by the fact that double heterozygous mice exhibit more severe sperm defects than do mice homozygous for a single mutation. Interestingly, sperm from transition protein-deficient mice are able to fertilize oocytes using ICSI if the sperm were isolated from the testis or caput epididymus, but are not capable of fertilization if isolated from the cauda epididymus (Suganuma *et al.*, 2005).

Replacement of transition proteins with protamines

Protamine 1 is translated as a mature protein of 50 amino acids, whereas protamine 2 is initially 103 amino acids and undergoes N-terminus cleavage to a mature protein of 57 amino acids (Figure 3) (Aoki and Carrell, 2003). Following translation, protamine 1 is immediately phosphorylated, primarily under the control of serine/arginine protein-specific kinase 1 (SRPK1) (Green *et al.*, 1994; Papoutsopoulou *et al.*, 1999). A protamine 2 intermediate protein is phosphorylated, largely under the control of Ca^{2+}/calmodulin-dependent protein kinase 4 (CAMK4). The phosphorylation of the P2 intermediate is requisite for binding of the protein to chromatin, which is required for final cleavage of the protein to its mature form of 57 amino acids.

Protamine phosphorylation is not only necessary for final processing of P2, but also for proper binding of the proteins to DNA. However, once bound to DNA, the protamines are de-phosphorylated. The de-phosphorylation appears to be essential for proper condensation of the chromatin, although some controversy exists (Gusse *et al.*, 1986; Aoki and Carrell, 2003). A recent study evaluating the effects of organophosphorous pesticides on sperm protamine phosphorylation showed that the resulting sperm had abnormal chromatin condensation with subsequent DNA damage (Pina-Guzman *et al.*, 2005). Abnormal phosphorylation may be relevant to human exposure to organophosphorous pesticides (Sanchez-Pena *et al.*, 2004; Pina-Guzman *et al.*, 2005, 2006).

The replacement of transition proteins with protamines induces a conformational change in the packaging of the chromatin. The chromatin forms loop domains, which are less than half the size of somatic cell histone loops, then forms toroidal structures, which have a 6–20-fold increase in packaging compaction (Ward and Coffey, 1991; Balhorn *et al.*, 2000). The mechanism by which protamines induce the conformational changes is not well understood (D'Auria *et al.*, 1993; Bianchi *et al.*, 1994; Fuentes-Mascorro *et al.*, 2000; Aoki and Carrell, 2003). P1 and P2 may bind to the major and minor groove of DNA or to the DNA surface by interacting electrostatically with phosphate residues (D'Auria *et al.*, 1993; Bianchi *et al.*, 1994; Balhorn *et al.*, 1999; Fuentes-Mascorro *et al.*, 2000).

Protamines are currently thought to be necessary for (i) condensing the male genome to generate a more compact and hydrodynamic nucleus, (ii) protecting the genetic message from nucleases, mutagens or damage from ROS or other factors, (iii) epigenetic modification during spermiogenesis and (iv) removing transcription factors and proteins to help reset the imprinting code in the oocyte (Oliva, 2006). Altering the sperm protamine content can disrupt any of the functions listed above.

Regulation of protamine expression and potential causes of abnormalities

The human sperm haploid genome encodes a single copy of human P1 and P2 genes which maps to chromosome 16p13.3 (Domenjoud *et al.*, 1991). This locus also contains the TP2 gene. The P1–P2–TP2 locus spans a 28.5 kb region in which the three genes are arranged in a linear array, presumably facilitating concurrent or co-ordinated gene expression (Schluter *et al.*, 1992; Choudhary *et al.*, 1995). The P1 gene is present in all mammalian species, whereas P2 is present in mouse, hamster, rat, stallion and man (Calvin, 1976; McKay *et al.*, 1985, 1986; Poccia, 1986; Bower *et al.*, 1987; de Yebra *et al.*, 1993). However, in some species, the *protamine 2* gene is present but the protein is absent (Maier *et al.*, 1990) . These findings suggest that *P1* is inherited from a common ancestor, since it is present in all species and that the *P2* gene may have been derived from the *P1* gene through divergent evolution. Alternatively, the *P1* and *P2* genes may have been inherited from a single common ancestor, but successive species have lost the ability to express *protamine 2* (Calvin, 1976; Oliva and Dixon, 1990).

Several factors have been postulated and studied as possible causes of P1/P2 deregulation. These factors are summarized in Table I and discussed in the following sections.

Gene polymorphisms

The protamine or transition protein genes could harbour mutations or polymorphisms that could induce conformational changes in the proteins, which could alter their incorporation into sperm chromatin. De Yebra *et al.* (1993) performed a preliminary mutational analysis of the protamine gene analysis in four patients with

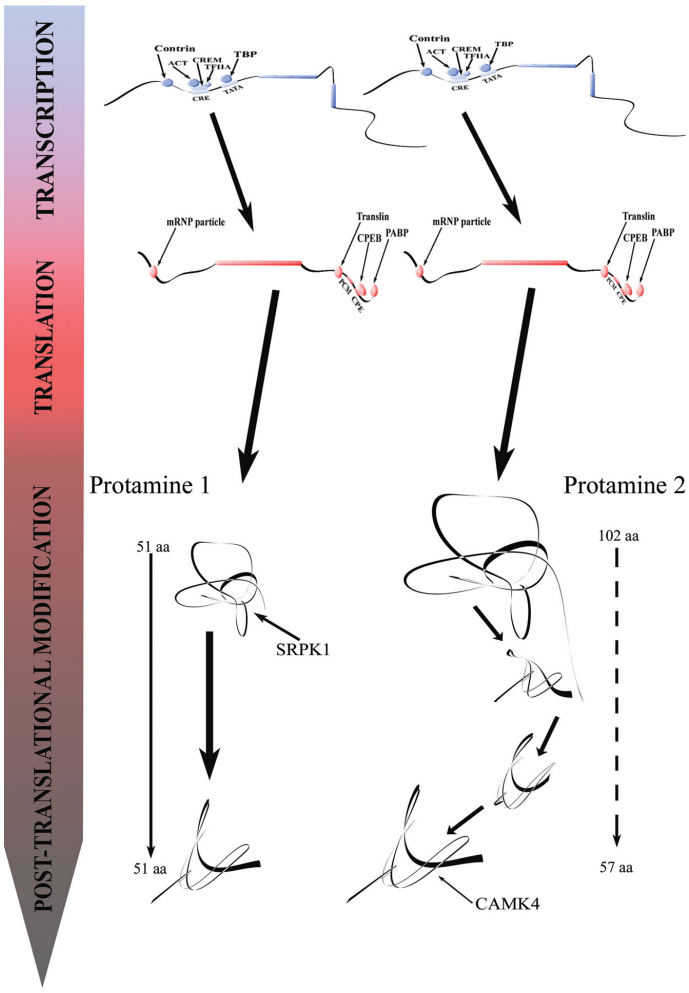


Figure 3. A schematic drawing of key regulators of expression of protamines 1 and 2 at the levels of genomic transcription, translation and post-translational modifications to yield the mature protamine products. The key regulating protein complexes are shown for each stage of protein maturation. CREM, cAMP response element modulator; ACT, activator of CREM in the testis; TFIIA, transcription factor II alpha; CRE, CREM response element; TBP, TATA box binding protein; PABP, poly-a binding protein; CPE, cytoplasmic polyadenylation element; CPEB, cytoplasmic polyadenylation element binding protein; PCM, polyadenlyation consensus motif; SRPK1, serine/arginine protein-specific kinase-1; CAMK4, calcium/calmodulin-dependent protein kinase IV.

markedly altered P1/P2 ratios with no mutations observed. Subsequently, Schlicker *et al.* (1994) screened 36 infertile patients with chromatin anomalies, but he failed to identify any mutations in the genes encoding P1, P2 or TP1. Tanaka *et al.* (2003) reported four synonymous polymorphisms in P1 and one SNP in P2 that generated a premature stop codon. The SNP in the P2 gene that induces translation termination may result in male infertility due to haploinsufficiency of P2.

Iguchi *et al.* (2006) sequenced the protamine genes in men exhibiting semen quality defects consistent with protamine abnormalities (i.e. sperm DNA damage). In their study, a heterozygous SNP which altered a highly conserved arginine residue was found in 10% (3/30) of the patients studied, but not seen in controls. This SNP converts one of the highly conserved arginines to a serine residue, therefore creating an RS sequence which can serve as a potential phosphorylation site for the enzyme SRPK1. Improper phosphorylation can substantially alter both DNA binding and protamine-to-protamine interaction in the sperm nucleus.

Recently, a larger patient population with known abnormal protamine ratios was screened to identify SNPs in the protamine and transition protein genes potentially responsible for the patients' altered protamine expression (Aoki *et al.*, 2006a). Fifteen SNPs were identified in this study (three SNPs in P1, seven in P2, two

Table I. Studies that have identified potential causes for P1/P2 dysregulation

Gene studies	
Protamines	Maier *et al.* (1990)
	Cho *et al.* (2003)
	Cho *et al.* (2001)
	Tanaka *et al.* (2003)
	Aoki *et al.* (2006)
	Iguchi *et al.* (2006)
	Queralt *et al.* (2003)
	Schnulle *et al.* (2004)
Transition proteins	Yu *et al.* (2000)
	Adham *et al.* (2001)
	Meistrich *et al.* (2003)
	Shirley *et al.* (2004)
	Suganuma *et al.* (2005)
	Zhao *et al.* (2004)
Protamine UTR	Our laboratory's unpublished results
	De Jonckheere *et al.* (1994)
Ancillary genes	
CAMK4/SRPK1	Wu *et al.* (2000)
AF5q31	Urano *et al.* (2005)
MARs	Wykes *et al.* (2003)
	Martins *et al.* (2004)
Transcription	
CREM	Steger (1999)
	Behr *et al.* (2000)
	Bleny *et al.* (1996)
	Weinbaur *et al.* (1998)
ACT /KIF17B	Macho *et al.* (2002)
	Kotaja *et al.* (2004)
MSY2	Yang *et al.* (2005b)
Storage	
Protamine 3′ UTR	Lee *et al.* (1995)
	Fajardo *et al.* (1994, 1997)
	Kwon and Hecht (1993)
Translation	
PRBP/TRBP	Zhong *et al.* (1999)
	Braun *et al.* (2000)
ACT /KIF17B	Macho *et al.* (2002)
	Kotaja *et al.* (2004)

A summary of studies investigating potential abnormalities that may underlie aberrant P1 and P2 expression. These include mutations in the protamine genes or in ancillary genes, defects in transcriptional regulatory mechanisms, alterations in RNA binding proteins or RNA binding protein sites and aberrant translational regulatory factors.

in TP1 and three in TP2); however, the frequencies of these SNPs were similar in protamine-deficient patients, severely infertile patients without protamine defects and fertile controls. The SNPs identified in this study included differences from the SNPs reported by Tanaka *et al.* (2003). It is important to note that the differences may reflect differences in the populations studied (European ancestry versus Japanese ancestry). Also, it is important to note that the P1 SNP reported by Iguchi et al. was not elevated in the study by Aoki *et al.* (2006a) (Iguchi *et al.*, 2006). In summary, these studies indicate that gene SNPs in the protamine and transition protein genes are not likely to be a common cause of protamine abnormalities.

In addition to evaluating the coding gene sequences, it is important to consider possible mutations in the upstream and downstream non-coding regions of the gene. De Jonckheere *et al.* (1994) examined potential mutations in the untranslated regions (UTRs) of the protamine genes and identified a candidate mutation in the GA repeat upstream of the transcriptional start site of P2. Recently, our laboratory identified 14 SNPs in the UTRs of P1 and P2 (Emery *et al.*, 2006, submitted for publication). Of greatest significance was a GC change 62 bp into the 3′-UTR of P2, which occurred in five infertility patients and in six of the Utah CEPH database patients. Interestingly, none of the men with confirmed normal protamine ratios, or the fertile sperm donors, carried this change. This G/C SNP was found in the sperm cDNAs and in the genomic sequence. We also noted the previously identified GA repeat in the genomic P2 5′-UTR. The variable length GA repeat length occurs with equal frequency in UGRP controls and in men with abnormal protamine expression. The most prevalent GA repeat lengths were 12, 15 and 18 bp.

Regulation of transcription

Protamines are expressed in the round spermatid stage, but protamine transcription and translation are temporarily uncoupled in the developing spermatid due to transcriptional and translational regulatory mechanisms (Calvin, 1976; Sassone-Corsi, 2002). Transcriptional regulation depends on potentiation of the genes via association with nuclear matrix attachment regions (MARs), and binding of *trans*-acting factors to the promoter region (Martins *et al.*, 2004; McCarrey *et al.*, 2005).

The 5′ and 3′ regions surrounding P1/P2/TP2 genes contain MARs which are *cis*-regulatory units involved in attachment of the DNA to the protein scaffolding of the nuclear matrix in an organized manner of loop domains which potentiates the genes for transcription dependent on other *trans*-regulatory factors (Martins *et al.*, 2004). The MARs are located at the linker sites of protamine toroids and contain repetitive elements which may, contrary to the usual paradigm, be highly methylated in the round spermatid stage in which the genes undergo transcription (Choi *et al.*, 1997; Shaman *et al.*, 2006). Kramer *et al.* (1997) scanned the P1–P2–TP2 locus in several oligozoospermic, infertile individuals and identified mutations in the sperm nuclear matrix of two of the five affected individuals. This laboratory later used transgenic analysis to evaluate the relationship between the nuclear matrix associations and the expression of protamine genes (Martins *et al.*, 2004). This study demonstrated that in the absence of the 5′-MAR and with the presence of only the 3′-MAR, protamine transcription was reduced but not ablated. This suggests that the 3′-MAR provides a protective role against silencing of the protamine genes and that synergy between the upstream and downstream MARS is required for the proper regulation of the protamine genes (Martins *et al.*, 2004).

Trans-regulatory factors of transcription act via the promoter region and include the TATA-box protein (TBP), cAMP response element modulator (CREM) and Y-box proteins (Steger *et al.*, 2000; Maclean and Wilkinson, 2005; Tanaka and Baba, 2005; DeJong, 2006). TBPs bind to the TATA sequence in conjunction with other factors to facilitate RNA polymerase 2 interaction with the gene sequence. TBP is not essential for gene expression, but is important in initiating transcription (Schmidt and Schibler, 1997; Kimmins *et al.*, 2004a). TBP over-expression occurs between 18 and 28 days in mice, which corresponds with the transcription of genes in haploid cells (Schmidt and Schibler, 1995). TBP-like factor (TLF) has a sequence similar to TBP and a similarly unique expression profile in the testis, but is unique in that it binds to TATA-less promoters such as pes-10 (Kaltenbach *et al.*, 2000). TLF is seen in the cytoplasm during

the early pachytene stage, then enters the nucleus and remains there throughout the round spermatid stage until it moves to the cytoplasm again in the elongating spermatid stage when transcription is terminated (Martianov *et al.*, 2002). TLF may act as both an activator and a repressor of transcription (Moore *et al.*, 1999). TLF null mice have abnormal heterochromatin organization, which may affect protamine replacement and chromatin condensation and induce apoptosis (Martianov *et al.*, 2001, 2002).

The transcription factor CREM is highly expressed in male germ cells (Delmas *et al.*, 1993) and is known to regulate the expression of several post-meiotic genes, such as the transition proteins and protamines, and is likely the key regulator of gene expression during spermatogenesis (Krausz and Sassone-Corsi, 2005; Hogeveen and Sassone-Corsi, 2006). Targeted disruption of the CREM gene blocks the differentiation program in the first step of spermiogenesis (Blendy *et al.*, 1996; Nantel *et al.*, 1996). These findings indicate a crucial role of CREM in post-meiotic germ cell differentiation, linking the action of hormonal stimuli to direct regulation of spermatogenesis genes (Sassone-Corsi, 1998).

The cAMP response element (CRE) is found in the promoter region of the protamine genes and it serves as a binding site for the transcriptional activator CREM tau in the testes. CREM tau is associated with the activator of CREM (ACT) and the testis-specific kinesin protein KIF 17B. In addition to serving as a molecular motor for mRNA transport, KIF 17B also acts as a co-factor to ACT, and hence as an activator for CREM-mediated transcription. Interestingly, all mRNAs isolated by co-immunoprecipitation with KIF 17B and testis–brain RNA-binding protein (TB RBP) are transcripts which are regulated through CREM activity (Chennathukuzhi *et al.*, 2003a). By transport of the ribonucleoprotein (RNP) to the cytoplasm, KIF 17 B/ACT is not available in the nucleus, and hence CREM-regulated transcription is repressed (Kimmins *et al.*, 2004b; Krausz and Sassone-Corsi, 2005). The linkage of KIF 17B to regulation of both CREM-mediated transcription and temporal regulation of translation indicates the major role this protein plays in the regulation of the expression of protamines and other key spermiogenesis proteins (Nagamori *et al.*, 2006; Sassone-Corsi, 2005).

Male mice lacking a functional *CREM* gene are sterile due to maturation arrest at the round spermatid stage (Blendy *et al.*, 1996; Nantel *et al.*, 1996). Mice with selective ACT deletion display a drastic decrease in the number of mature sperm and exhibit major defects in sperm head morphology (chromatin compaction and acrosome defects) and tail morphology (Kotaja *et al.*, 2004). Our laboratory recently identified potentially significant polymorphisms of the *ACT* gene in men with severe infertility (Christensen *et al.*, 2006), but studies in patients with known protamine expression abnormalities have not yet been completed.

The role of Y-box proteins in the regulation of protamine expression

Y-box proteins are important regulators of protein expression during spermatogenesis. They bind both DNA and RNA and, generally, up-regulate transcription and down-regulate translation. Several Y-box proteins have been identified in the germ cells of mice and humans (Braun, 1990; Tafuri *et al.*, 1993; Yiu and Hecht, 1997; Tekur *et al.*, 1999; Iuchi *et al.*, 2001) Among the Y-box proteins active during spermatogenesis is Contrin and its mouse orthologue MSY2, which are found in very high concentrations in male germ cells (Gu *et al.*, 1998; Tekur *et al.*, 1999). Two functions have been proposed for Contrin/MSY2 RNA-binding protein in regard to protamine expression (Yang *et al.*, 2005b). First, it serves as a co-activator for protamine transcription by binding to the Y box element sequence found in the gene promoter region. Second, it stabilizes the maternal and paternal transcripts in the cytoplasm to effect the temporal regulation of translation as described above. Yang et al. generated an MSY2 knockout model to define the function of MSY2 in mammalian development (Yang *et al.*, 2006). MSY2-null male mice were sterile, with severely amorphous and multinucleated spermatids. The sperm were observed in the seminiferous tubules, but not in the epididymus (Yang *et al.*, 2006).

Translin is another DNA/RNA-binding protein involved in translational regulation during spermatogenesis. There is a high degree of similarity between Translin and its mouse orthologue, TB RBP. Like Contrin, Translin binds to specific mRNAs in the testis, forming an RNP complex which is transported to the cytoplasm and adjacent cells through intercytoplasmic bridges (Morales *et al.*, 2002). TB RBP interacts with a protein termed Translin-associated factor X that has a nuclear localization signal, although TB RBP has a cytoplasmic localization signal. The ratio of the two proteins determines the actual localization of the complex (Cho *et al.*, 2004). KIF 17B is essential for proper movement of the Contrin and Translin-containing RNP particles to the cytoplasm and through adjacent cytoplasmic bridges (Chennathukuzhi *et al.*, 2003a). Mice lacking the TB RBP gene are able to sire offspring, but have reduced sperm production (Chennathukuzhi *et al.*, 2003b). The direct effect of TB RBP gene knock-out on protamine expression is not known (Kimmins *et al.*, 2004a; Sassone-Corsi, 2005).

Regulation of translation

Protamines 1 and 2 undergo translational control during spermiogenesis (Steger, 2001). Protamines are transcribed in the round spermatid stage, but the mRNAs are stored in translationally repressed RNPs in early haploid cells and activated in elongated spermatids (Kleene, 1989) . This stage-specific pattern of gene expression is essential for correct sequential nucleoprotein exchange and complete differentiation of round spermatids into mature spermatozoa. Delaying mRNA translation prevents new mRNA synthesis in elongated spermatids, precocious chromatin condensation and infertility (Lee *et al.*, 1995).

The mechanism by which translational regulation is operating is not entirely understood; nevertheless, a few regulatory sites in the 3′-UTR have been identified to have profound effects on translation. The Y-box proteins Contrin and Translin and the kinesin KIF 17B are essential for transport of the RNP to the cytoplasm and for delayed translation (Kwon and Hecht, 1991, 1993; Murray *et al.*, 1992; Fajardo *et al.*, 1994; Aoki *et al.*, 1995; Fajardo *et al.*, 1997; Iuchi *et al.*, 2001). Poly-A-binding protein also has two important roles in regulating translation. First, it protects the mRNA transcript from degradation, thereby preserving the transcript until translational repression is removed. Second, it serves as a repressor protein (Bernstein *et al.*, 1989). Protamine

translation is initiated by covalent modification of the RNP, which leads to the release of the mRNA from the RNP and the shortening of the poly-A-tail (Kleene, 2003).

One interesting protein which may be involved in translational regulation of the protamines is protamine 1 RNA-binding protein (PRBP) and its human orthologue TAR RNA-binding protein (TRBP) (Lee *et al.*, 1996; Siffroi *et al.*, 2001). PRBP is highly expressed in the differentiating male germ cell and contains two copies of a double-stranded, RNA-binding domain which binds to the 3′-UTR of protamine 1 (Lee *et al.*, 1996). Male mice with PRBP deletion are infertile and severely oligozoospermic due to failure of protamine transcripts to undergo translation, which results in delayed replacement of the transition proteins and subsequent failure of spermatogenesis (Zhong *et al.*, 1999). TRBP is a 43 kDa protein found in the cytoplasm of elongating spermatids. No studies have evaluated this protein in men with known protamine defects (Siffroi *et al.*, 2001).

Post-translation modifications

P1 is translated as a mature protein, whereas P2 is synthesized as a 103 aa precursor that undergoes proteolytic cleavage at the N-terminus to generate the mature form (Meistrich *et al.*, 1992; Oliva and Mezquita, 1982; Balhorn *et al.*, 1999). P1 is rapidly phosphorylated by SRPK1 after translation, whereas CAMK4 phosphorylates P2 protein subsequent to its proteolytic modifications and its binding to the DNA (Green *et al.*, 1994; Papoutsopoulou *et al.*, 1999). Phosphorylation of these proteins ensures their proper binding to DNA (Figure 3). Mutations or polymorphisms of SRPK1 or CAMK4 may be potential causes of abnormal protamine expression, but have yet to be evaluated. Targeted mutations in the CAMK4 gene in male mice resulted in infertility, a specific loss of P2 and retention of TP2, which suggests that the mature, phosphorylated P2 protein must interact with chromatin to displace TP2 (Wu *et al.*, 2000). Yoshii *et al.* (2005) have recently performed 2D electrophoresis and identified five bands for variants of protamine 1 and six bands for protamine 2 variants. This technique may be valuable in identifying post-translational defects of protamines in patients with known defects of protamine expression.

The relationship between abnormal protamine expression and spermatogenesis

Two caveats are relevant in considering the relationship between abnormal protamine expression and spermatogenesis. First, abnormal protamine expression is relatively common in male infertility patients, but rare in men with known fertility (Carrell and Liu, 2001). Second, many patients with abnormal protamine expression exhibit severe defects of semen quality, including oligozoospermia (Carrell and Liu, 2001; Aoki and Carrell, 2003; Aoki *et al.*, 2005a). The link between abnormal protamine expression and severely altered spermatogenesis is not intuitively obvious, but animal studies have also shown that spermatogenesis is severely altered when protamine expression is experimentally reduced (Zhong *et al.*, 1999).

One possible explanation for the possible link between altered protamine expression and severely reduced spermatogenesis is the hypothesis that protamine expression may act as a ‘checkpoint’ during spermiogenesis and that abnormal protamine expression leads to an increased level of apoptosis. This hypothesis is supported by the fact that protamine haploinsufficiency causes severe disturbances of spermatogenesis in the mouse and the fact that protamine alterations are associated with DNA damage which may result in the initiation of an apoptosis pathway (Cho *et al.*, 2001, 2003). Differences in microenvironments within the seminiferous tubules or incomplete (abortive) apoptosis may allow some sperm survival.

An alternative hypothesis may be that abnormal protamine expression is usually the result of an abnormal functioning of a regulator of transcription, translation or post-translational modifications that affects not only the protamines, but also a broad range of genes involved in spermatogenesis. Candidate regulators may include ACT, CREM, Translin, Contrin and KIF17B. Particularly attractive targets in this regard may be the Y-box proteins Translin and Contrin, and the associated kinesin, KIF 17B, because they are involved not only in transcriptional regulation, but also in transport to the cytoplasm and translational regulation, since the data indicate that patients with abnormal protamine expression often have higher than normal levels of transcript in mature sperm (Aoki *et al.*, 2006d). Other key regulators, such as phosphorylation mechanisms, could be responsible for a broad defect affecting spermatogenesis.

Conclusions

From studies performed to date, it is clear that the presence of an altered P1/P2 ratio is clinically relevant and portends a reduced fertility. Semen samples with altered P1/P2 ratios generally have other abnormalities, such as increased DNA damage, low sperm counts and reduced fertilizing capacity (Aoki *et al.*, 2006e; Oliva, 2006). Two over-riding questions should be addressed in future studies. First, what is the clinical relevance, beyond low sperm quality, of abnormal sperm protamine expression? and, second, what is the cause of abnormal protamine expression?

When undergoing IVF with ICSI, men with abnormal protamine expression have fertilization, implantation and pregnancy rates equal to patients undergoing IVF for other diagnoses, including obstructive azoospermia (Aoki *et al.*, 2006e). However, given the role of protamines in silencing and resetting the paternal genome, and given the controversial, but possible relationship between increased imprinting errors in offspring and ICSI, future studies are needed to evaluate the potential risks of undergoing ART with sperm of known protamine abnormalities (Allen and Reardon, 2005; Chang *et al.*, 2005; Ludwig, 2005). Those studies will likely include an evaluation of offspring derived from animal models and also include further human sperm studies, such as the imprinting status of specific genes. Additionally, preliminary studies indicate that the P1/P2 ratio may be a more sensitive and accurate predictor of sperm functional ability than other available sperm function assays (Aoki *et al.*, 2006e). This concept should be further analysed.

Second, abnormal protamine expression is an intriguing pathology due to its relationship with altered spermatogenesis. Future studies will include analysis of regulators of transcription and translation and post-translational modifications, areas that will likely yield an increased understanding of the cause of abnormal protamine expression, and may also provide important

information regarding spermatogenesis in general. Protamines may be an important marker in better understanding the key regulatory pathways of spermatogenesis, especially if abnormal protamine expression reflects a aberrant function of a key transcription of translation regulator or acts as a crucial part of a checkpoint pathway.

References

Adham IM, Nayernia K, Burkhardt-Gottges E, Topaloglu O, Dixkens C, Holstein AF and Engel W (2001) Teratozoospermia in mice lacking the transition protein 2 (Tnp2). Mol Hum Reprod 7,513–520.

Ahmadi A and Ng SC (1999) Fertilizing ability of DNA-damaged spermatozoa. J Exp Zool 284,696–704.

Aitken RJ and Baker MA (2004) Oxidative stress and male reproductive biology. Reprod Fertil Dev 16,581–588.

Allen C and Reardon W (2005) Assisted reproduction technology and defects of genomic imprinting. BJOG 112,1589–1594.

Ames BN, Shigenaga MK and Hagen TM (1993) Oxidants, antioxidants, and the degenerative diseases of aging. Proc Natl Acad Sci U S A 90,7915–7922.

Aoki VW and Carrell DT (2003) Human protamines and the developing spermatid: their structure, function, expression and relationship with male infertility. Asian J Androl 5,315–324.

Aoki K, Suzuki K, Sugano T, Tasaka T, Nakahara K, Kuge O, Omori A and Kasai M (1995) A novel gene, Translin, encodes a recombination hotspot binding protein associated with chromosomal translocations. Nat Genet 10,167–174.

Aoki VW, Liu L and Carrell DT (2005a) Identification and evaluation of a novel sperm protamine abnormality in a population of infertile males. Hum Reprod 20,1298–1306.

Aoki VW, Moskovtsev S, Willis J, Liu L, Mullen JBM and Carrell DT (2005b) DNA integrity is compromised in protamine-deficient sperm. J Androl 26,741–748.

Aoki VW, Christensen GL, Atkins JF and Carrell DT (2006a) Identification of novel polymorphisms in the nuclear protein genes and their reltionship with human sperm protamine deficiency and severe male infertility. Fertil Steril 86,1416–1422.

Aoki VW, Emery BR and Carrell DT (2006b) Global sperm DNA methylation is unaffected in protamine-deficient infertile males. Fertil Steril 86,1541–1543.

Aoki VW, Emery BR, Liu L and Carrell DT (2006c) Protamine levels vary between individual sperm cells of infertile human males and correlate with viability and DNA integrity. J Androl 27,890–898.

Aoki VW, Liu L and Carrell DT (2006d) A novel mechanism of protamine expression deregulation highlighted by abnormal protamine transcript retention in infertile human males with sperm protamine deficiency. Mol Hum Rep 12,41–50.

Aoki VW, Liu L, Jones KP, Hatasaka HH, Gibson M, Peterson CM and Carrell DT (2006e) Sperm protamine 1/protamine 2 ratios are related to IVF pregnancy rates and predictive of sperm fertilizing ability. Fertil Steril 86,1408–1415.

Aravindan GR, Bjordahl J, Jost LK and Evenson DP (1997) Susceptibility of human sperm to in situ DNA denaturation is strongly correlated with DNA strand breaks identified by single-cell electrophoresis. Exp Cell Res 236,231–237.

Ashwood-Smith MJ and Edwards RG (1996) DNA repair by oocytes. Mol Hum Reprod 2,46–51.

Balhorn R (1982) A model for the structure of chromatin in mammalian sperm. J Cell Biol 93,298–305.

Balhorn R, Reed S and Tanphaichitr N (1988) Aberrant protamine 1/protamine 2 ratios in sperm of infertile human males. Experientia 44,52–55.

Balhorn R, Cosman M, Thornton K, Krishnan VV, Corzett M, Bench G, Kramer C, Lee J, Hud NV and Allen MJ *et al.* (1999) In *The Male Gamete: From Basic Science to Clinical Applications*. Gagnon C (eds.). Cache River Press, Vienna, IL.

Balhorn R, Brewer L and Corzett M (2000) DNA condensation by protamine and arginine-rich peptides: analysis of toroid stability using single DNA molecules. Mol Reprod Dev 56,230–234.

Barroso G, Morshedi M and Oehninger S (2000) Analysis of DNA fragmentation, plasma membrane translocation of phosphatidylserine and oxidative stress in human spermatozoa. Hum Reprod 15,1338–1344.

Baskaran R and Rao MR (1991) Mammalian spermatid specific protein, TP2, is a zinc metalloprotein with two finger motifs. Biochem Biophys Res Commun 179,1491–1499.

Behr R and Weinbauer GF (2000) CREM activator and repressor isoforms in human testis: sequence variations and inaccurate splicing during impaired spermatogenesis. Mol Hum Reprod 6,967–972.

Belokopytova IA, Kostyleva EI, Tomilin AN and Vorob'ev VI (1993) Human male infertility may be due to a decrease of the protamine P2 content in sperm chromatin. Mol Reprod Dev 34,53–57.

Bianchi F, Rousseaux-Prevost R, Bailly C and Rousseaux J (1994) Interaction of human P1 and P2 protamines with DNA. Biochem Biophys Res Commun 201,1197–1204.

Blendy JA, Kaestner KH, Weinbauer GF, Nieschlag E and Schutz G (1996) Severe impairment of spermatogenesis in mice lacking the CREM gene. Nature 380,162–165.

Bower PA, Yelick PC and Hecht NB (1987) Both P1 and P2 protamine genes are expressed in mouse, hamster, and rat. Biol Reprod 37,479–488.

Braun RE (1990) Temporal translational regulation of the protamine 1 gene during mouse spermatogenesis. Enzyme 44,120–128.

Braun RE (2000) Temporal control of protein synthesis during spermatogenesis. Int J Androl 23 (Suppl 2), 92–94.

Brewer L, Corzett M and Balhorn R (2002) Condensation of DNA by spermatid basic nuclear proteins. Journal of Biological Chemistry 277,38895–38900.

Calvin HI (1976) Comparative analysis of the nuclear basic proteins in rat, human, guinea pig, mouse and rabbit spermatozoa. Biochimica et Biophysica Acta 434,377–389.

Candido EP and Dixon GH (1972) Trout testis cells. 3. Acetylation of histones in different cell types from developing trout testis. J Biol Chem 247,5506–5510.

Caron N, Veilleux S and Boissonneault G (2001) Stimulation of DNA repair by the spermatidal TP1 protein. Mol Reprod Dev 58,437–443.

Carrell D and Liu L (2001) Altered protamine 2 expression is uncommon in donors of known fertility, but common among men with poor fertilizing capacity, and may reflect other abnormalities of spermiogenesis. J Androl 22,604–610.

Chang AS, Moley KH, Wangler M, Feinberg AP and Debaun MR (2005) Association between Beckwith-Wiedemann syndrome and assisted reproductive technology. Fertil Steril 83,349–354.

Chennathukuzhi V, Morales CR, El-Alfy M and Hecht NB (2003a) The kinesin KIF17b and RNA-binding protein TB-RBP transport specific cAMP-responsive element modulator-regulated mRNAs in male germ cells. Proc Natl Acad Sci USA 100,15566–15571.

Chennathukuzhi V, Stein JM, Abel T, Donlon S, Yang S, Miller JP, Allman DM, Simmons RA and Hecht NB (2003b) Mice deficient for testis-brain RNA-binding protein exhibit a coordinate loss of TRAX, reduced fertility, altered gene expression in the brain, and behavioral changes. Mol Cell Biol 23,6419–6434.

Chevaillier P, Mauro N, Feneux D, Jouannet P and David G (1987) Anomalous protein complement of sperm nuclei in some infertile men. Lancet 2,806–807.

Cho C, Willis WD, Goulding EH, Jung-Ha H, Choi YC, Hecht NB and Eddy EM (2001) Haploinsufficiency of protamine-1 or -2 causes infertility in mice. Nat Genet 28,82–86.

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

Cho YS, Chennathukuzhi VM, Handel MA, Eppig J and Hecht NB (2004) The relative levels of translin-associated factor X (TRAX) and testis brain RNA-binding protein determine their nucleocytoplasmic distribution in male germ cells. J Biol Chem 279,31514–31523.

Choi YC, Aizawa A and Hecht NB (1997) Genomic analysis of the mouse protamine 1, protamine 2, and transition protein 2 gene cluster reveals hypermethylation in expressing cells. Mamm Genome 8,317–323.

Choudhary SK, Wykes SM, Kramer JA, Mohamed AN, Koppitch F, Nelson JE and Krawetz SA (1995) A haploid expressed gene cluster exists as a single chromatin domain in human sperm. Journal of Biological Chemistry 270,8755–8762.

Christensen GL, Wooding SP, Ivanov IP, Atkins JF and Carrell DT (2006) Sequencing and haplotype analysis of the activator of CREM in the testis (ACT) gene in populations of fertile and infertile males. Molec Hum Reprod 12,257–262.

Churikov D, Siino J, Svetlova M, Zhang K, Gineitis A, Morton Bradbury E and Zalensky A (2004a) Novel human testis-specific histone H2B encoded by the interrupted gene on the X chromosome. Genomics 84,745–56.

Churikov D, Zalenskaya IA and Zalensky AO (2004b) Male germline-specific histones in mouse and man. Cytogenet Genome Res 105,203–214.

Cisternas P and Moreno RD (2006) Comparative analysis of apoptotic pathways in rat, mouse, and hamster spermatozoa. Mol Reprod Dev 73,1318–1325.

Corzett M, Mazrimas J and Balhorn R (2002) Protamine 1: protamine 2 stoichiometry in the sperm of eutherian mammals. Mol Reprod Dev 61,519–527.

Courtens JL and Loir M (1981) Ultrastructural detection of basic nucleoproteins: alcoholic phosphotungstic acid does not bind to arginine residues. J Ultrastruct Res 74,322–326.

Dadoune JP (1995) The nuclear status of human sperm cells. Micron 26,323–345.

Dadoune JP (2003) Expression of mammalian spermatozoal nucleoproteins. Microsc Res Tech 61,56–75.

D'Auria G, Paolillo L, Sartorio R and Wurzburger S (1993) Structure and function of protamines: an 1H nuclear magnetic resonance investigation of the interaction of clupeines with mononucleotides. Biochim Biophys Acta 1162,209–216.

De Jonckheere J, Nelson JE, Ginsburg KA, Martin L and Krawetz SA (1994) GA repeat polymorphism at the PRM2 male fertility locus. Hum Mol Genet 3,1915.

DeJong J (2006) Basic mechanisms for the control of germ cell gene expression. Gene 366,39–50.

Delmas V, van der Hoorn F, Mellstrom B, Jegou B and Sassone-Corsi P (1993) Induction of CREM activator proteins in spermatids: down-stream targets and implications for haploid germ cell differentiation. Mol Endocrinol 7,1502–1514.

Dixon GH, Aiken JM, Jankowski JM, McKenzie D, Moir R and States JC (1986) Organization and evolution of the protamine genes of salmoind fishes. In Reeck GR, Goodwin GH and Puigdomenech P (eds.). *Chromosomal Proteins and Gene Expression*. Plenum Publishing Corp., New York, NY.

Domenjoud L, Kremling H, Burfeind P, Maier WM and Engel W (1991) On the expression of protamine genes in the testis of man and other mammals. Andrologia 23,333–337.

Fajardo MA, Butner KA, Lee K and Braun RE (1994) Germ cell-specific proteins interact with the 3′ untranslated regions of Prm-1 and Prm-2 mRNA. Dev Biol 166,643–653.

Fajardo MA, Haugen HS, Clegg CH and Braun RE (1997) Separate elements in the 3′ untranslated region of the mouse protamine 1 mRNA regulate translational repression and activation during murine spermatogenesis. Dev Biol 191,42–52.

Fenic I, Sonnack V, Failing K, Bergmann M and Steger K (2004) In vivo effects of histone-deacetylase inhibitor trichostatin-A on murine spermatogenesis. J Androl 25,811–818.

Fraga CG, Motchnik PA, Wyrobek AJ, Rempel DM and Ames BN (1996) Smoking and low antioxidant levels increase oxidative damage to sperm DNA. Mutat Res 351,199–203.

Fuentes-Mascorro G, Serrano H and Rosado A (2000) Sperm chromatin. Arch Androl 45,215–225.

Gatewood JM, Cook GR, Balhorn R, Bradbury EM and Schmid CW (1987) Sequence-specific packaging of DNA in human sperm chromatin. Science 236,962–964.

Gineitis AA, Zalenskaya IA, Yau PM, Bradbury EM and Zalensky AO (2000) Human sperm telomere-binding complex involves histone H2B and secures telomere membrane attachment. J Cell Biol 151,1591–1598.

Gorczyca W, Traganos F, Jesionowska H and Darzynkiewicz Z (1993) Presence of DNA strand breaks and increased sensitivity of DNA in situ to denaturation in abnormal human sperm cells: analogy to apoptosis of somatic cells. Exp Cell Res 207,202–205.

Govin J, Caron C, Rousseaux S and Khochbin S (2005) Testis-specific histone H3 expression in somatic cells. Trends Biochem Sci 30,357–359.

Green GR, Balhorn R, Poccia DL and Hecht NB (1994) Synthesis and processing of mammalian protamines and transition proteins. Mol Reprod Dev 37,255–263.

Grimes SR Jr and Henderson N (1984) Hyperacetylation of histone H4 in rat testis spermatids. Experimental Cell Research 152,91–97.

Grimes SR Jr, Platz RD, Meistrich ML and Hnilica LS (1975) Partial characterization of a new basic nuclear protein from rat testis elongated spermatids. Biochem Biophys Res Commun 67,182–189.

Gu W, Tekur S, Reinbold R, Eppig JJ, Choi YC, Zheng JZ, Murray MT and Hecht NB (1998) Mammalian male and female germ cells express a germ cell-specific Y-Box protein, MSY2. Biol Reprod 59,1266–1274.

Gusse M, Sautiere P, Belaiche D, Martinage A, Roux C, Dadoune JP and Chevaillier P (1986) Purification and characterization of nuclear basic proteins of human sperm. Biochimica et Biophysica Acta 884,124–134.

Hammadeh ME, Zeginiadov T, Rosenbaum P, Georg T, Schmidt W and Strehler E (2001) Predictive value of sperm chromatin condensation (aniline blue staining) in the assessment of male fertility. Arch Androl 46,99–104.

Hartman S, Bergmann M, Bohle RM, Weidner W and Steger K (2006) Genetic imprinting during impaired spermatogenesis. Mol Hum Reprod 12,407–411.

Hazzouri M, Rousseaux S, Mongelard F, Usson Y, Pelletier R, Faure AK, Vourc'h C and Sele B (2000) Genome organization in the human sperm nucleus studied by FISH and confocal microscopy. Mol Reprod Dev 55,307–315.

Hecht NB (1990) Regulation of 'haploid expressed genes' in male germ cells. J Reprod Fertil 88,679–693.

Hikim AP, Wang C, Leung A and Swerdloff RS (1995) Involvement of apoptosis in the induction of germ cell degeneration in adult rats after gonadotropin-releasing hormone antagonist treatment. Endocrinology 136,2770–2775.

Hogeveen KN and Sassone-Corsi P (2006) Regulation of gene expression in post-meiotic male germ cells: CREM-signalling pathways and male fertility. Hum Fertil (Camb) 9,73–79.

Hong L, Schroth GP, Matthews HR, Yau P and Bradbury EM (1993) Studies of the DNA binding properties of histone H4 amino terminus. Thermal denaturation studies reveal that acetylation markedly reduces the binding constant of the H4 "tail" to DNA. J Biol Chem 268,305–314.

Host E, Lindenberg S and Smidt-Jensen S (2000a) DNA strand breaks in human spermatozoa: correlation with fertilization in vitro in oligozoospermic men and in men with unexplained infertility. Acta Obstet Gynecol Scand 79,189–193.

Host E, Lindenberg S and Smidt-Jensen S (2000b) The role of DNA strand breaks in human spermatozoa used for IVF and ICSI. Acta Obstet Gynecol Scand 79,559–563.

Iguchi N, Yang S, Lamb DJ and Hecht NB (2006) An SNP in protamine 1: a possible genetic cause of male infertility? J Med Genet 43,382–384.

Iuchi Y, Kobayashi T, Kaneko T, Takahara M, Ogino T and Fujii J (2001) Expression of a Y-box protein, YB2/RYB-a, precedes protamine 2 expression during spermatogenesis in rodents. Mol Hum Reprod 7,1023–1031.

Kaltenbach L, Horner MA, Rothman JH and Mango SE (2000) The TBP-like factor CeTLF is required to activate RNA polymerase II transcription during C. elegans embryogenesis. Mol Cell 6,705–713.

Kemal Duru N, Morshedi M and Oehninger S (2000) Effects of hydrogen peroxide on DNA and plasma membrane integrity of human spermatozoa. Fertil Steril 74,1200–1207.

Kennedy BP and Davies PL (1980) Acid-soluble nuclear proteins of the testis during spermatogenesis in the winter flounder. Loss of the high mobility group proteins. J Biol Chem 255,2533–2539.

Kennedy BP and Davies PL (1981) Phosphorylation of a group of high molecular weight basic nuclear proteins during spermatogenesis in the winter flounder. J Biol Chem 256,9254–9259.

Kierszenbaum AL (2001) Transition nuclear proteins during spermiogenesis: unrepaired DNA breaks not allowed. Mol Reprod Dev 58,357–358.

Kimmins S, Kotaja N, Davidson I and Sassone-Corsi P (2004a) Testis-specific transcription mechanisms promoting male germ-cell differentiation. Reproduction 128,5–12

Kimmins S, Kotaja N, Fienga G, Kolthur US, Brancorsini S, Hogeveen K, Monaco L and Sassone-Corsi P (2004b) A specific programme of gene transcription in male germ cells. Reprod Biomed Online 8,496–500

Kistler WS, Noyes C, Hsu R and Heinrikson RL (1975) The amino acid sequence of a testis-specific basic protein that is associated with spermatogenesis. J Biol Chem 250,1847–1853.

Kistler WS, Henriksen K, Mali P and Parvinen M (1996) Sequential expression of nucleoproteins during rat spermiogenesis. Experimental Cell Research 225,374–381.

Kleene KC (1989) Poly(A) shortening accompanies the activation of translation of five mRNAs during spermiogenesis in the mouse. Development 106,367–373.

Kleene KC (2003) Patterns, mechanisms, and functions of translation regulation in mammalian spermatogenic cells. Cytogenet Genome Res 103,217–224.

Kodama H, Yamaguchi R, Fukuda J, Kasai H and Tanaka T (1997) Increased oxidative deoxyribonucleic acid damage in the spermatozoa of infertile male patients. Fertil Steril 68,519–524.

Kotaja N, De Cesare D, Macho B, Monaco L, Brancorsini S, Goossens E, Tournaye H, Gansmuller A and Sassone-Corsi P (2004) Abnormal sperm in mice with targeted deletion of the act (activator of cAMP-responsive element modulator in testis) gene. Proc Natl Acad Sci USA 101,10620–10625.

Kramer JA, Zhang S, Yaron Y, Zhao Y and Krawetz SA (1997) Genetic testing for male infertility: a postulated role for mutations in sperm nuclear matrix attachment regions. Genet Test 1,125–129.

Krausz C and Sassone-Corsi P (2005) Genetic control of spermiogenesis: insights from the CREM gene and implications for human infertility. Reprod Biomed Online 10,64–71.

Krawetz SA and Dixon GH (1988) Sequence similarities of the protamine genes: implications for regulation and evolution. J Mol Evol 27,291–297.

Kwon YK and Hecht NB (1991) Cytoplasmic protein binding to highly conserved sequences in the 3′ untranslated region of mouse protamine 2 mRNA, a translationally regulated transcript of male germ cells. Proc Natl Acad Sci USA 88,3584–3588.

Kwon YK and Hecht NB (1993) Binding of a phosphoprotein to the 3′ untranslated region of the mouse protamine 2 mRNA temporally represses its translation. Mol Cell Biol 13,6547–6557.

Lee K, Haugen HS, Clegg CH and Braun RE (1995) Premature translation of protamine 1 mRNA causes precocious nuclear condensation and arrests spermatid differentiation in mice. Proc Natl Acad Sci USA 92,12451–12455.

Lee K, Fajardo MA and Braun RE (1996) A testis cytoplasmic RNA-binding protein that has the properties of a translational repressor. Mol Cell Biol 16,3023–3034.

Le Lannic G, Arkhis A, Vendrely E, Chevaillier P and Dadoune JP (1993) Production, characterization, and immunocytochemical applications of monoclonal antibodies to human sperm protamines. Mol Reprod Dev 36,106–112.

Lewis JD, Song Y, de Jong ME, Bagha SM and Ausio J (2003) A walk though vertebrate and invertebrate protamines. Chromosoma 111,473–482.

Loir M and Lanneau M (1984) Structural function of the basic nuclear proteins in ram spermatids. J Ultrastruct Res 86,262–272.

Ludwig M (2005) Risk during pregnancy and birth after asssisted reproductive technologies: an integral view of the problem. Semin Reprod Med 23,363–370.

Ludwig M, Katalinic A, Gross S, Sutcliffe A, Varon R and Horsthemke B (2005) Increased prevalence of imprinting defects in patients with Angelman syndrome born to subfertile couples. J Med Genet 42,289–291.

Luerssen H, Maier WM, Hoyer-Fender S and Engel W (1989) The nucleotide sequence of rat transition protein 2 (TP2) cDNA. Nucleic Acids Res 17,3585.

Macho B, Brancorsini S, Fimia GM, Setou M, Hirokawa N and Sassone-Corsi P (2002) CREM-dependent transcription in male germ cells controlled by a kinesin. Science 298, 2388–2390.

Maclean JA II and Wilkinson MF (2005) Gene regulation in spermatogenesis. Curr Top Dev Biol 71,131–197.

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

Manicardi GC, Bianchi PG, Pantano S, Azzoni P, Bizzaro D, Bianchi U and Sakkas D (1995) Presence of endogenous nicks in DNA of ejaculated human spermatozoa and its relationship to chromomycin A3 accessibility. Biol Reprod 52,864–867.

Marcon L and Boissonneault G (2004) Transient DNA strand breaks during mouse and human spermiogenesis new insights in stage specificity and link to chromatin remodeling. Biol Reprod 70,910–918.

Martianov I, Fimia GM, Dierich A, Parvinen M, Sassone-Corsi P and 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.

Martianov I, Brancorsini S, Gansmuller A, Parvinen M, Davidson I and Sassone-Corsi P (2002) Distinct functions of TBP and TLF/TRF2 during spermatogenesis: requirement of TLF for heterochromatic chromocenter formation in haploid round spermatids. Development 129, 945–955.

Martins RP, Ostermeier GC and Krawetz SA (2004) Nuclear matrix interactions at the human protamine domain: a working model of potentiation. J Biol Chem 279,51862–51868.

McCarrey JR, Geyer CB and Yoshioka H (2005) Epigenetic regulation of testis-specific gene expression. Ann N Y Acad Sci 1061,226–242.

McKay DJ, Renaux BS and Dixon GH (1985) The amino acid sequence of human sperm protamine P1. Biosci Rep 5,383–391.

McKay DJ, Renaux BS and Dixon GH (1986) Human sperm protamines. Amino-acid sequences of two forms of protamine P2. Eur J Biochem 156,5–8.

McPherson SM and Longo FJ (1993) Nicking of rat spermatid and spermatozoa DNA: possible involvement of DNA topoisomerase II. Dev Biol 158,122–130.

Meistrich ML, Trostle-Weige PK, Lin R, Bhatnagar YM and Allis CD (1992) Highly acetylated H4 is associated with histone displacement in rat spermatids. Mol Reprod Dev 31,170–181.

Meistrich ML, Mohapatra B, Shirley CR and Zhao M (2003) Roles of transition nuclear proteins in spermiogenesis. Chromosoma 111,483–488.

Mengual L, Ballesca JL, Ascaso C and Oliva R (2003) Marked differences in protamine content and P1/P2 ratios in sperm cells from percoll fractions between patients and controls. J Androl 24,438–447.

Moore PA, Ozer J, Salunek M, Jan G, Zerby D, Campbell S and Lieberman PM (1999) A human TATA binding protein-related protein with altered DNA binding specificity inhibits transcription from multiple promoters and activators. Mol Cell Biol 19,7610–7620.

Morales CR, Lefrancois S, Chennathukuzhi V, El-Alfy M, Wu X, Yang J, Gerton GL and Hecht NB (2002) A TB-RBP and Ter ATPase complex accompanies specific mRNAs from nuclei through the nuclear pores and into intercellular bridges in mouse male germ cells. Dev Biol 246,480–494.

Morse-Gaudio M and Risley MS (1994) Topoisomerase II expression and VM-26 induction of DNA breaks during spermatogenesis in Xenopus laevis. J Cell Sci 107 (Pt 10), 2887–2898.

Murray MT, Schiller DL and Franke WW (1992) Sequence analysis of cytoplasmic mRNA-binding proteins of Xenopus oocytes identifies a family of RNA-binding proteins. Proc Natl Acad Sci USA 89,11–15.

Nagamori I, Yomogida K, Adams PD, Sassone-Corsi P and Nojima H (2006) Transcription factors, cAMP-responsive element modulator (CREM) and Tisp40, act in concert in postmeiotic transcriptional regulation. J Biol Chem 281,15073–15081.

Nantel F, Monaco L, Foulkes NS, Masquilier D, LeMeur M, Henriksen K, Dierich A, Parvinen M and Sassone-Corsi P (1996) Spermiogenesis deficiency and germ-cell apoptosis in CREM-mutant mice. Nature 380,159–162.

Nasr-Esfahani MH, Razavi S, Mozdarani H, Mardani M and Azvagi H (2004) Relationship between protamine deficiency with fertilization rate and incidence of sperm premature chromosomal condensation post-ICSI. Andrologia 36,95–100.

Oliva R (2006) Protamines and male infertility. Hum Reprod Update 12,417–435.

Oliva R and Mezquita C (1982) Histone H4 hyperacetylation and rapid turnover of its acetyl groups in transcriptionally inactive rooster testis spermatids. Nucleic Acids Res 10,8049–8059.

Oliva R and Mezquita C (1986) Marked differences in the ability of distinct protamines to disassemble nucleosomal core particles in vitro. Biochemistry 25,6508–6511.

Oliva R and Dixon GH (1990) Vertebrate protamine gene evolution I. Sequence alignments and gene structure. J Mol Evol 30,333–346.

Oliva R and Dixon GH (1991) Vertebrate protamine genes and the histone-to-protamine replacement reaction. Prog Nucleic Acid Res Mol Biol 40,25–94.

Oliva R, Bazett-Jones D, Mezquita C and Dixon GH (1987) Factors affecting nucleosome disassembly by protamines in vitro. Histone hyperacetylation and chromatin structure, time dependence, and the size of the sperm nuclear proteins. Journal of Biological Chemistry 262, 17016–17025.

Oosterhuis GJ, Mulder AB, Kalsbeek-Batenburg E, Lambalk CB, Schoemaker J and Vermes I (2000) Measuring apoptosis in human spermatozoa: a biological assay for semen quality? Fertil Steril 74,245–250.

Padron OF, Brackett NL, Sharma RK, Lynne CM, Thomas AJ Jr and Agarwal A (1997) Seminal reactive oxygen species and sperm motility and morphology in men with spinal cord injury. Fertil Steril 67,1115–1120.

Papoutsopoulou S, Nikolakaki E, Chalepakis G, Kruft V, Chevaillier P and Giannakouros T (1999) SR protein-specific kinase 1 is highly expressed in testis and phosphorylates protamine 1. Nucleic Acids Res 27,2972–2980.

Pentikainen V, Erkkila K and Dunkel L (1999) Fas regulates germ cell apoptosis in the human testis in vitro. Am J Physiol 276,pE310–E316.

Perry AC, Wakayama T, Kishikawa H, Kasai T, Okabe M, Toyoda Y and Yanagimachi R (1999) Mammalian transgenesis by intracytoplasmic sperm injection. Science 284,1180–1183.

Pina-Guzman B, Solis-Heredia MJ and Quintanilla-Vega B (2005) Diazinon alters sperm chromatin structure in mice by phosphorylating nuclear protamines. Toxicol Appl Pharmacol 202,189–198.

Pina-Guzman B, Solis-Heredia MJ, Rojas-Garcia AE, Uriostegui-Acosta M and Quintanilla-Vega B (2006) Genetic damage caused by methyl-parathion in mouse spermatozoa is related to oxidative stress. Toxicol Appl Pharmacol 216,216–224.

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

Poccia D (1986) Remodeling of nucleoproteins during gametogenesis, fertilization, and early development. Int Rev Cytol 105,1–65.

Risley MS, Einheber S and Bumcrot DA (1986) Changes in DNA topology during spermatogenesis. Chromosoma 94,217–227.

Roca J and Mezquita C (1989) DNA topoisomerase II activity in nonreplicating, transcriptionally inactive, chicken late spermatids. EMBO J 8,1855–1860.

Rodriguez I, Ody C, Araki K, Garcia I and Vassalli P (1997) An early and massive wave of germinal cell apoptosis is required for the development of functional spermatogenesis. EMBO J 16,2262–2270.

Sailer BL, Jost LK and Evenson DP (1995) Mammalian sperm DNA susceptibility to in situ denaturation associated with the presence of DNA strand breaks as measured by the terminal deoxynucleotidyl transferase assay. J Androl 16,80–87.

Sakkas D, Manicardi G, Bianchi PG, Bizzaro D and Bianchi U (1995) Relationship between the presence of endogenous nicks and sperm chromatin packaging in maturing and fertilizing mouse spermatozoa. Biol Reprod 52,1149–1155.

Sakkas D, Mariethoz E and St John JC (1999) Abnormal sperm parameters in humans are indicative of an abortive apoptotic mechanism linked to the Fas-mediated pathway. Exp Cell Res 251,350–355.

Saleh RA and Agarwal A (2002) Oxidative stress and male infertility: from research bench to clinical practice. J Androl 23,737–752.

Sanchez-Pena LC, Reyes BE, Lopez-Carrillo L, Recio R, Moran-Martinez J, Cebrian ME and Quintanilla-Vega B (2004) Organophosphorous pesticide exposure alters sperm chromatin structure in Mexican agricultural workers. Toxicol Appl Pharmacol 196,108–113.

Sassone-Corsi P (1998) Coupling gene expression to cAMP signalling: role of CREB and CREM. Int J Biochem Cell Biol 30,27–38.

Sassone-Corsi P (2002) Unique chromatin remodeling and transcriptional regulation in spermatogenesis. Science 296,2176–2178.

Sassone-Corsi P (2005) Transcription factors governing male fertility. Andrologia 37,228–229.

Schlicker M, Schnulle V, Schneppel L, Vorob'ev VI and Engel W (1994) Disturbances of nuclear condensation in human spermatozoa: search for mutations in the genes for protamine 1, protamine 2 and transition protein 1. Hum Reprod 9,2313–2317.

Schluter G, Kremling H and Engel W (1992) The gene for human transition protein 2: nucleotide sequence, assignment to the protamine gene cluster, and evidence for its low expression. Genomics 14,377–383.

Schmidt EE and Schibler U (1995) High accumulation of components of the RNA polymerase II transcription machinery in rodent spermatids. Development 121,2373–2383.

Schmidt EE and Schibler U (1997) Developmental testis-specific regulation of mRNA levels and mRNA translational efficiencies for TATA-binding protein mRNA isoforms. Dev Biol 184,138–149.

Shaman JA, Prisztoka R and Ward WS (2006) Topoisomerase IIB and an extracellular nuclease interact to digest sperm DNA in an apoptotic-like manner. Biol Reprod 75,741–748.

Shirley CR, Hayashi S, Mounsey S, Yanagimachi R and Meistrich ML (2004) Abnormalities and reduced reproductive potential of sperm from Tnp1- and Tnp2-null double mutant mice. Biol Reprod 71, 1220–1229.

Siffroi JP, Pawlak A, Alfonsi MF, Troalen F, Guellaen G and Dadoune JP (2001) Expression of the TAR RNA binding protein in human testis. Mol Hum Reprod 7,219–225.

Sikka SC (2001) Relative impact of oxidative stress on male reproductive function. Curr Med Chem 8,851–862.

Silber SJ (1995) What forms of male infertility are there left to cure? Hum Reprod 10,503–504.

Singh J and Rao MR (1988) Interaction of rat testis protein, TP, with nucleosome core particle. Biochem Int 17,701–710.

Smith A and Haaf T (1998) DNA nicks and increased sensitivity of DNA to fluorescence in situ end labeling during functional spermiogenesis. Biotechniques 25,496–502.

Sonnack V, Failing K, Bergmann M and Steger K (2002) Expression of hyperacetylated histone H4 during normal and impaired human spermatogenesis. Andrologia 34,384–390.

Sotolongo B, Huang TT, Isenberger E and Ward WS (2005) An endogenous nuclease in hamster, mouse, and human spermatozoa cleaves DNA into loop-sized fragments. J Androl 26,272–280.

Steger K (1999) Transcriptional and translational regulation of gene expression in haploid spermatids. Anat Embryol (Berl) 199,471–487.

Steger K, Pauls K, Klonisch T, Franke FE and Bergmann M (2000) Expression of protamine-1 and -2 mRNA during human spermiogenesis. Mol Hum Reprod 6,219–225.

Steger K (2001) Haploid spermatids exhibit translationally repressed mRNAs. Anat Embryol (Berl) 203,323–334.

Suganuma R, Yanagimachi R and Meistrich ML (2005) Decline in fertility of mouse sperm with abnormal chromatin during epididymal passage as revealed by ICSI. Hum Reprod 20,3101–3108.

Sun JG, Jurisicova A and Casper RF (1997) Detection of deoxyribonucleic acid fragmentation in human sperm: correlation with fertilization in vitro. Biol Reprod 56,602–607.

Szczygiel MA and Ward WS (2002) Combination of dithiothreitol and detergent treatment of spermatozoa causes paternal chromosomal damage. Biol Reprod 67,1532–1537.

Tafuri SR, Familari M and Wolffe AP (1993) A mouse Y box protein, MSY1, is associated with paternal mRNA in spermatocytes. J Biol Chem 268, 12213–12220.

Tanaka H and Baba T (2005) Gene expression in spermiogenesis. Cell Mol Life Sci 62,344–354.

Tekur S, Pawlak A, Guellaen G and Hecht NB (1999) Contrin, the human homologue of a germ-cell Y-box-binding protein: cloning, expression, and chromosomal localization. J Androl 20,135–144.

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

Tanaka H, Matsuoka Y, Onishi M, Kitamura K, Miyagawa Y, Nishimura H, Tsujimura A, Okuyama A and Nishimune Y (2006) Expression profiles and single-nucleotide polymorphism analysis of human HANP1/H1T2 encoding a histone H1-like protein. Int J Androl 29,353–359.

Tanaka H, Miyagawa Y, Tsujimura A, Matsumiya K, Okuyama A and Nishimune Y (2003) Single nucleotide polymorphisms in the protamine-1 and -2 genes of fertile and infertile human male populations. Mol Hum Reprod 9,69–73.

Torregrosa N, Dominguez-Fandos D, Camejo MI, Shirley CR, Meistrich ML, Ballesca JL and Oliva R (2006) Protamine 2 precursors, protamine 1/protamine 2 ratio, DNA integrity and other sperm parameters in infertile patients. Hum Reprod 21,2084–2089.

Twigg J, Fulton N, Gomez E, Irvine DS and Aitken RJ (1998a) Analysis of the impact of intracellular reactive oxygen species generation on the structural and functional integrity of human spermatozoa: lipid peroxidation, DNA fragmentation and effectiveness of antioxidants. Hum Reprod 13,1429–1436.

Twigg J, Irvine DS, Houston P, Fulton N, Michael L and Aitken RJ (1998b) Iatrogenic DNA damage induced in human spermatozoa during sperm preparation: protective significance of seminal plasma. Mol Hum Reprod 4,439–445.

Umehara T and Horikoshi M (2003) Transcription initiation factor IID-interactive histone chaperone CIA-II implicated in mammalian spermatogenesis. J Biol Chem 278,35660–35667.

Verpoest W and Tournaye H (2006) ICSI: Hype or hazard? Hum Fertil 9,81–92.

Ward WS and Coffey DS (1991) DNA packaging and organization in mammalian spermatozoa: comparison with somatic cells. Biol Reprod 44,569–574.

Weinbauer GF, Behr R, Bergmann M and Nieschlag E (1998) Testicular cAMP responsive element modulator (CREM) protein is expressed in round spermatids but is absent or reduced in men with round spermatid maturation arrest. Mol Hum Reprod 4,9–15.

Wilkins MH (1956) Physical studies of the molecular structure of deoxyribose nucleic acid and nucleoprotein. Cold Spring Harb Symp Quant Biol 21,75–90.

Wu JY, Ribar TJ, Cummings DE, Burton KA, McKnight GS and Means AR (2000) Spermiogenesis and exchange of basic nuclear proteins are impaired in male germ cells lacking Camk4. Nat Genet 25,448–452.

Wykes SM and Krawetz SA (2003) The structural organization of sperm chromatin. Journal of Biological Chemistry 278,29471–29477.

Yang J, Medvedev S, Reddi PP, Schultz RM and Hecht NB (2005a) The DNA/RNA-binding protein MSY2 marks specific transcripts for cytoplasmic storage in mouse male germ cells. Proc Natl Acad Sci USA 102,1513–1518.

Yang J, Medvedev S, Yu J, Tang LC, Agno JE, Matzuk MM, Schultz RM and Hecht NB (2005b) Absence of the DNA-/RNA-binding protein MSY2 results in male and female infertility. Proc Natl Acad Sci USA 102,5755–5760.

Yang J, Medvedev S, Yu J, Schultz RM and Hecht NB (2006) Deletion of the DNA/RNA-binding protein MSY2 leads to post-meiotic arrest. Mol Cell Endocrinol 250,20–24.

de Yebra L, Ballesca JL, Vanrell JA, Bassas L and Oliva R (1993) Complete selective absence of protamine P2 in humans. J Biol Chem 268,10553–10557.

de Yebra L, Ballesca JL, Vanrell JA, Corzett M, Balhorn R and Oliva R (1998) Detection of P2 precursors in the sperm cells of infertile patients who have reduced protamine P2 levels. Fertil Steril 69,755–759.

Yiu GK and Hecht NB (1997) Novel testis-specific protein-DNA interactions activate transcription of the mouse protamine 2 gene during spermatogenesis. J Biol Chem 272,26926–26933.

Yoshii T, Kuji N, Komatsu S, Iwahashi K, Tanaka Y, Yoshida H, Wada A and Yoshimura Y (2005) Fine resolution of human sperm nucleoproteins by two-dimensional electrophoresis. Mol Hum Reprod 11, 677–681.

Yu YE, Zhang Y, Unni E, Shirley CR, Deng JM, Russell LD, Weil MM, Behringer RR and Meistrich ML (2000) Abnormal spermatogenesis and reduced fertility in transition nuclear protein 1-deficient mice. Proceedings of the National Academy of Sciences of the United States of America 97,4683–4688.

Zalensky AO, Siino JS, Gineitis AA, Zalenskaya IA, Tomilin NV, Yau P and Bradbury EM (2002) Human testis/sperm-specific histone H2B (hTSH2B). Molecular cloning and characterization. J Biol Chem 277,43474–43480.

Zhang X, San Gabriel M and Zini A (2006) Sperm nuclear histone to protamine ratio in fertile and infertile men: evidence of heterogeneous subpopulations of spermatozoa in the ejaculate. J Androl 27,414–420.

Zhao M, Shirley CR, Yu YE, Mohapatra B, Zhang Y, Unni E, Deng JM, Arango NA, Terry NH and Weil MM *et al.* (2001) Targeted disruption of the transition protein 2 gene affects sperm chromatin structure and reduces fertility in mice. Mol Cell Biol 21,7243–7255.

Zhao M, Shirley CR, Mounsey S and Meistrich ML (2004) Nucleoprotein transitions during spermiogenesis in mice with transition nuclear protein Tnp1 and Tnp2 mutations. Biol Reprod 71,1016–1025.

Zhong J, Peters AH, Lee K and Braun RE (1999) A double-stranded RNA binding protein required for activation of repressed messages in mammalian germ cells. Nat Genet 22,171–174.

Submitted on September 15, 2006; resubmitted on October 20, 2006; accepted on November 7, 2006

CHAPTER 3

IDENTIFICATION OF GENETIC VARIATION IN THE 5' AND 3' NON-CODING REGIONS OF THE PROTAMINE GENES IN PATIENTS WITH PROTAMINE DEREGULATION

Published in the *Archives of Andrology: Journal of Reproductive Systems*

Authors

Sue Hammoud

Benjamin R. Emery

Vincent W. Aoki

Douglas T. Carrell

Archives of Andrology: Journal of Reproductive Systems, 53:267–274, 2007

ISSN: 0148-5016 print/1521-0375 online
DOI: 10.1080/01485010701569890

informa
healthcare

Research Article

Identification of Genetic Variation in the 5′ and 3′ Non-coding Regions of the Protamine Genes in Patients with Protamine Deregulation

Sue Hammoud*, B.S., and Benjamin R. Emery*, B.S.
Andrology and IVF Laboratories, Division of Urology, Department of Surgery, School of Medicine, University of Utah, Salt Lake City, Utah, USA and Department of Physiology, School of Medicine, University of Utah, Salt Lake City, Utah, USA

Vincent W. Aoki, PhD
Andrology and IVF Laboratories, Division of Urology, Department of Surgery, School of Medicine, University of Utah, Salt Lake City, Utah, USA

Douglas T. Carrell, PhD
Andrology and IVF Laboratories, Division of Urology, Department of Surgery, School of Medicine, University of Utah, Salt Lake City, Utah, USA; Department of Physiology, School of Medicine, University of Utah, Salt Lake City, Utah, USA and Department of Obstetrics and Gynecology, School of Medicine, University of Utah, Salt Lake City, Utah, USA

Deregulation of sperm nuclear protamine ratio (P1/P2) has been shown to correlate with male factor infertility in humans, but the cause of this abnormal protein expression has yet to be identified. Recent studies have shown that there is little genetic variability in the coding regions of either of the protamine gene sequences. However, these studies did not investigate the 5′ or 3′ non-coding regions of these genes for mutations that might account for changes in the transcriptional or translational regulation of the protamines.

In an effort to determine if genetic variation in these non-coding regions may account for aberrant protamine expression, we have sequenced the 5′ and 3′ untranslated regions (UTRs) of both protamine 1 (P1) and protamine 2 (P2) genes in a population of infertile men with protamine deregulation, men presenting for infertility work-up with normal protamine ratios, and a population of unrelated, fertile men from the Utah Genetic Reference Project (UGRP). This analysis has identified 14 single nucleotide polymorphisms (SNPs), of which 13 were novel SNPs in the UTRs of P1 and P2, and verified the existence of a variable length repeat (VLR), GA_n, in the P2 5′ region. The SNP frequencies and VLR allelic frequencies did not achieve statistical significance between the populations, however, one of the SNPs identified in the 3′ UTR of protamine 2 was found at a low frequency in the abnormal protamine patients, but was completely absent in men with verified normal protamine ratio and donors of known fertility.

In conclusion, a number of SNPs have been reported in the protamine genes and the untranslated regions, however, these gene variants do not appear to be responsible for protamine deficiency. Hence, the underlying cause for aberrant protamine expression may possibly be due to abnormalities in candidate spermatogenic transcriptional/translational regulators, post-translational modifiers, or as-of-yet unidentified factors affecting the testicular environment.

KEYWORDS Infertility, Protamine, Spermatogenesis, Transcription, Translation

INTRODUCTION

During spermiogenesis, the sperm nucleus undergoes genome-wide reorganization, which involves the removal of histones and their replacement by various nuclear proteins, including highly positively charged protamines [Wouters-Tyrou et al. 1998; Sassone-Corsi 2002]. Humans express two types of protamines, protamine 1 (P1) and protamine 2 (P2), expressed in the post-meiotic haploid spermatid and stored in messenger ribonuclear protein (mRNP) particles [Steger et al. 1998; Steger 1999], and translated in the elongating spermatid [Steger 2001; Steger et al. 2002]. The temporal uncoupling of transcription and translation prevents precocious chromatin condensation and spermatogenic arrest [Lee et al. 1995].

Biochemical analysis of human sperm from proven fertile donors and infertile males suggests that the relative proportion of protamines (P1/P2) bound to DNA is associated with male infertility. The mean P1/P2 ratio in human sperm nuclei is approximately 1.0 [Balhorn et al. 1999; Carrell and Liu 2001; Oliva 2006]. A number of studies have described infertile male populations with an elevated P1/P2 [Balhorn et al. 1988; Belokopytova et al. 1993; Aoki et al. 2005a], non-detectable levels of P2 [Chevaillier et al. 1987; Balhorn et al. 1988; Belokopytova et al. 1993; de Yebra et al. 1998; Carrell and Liu 2001; Aoki et al. 2005a; Oliva 2006], or a diminished P1/P2 [Aoki et al. 2005a]. Taken together these studies indicate abnormal protamine stoichiometry derives from aberrant expression of either P1 or P2.

Several groups evaluated the coding sequences of protamine 1, 2 and transition protein 1, 2 to identify specifically relevant alterations that may possibly explain the cause of protamine deregulation [de Yebra and Oliva 1993; Schlicker et al. 1994; Tanaka et al. 2003; Miyagawa et al. 2005; Aoki et al. 2006a; Iguchi et al. 2006]. Generally, there was no correlation between the absence of protamines and mutations within the coding regions of the protamine genes or transition protein genes. These findings suggest that genetic sequence aberrations are an unlikely cause for protamine deficiency. Subsequent studies reported that patients with an elevated P1/P2 ratio under-expressed P2 protein and retained higher levels of protamine 2 transcript when compared to patients with normal or high P2 protein [Aoki et al. 2006c; Torregrosa et al. 2006]. However, patients with a diminished P1/P2 ratio under-expressed P1 protein and over-expressed P1 transcript [Aoki et al. 2006c; Torregrosa et al. 2006]. Previous studies have indicated that patients with an elevated protamine 2 transcript had higher levels of DNA damage [Torregrosa et al. 2006], whereas patients with increased protamine 1 transcript retention had low sperm motility [Lambard et al. 2004; Aoki et al. 2006c]. This variation in transcript retention may be explained by reduced protamine transcription, altered translation of the transcript or simply altered retention of unprocessed transcripts, but this is inconclusive as to the cause of altered transcript retention. Nevertheless, there may be defective regulatory element in the protamine genes, which cause a change in protein expression. It is possible that this regulatory element is related to the binding sites in the 5′ or 3′ untranslated regions (UTR) of the protamine genes.

*Both authors contributed equally.

Abbreviations: UGRP: Utah Genetic Reference Program; UTRs: untranslated regions; VLR: variable length repeat; SNP: Single Nucleotide Polymorphism.

Received 19 February 2007; accepted 03 April 2007.

Address correspondence to Douglas T. Carrell, Ph.D., University of Utah IVF and Andrology Laboratories, 675 Arapeen Drive, Suite 205, Salt Lake City, UT, USA, 84108. E-mail: douglas.carrell@hsc.utah.edu

Several studies have emphasized the importance of the UTRs of the protamine genes and their role in transcriptional and translational regulation [Braun 1990; Tamura et al. 1992; Schmidt et al. 1997; Schmid et al. 2001]. These studies identify several trans-acting factors that bind to certain motifs within these regions in order to modulate gene expression. However, the genomic sequence of the untranslated region of the protamine genes has not been evaluated. Single nucleotide polymorphisms in the conserved 5′ and 3′ UTRs may alter the binding dynamics of trans-acting factors. Consequently, this may interrupt protamine transcription, promote premature translation, or delay translational activation. The objective of this study was to screen a population composed of severely infertile men with identified aberrations in the P1/P2 ratio, individuals with a normal protamine ratio, fertile donors, and men from the Utah Genetic Reference Program (UGRP) for mutations in the untranslated regions of P1 and P2.

RESULTS

Four primer sets were used to analyze the 5′ and 3′ untranslated regions of protamine 1 and 2 in 315 study participants. Direct sequencing of the PCR amplified DNA identified 14 SNPs in the untranslated regions of both protamine genes, 1 of which was previously reported (rs 2301365) (Table 1), and one variable length repeat (VLR) in 5′ UTR of protamine 2 (Figure 1). The SNP frequencies were compared among the different groups using chi square test to determine if there was a significant difference in the occurrence of SNPs. The frequencies of SNPs reported were not significantly different between infertile men with an abnormal P1/P2 ratio, men with normal P1/P2 ratio, and men from the Utah Genetic Reference Program (UGRP). Of greatest interest was a G/C change 62 bp into the 3′ UTR of P2, which occurred in five infertile heterozygous men with an abnormal P1/P2 and 6 heterozygotes from the UGRP. Interestingly, none of the men with confirmed normal protamine ratios carried this change. This G/C SNP was confirmed in cDNA samples of men with abnormal P1/P2 ratios, but absent in all fertile donors tested.

The P2 5′ GA repeat lengths were found at relatively equal frequency among the abnormal P1/P2 ratio, normal P1/P2 ratio, and UGRP patients (Figure 1). The most common GA repeat length was 29 base pairs long. Chi square analysis was used to determine whether the observed allelic heterozygosity may vary among the different groups, however, significant difference was not achieved. The observed percentages for allelic hetrozygosity were: 63% in UGRP, 73% in normal P1/P2 ratio patients, and 80% in abnormal P1/P2 ratio. These percentages fell within the normal distribution of Hardy-Weinberg

TABLE 1 **Identified Single Nucleotide Polymorphisms in the Untranslated Regions of the Protamine Genes**

		Allele frequency in the study populations			
	Single nucleotide polymorphism (SNP)	UGRP	Donors (cDNA)	Normal P1/P2	Abnormal P1/P2
1	P1 3′ (+86bp) tgccaC/Tcatcca	0.000	0.000	0.021	0.000
2	P1 3′ (+98bp) tccaA/Gtaaaa	0.036	0.000	0.000	0.000
3	P1 5′ (−274bp) tcttG/Ttat	0.000	N/A	0.011	0.000
4	P1 5′ (−247bp) ggcagC/Accc	0.000	N/A	0.005	0.013
5	P1 5′ (−236bp) actcG/Agggg	0.000	N/A	0.011	0.031
6	P1 5′ (−189bp) caggC/Acacc	0.263	N/A	0.255	0.208
7	P1 5′ (−112bp) caggC/Tcgca	0.000	N/A	0.011	0.000
8	P1 5′ (−105bp) agaG/Cctggc	0.000	N/A	0.005	0.009
9	P1 5′ (−91bp) ctggC/Tccct	0.000	N/A	0.005	0.000
10	P2 3′ (+62bp) aagtG/Caggc	0.027	0.000	0.000	0.020
11	P2 5′ (−392bp) atgtG/Acgta	0.043	N/A	0.053	0.044
12	P2 5′ (−389bp) gcgT/Caccc	0.112	N/A	0.147	0.149
13	P2 5′ (−371bp) gggaC/Gatga	0.048	N/A	0.059	0.022
14	P2 5′ (−226bp) agttG/Atgac	0.016	N/A	0.035	0.026

The SNP locations are numbered according to their positions in the genomic sequence: upstream of the first start codon (negative value) and downstream of the termination codon (positive value).

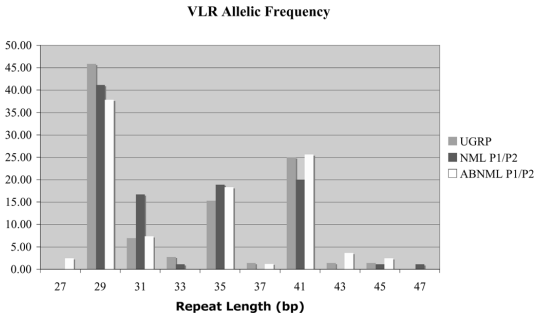


FIGURE 1 **(GA)n Variable Length Repeat (VLR) and Allelic frequency percentage in three test groups.**

equilibrium as evaluated using the genetic population software package GenALEx 6.0 (Peakall and Smouse 2006).

The majority of reported SNPs in this study are novel SNPs (Table 1), however, a few additional SNPs were not observed in our patient population. Two SNPs, a C/T change and a C/G change (rs11544791 and rs1126926), were noted in the untranslated regions of protamine 1 5′ and 3′ and were both identified in cDNA samples [Lee et al. 1987]. In addition, four SNPs were reported in protamine 2, two were found in genomic sequence (T/C and G/C) (rs 424908 and rs 452495) and the other two were in cDNA samples (A/T and C/T) (rs 1042801 and rs11545258) (Domenjoud et al. 1988; Kramer and Krawetz 1996). Bioinformatic analysis failed to localize any of the newly identified SNPs to transcriptional or translational factor binding sites.

DISCUSSION

The study objective was to determine whether mutations in the untranslated regions of the protamine genes account for aberrant transcription and translation regulation of these genes. This study identified 14 SNPs by direct sequencing in both the forward and reverse directions. Of greatest significance was a G/C change in 3′ UTR of P2 of abnormal protamine patients which did not localize to a known regulatory binding region, but it may be associated with an unknown mRNA regulator-binding site given that it did not appear in known fertile donors or patients with a verified normal protamine ratio, even in a heterozygous manner. All identified SNPs were present in similar frequencies in normal P1/P2, fertile donors, UGRP, and protamine deficient patients. These SNPs may possibly achieve significance if a greater sample size was evaluated.

The variable GA repeat upstream of the transcriptional initiation start site was reported from Ceph families in earlier studies [De Jonckheere et al. 1994; Schnulle et al. 1994]. Both studies, as well as ours, consistently reported several GA repeat lengths, the most common length reported in all populations tested is 29bp, which suggests that the variability in the untranslated region is unlikely to account for the diminished expression of P2 in infertile patients. From these data we can conclude that the screened populations have a significant number of repeat lengths present, but the repeat variability does not reflect a direct association with protamine deregulation.

Understanding the cause of protamine deregulation has been of great interest. Indeed, many investigators have attempted to identify any significant mutations in the protamine coding regions [de Yebra and Oliva 1993; Schlicker et al. 1994; Aoki et al. 2006a]. However, Tanaka, et al. in 2003 identified one azoospermic patient with a premature stop codon in P2 [Tanaka et al. 2003]. The SNP occurred at very low frequency (1/153 patients), however, it was absent from the fertile patient population screened. This SNP in the P2 gene should induce translation termination, which may result in male infertility due to haploinsufficiency of P2.

Later, Iguchi et al. sequenced the protamine genes in men exhibiting semen quality defects consistent with protamine abnormalities (i.e. sperm DNA damage, abnormal morphology, and normal sperm concentration) [Iguchi et al. 2006]. In their study, a heterozygous SNP which altered a highly conserved arginine residue in P1 was found in 10% (3/30) of the patients studied, but not seen in controls or reported in the literature. This SNP converts a highly conserved arginine to a serine residue, therefore creating a RS sequence which can serve as a potential phosphorylation site for the enzyme serine arginine protein kinase 1 (SRPK1). The P1/P2 ratio in the patient carrying this SNP is unknown. These findings can be a potential explanation for the occurrence of a small proportion of the abnormalities seen in protamine deregulation. However, the observations in both these studies were infrequent and limited to Asian populations.

The data from this current study is consistent with previously published data, which indicates that the genetic variability of the protamine gene locus is evident, even within the 5′ and 3′ non-coding regions (i.e. the low frequency +62 bp G/C P2 3′ SNP). Nevertheless, these changes do not account for a large percent of protamine deregulation cases. Given this conclusion, the over-riding question would be: what is the cause of protamine deregulation.

Two factors are worth considering when addressing protamine deregulation. First, abnormal protamine expression is strikingly common in infertility patients. Second, patients with abnormal protamine expression generally have a low sperm count, decreased sperm motility and morphology, increased sperm DNA damage [Mengual et al. 2003; Aoki et al. 2006d]. These clinical findings may suggest a defect in a more general spermatogenic-regulatory mechanism affecting a wide range of spermatogenic-specific genes, including the protamines. The indirect association between abnormal protamine expression and severely altered spermatogenesis may not be intuitively obvious, but animal studies have also shown that spermatogenesis is severely altered when protamine expression is experimentally reduced [Zhong et al. 1999; Aoki et al. 2005]. In contrast to human data, haploinsufficient mice show a homogenous population of severely affected sperm, while infertile male patients with protamine deregulation exhibit a truly heterogenous population of sperm [Cho et al. 2001; Aoki et al. 2005; Aoki et al. 2006b; Zhang et al. 2006]. This may be due to focal disruptions in the tubular structure of the testis or disruption of upstream regulators with an incomplete phenotypic penetrance.

A possible explanation for the link between altered protamine expression and severely reduced spermatogenesis is that abnormal protamine expression is the result of an abnormal functioning of a regulator of transcription, translation, or post-translational modifications that would affect not only the protamines, but a broad range of genes involved in spermatogenesis [Carrell et al. 2007]. Candidate regulators are trans regulatory factors of transcription that act via the promoter region such as cAMP response modulator (CREM) and Contrin. The testis-specific isoform of CREM, highly expressed in male germ cells [Delmas et al. 1993], is known to regulate the expression of several post-meiotic genes, such as the transition proteins and protamines [Krausz and Sassone-Corsi 2005; Hogeveen and Sassone-Corsi 2006].

The CREM dependent pathway is associated with other factors such as ACT, KIF17B, and Translin, whereas the Contrin pathway is not yet fully elucidated [Herbert and Hecht 1999; Yu et al. 2002; Hogeveen and Sassone-Corsi 2006]. Two functions have been proposed for Contrin, the human MSY2 RNA binding protein, in regards to protamine expression [Yang et al. 2005a; Yang et al. 2007]. First, it serves as a co-activator for protamine transcription by binding to the Y box element sequence found in the gene promoter region. Second, it stabilizes paternal transcripts in the cytoplasm to effect the temporal regulation of translation as described above.

Mice lacking any of the listed proteins are usually infertile and express a wide range of sperm abnormalities. Therefore, these proteins appear particularly attractive targets for abnormal protamine expression [Blendy et al. 1996; Nantel et al. 1996; Kotaja et al. 2004; Yang et al. 2005b; Yang et al. 2006]. Other key regulators, such as phosphorylation mechanisms, could also be responsible for a broad defect in spermatogenesis.

In conclusion, future studies will include analysis of regulators of transcription, translation, and post-translational modifications. These areas will possibly provide a better understanding of the underlying causes of abnormal protamine expression, and help clarify important information regarding spermatogenesis in general. Protamines may be an important marker in better understanding key regulatory pathways of spermatogenesis, especially if abnormal protamine expression reflects an aberrant function of a key transcriptional or translational regulator.

MATERIAL AND METHODS

Study Population

University of Utah Institutional Review Board approval was obtained for all aspects of this study. A total of 315 patients were recruited for the 5′ and 3′ UTR screening. These patients were classified into four groups: group one contained 123 infertile men with an abnormal P1/P2 ratio, group 2 contained 96 patients with a normal P1/P2 ratio, group 3 includes 30 fertile donors, used to confirm the GA

polymorphism in P2 5 UTR, and group 4 contains 96 DNA samples from the UGRP. The later group was selected based on unrelatedness and known paternity. The cut-off values for abnormal (low <0.8 or high >1.2) or normal (0.8–1.2) P1/P2 ratio was determined in a previous study in our lab [Aoki et al. 2005a]. The abnormal protamine group contains men with severely abnormal semen parameters and are all teratoozospermic.

Evaluation of Sperm P1/P2 Ratio

Study participants with an abnormal protamine ratio were identified using nuclear protein extraction, gel electrophoresis, and densitometry analysis. Sperm nuclear proteins were extracted from cryopreserved semen aliquots as previously described [Carrell and Liu 2001]. Gel electrophoresis reagents were obtained from BioRad Laboratories (Herculles, CA, USA). All other reagents were obtained from Sigma Chemical Company (St Louis, MO, USA).

Briefly, neat cryopreserved semen samples were centrifuged (500 × g, 5 minutes) and washed in 1 mM phenylmethylsulfonylfluoride (PMSF). The pellet is resuspended in Tris buffer (100 mM) containing EDTA (20 mM) and PMSF (1 mM, pH 8), 6 M Guanidine and 575 mM dithiothreitol, and 522 mM sodium iodoacetate. The suspension was kept at room temperature in the dark for 30 minutes and mixed with 100% ethanol. The ethanol wash was repeated and then resuspended in 0.5 M HCL, incubated for 15 minutes at 37°C. The nuclear proteins were precipitated by the addition of 100% trichloroacetic acid (at a final concentration of 20%) to the supernatant. The solution was incubated (4°C for 5 minutes), centrifuged (12,000 × g, 10 minutes), and washed twice in 1% 2-mercaptoethanol in acetone. The final pellet was dried and stored at −20°C for electrophoresis. Acetic acid urea gel electrophoresis was used to evaluate the intensity of the protamine bands, as previously described. Band intensity was measured using the NIH Image J software.

Screening the Untranslated Regions of the Protamine Genes

Venous blood was obtained using standard phlebotomy techniques and genomic DNA extraction was done using Puregene DNA extraction kit (Minneapolis, MN). The cDNA samples were prepared from previously extracted RNA samples using Trizol® Reagents (Invitrogen, CA, USA) and reverse transcribed using MMLV RT enzyme (Promega, WI, USA). Primer sets were designed and optimized to amplify cDNA and genomic DNA untranslated regions of the protamine genes with standard polymerase chain reaction (PCR) techniques. Primer pairs are as described (Table 2). Primary PCR products were cleaned using Exosap-it, and sequenced in the forward and reverse direction.

Samples were sequenced using ABI 3700 capillary sequencer. Sequence traces were assembled using the Phrap software and analyzed for significant changes using Phred and Consed. Phred assigns a quantitative value to quality of each sequenced base. This base quality provides a probabilistic estimate of the correctness of the base call. The sequences were assembled using Phrap program.

Potential mutations were identified using Consed, which has the ability to search for high quality base discrepancies in the assembled sequence. Visual analysis was also used to confirm identified polymorphisms and potential mutations. The nucleotide base changes are reported alongside the location of

TABLE 2 Primer Sequences

Region	Primer sequence	Temperature
P1 5′F	TGGTGCATGTCTGTAGTCCAACTACTG	61
P1 5′R	ACTTACTCATGGCTCTCCTCCGT	
P1 3′F	GCAGATATTACCGCCAGAGACA	64
P1 3′R	TCAAGAACAAGGAGAGAAGAGTGG	
P2 5′F	GGTAGAGGCTGCTATGATCCATGATTGC	64
P2 5′R	ATGCAACTGCTGCCTGTACA	
P2 3′F	CCACCTGACAAAAGCTCCAG	64
P2 3′R	AGCCAGGTTTGTGTGATTCG	

the SNPs. The frequencies of novel SNPs were compared by chi-square analysis.

GA Repeat

The lengths of the GA repeat located 5′ of the P2 gene was identified in men from three groups, men with abnormal P1/P2 ratios (n = 41), normal P1/P2 ratios (n = 45), and men from the UGRP (n = 45). The genotyping followed standard amplified fragment length polymorphism (AFLP) genomic analysis protocols. Briefly, PCR was performed using a FAM-labeled reverse primer and forward primer with a GTTTCT genotyping enhancement element flanking the P2 5′ GA_n. The PCR product was diluted into 96-well Applied Biosystems trays, processed, and analyzed on an Applied Biosystems 3130 × l Genetic Analyzer at a university core facility. The fragment length was determined using Applied Biosystems GeneMapper v3.7 and then converted to the repeat length by comparison to direct sequencing of a known homozygous individual.

ACKNOWLEDGMENTS

The authors would like to thank Dr. Robert Weiss and colleges for their technical advice and expertise in sequence analysis.

REFERENCES

Aoki, V. W., Christensen, G. L., Atkins, J. F. and Carrell, D. T. (2006a). Identification of novel polymorphisms in the nuclear protein genes and their relationship with human sperm protamine deficiency and severe male infertility. Fertil Steril **86**(5):1416–1422.

Aoki, V. W., Emery, B. R., Liu, L. and Carrell, D. T. (2006b). Protamine levels vary between individual sperm cells of infertile human males and correlate with viability and DNA integrity. J Androl **27**(6):890–898.

Aoki, V. W., Liu, L. and Carrell, D. T. (2005a). Identification and evaluation of a novel sperm protamine abnormality in a population of infertile males. Hum Reprod **20**(5):1298–1306.

Aoki, V. W., Liu, L. and Carrell, D. T. (2006c). A novel mechanism of protamine expression deregulation highlighted by abnormal protamine transcript retention in infertile human males with sperm protamine deficiency. Molecular Human Reproduction **12**(1):41–50.

Aoki, V. W., Liu, L., Jones, K. P., Hatasaka, H. H., Gibson, M., Peterson, C. M. and Carrell, D. T. (2006d). Sperm protamine 1/protamine 2 ratios are related to in vitro fertilization pregnancy rates and predictive of fertilization ability. Fertil Steril **86**(5):1408–1415.

Aoki, V. W., Moskovtsev, S. I., Willis, J., Liu, L., Mullen, J. B. and Carrell, D. T. (2005). DNA integrity is compromised in protamine-deficient human sperm. J Androl **26**(6):741–748.

Balhorn, R., Cosman, M., Thornton, K., Krishnan, V. V., Corzett, M., Bench, G., Kramer, C., Lee, J., Hud, N. V., Allen, M. J., Prieto, C., Meyer-Ilse, W., Brown, J. T., Kirz, J., Zhang, X., Baradbury, E. M., Maki, G., Braun, R. E. and Breed, W. G. (1999). Protamine mediated condensation of DNA in mammalian sperm. The Male Gamete: From Basic Science to Clinical Applications. Gagnon, C. Vienna, IL, Cache River Press, pp 55–70.

Balhorn, R., Reed, S. and Tanphaichitr, N. (1988). Aberrant protamine 1/protamine 2 ratios in sperm of infertile human males. Experientia **44**(1):52–55.

Belokopytova, I. A., Kostyleva, E. I., Tomilin, A. N. and Vorob'ev, V. I. (1993). Human male infertility may be due to a decrease of the protamine P2 content in sperm chromatin. Mol Reprod Dev **34**(1):53–57.

Blendy, J. A., Kaestner, K. H., Weinbauer, G. F., Nieschlag, E. and Schutz, G. (1996). Severe impairment of spermatogenesis in mice lacking the CREM gene. Nature **380**(6570):162–165.

Braun, R. E. (1990). Temporal translational regulation of the protamine 1 gene during mouse spermatogenesis. Enzyme **44**(1–4):120–128.

Carrell, D. T., Emery, B. R. and Hammoud, S. (2007). Altered protamine expression and diminished spermatogenesis: What is the link? Hum Reprod Update **13**(3):313–327.

Carrell, D. and Liu, L. (2001). Altered protamine 2 expression is uncommon in donors of known fertility, but common among men with poor fertilizing capacity, and may reflect other abnormalities of spermiogenesis. J Androl **22**(4):604–610.

Chevaillier, P., Mauro, N., Feneux, D., Jouannet, P. and David, G. (1987). Anomalous protein complement of sperm nuclei in some infertile men. Lancet **2**(8562):806–807.

Cho, C., Willis, W. D., Goulding, E. H., Jung-Ha, H., Choi, Y. C., Hecht, N. B. and Eddy, E. M. (2001). Haploinsufficiency of protamine-1 or -2 causes infertility in mice. Nat Genet **28**(1):82–86.

De Jonckheere, J., Nelson, J. E., Ginsburg, K. A., Martin, L. and Krawetz, S. A. (1994). GA repeat polymorphism at the PRM2 male fertility locus. Hum Mol Genet **3**(10):1915.

de Yebra, L., Ballesca, J. L., Vanrell, J. A., Corzett, M., Balhorn, R. and Oliva, R. (1998). Detection of P2 precursors in the sperm cells of infertile patients who have reduced protamine P2 levels. Fertil Steril **69**(4):755–759.

de Yebra, L. and Oliva, R. (1993). Rapid analysis of mammalian sperm nuclear proteins. Anal Biochem **209**(1):201–203.

Delmas, V., van der Hoorn, F., Mellstrom, B., Jegou, B. and Sassone-Corsi, P. (1993). Induction of CREM activator proteins in spermatids: Downstream targets and implications for haploid germ cell differentiation. Mol Endocrinol **7**(11):1502–1514.

Domenjoud, L., Fronia, C., Uhde, F. and Engel, W. (1988). Sequence of human protamine 2 cDNA. Nucleic Acids Res **16**(15):7733.

Herbert, T. P. and Hecht, N. B. (1999). The mouse Y-box protein, MSY2, is associated with a kinase on non-polysomal mouse testicular mRNAs. Nucleic Acids Res **27**(7):1747–1753.

Hogeveen, K. N. and Sassone-Corsi, P. (2006). Regulation of gene expression in post-meiotic male germ cells: CREM-signalling pathways and male fertility. Hum Fertil (Camb) **9**(2):73–79.

Iguchi, N., Yang, S., Lamb, D. J. and Hecht, N. B. (2006). An SNP in protamine 1: A possible genetic cause of male infertility? J Med Genet **43**(4):382–384.

Kotaja, N., De Cesare, D., Macho, B., Monaco, L., Brancorsini, S., Goossens, E., Tournaye, H., Gansmuller, A. and Sassone-Corsi, P. (2004). Abnormal sperm in mice with targeted deletion of the act (activator of cAMP-responsive element modulator in testis) gene. Proc Natl Acad Sci USA **101**(29):10620–10625.

Kramer, J. A. and Krawetz, S. A. (1996). Nuclear matrix interactions within the sperm genome. J Biol Chem **271**(20):11619–11622.

Krausz, C. and Sassone-Corsi, P. (2005). Genetic control of spermiogenesis: Insights from the CREM gene and implications for human infertility. Reprod Biomed Online **10**(1):64–71.

Lambard, S., Galeraud-Denis, I., Martin, G., Levy, R., Chocat, A. and Carreau, S. (2004). Analysis and significance of mRNA in human ejaculated sperm from normozoospermic donors: Relationship to sperm motility and capacitation. Mol Hum Reprod **10**(7):535–541.

Lee, C. H., Hoyer-Fender, S. and Engel, W. (1987). The nucleotide sequence of a human protamine 1 cDNA. Nucleic Acids Res **15**(18):7639.

Lee, K., Haugen, H. S., Clegg, C. H. and Braun, R. E. (1995). Premature translation of protamine 1 mRNA causes precocious nuclear condensation and arrests spermatid differentiation in mice. Proceedings of the National Academy of Sciences of the United States of America **92**(26):12451–12455.

Mengual, L., Ballesca, J. L., Ascaso, C. and Oliva, R. (2003). Marked differences in protamine content and P1/P2 ratios in sperm cells from percoll fractions between patients and controls. J Androl **24**(3):438–447.

Miyagawa, Y., Nishimura, H., Tsujimura, A., Matsuoka, Y., Matsumiya, K., Okuyama, A., Nishimune, Y. and Tanaka, H. (2005). Single-nucleotide polymorphisms and mutation analyses of the TNP1 and TNP2 genes of fertile and infertile human male populations. J Androl **26**(6):779–786.

Nantel, F., Monaco, L., Foulkes, N. S., Masquilier, D., LeMeur, M., Henriksen, K., Dierich, A., Parvinen, M. and Sassone-Corsi, P. (1996). Spermiogenesis deficiency and germ-cell apoptosis in CREM-mutant mice. Nature **380**(6570):159–162.

Oliva, R. (2006). Protamines and male infertility. Hum Reprod Update **12**(4):417–435.

Peakall, R. and Smouse, P. E. (2006). Genalex 6: Genetic analysis in Excel. Molecular Ecology Notes **1**:288–295.

Sassone-Corsi, P. (2002). Unique chromatin remodeling and transcriptional regulation in spermatogenesis. Science **296**(5576):2176–2178.

Schlicker, M., Schnulle, V., Schneppel, L., Vorob'ev, V. I. and Engel, W. (1994). Disturbances of nuclear condensation in human spermatozoa: Search for mutations in the genes for protamine 1, protamine 2 and transition protein 1. Hum Reprod **9**(12):2313–2317.

Schmid, C., Heng, H. H., Rubin, C., Ye, C. J. and Krawetz, S. A. (2001). Sperm nuclear matrix association of the PRM1 → PRM2 → TNP2 domain is independent of Alu methylation. Mol Hum Reprod **7**(10):903–911.

Schmidt, E. E., Ohbayashi, T., Makino, Y., Tamura, T. and Schibler, U. (1997). Spermatid-specific overexpression of the TATA-binding protein gene involves recruitment of two potent testis-specific promoters. J Biol Chem **272**(8):5326–5334.

Schnulle, V., Schlicker, M. and Engel, W. (1994). A (GA)n repeat polymorphism in the human protamine 2 (PRM 2) gene. Hum Mol Genet **3**(8):1445.

Steger, K. (1999). Transcriptional and translational regulation of gene expression in haploid spermatids. Anat Embryol (Berl) **199**(6):471–487.

Steger, K. (2001). Haploid spermatids exhibit translationally repressed mRNAs. Anat Embryol (Berl) **203**(5):323–334.

Steger, K., Fink, L., Klonisch, T., Bohle, R. M. and Bergmann, M. (2002). Protamine-1 and -2 mRNA in round spermatids is associated with RNA-binding proteins. Histochem Cell Biol **117**(3):227–234.

Steger, K., Klonisch, T., Gavenis, K., Drabent, B., Doenecke, D. and Bergmann, M. (1998). Expression of mRNA and protein of nucleoproteins during human spermiogenesis. Mol Hum Reprod **4**(10):939–945.

Tamura, T., Makino, Y., Mikoshiba, K. and Muramatsu, M. (1992). Demonstration of a testis-specific trans-acting factor Tet-1 in vitro that binds to the promoter of the mouse protamine 1 gene. J Biol Chem **267**(7):4327–4332.

Tanaka, H., Miyagawa, Y., Tsujimura, A., Matsumiya, K., Okuyama, A. and Nishimune, Y. (2003). Single nucleotide polymorphisms in the protamine-1 and -2 genes of fertile and infertile human male populations. Mol Hum Reprod **9**(2):69–73.

Torregrosa, N., Dominguez-Fandos, D., Camejo, M. I., Shirley, C. R., Meistrich, M. L., Ballesca, J. L. and Oliva, R. (2006). Protamine 2 precursors, protamine 1/protamine 2 ratio, DNA integrity and other sperm parameters in infertile patients. Hum Reprod **21**(8):2084–2089.

Wouters-Tyrou, D., Martinage, A., Chevaillier, P. and Sautiere, P. (1998). Nuclear basic proteins in spermiogenesis. Biochimie **80**(2):117–128.

Yang, J., Medvedev, S., Reddi, P. P., Schultz, R. M. and Hecht, N. B. (2005a). The DNA/RNA-binding protein MSY2 marks specific transcripts for cytoplasmic storage in mouse male germ cells. Proc Natl Acad Sci USA **102**(5):1513–1518.

Yang, J., Medvedev, S., Yu, J., Schultz, R. M. and Hecht, N. B. (2006). Deletion of the DNA/RNA-binding protein MSY2 leads to post-meiotic arrest. Mol Cell Endocrinol **250**(1–2):20–24.

Yang, J., Medvedev, S., Yu, J., Tang, L. C., Agno, J. E., Matzuk, M. M., Schultz, R. M. and Hecht, N. B. (2005b). Absence of the DNA/RNA-binding protein MSY2 results in male and female infertility. Proc Natl Acad Sci USA **102**(16):5755–5760.

Yang, J., Morales, C. R., Medvedev, S., Schultz, R. M. and Hecht, N. B. (2007). In the absence of the mouse DNA/RNA-binding protein MSY2, messenger RNA instability leads to spermatogenic arrest. Biol Reprod **76**(1):48–54.

Yu, J., Hecht, N. B. and Schultz, R. M. (2002). RNA-binding properties and translation repression in vitro by germ cell-specific MSY2 protein. Biol Reprod **67**(4):1093–1098.

Zhang, X., San Gabriel, M. and Zini, A. (2006). Sperm nuclear histone to protamine ratio in fertile and infertile men: Evidence of heterogeneous subpopulations of spermatozoa in the ejaculate. J Androl **27**(3):414–420.

Zhong, J., Peters, A. H., Lee, K. and Braun, R. E. (1999). A double-stranded RNA binding protein required for activation of repressed messages in mammalian germ cells. Nat Genet **22**(2):171–174.

CHAPTER 4

THE HUMAN SPERM EPIGENOME AND ITS POTENTIAL ROLE IN EMBRYONIC DEVELOPMENT

Published in *Molecular Human Reproduction*

Authors

Douglas T. Carrell

Saher Sue Hammoud

Molecular Human Reproduction, Vol.16, No.1 pp. 37–47, 2010
Advanced Access publication on November 11, 2009 doi:10.1093/molehr/gap090

MHR Basic science of reproductive medicine

NEW RESEARCH HORIZON ***Review***

The human sperm epigenome and its potential role in embryonic development

Douglas T. Carrell[1] and Saher Sue Hammoud

IVF and Andrology Laboratories, Departments of Surgery (Urology), Obstetrics and Gynecology, and Physiology, 675 S. Arapeen Dr, Suite #205, University of Utah School of Medicine, Salt Lake City, UT 84108, USA

[1]Correspondence address. Tel: +1-801-587-3772; Fax: +1-801-581-6127; E-mail: douglas.carrell@hsc.utah.edu

ABSTRACT: Along with many of the genome-wide transitions in chromatin composition throughout spermatogenesis, epigenetic modifications on histone tails and DNA are continuously modified to ensure stage specific gene expression in the maturing spermatid. Recent findings have suggested that the repertoire of epigenetic modifications in the mature sperm may have a potential role in the developing embryo and alterations in the epigenetic profile have been associated with infertility. These changes include DNA demethylation and the retention of modified histones at important developmental, signaling and micro-RNA genes, which resemble the epigenetic state of an embryonic stem cell. This review assesses the significance of epigenetic changes during spermatogenesis, and provides insight on recent associations made between altered epigenetic profiles in the mature sperm and its relationship to infertility.

Key words: methylation / histone / *in vitro* fertilization / embryogenesis / epigenetic

Introduction

In vitro fertilization (IVF) and other assisted reproductive technologies (ART) have accounted for ~3 million births since the world's first IVF baby was born in 1978 (Cohen, 1978). Reports examining the long-term health consequences of these babies are limited; however, follow-up studies have reported increased intrauterine growth restriction and lower birth weights in singletons conceived by IVF compared with natural conceptions (Steel and Sutcliffe, 2009). More recently, increased perinatal mortality, congenital anomalies and epigenetic abnormalities have been reported to be associated with IVF (Seif *et al.*, 2006; Allen *et al.*, 2008; Kalra and Molinaro, 2008; Reefhuis *et al.*, 2009; Steel and Sutcliffe, 2009).

The underlying causes of increased anomalies in IVF offspring are unknown, but alterations in the normal epigenetic state of gametes of severely infertile patients undergoing IVF has been proposed as one potential contributor (Cutfield *et al.*, 2007; Lim *et al.*, 2009; Manipalviratn *et al.*, 2009). Epigenetic modifications are covalent modifications present on either the DNA itself or to the proteins that are closely associated with DNA (histones in somatic cells and histones and protamines in sperm), both of which are important in modifying gene expression without changing the genetic code itself. These modifications comprise what is commonly referred to as the epigenome, which in somatic cells regulates cellular fate and function (Bernstein *et al.*, 2002, 2005, 2006; Li, 2002; Jones and Baylin, 2007). It is now well understood that the epigenome can become disrupted or altered, which may contribute significantly to the onset of epigenetic changes observed in many diseases and may be causative of some diseases (Feinberg, 2007; Jones and Baylin, 2007).

Recent studies have demonstrated that sperm have unique and potentially important epigenetic modifications. This brief review describes chromatin and epigenetic changes throughout spermatogenesis, their potential role in normal embryonic development, and their implications in male infertility.

Histone Modifications During Spermatogenesis

Male germ cells undergo unique and extensive chromatin and epigenetic remodeling soon after their specification (determination to become a spermatocyte) and during the differentiation process to become a mature spermatozoon (Seki *et al.*, 2005). Although the mechanisms regulating and orchestrating specification and spermiogenesis remain poorly understood, some progress has been made in elucidating the changes associated with the complex cellular changes. During mitosis and meiosis, male germ cell DNA is packaged in nucleosomes, comprised of histone 2A (H2A), histone 2B (H2B), histone 3 (H3) and histone 4 (H4), all of which are susceptible to covalent modifications, such as methylation, acetylation, ubiquitination

and phosphorylation. Each of these chemical modifications to histones works alone or in concert to influence gene repression and/or activation (Fig. 1).

Histone methylation on lysine (K) residues of H3 or H4 can promote gene activation and/or repression (Lachner and Jenuwein, 2002; Suganuma and Workman, 2008). Monomethylation, dimethylation and trimethylation modifications of H3K4, H3K9 or H3K27 display tightly controlled temporal expression and ensure proper progression through spermatogenesis (Khalil *et al*., 2004; Godmann *et al*., 2007; Payne and Braun, 2006). The level of H3K4 methylation peaks in the spermatogonial stem cell stage (Fig. 2), and a targeted loss of H3K4 methylation, caused by reduction of Mll2 activity (an H3K4 methyl transferase), results in a dramatic reduction in the number of spermatocytes (Table I), suggesting that H3K4 methylation is essential

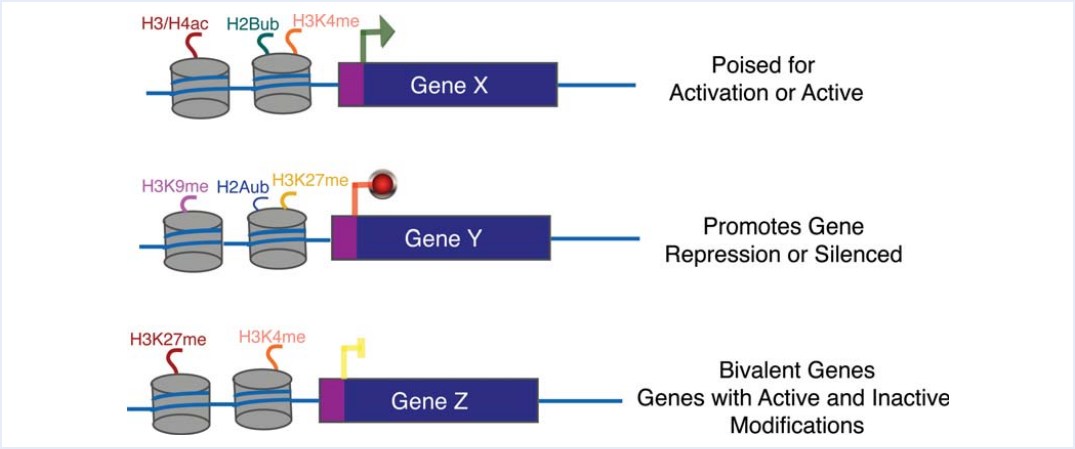


Figure 1 Chromatin modifications determine gene state. Histone modifications promote either gene activation or repression; however, in embryonic stem cells and sperm a subset of genes are commonly associated with both active and inactive marks. ac = acetylation, me = methylation, ub = ubiquitination.

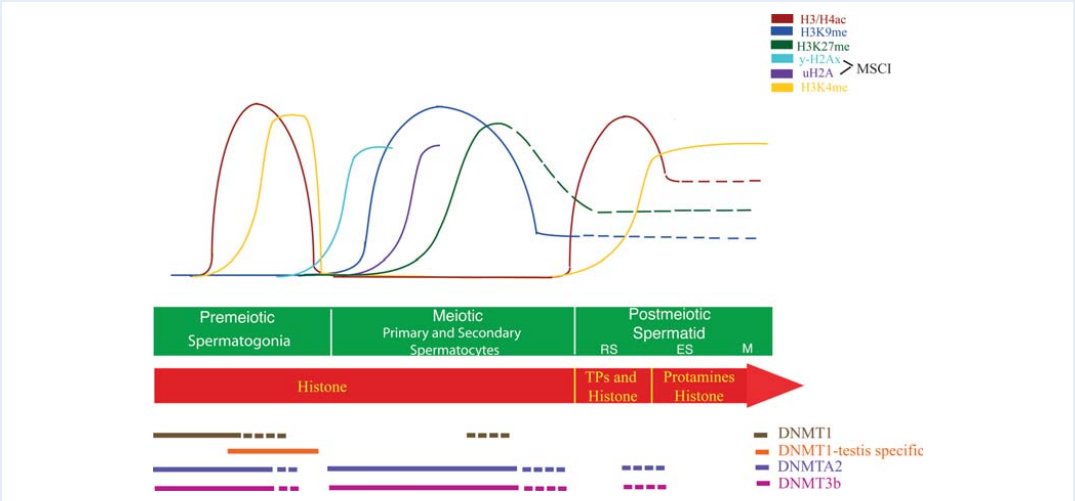


Figure 2 Hypothetical expression profiles of histone modifications and DNA methyltransferases during spermatogenesis. In the top panel, solid lines indicate investigated histone level. In the bottom panel dashed lines indicate low expression. Ac = acetylation, me = methylation, TPs = transition proteins, Ax = phosphorylation, MSCI = meiotic sex chromosome inactivation, DNMT = DNA methyl transferase, RS = round spermatid, ES = elongated spermatid, M = mature spermatid.

Table I A Summary of chromatin modifiers and their associated functions

Chromatin modifier	Function
DNMT1	Maintenance DNA methyltransferase
DNMT3a	*de-novo* DNA methyltransferase
DNMT3b	*de-novo* DNA methyltransferase
DNMT3L	*de-novo* DNA methyltransferase (no catalytic activity)
MII	H3K4 methyl transferase
JHDM2A	H3K9 demethylase
HAT	Histone acetyl transferase
HDACs	Histone deacetylase
LSD1/KDM	H3K4 demethylase

for the exit from the stem cell stage and commitment to become a spermatocyte (Glaser *et al*., 2009). In contrast, H3K9 methylation and H3K27methylation are low in the stem cell and increase during meiosis (Fig. 2), persisting long after meiosis is complete, presumably to ensure gene-silencing (Payne and Braun, 2006). The role ascribed to each of these modifications has been primarily characterized by immunofluorescence data with no gene specific localization. Methylation on lysine 9 of histone H3 (H3K9me) is associated with the sex chromosomes, euchromatin and heterochromatin in the late pachytene stage, however, the levels of H3K9 methylation drop upon completion of meiosis. This reduction in H3K9me is concurrently associated with an increase in H3K4me levels (Fig. 2) (Glaser *et al*., 2009).

The timing of establishment and removal of methylation marks is critical to normal spermatogenesis, as demonstrated by numerous transgenic animal models. Loss of LSD1/KDM1 (H3K4me demethylase) during mid to late meiosis in *Caenorhabditis elegans* results in germ cell apoptosis and progressive sterility that is maintained through many generations (Shi *et al*., 2004; Lee *et al*., 2005; Katz *et al*., 2009) (Table I). Similarly, removal of H3K9me at the end of meiosis is essential for the completion of spermatogenesis (Okada *et al*., 2007). Targeted disruption of the H3K9 demethylase JHDM2A (JmjC domain containing histone demethylase 2A, also known as JMJD1A) (Table I) results in complete loss of protamine 1 (PRM1) and transition protein 1 (TNP1) expression, defective chromatin condensation, and infertility (Okada *et al*., 2007). These studies show that methylation acts through various mechanisms to guide spermatogenesis (Table I).

Histone acetylation of lysine residues is dynamically regulated by histone acetyltransferases (HATs) and histone deacetylases (HDACs), both of which are shown to be essential for spermatogenesis (Christensen *et al*., 1984; Grimes and Henderson, 1984; Hazzouri *et al*., 2000; Lahn *et al*., 2002; Sonnack *et al*., 2002; Fenic *et al*., 2004, 2008; An, 2007, Kurtz *et al*., 2007). Histone acetylation relaxes chromatin and promotes polymerase II (Pol II) gene transcription, whereas deacetylation is associated with gene silencing (Jenuwein and Allis, 2001). Acetylation levels on both H3 and H4 are high in the stem cell phase and are completely removed during meiosis (Fig. 2). Re-acetylation of H4 happens in the elongating spermatid and is known to be a prerequisite for the histone-to-protamine exchange process (Fig. 2) (Hazzouri *et al*., 2000). However, recent studies show that both H3 and H4 acetylation occur simultaneously in the elongating stage, possibly extending to H2A and H2B acetylation as well (Nair *et al*., 2008). These data raise the question whether all histones may need to be acetylated to ensure proper histone to protamine exchange in maturing sperm. Enzymes involved in H4 hyperacetylation in the round spermatid are unknown, however, two candidates have emerged, testis specific chromodomain protein (CDY) and HAT (monocytic leukemia) 4 (MYST4). Both of these acetyltransferases are expressed during the maturing spermatid stage, localize to the nucleus, and have been shown to have potent H4 acetylase activity (Lahn *et al*., 2002; McGraw *et al*., 2007). H3 acetylation in the elongating spermatid was shown to be Pygopus 2 (Pygo 2) dependent. Pygo 2 has an evolutionarily conserved plant homeodomain (PHD) finger domain that binds trimethylated H3K4 and facilitates H3 acetylation (Nair *et al*., 2008). Evidence for H2B acetylation has been recently described by mass spectrometery studies, however, very little is known about its function during spermatogenesis or the enzymes required for its acetylation (Lu *et al*., 2009).

Although acetylation is broad, studies using HDAC inhibitors have demonstrated that the acetylation process is also necessarily specific. The effect of HDAC inhibitors on spermatogenesis is poorly defined, but a few studies have shown that although treating mice with HDAC inhibitors did not result in hyperacetylation, it did cause severe infertility (Fenic *et al*., 2004). Trichostatin A (TSA) treated animals had no evidence of H4 hypercetylation in the round spermatid, but the number of spermatids was significantly reduced (Fenic *et al*., 2004, 2008). The inability to detect the hyperacetylation following TSA treatment maybe due to an increase in apoptosis in cells with abnormal acetylation levels, or due to a compensatory mechanism involving alternative HDACs that are insensitive to TSA (Pivot-Pajot *et al*., 2003).

Histone phosphorylation occurs at serine residues of all core histones and is generally associated with gene activation (Berger, 2002). However, H2Ax phosphorylation (also known as γH2Ax) in germ cells confers the formation of X/Y sex body during spermatogenesis and is a marker for telomere clustering and double stranded breaks (Fernandez-Capetillo *et al*., 2003a, b). H2Ax phosphorylation is dependent on the ataxia telangiectasia DNA repair and Rad3 related protein ATR, and on the tumor suppressor BRCA1. Together γH2Ax, ATR and BRCA1 initiate meiotic sex chromosome inactivation (MSCI), but to maintain MSCI throughout the pachytene stage there are many other epigenetic modifications including ubiquinated H2A that are localized to the XY body, however, the exact function each performs are unknown (Hoyer-Fender, 2003).

The effect of ubiquination varies depending on the core histone modified: ubiquination of H2A associates with transcriptional repression, whereas, mono-ubiquination of H2B is linked to transcriptional activation in sperm (Baarends *et al*., 2005; Zhu *et al*., 2005). In male germ cells, recruitment of ubiquinated H2A to the sex body and telomeres occurs long after γH2Ax incorporation (Fig. 2), which indicates that H2A-ubiquination may be involved in maintaining silencing in the inactive chromatin, but not establishing MSCI.

These brief descriptions of histone modifications during spermatogenesis (data summarized in Fig. 2) demonstrate the varied

ways in which epigenetic modification regulate spermatogenesis. Although most histones are replaced with protamines during the elongating spermatid stage, some of the modified nucleosomes escape the histone to protamine transition and as a result are retained in mature sperm, suggesting that these retained nucleosomes may also play a role in the paternal contribution to the embryo.

The Role of Paternal Histones in the Epigenetic Control of Embryogenesis

A hallmark of spermiogenesis is the widespread changes in chromatin structure during spermiogenesis, including the exchange of most canonical histones for protamines (Ward and Coffey, 1991). Protamines are small basic proteins that bind DNA to form toroids; tightly packed structures which compact the genome beyond what is attainable by nucleosomes. The high-level of compaction is an essential attribute for genome transport in the mature sperm head (Balhorn *et al*., 2000). The histone to protamine exchange process is incomplete, with a small percentage (5–15%) of the genome bound to nucleosomes (Tanphaichitr *et al*., 1978; Wykes and Krawetz, 2003). The replacement of somatic histones by protamines is important for nuclear chromatin compaction, sperm maturation and fertility (Gatewood *et al*., 1987, 1990; Balhorn *et al*., 1988; de Yebra *et al*., 1998; Corzett *et al*., 2002; Aoki *et al*., 2004, 2005, 2006a, b, c; Hammoud *et al*., 2009a, b). In humans the relative proportion of protamine-1 (P1) to protamine-2 (P2) is strictly regulated at approximately a 1:1 ratio and alterations in the P1/P2 ratio are very rare in fertile men and relatively common in infertile men (Chevaillier *et al*., 1987; Balhorn *et al*., 1988, 1999; Belokopytova *et al*., 1993; de Yebra *et al*., 1998; Carrell and Liu, 2001; Corzett *et al*., 2002; Aoki *et al*., 2005, Oliva, 2006; Carrell *et al*., 2007). Reports from many labs have shown that changes in the P1/P2 ratio are not only associated with altered sperm quality, but also associated with decreased embryo quality and IVF outcome compared with infertile patients with a normal P1/P2 ratio (Aoki *et al*., 2006a, b, c; Depa-Martynow *et al*., 2007). These preliminary associations suggest that protamines and histones may have a greater role during the preimplantation embryo development than previously expected.

The retained nucleosomes are comprised of canonical histones (H2A, H2B, H3, H4) as well as a testes-specific histone variant (tH2B) (Gatewood *et al*., 1990; Kimmins and Sassone-Corsi, 2005). Until recently, the role for the retained nucleosomes was unknown, but it was speculated that the retained nucleosomes were either remnants of incomplete histone to protamine replacement, or that they may have a biologically significant role during early embryogenesis. In recent work from our laboratory, we have shown that nucleosomes retained in sperm are not simply randomly distributed remnants of inefficient protamine replacement, but are instead significantly enriched at many loci important for embryo development, including genes of key embryonic transcription factors and signaling pathway proteins. Histones were also significantly enriched at the promoters of miRNAs and imprinted genes (addressed in more detail below) (Hammoud *et al*., 2009a, b). These findings challenge the widely assumed notion that the paternal genome provides little in epigenetic contributions beyond a small set of paternally imprinted genes and a modest repertoire of packaged RNA, due to the repackaging of the vast majority of the genome by protamine.

The identification of retained nucleosomes at key developmental genes was striking, but to have any potential paternal contribution to the developing embryo secondary modifications on the retained nucleosomes (such as the modification discussed above) may be key to differentiate the paternally poised genes from all other genes that have acquired acetylated histones following protamine displacement soon after fertilization. To better understand the chromatin landscape at histone-associated developmental, signaling and miRNA genes three chromatin attributes were tested: histone variants, histone modifications and DNA methylation. We hypothesized that consistent, orderly and biologically relevant patterns in histone modifications, variants and/or DNA methylation could imply a programmatic marking, such as 'poising' genes for activation during early embryogenesis, as opposed to a random and non-biologically relevant inefficiency in protamine replacement and epigenetic marking.

The first logical candidate tested for paternal genome poising was the testis specific H2B variant (tH2B) which is incorporated late in spermatogenesis and comprises a large percentage of retained histones (Hammoud *et al*., 2009a, b). Analysis of tH2B distribution throughout the sperm genome revealed a very significant enrichment of this histone variant at genes for ion channels and genes involved in spermiogenesis, but not at promoters for developmental genes. A second variant in spermatogenesis that has been recently implicated in gene poising in other cell types, such as embryonic stem (ES) cells, is the histone variant H2Az. H2Az in ES cells were shown to be a key regulator of chromatin function and associated with targets of the Polycomb complex at genes essential for ES cell differentiation (Creyghton *et al*., 2008). In contrast to its role in ES cells, H2Az in sperm was enriched at pericentric heterochromatin, which is consistent with prior immunostaining studies (Rangasamy *et al*., 2003). Data from these two variants suggested that canonical histones might be an alternative at developmental genes.

Chromatin immunoprecipitation (ChIP) followed by either microarray or deep sequencing analysis clearly shows that modified canonical nucleosomes reveal attributes of a remnant spermatogenesis program (primarily enriched with H3K4me3), as well as a future developmental program (enriched with H3K27me3, H3K4me2 and H3K4me3) (Hammoud *et al*., 2009a, b). GO term analyses for H3K4me3 yielded genes important for changing nuclear architecture, RNA metabolism, spermatogenesis and a selected number of transcription factors important for embryonic development. However, the majority of the developmental and signaling transcription factors were significantly enriched with H3K27me3 and H3K4me2. Interestingly, many of the developmental promoters harboring an activation mark such as H3K4me3 also retained a silencing mark H3K27me3 (Bernstein *et al*., 2002, 2006), resembling the bivalently marked (H3K4me3 and H3K27me3) promoters commonly seen in ES cells that are typically silenced prior to ES cell differentiation, but necessary for embryonic differentiation (Fig. 1). Furthermore, modified canonical nucleosomes extended beyond known gene promoters to promoters of non-coding RNAs, miRNAs and imprinted loci. No clear pattern was seen at many of the miRNA and non-coding RNA promoters, but this is primarily due to the limited knowledge on the functional role of these miRNAs in development. Taken together, these data reveal extensive

histone modification patterns, and significant similarities to patterns observed in ES cells that may indicate a significant role for sperm epigenetic marking in the establishment of embryonic totipotency.

Significance of DNA methylation in the paternal germline

Packaging and transcriptional control of DNA in eukaryotes is by in large governed by the highly conserved role of histones, however, in higher organisms DNA methylation has been shown to have an essential role in normal embryonic development, regulating gene expression, X chromosome inactivation, genomic imprinting and silencing of endogenous retroviruses (Jaenisch and Jahner, 1984; Surani, 1998; Ng and Bird, 1999). DNA methylation occurs primarily at cytosine residues in a CpG context and is catalyzed by two important classes of DNA methyltransferases (DNMTs): maintenance DNA methyltaransferase (*Dnmt1*) and *de novo* DNA methyltransferases (*Dnmt3a*, *Dnmt3b*, *Dnmt3l*) (Fig. 2, Table I) (Eden and Cedar, 1994).

Dnmt1 is the major methyltransferase in somatic cells. It has a preference for hemimethylated DNA and is critical for the maintenance of DNA methylation patterns in replicating cells (Bestor, 1992; Lei *et al*., 1996) (Table I). Mice homozygous for a targeted partial deletion of *Dnmt1* or complete loss of function of *Dnmt1* have retarded growth and die by mid-gestation (Li *et al*., 1992, 1996). *Dnmt1* deficient embryos have less than 5% of the normal levels of cytosine methylation, regain biallelic expression at imprinted loci and have ectopic expression of *Xist* and retrotransposons (Li *et al*., 1993; Panning and Jaenisch, 1996; Walsh *et al*., 1998; Goll and Bestor, 2005).

The second class of DNMTs are *de novo* methyltransferases (*Dnmt3a*, *3b*, *3l*) and are essential for establishing new DNA methylation patterns during development (Table I). Embryos lacking *Dnmt3a* or *3b* lost all *de novo* methylation capabilities in ES cells and early embryos (resulted in embryonic lethality), but had no effect on the maintenance of imprinted loci (Hsieh, 1999; Lyko *et al*., 1999; Okano *et al*., 1999). Dnmt3a and b were required for methylating centromeric loci and imprinted genes. *Dnmt3l* is closely related in sequence to *DNMT3A* and *B*, but lacks the catalytic domain. *Dnmt3l* mediates *de novo* methylation by stimulating the catalytic activity of *DNMT3A2*, an isoform variant of DNMT3A (Hata *et al*., 2002; Suetake *et al*., 2004). *Dnmt3A2* along with its cofactor *Dnmt3l* establishes locus-specific DNA methylation of paternal imprints prior to meiosis in spermatogenesis (Bourc'his *et al*., 2001; Kaneda *et al*., 2004). Males' haploinsufficient for *Dnmt3l* are phenotypically normal and fertile, but have subtle changes in methylation and chromatin state of the genome in pre-meiotic spermatogonia (Bourc'his and Bestor, 2004; Webster *et al*., 2005). The phenotype for DNMT3L null male germ cells is significantly different in male versus female germ cells. Male germ cells fail to methylate LINE-1 (long interspersed elements) and IAP (intracisternal A particles) classes of retrotransposons, have severe asynapsis at meiotic prophase, and undergo apoptosis of all germ cells before pachytene (Bourc'his and Bestor, 2004). Whereas, methylation patterns at the small number of paternally methylated DMRs are almost normal, suggesting some functional redundancy between DNMTs (Bourc'his and Bestor, 2004; Webster *et al*., 2005). Strikingly, in the oocyte DNMT3L deficient germ cells carried out normal meiosis and methylation at repeat sequences, but the obvious methylation defect was limited to maternally imprinted loci. The differences observed between male and female knockouts are intriguing and raise many questions regarding targeting mechanism and differential regulation between sexes.

In addition to the DNMT gene knockout or haploinsufficiency studies used to assess the functional significance of DNA methylation in the germ line and early embryo, animals treated with DNMT inhibitors such as 5-aza-2'-deoxyCytidine showed altered gene expression patterns and loss of methylation in the germline. The severity of the phenotype observed in the treatment group was heavily dependent on the duration of treatment: short-term exposure in mice and rats decreased fertility (Seifertova *et al*., 1976; Raman and Narayan, 1995),whereas, prolonged treatment of 5-azacytidine (11weeks) in male rats resulted in a dose-dependent reductions in testis, epididymal weights and sperm counts, increase in germ cell apoptosis, and a significant increase in preimplantation loss (Doerksen and Trasler, 1996; Doerksen *et al*., 2000). These findings suggest that sperm DNA methylation plays a critical role in the differentiation of spermatogonia and early embryo viability, however, this appears counterintuitive since DNA methylation patterns are erased and reestablished immediately after fertilization and once again when primordial germ cells (PGCs) reach the genital ridge (Reik *et al*., 2001; Hajkova *et al*., 2002, Seki *et al*., 2005, 2007).

In humans, the relationship between bulk methylation levels with respect to IVF outcome was recently examined in one study. No significant correlation was made between bulk DNA methylation levels and the fertilization rate or embryo quality, but a lower 5-methyl cytosine signal ($<$555 AU) intensity correlated with a lower pregnancy (8.3 versus 33.3%) rates (Benchaib *et al*., 2005). These preliminary associations in humans are interesting, but interpretation and implications of measuring bulk methylation levels are limited and provide very little understanding of causality or the programs perturbed (activation of retrotransposons, changes at imprinted genes, etc.), and more genome-wide approaches are needed.

Sperm DNA Methylation Profiles and its Role in the Paternal Germline

Germ cells undergo extensive epigenetic reprogramming during proliferation and migration to the genital ridge (Seki *et al*., 2005, 2007). DNA methylation is erased and re-established in a sex and sequence specific manner during gametogenesis (Reik *et al*., 2001; Hajkova *et al*., 2002). The timing for methylation reestablishment differs between sexes and is continual for certain gene classes (Trasler, 2006). In males, *de novo* genomic methylation begins prenatally (prospermatogonia) at imprinted loci and repetitive elements with a general consensus being, that methylation patterns are completed by the end of pachytene stage of meiosis (Oakes *et al*., 2007a, b). However, exceptions have been reported at a few gene promoters that are expressed early in spermatogenesis (Pgk-2, ApoA1 and Oct-3/4) but are silenced in the maturing spermatid by gradually acquiring methylation in postmeiotic spermatocytes (Ariel *et al*., 1991, 1994).

Recent genome-wide methylation studies have indicated that the sperm epigenome differs markedly from that of somatic cells, but is very similar to ES cells and embryonic germ cells (EG) (Eckhardt *et al*., 2006, Oakes *et al*., 2007a, b; Weber *et al*., 2007; Farthing

et al., 2008). Examining sperm DNA methylation of chromosomes 6, 20 and 22 using restriction landmark genomic studies showed that many loci were differentially methylated between sperm and somatic cells. Extending methylation analysis genome-wide to all promoters of the human genome showed that the promoters that were differentially methylated in sperm and somatic tissue (or acquire methylation upon differentiation) were promoters with 'weak' CpG islands. More recently, we have shown in our gene ontology analysis that the hypomethylated promoters in the mature sperm are the promoters of developmental transcription and signaling factors. Interestingly, the DNA hypomethylated promoters in mature sperm greatly overlapped with the developmental promoters bound by the self renewal network transcription factors in human ES cells (e.g. OCT4, SOX2, NANOG, KLF4 and FOXD3 proteins) (Boyer *et al*., 2005). In ES cells, these pluripotency proteins promote self renewal and also work with repressive polycomb complexes (PRC2) to help repress a large set of developmental regulators (including *HOX* genes) to prevent differentiation (Cao *et al*., 2002; Bernstein *et al*., 2006; Lee *et al*., 2006; Muller and Kassis, 2006; Schwartz *et al*., 2006; Takahashi *et al*., 2007; Tanay *et al*., 2007; Wernig *et al*., 2007; Kopp *et al*., 2008). This overlap suggested that pluripotency or polycomb complex factors might be involved in the establishment and maintenance of sperm DNA methylation poising in sperm. Unfortunately, pluripotency and polycomb factors were not detected in the mature sperm, in fact many of the pluripotency promoters including several key members of the self-renewal network (*OCT4*, *NANOG*, *FOXD3*) themselves acquire methylation throughout spermatogenesis, whereas their developmental target genes remain hypomethylated, consistent with recent studies in mice (Down *et al*., 2008; Farthing *et al*., 2008; Illingworth *et al*., 2008; Mohn *et al*., 2008). These findings show that genes encoding early developmental transcription factors as well as signaling proteins are DNA hypomethylated and histone bound. Furthermore, developmental promoters are selectively methylated during development, which may help commit cells to differentiation decisions. Histone retention and DNA demethylation contribute to a poised state that ensures transcriptional competence and activation of developmental regulators in the early embryo.

Epigenetic alterations and male infertility

The incidence of infertility has been rising, currently affecting one in every seven couples in the western population. Male infertility is responsible for roughly half of the cases of infertility. The underlying cause of male infertility is unexplained in ~50% infertile men, and genetic causes have been proposed to be likely (Carrell, 2008a, b; Matzuk and Lamb, 2008). Several studies have explored possible genetic causes using mouse models (O'Bryan and de Kretser, 2006), candidate gene sequencing (Miyamoto *et al*., 2003; Aoki *et al*., 2006a, b, c; Hammoud *et al*., 2007, 2009a, b), and recently the first genome-wide association study in oligozoospermic and azoospermic men has been completed (Aston and Carrell, 2009). However, these studies have revealed that single gene polymorphisms are not likely to be the cause of most cases of male infertility, but male infertility is likely to be a multifactorial disease. Similar to other complex diseases, such as cancer, epigenetic alterations may be a component contributing to infertility.

As described above, recent studies have shown that epigenetic modifications in sperm (both histone modifications and DNA methylation) appear to poise the paternal genome to participate in early embryogenesis. Additionally, several studies indicate that DNA methylation is altered, in at least imprinted genes, oligozoospermic men and men with improper histone to protamine replacement (Marques *et al*., 2004, 2008; Bowdin *et al*., 2007; Doornbos *et al*., 2007; Hammoud, in press; Kobayashi *et al*., 2007, 2009). These observations beg the question of whether methylation defects, of both imprinted and non-imprinted genes, as well as other epigenetic defects (such as histone localization or modifications in the mature sperm), may play an important role in the development and growth of ART offspring (Manipalviratn *et al*., 2009). If epigenetic profiles of the mature sperm are critical, then alterations in epigenetic patterns in infertile males can provide a logic for the increased risk for preterm birth, low birthweight, congenital anomalies, perinatal mortality, and several other pregnancy-related complications seen at a higher frequency in babies conceived by IVF (Hansen *et al*., 2002; Kalra and Molinaro, 2008). This growing field of epigenetics in early gametes and embryos may be of benefit in understanding such observations (Carrell, 2008a, b). Current studies in our laboratory are focusing on genome-wide changes in histone localization as well as DNA methylation (discussed below) in male partners of recurrent pregnancy loss patients, repeated failed IVF patients and infertile males with an altered histone to protamine ratio.

Current findings showing that epigenetics patterns in germline are extensive and of potential significance, only strengthens further the previous associations that showed abnormal methylation of imprinted genes (genes expressed in a parent-of-origin manner), in the gametes of some infertile men or babies conceived by IVF (DeBaun *et al*., 2003; Gosden *et al*., 2003; Maher, 2005). A gain or loss of expression of imprinted genes has been implicated in many diseases (Jaenisch and Bird, 2003; Seitz *et al*., 2004; Morgan *et al*., 2005; Royo *et al*., 2006) including Beckwith–Weidemann syndrome (BWS) and Angelman's Syndrome (AS), both of which have been significantly correlated with IVF babies (DeBaun *et al*., 2003; Gosden *et al*., 2003; Maher, 2005).

Whether the increased incidence of imprinting abnormalities in IVF babies arises from *in vitro* manipulations of embryos or gametes, are due to ovulation induction medications, or are inherited from the gametes of infertile patients is unclear (Cummins and Jequier, 1994; de Kretser, 1995; Edwards and Ludwig, 2003; Marques *et al*., 2004, 2008; Bowdin *et al*., 2007, Doornbos *et al*., 2007, Kobayashi *et al*., 2007, 2009; Laprise, 2009). Support for all former hypotheses have been provided but this review will focus on one aspect, preexistent methylation alterations in the gametes of infertility patients (Cummins and Jequier, 1994; de Kretser, 1995; Edwards and Ludwig, 2003; Marques *et al*., 2004, 2008; Bowdin *et al*., 2007, Doornbos *et al*., 2007, Kobayashi *et al*., 2007, 2009; Laprise, 2009). This finding first surfaced a few years ago and showed that DNA methylation patterns at paternally imprinted loci are altered in the sperm of severely oligozoospermic patients (Marques *et al*., 2004, 2008). Subsequently in 2007, Houshdaran *et al*. reported that a broader alteration in DNA methylation in sperm is seen at a handful of imprinted loci, CpG islands upstream of gene promoters, and a few repetitive elements in infertile patients with poor semen parameters (Houshdaran *et al*., 2007). Furthermore, our lab has shown

that methylation alterations extend beyond severely oligozoospermic patients to patients with relatively normal sperm counts but with abnormal chromatin packaging, defined by an altered P1/P2 ratio (Hammoud, in press). Interestingly, in the small number of patients tested with either oligozoospermia or abnormal protamine expression we observed that methylation alterations varied between the two different etiologies of infertility. For instance oligozoospermic patients were hypermethylated at *MEST*, an imprinted gene associated with Silver Russell Syndrome (SRS), whereas abnormal protamine patients had significant changes at *LIT1* and small nuclear ribonucleoprotein polypeptide N (*SNRPN)*, genes that may be associated with cases of transient neonatal diabetes milletus (TNDM) and alternative splicing (AS). These findings suggest that the risk of transmitting epigenetic alterations may vary with the classification of infertility; however, it is important to note that not all patients or alleles were affected to a similar extent. The differences in the degree of methylation within some genes or alleles compared with others raises an important questions for future studies: whether there is a variable risk to the different CpGs and whether abnormal methylation has a threshold level for conferring disease risk in the embryo or is a gradual continuum.

Pre-existent methylation alterations in the gametes of infertile patients pose a risk for transgenerational epigenetic inheritance. Evidence for transgenerational epigenetic inheritance remains controversial in humans (Oswald *et al*., 2000; Reik *et al*., 2003; Morgan *et al*., 2005), however, considerable evidence for transgenerational epigenetic inheritance in mice has become apparent at intracisternal A particles (IAPs) and imprinted genes(Bultman *et al*., 1992; Morgan *et al*., 1999; Rakyan *et al*., 2003). In humans, evidence for germline epigenetic inheritance has come almost exclusively from epidemiological studies. The strongest evidence for germline epigenetic inheritance comes from the work of Horsthemke and colleagues (Buiting *et al*., 2003) where they have shown that the presence of epimutations and not genetic mutations, at the *SNRPN–SNURF upstream reading frame* locus was inherited from the paternal grandmother (Buiting *et al*., 2003). Furthermore, Kagami *et al*. (2007) showed that defective methylation at the DMR of *MEST* in sperm may have been inherited by an ART born baby with SRS. More recently, Kobayashi *et al.* showed abnormal DNA methylation at many imprinted loci in 17 of 78 assisted reproductive technology (ART) embryos (21.8%) tested. Although some of the imprinting errors identified may have risen during the ART process, however, in seven cases hypomethylation at H19 and GTL2 was present both in the sperm and in the embryo, suggesting that abnormal hypomethylation may be paternally inherited (Kobayashi *et al*., 2009). In summary, human transgenerational epigenetic inheritance is uncertain, but is possible if altered DNA methylation is inefficiently cleared between generations or if methylation pattern are not properly reestablished due to refractory elements such as retrotransposons or certain histone modifications in the embryo and PGCs.

Structural and *in vitro* data show that certain modified nucleosmes such as H3K4 methylation can deter DNA methylation in mice (Ooi *et al*., 2007). This interdependency or cooperativity relationship can have potential implications in reprogramming, especially at imprinted loci. Recent studies have shown that maternally and paternally imprinted alleles retain differential histone modifications (methylation and acetylation) to promote either allele activation or repression in somatic cells (Fournier *et al*., 2002; Delaval and Feil, 2004; Delaval *et al*., 2007). Consistent with the findings in somatic cells, imprinted genes (imprints are established in the gametes) in human sperm used similar poising mechanisms: H3K4me3 associated with many of the paternally expressed DMRs (Hammoud *et al*., 2009a, b), whereas, maternally imprinted (paternally repressed) loci lacked H3K4me3 and had moderate levels of H3K9me3, a repressive chromatin signature, residing at a few tested loci by qPCR. The presence of modified nucleosomes in the germline may serve as an epigenetic cellular memory to help reestablish and maintain parent of origin identity. However, in the cases of male infertility with altered histone retention this may be problematic at the time of reprogramming if the retained nucleosomes in the mature sperm are improperly placed or modified. Whether this is one of the underlying factors that contributes to the poor embryo outcome in patients with an abnormal histone to protamine ratio is unknown.

Conclusion and future directions

Sperm chromatin state is highly dynamic and retains important chromatin attributes that help facilitate the proper progression of spermatogenesis as well as being a potential contributor to early developmental processes. This continual epigenetic remodeling state may make sperm cells susceptible to impediments of environmental factors, aging process, or diseases such as infertility, but the ramifications of the altered chromatin states in the germ-line are not entirely known. Future studies are needed to establish perdurance of paternally retained modified nucleosomes in the early embryo, and their potential effects if abnormally retained.

References

Allen C, Bowdin S, Harrison RF, Sutcliffe AG, Brueton L, Kirby G, Kirkman-Brown J, Barrett C, Reardon W, Maher E. Pregnancy and perinatal outcomes after assisted reproduction: a comparative study. *Ir J Med Sci* 2008;**177**:233–241.

An W. Histone acetylation and methylation: combinatorial players for transcriptional regulation. *Subcell Biochem* 2007;**41**:351–369.

Aoki VW, Wilcox AL, Thorp C, Hamilton BD, Carrell DT. Improved *in vitro* fertilization embryo quality and pregnancy rates with intracytoplasmic sperm injection of sperm from fresh testicular biopsy samples vs. frozen biopsy samples. *Fertil Steril* 2004;**82**:1532–1535.

Aoki VW, Liu L, Carrell DT. Identification and evaluation of a novel sperm protamine abnormality in a population of infertile males. *Hum Reprod* 2005;**20**:1298–1306.

Aoki VW, Christensen GL, Atkins JF, Carrell DT. Identification of novel polymorphisms in the nuclear protein genes and their relationship with human sperm protamine deficiency and severe male infertility. *Fertil Steril* 2006a;**86**:1416–1422.

Aoki VW, Emery BR, Liu L, Carrell DT. Protamine levels vary between individual sperm cells of infertile human males and correlate with viability and DNA integrity. *J Androl* 2006b;**27**:890–898.

Aoki VW, Liu L, Jones KP, Hatasaka HH, Gibson M, Peterson CM, Carrell DT. Sperm protamine 1/protamine 2 ratios are related to *in vitro* fertilization pregnancy rates and predictive of fertilization ability. *Fertil Steril* 2006c;**86**:1408–1415.

Ariel M, McCarrey J, Cedar H. Methylation patterns of testis-specific genes. *Proc Natl Acad Sci USA* 1991;**88**:2317–2321.

Ariel M, Cedar H, McCarrey J. Developmental changes in methylation of spermatogenesis-specific genes include reprogramming in the epididymis. *Nat Genet* 1994;**7**:59–63.

Aston KI, Carrell DT. Genome-wide Study of Single Nucleotide Polymorphisms Associated with Azoospermia and Severe Oligozoospermia. *J Androl* 2009;**30**:711–725.

Baarends WM, Wassenaar E, van der Laan R, Hoogerbrugge J, Sleddens-Linkels E, Hoeijmakers JH, de Boer P, Grootegoed JA. Silencing of unpaired chromatin and histone H2A ubiquitination in mammalian meiosis. *Mol Cell Biol* 2005;**25**:1041–1053.

Balhorn R, Reed S, Tanphaichitr N. Aberrant protamine 1/protamine 2 ratios in sperm of infertile human males. *Experientia* 1988;**44**:52–55.

Balhorn R, Cosman M, Thornton K, Krishnan VV, Corzett M, Bench G, Kramer C, Lee J, Hud NV, Allen MJ *et al.* Protamine mediated condensation of DNA in mammalian sperm. In Gagnon C (ed). *The Male Gamete: From Basic Science to Clinical Applications*. Vienna, IL: Cache River Press, 1999, 55–70.

Balhorn R, Brewer L, Corzett M. DNA condensation by protamine and arginine-rich peptides: analysis of toroid stability using single DNA molecules. *Mol Reprod Dev* 2000;**56**:230–234.

Belokopytova IA, Kostyleva EI, Tomilin AN, Vorob'ev VI. Human male infertility may be due to a decrease of the protamine P2 content in sperm chromatin. *Mol Reprod Dev* 1993;**34**:53–57.

Benchaib M, Braun V, Ressnikof D, Lornage J, Durand P, Niveleau A, Guerin JF. Influence of global sperm DNA methylation on IVF results. *Hum Reprod* 2005;**20**:768–773.

Berger SL. Histone modifications in transcriptional regulation. *Curr Opin Genet Dev* 2002;**12**:142–148.

Bernstein BE, Humphrey EL, Erlich RL, Schneider R, Bouman P, Liu JS, Kouzarides T, Schreiber SL. Methylation of histone H3 Lys 4 in coding regions of active genes. *Proc Natl Acad Sci USA* 2002; **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 *et al.* Genomic maps and comparative analysis of histone modifications in human and mouse. *Cell* 2005;**120**:169–181.

Bernstein BE, Mikkelsen TS, Xie X, Kamal M, Huebert DJ, Cuff J, Fry B, Meissner A, Wernig M, Plath K *et al.* A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell* 2006; **125**:315–326.

Bestor TH. Activation of mammalian DNA methyltransferase by cleavage of a Zn binding regulatory domain. *EMBO J* 1992;**11**:2611–2617.

Bourc'his D, Bestor TH. Meiotic catastrophe and retrotransposon reactivation in male germ cells lacking Dnmt3L. *Nature* 2004; **431**:96–99.

Bourc'his D, Xu GL, Lin CS, Bollman B, Bestor TH. Dnmt3L and the establishment of maternal genomic imprints. *Science* 2001; **294**:2536–2539.

Bowdin S, Allen C, Kirby G, Brueton L, Afnan M, Barratt C, Kirkman-Brown J, Harrison R, Maher ER, Reardon W. A survey of assisted reproductive technology births and imprinting disorders. *Hum Reprod* 2007;**22**:3237–3240.

Boyer LA, Lee TI, Cole MF, Johnstone SE, Levine SS, Zucker JP, Guenther MG, Kumar RM, Murray HL, Jenner RG *et al.* Core transcriptional regulatory circuitry in human embryonic stem cells. *Cell* 2005;**122**:947–956.

Buiting K, Gross S, Lich C, Gillessen-Kaesbach G, el-Maarri O, Horsthemke B. Epimutations in Prader-Willi and Angelman syndromes: a molecular study of 136 patients with an imprinting defect. *Am J Hum Genet* 2003;**72**:571–577.

Bultman SJ, Michaud EJ, Woychik RP. Molecular characterization of the mouse agouti locus. *Cell* 1992;**71**:1195–1204.

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

Carrell DT, Liu L. Altered protamine 2 expression is uncommon in donors of known fertility, but common among men with poor fertilizing capacity, and may reflect other abnormalities of spermiogenesis. *J Androl* 2001;**22**:604–610.

Carrell DT, Emery BR, Hammoud S. Altered protamine expression and diminished spermatogenesis: what is the link? *Hum Reprod Update* 2007;**13**:313–327.

Carrell DT. Contributions of spermatozoa to embryogenesis: assays to evaluate their genetic and epigenetic fitness. *Reprod Biomed Online* 2008a;**16**:474–484.

Carrell DT. Elucidating the genetics of male infertility: understanding transcriptional and translational regulatory networks involved in spermatogenesis. *Int J Androl* 2008b;**31**:455–456.

Chevaillier P, Mauro N, Feneux D, Jouannet P, David G. Anomalous protein complement of sperm nuclei in some infertile men. *Lancet* 1987;**2**:806–807.

Christensen ME, Rattner JB, Dixon GH. Hyperacetylation of histone H4 promotes chromatin decondensation prior to histone replacement by protamines during spermatogenesis in rainbow trout. *Nucleic Acids Res* 1984;**12**:4575–4592.

Cohen ME. The 'brave new baby' and the law: fashioning remedies for the victims of *in vitro* fertilization. *Am J Law Med* 1978;**4**:319–336.

Corzett M, Mazrimas J, Balhorn R. Protamine 1: protamine 2 stoichiometry in the sperm of eutherian mammals. *Mol Reprod Dev* 2002;**61**:519–527.

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

Cummins JM, Jequier AM. Treating male infertility needs more clinical andrology, not less. *Hum Reprod* 1994;**9**:1214–1219.

Cutfield WS, Hofman PL, Mitchell M, Morison IM. Could epigenetics play a role in the developmental origins of health and disease? *Pediatr Res* 2007; **61**:68R–75R.

de Kretser DM. The potential of intracytoplasmic sperm injection (ICSI) to transmit genetic defects causing male infertility. *Reprod Fertil Dev* 1995; **7**:137–141; discussion 141–2.

de Yebra L, Ballesca JL, Vanrell JA, Corzett M, Balhorn R, Oliva R. Detection of P2 precursors in the sperm cells of infertile patients who have reduced protamine P2 levels. *Fertil Steril* 1998;**69**:755–759.

DeBaun MR, Niemitz EL, Feinberg AP. Association of *in vitro* fertilization with Beckwith–Wiedemann syndrome and epigenetic alterations of LIT1 and H19. *Am J Hum Genet* 2003;**72**:156–160.

Delaval K, Feil R. Epigenetic regulation of mammalian genomic imprinting. *Curr Opin Genet Dev* 2004;**14**:188–195.

Delaval K, Govin J, Cerqueira F, Rousseaux S, Khochbin S, Feil R. Differential histone modifications mark mouse imprinting control regions during spermatogenesis. *EMBO J* 2007;**26**:720–729.

Depa-Martynow M, Kempisty B, Lianeri M, Jagodzinski PP, Jedrzejczak P. Association between fertilin beta, protamines 1 and 2 and spermatid-specific linker histone H1-like protein mRNA levels, fertilization ability of human spermatozoa, and quality of preimplantation embryos. *Folia Histochem Cytobiol* 2007;**45**:S79–S85.

Doerksen T, Trasler JM. Developmental exposure of male germ cells to 5-azacytidine results in abnormal preimplantation development in rats. *Biol Reprod* 1996;**55**:1155–1162.

Doerksen T, Benoit G, Trasler JM. Deoxyribonucleic acid hypomethylation of male germ cells by mitotic and meiotic exposure to 5-azacytidine is associated with altered testicular histology. *Endocrinology* 2000; **141**:3235–3244.

Doornbos ME, Maas SM, McDonnell J, Vermeiden JP, Hennekam RC. Infertility, assisted reproduction technologies and imprinting disturbances: a Dutch study. *Hum Reprod* 2007;**22**:2476–2480.

Down TA, Rakyan VK, Turner DJ, Flicek P, Li H, Kulesha E, Graf S, Johnson N, Herrero J, Tomazou EM *et al.* A Bayesian deconvolution strategy for immunoprecipitation-based DNA methylome analysis. *Nat Biotechnol* 2008;**26**:779–785.

Eckhardt F, Lewin J, Cortese R, Rakyan VK, Attwood J, Burger M, Burton J, Cox TV, Davies R, Down TA *et al.* DNA methylation profiling of human chromosomes 6, 20 and 22. *Nat Genet* 2006;**38**:1378–1385.

Eden S, Cedar H. Role of DNA methylation in the regulation of transcription. *Curr Opin Genet Dev* 1994;**4**:255–259.

Edwards RG, Ludwig M. Are major defects in children conceived *in vitro* due to innate problems in patients or to induced genetic damage? *Reprod Biomed Online* 2003;**7**:131–138.

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

Feinberg AP. Phenotypic plasticity and the epigenetics of human disease. *Nature* 2007;**447**:433–440.

Fenic I, Sonnack V, Failing K, Bergmann M, Steger K. *In vivo* effects of histone-deacetylase inhibitor trichostatin-A on murine spermatogenesis. *J Androl* 2004;**25**:811–818.

Fenic I, Hossain HM, Sonnack V, Tchatalbachev S, Thierer F, Trapp J, Failing K, Edler KS, Bergmann M, Jung M *et al.* In vivo application of histone deacetylase inhibitor trichostatin-a impairs murine male meiosis. *J Androl* 2008;**29**:172–185.

Fernandez-Capetillo O, Liebe B, Scherthan H, Nussenzweig A. H2AX regulates meiotic telomere clustering. *J Cell Biol* 2003a;**163**:15–20.

Fernandez-Capetillo O, Mahadevaiah SK, Celeste A, Romanienko PJ, Camerini-Otero RD, Bonner WM, Manova K, Burgoyne P, Nussenzweig A. H2AX is required for chromatin remodeling and inactivation of sex chromosomes in male mouse meiosis. *Dev Cell* 2003b;**4**:497–508.

Fournier C, Goto Y, Ballestar E, Delaval K, Hever AM, Esteller M, Feil R. Allele-specific histone lysine methylation marks regulatory regions at imprinted mouse genes. *EMBO J* 2002;**21**:6560–6570.

Gatewood JM, Cook GR, Balhorn R, Bradbury EM, Schmid CW. Sequence-specific packaging of DNA in human sperm chromatin. *Science* 1987;**236**:962–964.

Gatewood JM, Cook GR, Balhorn R, Schmid CW, Bradbury EM. Isolation of four core histones from human sperm chromatin representing a minor subset of somatic histones. *J Biol Chem* 1990;**265**:20662–20666.

Glaser S, Lubitz S, Loveland KL, Ohbo K, Robb L, Schwenk F, Seibler J, Roellig D, Kranz A, Anastassiadis K *et al.* The histone 3 lysine 4 methyltransferase, Mll2, is only required briefly in development and spermatogenesis. *Epigenetics Chromatin* 2009;**2**:5.

Godmann M, Auger V, Ferraroni-Aguiar V, Di Sauro A, Sette C, Behr R, Kimmins S. Dynamic regulation of histone H3 methylation at lysine 4 in mammalian spermatogenesis. *Biol Reprod* 2007;**77**:754–764.

Goll MG, Bestor TH. Eukaryotic cytosine methyltransferases. *Annu Rev Biochem* 2005;**74**:481–514.

Gosden R, Trasler J, Lucifero D, Faddy M. Rare congenital disorders, imprinted genes, and assisted reproductive technology. *Lancet* 2003; **361**:1975–1977.

Grimes SR Jr, Henderson N. Hyperacetylation of histone H4 in rat testis spermatids. *Exp Cell Res* 1984;**152**:91–97.

Hajkova P, Erhardt S, Lane N, Haaf T, El-Maarri O, Reik W, Walter J, Surani MA. Epigenetic reprogramming in mouse primordial germ cells. *Mech Dev* 2002;**117**:15–23.

Hammoud S, Emery BR, Aoki VW, Carrell DT. Identification of genetic variation in the 5' and 3' non-coding regions of the protamine genes in patients with protamine deregulation. *Arch Androl* 2007; **53**:267–274.

Hammoud S, Emery BR, Dunn D, Weiss RB, Carrell DT. Sequence alterations in the YBX2 gene are associated with male factor infertility. *Fertil Steril* 2009a;**91**:1090–1095.

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

Hammoud S, Purwar J, Pflueger C, Cairns BR, Carrell DT. Alterations in sperm DNA methylation patterns at imprinted loci in two classes of infertility. *Fertil Steril* 2009; Epub ahead of print October 30.

Hansen M, Kurinczuk JJ, Bower C, Webb S. The risk of major birth defects after intracytoplasmic sperm injection and *in vitro* fertilization. *N Engl J Med* 2002;**346**:725–730.

Hata K, Okano M, Lei H, Li E. Dnmt3L cooperates with the Dnmt3 family of *de novo* DNA methyltransferases to establish maternal imprints in mice. *Development* 2002;**129**:1983–1993.

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

Houshdaran S, Cortessis VK, Siegmund K, Yang A, Laird PW, Sokol RZ. Widespread epigenetic abnormalities suggest a broad DNA methylation erasure defect in abnormal human sperm. *PLoS One* 2007; **2**:e1289.

Hoyer-Fender S. Molecular aspects of XY body formation. *Cytogenet Genome Res* 2003;**103**:245–255.

Hsieh CL. In vivo activity of murine *de novo* methyltransferases, Dnmt3a and Dnmt3b. *Mol Cell Biol* 1999;**19**:8211–8218.

Illingworth R, Kerr A, Desousa D, Jorgensen H, Ellis P, Stalker J, Jackson D, Clee C, Plumb R, Rogers J *et al.* A novel CpG island set identifies tissue-specific methylation at developmental gene loci. *PLoS Biol* 2008; **6**:e22.

Jaenisch R, Bird A. Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. *Nat Genet* 2003;**33**:245–254.

Jaenisch R, Jahner D. Methylation, expression and chromosomal position of genes in mammals. *Biochim Biophys Acta* 1984;**782**:1–9.

Jenuwein T, Allis CD. Translating the histone code. *Science* 2001; **293**:1074–1080.

Jones PA, Baylin SB. The epigenomics of cancer. *Cell* 2007;**128**:683–692.

Kagami M, Nagai T, Fukami M, Yamazawa K, Ogata T. Silver-Russell syndrome in a girl born after *in vitro* fertilization: partial hypermethylation at the differentially methylated region of PEG1/ MEST. *J Assist Reprod Genet* 2007;**24**:131–136.

Kalra SK, Molinaro TA. The association of *in vitro* fertilization and perinatal morbidity. *Semin Reprod Med* 2008;**26**:423–435.

Kaneda M, Okano M, Hata K, Sado T, Tsujimoto N, Li E, Sasaki H. Essential role for *de novo* DNA methyltransferase Dnmt3a in paternal and maternal imprinting. *Nature* 2004;**429**:900–903.

Katz DJ, Edwards TM, Reinke V, Kelly WG. A *C. elegans* LSD1 demethylase contributes to germline immortality by reprogramming epigenetic memory. *Cell* 2009;**137**:308–320.

Khalil AM, Boyar FZ, Driscoll DJ. Dynamic histone modifications mark sex chromosome inactivation and reactivation during mammalian spermatogenesis. *Proc Natl Acad Sci USA* 2004;**101**:16583–16587.

Kimmins S, Sassone-Corsi P. Chromatin remodelling and epigenetic features of germ cells. *Nature* 2005;**434**:583–589.

Kobayashi H, Sato A, Otsu E, Hiura H, Tomatsu C, Utsunomiya T, Sasaki H, Yaegashi N, Arima T. Aberrant DNA methylation of imprinted loci in sperm from oligospermic patients. *Hum Mol Genet* 2007;**16**:2542–2551.

Kobayashi H, Hiura H, John RM, Sato A, Otsu E, Kobayashi N, Suzuki R, Suzuki F, Hayashi C, Utsunomiya T *et al.* DNA methylation errors at

imprinted loci after assisted conception originate in the parental sperm. *Eur J Hum Genet.* 2009; Epub ahead of print May 27.

Kopp JL, Ormsbee BD, Desler M, Rizzino A. Small increases in the level of Sox2 trigger the differentiation of mouse embryonic stem cells. *Stem Cells* 2008;**26**:903–911.

Kurtz K, Martinez-Soler F, Ausio J, Chiva M. Acetylation of histone H4 in complex structural transitions of spermiogenic chromatin. *J Cell Biochem* 2007;**102**:1432–1441.

Lachner M, Jenuwein T. The many faces of histone lysine methylation. *Curr Opin Cell Biol* 2002;**14**:286–298.

Lahn BT, Tang ZL, Zhou J, Barndt RJ, Parvinen M, Allis CD, Page DC. Previously uncharacterized histone acetyltransferases implicated in mammalian spermatogenesis. *Proc Natl Acad Sci USA* 2002;**99**: 8707–8712.

Laprise SL. Implications of epigenetics and genomic imprinting in assisted reproductive technologies. *Mol Reprod Dev* 2009;**76**:1006–1018.

Lee MG, Wynder C, Cooch N, Shiekhattar R. An essential role for CoREST in nucleosomal histone 3 lysine 4 demethylation. *Nature* 2005;**437**:432–435.

Lee TI, Jenner RG, Boyer LA, Guenther MG, Levine SS, Kumar RM, Chevalier B, Johnstone SE, Cole MF, Isono K *et al.* Control of developmental regulators by Polycomb in human embryonic stem cells. *Cell* 2006;**125**:301–313.

Lei H, Oh SP, Okano M, Juttermann R, Goss KA, Jaenisch R, Li E. *De novo* DNA cytosine methyltransferase activities in mouse embryonic stem cells. *Development* 1996;**122**:3195–3205.

Li E. Chromatin modification and epigenetic reprogramming in mammalian development. *Nat Rev Genet* 2002;**3**:662–673.

Li E, Bestor TH, Jaenisch R. Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell* 1992; **69**:915–926.

Li E, Beard C, Forster AC, Bestor TH, Jaenisch R. DNA methylation, genomic imprinting, and mammalian development. *Cold Spring Harb Symp Quant Biol* 1993;**58**:297–305.

Lim D, Bowdin SC, Tee L, Kirby GA, Blair E, Fryer A, Lam W, Oley C, Cole T, Brueton LA *et al.* Clinical and molecular genetic features of Beckwith–Wiedemann syndrome associated with assisted reproductive technologies. *Hum Reprod* 2009;**24**:741–747.

Lu S, Xie YM, Li X, Luo J, Shi XQ, Hong X, Pan YH, Ma X. Mass spectrometry analysis of dynamic post-translational modifications of TH2B during spermatogenesis. *Mol Hum Reprod* 2009;**15**:373–378.

Lyko F, Ramsahoye BH, Kashevsky H, Tudor M, Mastrangelo MA, Orr-Weaver TL, Jaenisch R. Mammalian (cytosine-5) methyltransferases cause genomic DNA methylation and lethality in Drosophila. *Nat Genet* 1999;**23**:363–366.

Maher ER. Imprinting and assisted reproductive technology. *Hum Mol Genet* 2005;**14**:R133–R138.

Manipalviratn S, DeCherney A, Segars J. Imprinting disorders and assisted reproductive technology. *Fertil Steril* 2009;**91**:305–315.

Marques CJ, Carvalho F, Sousa M, Barros A. Genomic imprinting in disruptive spermatogenesis. *Lancet* 2004;**363**:1700–1702.

Marques CJ, Costa P, Vaz B, Carvalho F, Fernandes S, Barros A, Sousa M. Abnormal methylation of imprinted genes in human sperm is associated with oligozoospermia. *Mol Hum Reprod* 2008;**14**:67–74.

Matzuk MM, Lamb DJ. The biology of infertility: research advances and clinical challenges. *Nat Med* 2008;**14**:1197–1213.

McGraw S, Morin G, Vigneault C, Leclerc P, Sirard MA. Investigation of MYST4 histone acetyltransferase and its involvement in mammalian gametogenesis. *BMC Dev Biol* 2007;**7**:123.

Miyamoto T, Hasuike S, Yogev L, Maduro MR, Ishikawa M, Westphal H, Lamb DJ. Azoospermia in patients heterozygous for a mutation in SYCP3. *Lancet* 2003;**362**:1714–1719.

Mohn F, Weber M, Rebhan M, Roloff TC, Richter J, Stadler MB, Bibel M, Schubeler D. Lineage-specific polycomb targets and *de novo* DNA methylation define restriction and potential of neuronal progenitors. *Mol Cell* 2008;**30**:755–766.

Morgan HD, Sutherland HG, Martin DI, Whitelaw E. Epigenetic inheritance at the agouti locus in the mouse. *Nat Genet* 1999; **23**:314–318.

Morgan HD, Santos F, Green K, Dean W, Reik W. Epigenetic reprogramming in mammals. *Hum Mol Genet* 2005;**14**:R47–R58.

Muller J, Kassis JA. Polycomb response elements and targeting of Polycomb group proteins in Drosophila. *Curr Opin Genet Dev* 2006; **16**:476–484.

Nair M, Nagamori I, Sun P, Mishra DP, Rheaume C, Li B, Sassone-Corsi P, Dai X. Nuclear regulator Pygo2 controls spermiogenesis and histone H3 acetylation. *Dev Biol* 2008;**320**:446–455.

Ng HH, Bird A. DNA methylation and chromatin modification. *Curr Opin Genet Dev* 1999;**9**:158–163.

Oakes CC, La Salle S, Smiraglia DJ, Robaire B, Trasler JM. Developmental acquisition of genome-wide DNA methylation occurs prior to meiosis in male germ cells. *Dev Biol* 2007a;**307**:368–379.

Oakes CC, La Salle S, Smiraglia DJ, Robaire B, Trasler JM. A unique configuration of genome-wide DNA methylation patterns in the testis. *Proc Natl Acad Sci USA* 2007b;**104**:228–233.

O'Bryan MK, de Kretser D. Mouse models for genes involved in impaired spermatogenesis. *Int J Androl* 2006;**29**:76–89; discussion 105–108.

Okada Y, Scott G, Ray MK, Mishina Y, Zhang Y. Histone demethylase JHDM2A is critical for Tnp1 and Prm1 transcription and spermatogenesis. *Nature* 2007;**450**:119–123.

Okano M, Bell DW, Haber DA, Li E. DNA methyltransferases Dnmt3a and Dnmt3b are essential for *de novo* methylation and mammalian development. *Cell* 1999;**99**:247–257.

Oliva R. Protamines and male infertility. *Hum Reprod Update* 2006; **12**:417–435.

Ooi SK, Qiu C, Bernstein E, Li K, Jia D, Yang Z, Erdjument-Bromage H, Tempst P, Lin SP, Allis CD *et al.* DNMT3L connects unmethylated lysine 4 of histone H3 to *de novo* methylation of DNA. *Nature* 2007; **448**:714–717.

Oswald J, Engemann S, Lane N, Mayer W, Olek A, Fundele R, Dean W, Reik W, Walter J. Active demethylation of the paternal genome in the mouse zygote. *Curr Biol* 2000;**10**:475–478.

Panning B, Jaenisch R. DNA hypomethylation can activate Xist expression and silence X-linked genes. *Genes Dev* 1996;**10**:1991–2002.

Payne C, Braun RE. Histone lysine trimethylation exhibits a distinct perinuclear distribution in Plzf-expressing spermatogonia. *Dev Biol* 2006;**293**:461–472.

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**:5354–5365.

Rakyan VK, Chong S, Champ ME, Cuthbert PC, Morgan HD, Luu KV, Whitelaw E. Transgenerational inheritance of epigenetic states at the murine Axin(Fu) allele occurs after maternal and paternal transmission. *Proc Natl Acad Sci USA* 2003;**100**:2538–2543.

Raman R, Narayan G. 5-Aza deoxyCytidine-induced inhibition of differentiation of spermatogonia into spermatocytes in the mouse. *Mol Reprod Dev* 1995;**42**:284–290.

Rangasamy D, Berven L, Ridgway P, Tremethick DJ. Pericentric heterochromatin becomes enriched with H2A.Z during early mammalian development. *EMBO J* 2003;**22**:1599–1607.

Reefhuis J, Honein MA, Schieve LA, Correa A, Hobbs CA, Rasmussen SA. Assisted reproductive technology and major structural birth defects in the United States. *Hum Reprod* 2009;**24**:360–366.

Reik W, Dean W, Walter J. Epigenetic reprogramming in mammalian development. *Science* 2001;**293**:1089–1093.

Reik W, Santos F, Mitsuya K, Morgan H, Dean W. Epigenetic asymmetry in the mammalian zygote and early embryo: relationship to lineage commitment? *Philos Trans R Soc Lond B Biol Sci* 2003;**358**:1403–1409; discussion 1409.

Royo H, Bortolin ML, Seitz H, Cavaille J. Small non-coding RNAs and genomic imprinting. *Cytogenet Genome Res* 2006;**113**:99–108.

Schwartz YB, Kahn TG, Nix DA, Li XY, Bourgon R, Biggin M, Pirrotta V. Genome-wide analysis of Polycomb targets in Drosophila melanogaster. *Nat Genet* 2006;**38**:700–705.

Seif MM, Edi-Osagie EC, Farquhar C, Hooper L, Blake D, McGinlay P. Assisted hatching on assisted conception (IVF & ICSI). *Cochrane Database Syst Rev* 2006; CD001894.

Seifertova M, Vesely J, Cihak A. Enhanced mortality in offsprings of male mice treated with 5-azacytidine prior to mating. Morphological changes in testes. *Neoplasma* 1976;**23**:53–60.

Seitz H, Royo H, Bortolin ML, Lin SP, Ferguson-Smith AC, Cavaille J. A large imprinted microRNA gene cluster at the mouse Dlk1-Gtl2 domain. *Genome Res* 2004;**14**:1741–1748.

Seki Y, Hayashi K, Itoh K, Mizugaki M, Saitou M, Matsui Y. Extensive and orderly reprogramming of genome-wide chromatin modifications associated with specification and early development of germ cells in mice. *Dev Biol* 2005;**278**:440–458.

Seki Y, Yamaji M, Yabuta Y, Sano M, Shigeta M, Matsui Y, Saga Y, Tachibana M, Shinkai Y, Saitou M. Cellular dynamics associated with the genome-wide epigenetic reprogramming in migrating primordial germ cells in mice. *Development* 2007;**134**:2627–2638.

Shi Y, Lan F, Matson C, Mulligan P, Whetstine JR, Cole PA, Casero RA. Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. *Cell* 2004;**119**:941–953.

Sonnack V, Failing K, Bergmann M, Steger K. Expression of hyperacetylated histone H4 during normal and impaired human spermatogenesis. *Andrologia* 2002;**34**:384–390.

Steel AJ, Sutcliffe A. Long-term health implications for children conceived by IVF/ICSI. *Hum Fertil (Camb)* 2009;**12**:21–27.

Suetake I, Shinozaki F, Miyagawa J, Takeshima H, Tajima S. DNMT3L stimulates the DNA methylation activity of Dnmt3a and Dnmt3b through a direct interaction. *J Biol Chem* 2004;**279**: 27816–27823.

Suganuma T, Workman JL. Crosstalk among Histone Modifications. *Cell* 2008;**135**:604–607.

Surani MA. Imprinting and the initiation of gene silencing in the germ line. *Cell* 1998;**93**:309–312.

Takahashi K, Tanabe K, Ohnuki M, Narita M, Ichisaka T, Tomoda K, Yamanaka S. Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell* 2007;**131**:861–872.

Tanay A, O'Donnell AH, Damelin M, Bestor TH. Hyperconserved CpG domains underlie Polycomb-binding sites. *Proc Natl Acad Sci USA* 2007;**104**:5521–5526.

Tanphaichitr N, Sobhon P, Taluppeth N, Chalermisarachai P. Basic nuclear proteins in testicular cells and ejaculated spermatozoa in man. *Exp Cell Res* 1978;**117**:347–356.

Trasler JM. Gamete imprinting: setting epigenetic patterns for the next generation. *Reprod Fertil Dev* 2006;**18**:63–69.

Walsh CP, Chaillet JR, Bestor TH. Transcription of IAP endogenous retroviruses is constrained by cytosine methylation. *Nat Genet* 1998; **20**:116–117.

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

Weber M, Hellmann I, Stadler MB, Ramos L, Paabo S, Rebhan M, Schubeler D. Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. *Nat Genet* 2007;**39**:457–466.

Webster KE, O'Bryan MK, Fletcher S, Crewther PE, Aapola U, Craig J, Harrison DK, Aung H, Phutikanit N, Lyle R *et al.* Meiotic and epigenetic defects in Dnmt3L-knockout mouse spermatogenesis. *Proc Natl Acad Sci USA* 2005;**102**:4068–4073.

Wernig M, Meissner A, Foreman R, Brambrink T, Ku M, Hochedlinger K, Bernstein BE, Jaenisch R. *In vitro* reprogramming of fibroblasts into a pluripotent ES-cell-like state. *Nature* 2007;**448**:318–324.

Wykes SM, Krawetz SA. The structural organization of sperm chromatin. *J Biol Chem* 2003;**278**:29471–29477.

Zhu B, Zheng Y, Pham AD, Mandal SS, Erdjument-Bromage H, Tempst P, Reinberg D. Monoubiquitination of human histone H2B: the factors involved and their roles in HOX gene regulation. *Mol Cell* 2005;**20**:601–611.

Submitted on September 23, 2009; resubmitted on October 14, 2009; accepted on October 15, 2009

CHAPTER 5

DISTINCTIVE CHROMATIN IN HUMAN SPERM PACKAGES GENES FOR EMBRYO DEVELOPMENT

Published in *Nature*

Authors

Saher Sue Hammoud

David A. Nix

Haiying Zhang

Jahnvi Purwar

Douglas T. Carrell

Bradley R. Cairns

Nature, 2009, Vol. 460, pages 473-479.

doi:10.1038/nature08162 nature

ARTICLES

Distinctive chromatin in human sperm packages genes for embryo development

Saher Sue Hammoud[1,2], David A. Nix[3], Haiying Zhang[1], Jahnvi Purwar[1], Douglas T. Carrell[2] & Bradley R. Cairns[1]

Because nucleosomes are widely replaced by protamine in mature human sperm, the epigenetic contributions of sperm chromatin to embryo development have been considered highly limited. Here we show that the retained nucleosomes are significantly enriched at loci of developmental importance, including imprinted gene clusters, microRNA clusters, *HOX* gene clusters, and the promoters of stand-alone developmental transcription and signalling factors. Notably, histone modifications localize to particular developmental loci. Dimethylated lysine 4 on histone H3 (H3K4me2) is enriched at certain developmental promoters, whereas large blocks of H3K4me3 localize to a subset of developmental promoters, regions in *HOX* clusters, certain noncoding RNAs, and generally to paternally expressed imprinted loci, but not paternally repressed loci. Notably, trimethylated H3K27 (H3K27me3) is significantly enriched at developmental promoters that are repressed in early embryos, including many bivalent (H3K4me3/H3K27me3) promoters in embryonic stem cells. Furthermore, developmental promoters are generally DNA hypomethylated in sperm, but acquire methylation during differentiation. Taken together, epigenetic marking in sperm is extensive, and correlated with developmental regulators.

During spermiogenesis canonical histones are largely exchanged for protamines[1,2], small basic proteins that form tightly packed DNA structures important for normal sperm function[3]. We find about 4% of the haploid genome retained in nucleosomes (Supplementary Fig. 1a). The rare retained nucleosomes in sperm consist of either canonical or histone variant proteins, including a testes-specific histone H2B (TH2B) with an unknown specialized function[4,5]. Their presence may simply be due to inefficient protamine replacement, leading to a low random distribution genome-wide with no impact in the embryo. Alternatively, these retained nucleosomes, along with attendant modifications, might be enriched at particular genes/loci. This latter possibility would raise the possibility for programmatic retention for an epigenetic function in the embryo. To address these questions, we localized the nucleosomes retained in mature sperm from fertile donors using high-resolution genomic approaches.

Developmental loci bear nucleosomes

To address donor variability, we examined nucleosome retention in a single donor (D1) and/or a pool of four donors (donor pool). Sperm chromatin was separated into protamine-bound and histone-bound fractions. In brief, mononucleosomes were isolated (>95% yield) by sequential MNase digestion and sedimentation (Supplementary Fig. 1b–e). This mononucleosome pool was used for chromatin immunoprecipitation (ChIP; to select modified nucleosomes), or the DNA was isolated from the mononucleosome pool to represent all nucleosomes. Purified DNA was subjected to high-throughput sequencing (Illumina GAII), or alternatively, was labelled and hybridized to a high-density promoter-tiling array (9 kilobase (kb) tiled; Supplementary Fig. 2, schematic).

Our initial array approach examined three replicas of D1 (pairwise average $R^2 = 0.85$). Notably, Gene Ontology analysis revealed nucleosomes significantly enriched at promoters that guide embryonic development—primarily developmental transcription factors and signalling molecules (Gene Ontology term false discovery rate (FDR) < 0.01; Box 1 and Supplementary Table 1; for all extended Gene Ontology categories see Supplementary Tables and Supplementary Data Set 1). To conduct genome-wide profiling, we performed high-throughput sequencing of nucleosomes from D1 or the donor pool. Regions significantly enriched for histone relative to the input control (sheared total sperm DNA) were identified using a 300-base-pair (bp) window metric[6]. For display, we depict the normalized difference score and FDR window scores (Fig. 1a, FDR transformation ($-10 \log_{10}$ (q-value FDR)), $20 = 0.01$, $25 = 0.003$, $30 = 0.001$, and $40 = 0.0001$). Histone-enriched loci for one individual (D1) were well correlated with a donor pool ($r = 0.7$). Globally, 76% of the top 9,841 histone-enriched regions (FDR 40 cutoff) intersect genic regions, whereas the expected intersection given random distribution is 36% ($P < 0.001$).

Interestingly, sequencing of D1 or the donor pool revealed significant (FDR < 0.001) histone retention at many loci important for embryo development, including embryonic transcription factors and signalling pathway components (Box 1, Supplementary Tables 2 and 3). We show this enrichment at *HOX* loci (Fig. 1, Supplementary Fig. 3), but also observe this at stand-alone developmental transcription factors (Supplementary Fig. 4) and signalling factors (Supplementary Fig. 5). An FDR of 60 yields 4,556 genes, of which 1,683 are grouped with developmental Gene Ontology categories (2,848 total developmental genes). The magnitude of nucleosome enrichment at developmental loci is modest, with high significance provided by a moderate average increase at a large number of loci. Histones are also significantly enriched at the promoters of microRNAs (miRNAs) ($P < 0.05$; Supplementary Fig. 6) and at the class of imprinted genes ($P < 0.0001$; Fig. 2), addressed in detail later. Selected loci were tested and confirmed by quantitative PCR (qPCR; Supplementary Fig. 7a–e). Outside of these enriched regions, we observe sequencing reads at low levels distributed genome-wide (for example, Figs 1a and 2a), an observation consistent with low levels of nucleosomes genome-wide, although contributions from non-nucleosomal contamination cannot be ruled out.

[1]Howard Hughes Medical Institute, Department of Oncological Sciences, and Huntsman Cancer Institute, [2]IVF and Andrology Laboratories, Departments of Surgery, Obstetrics and Gynecology, and Physiology, [3]Research Informatics and Bioinformatics Core Facility, Huntsman Cancer Institute, University of Utah School of Medicine, Salt Lake City, Utah 84112, USA.

Box 1 | Developmental genes are associated with particular chromatin attributes in human sperm

GoMiner was used to identify enriched categories, and all categories displayed have an FDR < 0.01. The top five general categories are listed, after omitting nearly identical/redundant classes. An expanded gene ontology table with the unfiltered top 30–60 categories, the total genes, number of changed genes, enrichment, and FDR are provided in the Supplementary Information.

Nucleosomes, Array D1
(1) Sequence-specific DNA binding; (2) multicellular organismal development; (3) regulation of transcription; (4) developmental process; (5) regulation of metabolic process.

Nucleosomes, Illumina GAII pooled donors
(1) Transcription factor activity; (2) cell fate commitment; (3) WNT receptor signalling; (4) neuron development; (5) embryonic development.

H3K4me2, Array D1
(1) Multicellular organismal development; (2) developmental process; (3) sequence-specific DNA binding; (4) anatomical structure development; (5) system development.

H3K4me3, Array D1
(1) mRNA processing; (2) RNA binding; (3) cell cycle; (4) transcription; (5) RNA splicing.

H3K4me3, Illumina GAII pooled donors
(1) RNA splicing; (2) translation; (3) cell cycle; (4) RNA metabolic process; (5) transcription.

H3K27me3, Illumina GAII pooled donors
(1) WNT receptor signalling; (2) embryonic organ development and morphogenesis; (3) cell fate commitment; (4) neuron differentiation; (5) sequence-specific DNA binding.

DNA hypomethylated promoters D1 and D2
(1) Embryonic development; (2) multicellular organismal development; (3) system development; (4) RNA biosynthetic process; (5) transcription factor activity.

DNA methylated promoters omitting CpG islands, array
(1) Transcription; (2) RNA biosynthetic process; (3) regulation of transcription; (4) embryonic development; (5) embryo morphogenesis.

Protamine occupancy (two replicas, $R^2 = 0.89$, arrays only) yielded 7,151 enriched regions (>2.5-fold), but failed to identify any enriched Gene Ontology term categories, although a few segments of the Y chromosome were notably enriched (including the testis-specific *TSPY* genes, data not shown). Regions of histone enrichment did not exclude protamine, consistent with a nucleosome-protamine mixture existing even at histone-enriched loci. However, as protamine fragments averaged ~750 bp, protamine depletion would have to be extensive (regions >2 kb) to be apparent on our arrays. Taken together, nucleosomes are significantly enriched in sperm at genes important for embryonic development, with transcription factors the most enriched class.

Localization of modified nucleosomes

Because histones replace protamines genome-wide at fertilization[7,8], unmodified histones retained in sperm would seem insufficient to influence gene regulation in embryos. Therefore, we examined three further chromatin properties in sperm: (1) histone variants, (2) histone modifications, and (3) DNA methylation. ChIP combined with promoter microarray analysis (termed ChIP-chip) of TH2B (two replicas, $R^2 = 0.93$) shows 0.3% of gene promoters with relatively high levels of TH2B (>twofold enrichment). Gene Ontology analysis showed significant (FDR < 0.06) enrichment at genes important for sperm biology, capacitation and fertilization (Supplementary Table 4), but not at developmental categories. ChIP sequencing (ChIP-seq) analysis with H2A.Z nucleosomes (at standard conditions, 150–250 mM salt) did not show significant enriched Gene Ontology categories, with high enrichment limited to pericentric heterochromatin (Supplementary Fig. 8), consistent with prior immunostaining[9].

Modified nucleosomes were localized by performing ChIP on mononucleosomes, followed by either array analysis or sequencing (Supplementary Fig. 2, schematic). We normalized the data set for each modification to the data set derived from input mononucleosomes, determined enriched regions (array >twofold; sequencing FDR 40), found the nearest neighbouring gene, and performed Gene Ontology analysis. In somatic cells, H3K4me2 is correlated with euchromatic regions. In sperm, H3K4me2 was enriched at many promoters, and at significant levels at promoters for developmental transcription factors (two replicas $R^2 = 0.94$; Gene Ontology term FDR < 0.06; Box 1 and Supplementary Table 5). In somatic cells, H3K4me3 is localized to: (1) the transcription start sites (TSS) of active genes, (2) genes bearing 'poised' RNA polymerase II (Pol II), and (3) the proximal promoter of inactive developmental regulators in embryonic stem (ES) cells—promoters that also bear the silencing mark H3K27me3 (refs 10, 11), and thus termed bivalent. Mature sperm are transcriptionally inert, and Pol II protein levels are barely detectable (data not shown), so the high H3K4me3 levels we observed in sperm chromatin (Supplementary Fig. 1f) seemed surprising. H3K4me3 was localized by both ChIP-chip (three replicas, $R^2 = 0.96$) and ChIP-seq. The raw data sets were similar ($r = 0.7$) and the thresholded data sets were very similar (array twofold; sequencing, FDR 40; 96% intersection, $P < 0.001$). With both data sets, simple inspection showed small peaks at many 5′ gene ends, with high levels and broader blocks at a subset of genes (that is, *HOX* loci; Fig. 1 and Supplementary Fig. 3). Gene Ontology term analyses with either data set yielded genes that are important for changing nuclear architecture, RNA metabolism, spermatogenesis, and also selected transcription factors important for embryonic development (FDR < 0.01, Box 1, Supplementary Tables 6 and 7 and Supplementary Fig. 9). H3K4me3 at genes related to nuclear architecture and spermatogenesis can presumably be attributed to their prior activation during gametogenesis. RNA metabolism occurs both in gametogenesis and the early embryo, so attribution to a prior program as opposed to a potential poising for a future program cannot be unambiguously attributed. However, several transcription and signalling factors of importance in embryo development exhibited high levels and a broad distribution of H3K4me3, including *EVX1/2*, *ID1*, *STAT3*, *KLF5*, *FGF9*, *SOX7/9*, certain *HOX* genes, and certain noncoding RNAs (Fig. 1 and Supplementary Figs 3 and 6).

Interestingly, ChIP-seq analysis showed significant levels of H3K27me3 at developmental promoters in sperm (Box 1, Fig. 1b, Supplementary Table 8 and Supplementary Figs 3 and 4), and overlapped significantly with H3K27me3-occupied genes in ES cells ($P < 0.01$), which are silent before differentiation. Furthermore, bivalent genes (bearing H3K4me3 and H3K27me3) in ES cells had a significant overlap with bivalent genes in sperm (FDR < 0.001 for each mark). Of the 1,999 genes identified as bivalent in ES cells, 861 were bivalent in sperm ($P < 0.01$; Supplementary Table 9). Also notable but not explored further were many blocks of high H3K4me3 or H3K27me3 in regions lacking annotation (Fig. 1a, oval). Furthermore, H3K9me3 was not detected at the small set developmental promoters tested, but was high at pericentric regions (qPCR only, Supplementary Fig. 7d). Taken together, our results demonstrate extensive histone modification patterns in sperm, and significant similarities to patterns observed in ES cells.

DNA methylation profiles

DNA methylation profiles examined two fertile donors (D2 and D4) using a methylated DNA immunoprecipitation (MeDIP) procedure

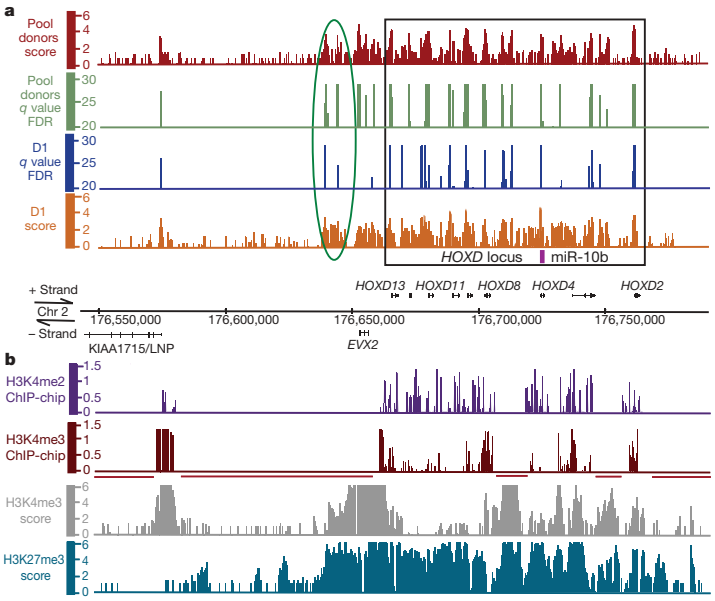


Figure 1 | Profiling of nucleosomes and their modifications at *HOXD*. For high-throughput sequencing, we show the mapped sequencing reads from D1 or a donor pool (red or orange bars, respectively; normalized difference score), and their significance (green or blue bars; FDR of 20 is <1% and FDR of 30 is <0.1%). **a**, The *HOXD* locus (black box) and an uncharacterized flanking locus (green oval). **b**, Profiling of nucleosome modifications at *HOXD* (in part **a**). The *y* axis is signal intensity ($\log_2$, for ChIP-chip), or the normalized difference score for sequencing. The regions not tiled on the array are underlined in red. Chr, chromosome.

and promoter arrays (individual replicates average D2 $R^2 = 0.97$ and D4 $R^2 = 0.89$). Their methylation patterns were highly similar (pairwise $R^2 = 0.86$), and extensive qPCR validated our array threshold (Supplementary Fig. 7e). Gene Ontology analysis of genes with pronounced DNA hypomethylation yielded transcription and signalling factors that guide embryo development (FDR < 0.05; Box 1 and Supplementary Table 10) including *HOX* loci (Fig. 3, blue bars, and Supplementary Figs 4 and 10). Hypomethylation also overlapped very significantly with histone-enriched promoters ($P < 0.02$; Supplementary Table 11). Bisulphite sequencing verified the MeDIP results, revealing extensive hypomethylation at developmental promoters in sperm (Supplementary Fig. 10b, c).

Notably, DNA-hypomethylated promoters in mature sperm overlap greatly with developmental promoters bound by the self-renewal network of transcription factors in human ES cells (for example, OCT4 (also known as POU5F1), SOX2, NANOG, KLF4 and FOXD3 proteins[12]; intersection of OCT4 protein occupancy and DNA hypomethylation, $P < 0.01$). In ES cells, these proteins promote self-renewal and also work with repressive polycomb complexes (PRC2; containing core component SUZ12) to help repress a large set of developmental regulators (including *HOX* genes) to prevent differentiation[10,13–20]. However, the hypomethylation of developmental genes in sperm is extensive (Fig. 3 and Supplementary Fig. 4). In fact, when CpG islands are omitted from the data sets, Gene Ontology term analysis of hypomethylated promoters still yields developmental genes (Box 1 and Supplementary Table 12). Notably, many of these developmental genes become methylated after differentiation; differential analysis of sperm and primary human fibroblasts (MeDIP, two replicas $R^2 = 0.86$) showed that many promoters occupied by PRC2 in human ES cells acquire methylation in fibroblasts (FDR < 0.01, Supplementary Tables 13 and 14; *HOXD* illustrated in Fig. 3, Supplementary Figs 4 and 5). Furthermore, the promoters driving several key members of the self-renewal network are themselves markedly hypermethylated in sperm (*OCT4*, *NANOG* and *FOXD3*, bisulphite sequencing in Supplementary Fig. 10c), whereas their developmental target genes are hypomethylated (bisulphite sequencing in Supplementary Fig. 10b), consistent with recent studies in mice[21–24].

Attributes of *HOX* clusters and miRNAs

Nucleosome enrichment was clear across *HOX* loci and proximal flanking regions, but falls off precipitously outside (*HOXD*, Fig. 1a; *HOXA*, Supplementary Fig. 3a). Histone-enriched *HOXD* regions with a single donor (D1) were largely shared with the donor pool (Fig. 1a; D1 versus donor pool, $r = 0.7$). Notably, retained nucleosomes have regional covalent modifications. For example, distinct and very large (5–20 kb) blocks of H3K4me3 are clearly observed at all *HOX* loci, and also at certain imprinted genes (addressed later). At *HOXD*, high H3K4me3 extends for ~20 kb, encompassing all of *EVX2* and extending to the 3′ region of *HOXD13* (Fig. 1b). Remarkably, a similar profile is observed at the related *HOXA* locus (Supplementary Fig. 3a). At *HOXD* a second block of H3K4me3 is observed in the region between *HOXD4* and *HOXD8* (Fig. 1b), a region that encodes several noncoding RNAs expressed during development. This region represents a marked difference from the chromatin status in ES cells; in ES cells *HOXD8–D11* are all bivalent. The distribution of H3K4me2 (determined from two replicas of D1) is clearly different from H3K4me3 at *HOX* loci (Fig. 1b and Supplementary Fig. 3). For example, at *HOXD*, H3K4me2 is enriched in *HOXD8–D11*, a region deficient in H3K4me3 (Fig. 1b). Notably, high H3K27me3 encompasses all *HOX* loci and their proximal flanking regions. In contrast, high levels of H3K9me (a mark of heterochromatin; Supplementary Fig. 7d) or H2A.Z were not detected at the *HOX* loci tested.

Histones are enriched at many miRNAs, especially miRNA clusters (Supplementary Fig. 6). For example, 16 of the 29 miRNA clusters on autosomes were significantly enriched ($P < 0.05$). Clusters include those bearing *let7e*, *mir-17*, *mir-15a*, *mir-96*, *mir-135b* and *mir-10a/*

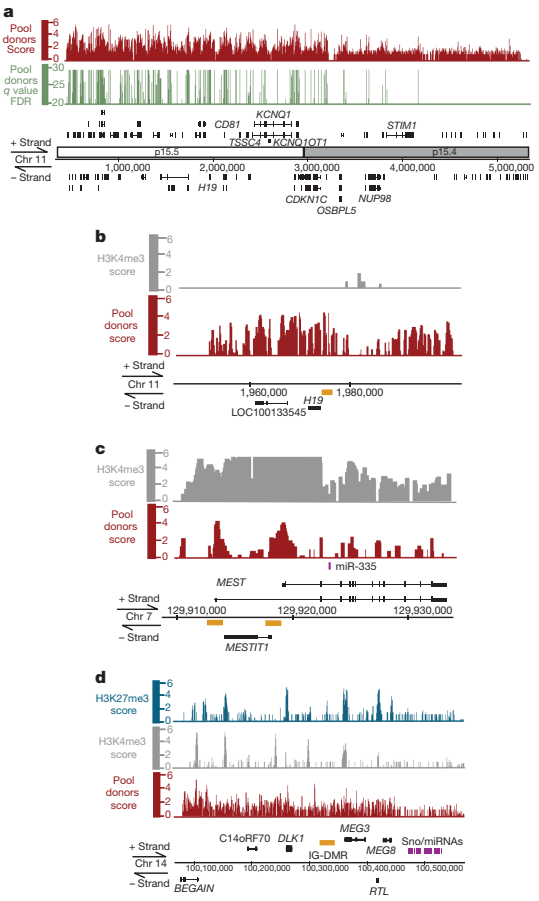


Figure 2 | Nucleosome enrichment at imprinted gene clusters, with high H3K4me3 at paternally expressed noncoding RNAs, and paternally demethylated regions. **a**, Histone enrichment at the 11p15.5 imprinted cluster (ending near *OSBPL5*), but not in the adjacent region. **b**, **c**, An expanded view of the DMRs (yellow rectangles) of *H19* (paternally methylated) (**b**) and *MEST* (paternally demethylated) (**c**). **d**, Moderate H3K4me3 at the promoters of the paternally expressed genes *BEGAIN*, *DLK1* and *RTL*, and the lack of H3K4me3 at the methylated intergenic-differentially methylated region (IG-DMR) of *MEG3* in sperm. Notably, both H3K4me3 and H3K27me3 reside at the promoter of *MEG3*, which later acquires DNA methylation in the embryo. Sno, small nucleolar.

b, as well as the stand-alone miRNAs *mir-153-1*, *mir-488* and *mir-760*. Notably, many histone-occupied miRNAs are associated with embryonic development[25] ($P < 0.01$), and their promoters were largely hypomethylated (Supplementary Fig. 10d). Furthermore, 7 of the 12 miRNAs on autosomes that are occupied by OCT4, NANOG and SOX2 in human ES cells[17] are also significantly occupied by histone (from pooled sequencing data). However, we do not at present understand the logic for their modification status; certain miRNA clusters have high histone and bivalent status, whereas others lack either modification (Supplementary Fig. 6).

Attributes of primary and secondary imprinted genes

Nucleosomes are significantly enriched at most imprinted genes in sperm, but at both paternally and maternally expressed loci. However, we observe marked specificity of H3K4me3 localization, with high and broad levels present at genes and noncoding RNAs that are paternally expressed. Locus 11p15.5 (Fig. 2a) is a large imprinted cluster with *IGF2*, *H19* and *KCNQ1* and several miRNAs. Here, increased levels of histone are present throughout the imprinted region (up to *OSBPL5*), but not in the large adjacent region lacking imprinted genes (Fig. 2a). Notably, the paternally silenced *H19* locus upstream of *KCNQ1* has a methylated DMR (Supplementary Fig. 10a) that lacks H3K4me3 (Fig. 2b). In contrast, *MEST* (a paternally expressed gene) has high H3K4me3 that extends from its promoter and first exon (containing the demethylated differentially methylated region (DMR); Fig. 2c and Supplementary Fig. 10a) through the second exon. The antisense noncoding RNA *MESTIT1* (also paternally expressed) is transcribed from the first intron, and is also very high in H3K4me3 (Fig. 2c). Furthermore, the promoter region of the paternally expressed antisense noncoding RNA *KCNQ1OT1* displays H3K4me3 (Fig. 2a and data not shown), and the DMR is DNA demethylated (Supplementary Fig. 10a). Several other examples of paternally expressed loci with blocks of H3K4me3 are provided in Supplementary Fig. 11, including *PEG3*, the noncoding RNAs *AIRN* (antisense to *IGF2R*) and *GNASAS* (antisense to *GNAS*). In contrast, genes flanking *KCNQ1* that are repressed by the noncoding RNA *KCNQOT1* (such as *OSBPL5*, *TSSC4* and *CD81*; Fig. 2a, expanded in Supplementary Fig. 11) contain histone, but lack H3K4me3. Notably, several paternally silenced genes (bearing DNA methylation) bore moderate (2–3-fold) enrichment of H3K9me3, a mark absent at paternally expressed genes (Supplementary Fig. 7d).

The 14q32.33 region (*DLK-DIO3*) is complex and interesting; paternally expressed genes such as *DLK1* and *RTL1* have moderate levels of H3K4me3 in their promoters, and the imprinting control locus (IG-DMR) lacks H3K4me3 (Fig. 2d) and is DNA methylated[26–28]. Notably, the promoter of *MEG3* (also known as *GTL2*; just downstream of the IG-DMR) lacks DNA methylation in sperm, but acquires DNA methylation in the embryo[26–28], termed secondary imprinting. Notably, the *MEG3* promoter region that later acquires DNA methylation initially bears both H3K4me3 and H3K27me3 in sperm; it is bivalent. One interpretation is that for mature sperm and early embryos, H3K4me3 prevents DNA methylation while H3K27me3 promotes silencing, with subsequent H3K4me removal enabling tissue-specific DNA methylation and secondary imprinting. Furthermore, our examination of the X chromosome inactivation centre showed an apparent bivalent status (and DNA hypomethylation) at the TSS of the *XIST* noncoding RNA, but not at *TSIX*,

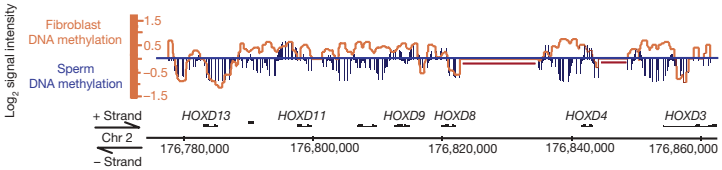


Figure 3 | Developmental promoters in sperm lack DNA methylation, but acquire methylation during development. DNA methylation of the *HOXD* locus in the mature sperm (blue bars) or primary fibroblasts (orange line overlay). The *y* axis is the signal intensity ($\log_2$) and the *x* axis is the annotated physical map (HG17). The regions not tiled on the array are underlined in red.

although future studies are required to determine whether these marks influence the regulation of this locus in the embryo (Supplementary Figs 6 and 10d; note that sequence reads on the X chromosome are half that on autosomes, as it is only present in 50% of sperm).

Modifications and expression timing

Transcriptome analysis has been performed in 4-cell and 8-cell human embryos, with 29 or 65 messenger RNAs identified as enriched, respectively[29]. Notably, genes in sperm bearing H3K4me3 but not H3K27me3 correlated with genes expressed at the 4-cell stage (14 out of 24, $P = 0.059$). Also, genes bearing high H3K4me2 were significantly enriched at genes expressed in the 4–8-cell stage (23 out of 49, $P < 0.02$; only 49 tiled on our array). In contrast, no significant correlation was observed with H3K27me3, which instead associates with transcription factors required for differentiation and organogenesis (discussed earlier). Furthermore, we verified by qPCR the presence of H3K4me2 or H3K4me3 at a subset of these stage-specific gene promoters (Supplementary Fig. 12). Thus, these findings reveal correlations of H3K4me2/3 enrichment, but not H3K27 enrichment, with early expression.

Conclusion

We provide several lines of evidence that the parental genome is packaged and covalently modified in a manner consistent with influencing embryo development. Previous analyses of DNA methylation in sperm identified hypomethylated promoters[23,24,30,31], showed similarities to the pattern in ES cells[24,31], and overlap between PRC2 and CpG islands[15,17,21,22]. We add that hypomethylated developmental promoters in human sperm overlap significantly with developmental promoters (in ES cells) occupied by the self-renewal network. Also, the promoters that acquire methylation in fibroblasts are primarily developmental transcription factors that are bound by PRC2 in human ES cells, consistent with recent work linking PRC2 to DNA methylation in development and neuronal differentiation in mice[21,32,33]. Thus, components of the self-renewal network emerge as candidates for helping to direct DNA hypomethylation in the germ line, and also to guide DNA hypermethylation to particular loci during differentiation, possibly to help 'lock in' differentiation decisions, although this remains to be tested.

The central findings of our work involve the significant enrichment of modified nucleosomes in the sperm genome at genes for embryo development, and a specificity to their modification patterns that might be instructive for the regulation of developmental genes, noncoding RNAs and imprinted loci. For example, histone retention and modification were clear at *HOX* loci and most of the targets of the self-renewal network in ES cells. One key concept in ES cell chromatin is the prevalence of developmental promoters with a bivalent status—bearing both H3K27me3 and H3K4me3 (ref. 10). Many promoters bivalent in ES cells are also bivalent in sperm, although some bear only H3K27me3 in sperm. Notably, H3K27me3 covers essentially all of the four *HOX* loci in sperm, whereas H3K4me3 is present in large blocks at only a subset of locations in *HOX* loci. Our work also provides correlations between H3K4me, but not H3K27me, and early expression in the embryo. In contrast, protamine-enriched loci did not show any significant Gene Ontology categories. However, there were certain segments of the Y chromosome with protamine enrichment, including the testis-specific *TSPY* genes, although the significance is not known.

We also find histones enriched at imprinted gene clusters, and a notable correlation between H3K4me3 and paternally expressed noncoding RNAs and genes; loci that lack DNA methylation in sperm. In contrast, maternally expressed noncoding RNAs/genes, and especially paternally methylated regions, lack H3K4me3 and (for the selected genes tested) contain moderate H3K9me3. Consistent with these observations, recent structural and *in vitro* data show that H3K4 methylation deters DNA methylation by DNMT3A2 and DNMT3L in mice[34]. However, experiments in model organisms are needed to address whether the modification patterns we report influence imprinting patterns *in vivo*. Taken together, we reveal chromatin features in sperm that may contribute to totipotency, developmental decisions and imprinting patterns, and open new questions about whether ageing and lifestyle affects chromatin in a manner that impacts fertility or embryo development.

METHODS SUMMARY

Biological samples. Sperm samples were obtained from four men of known fertility attending the University of Utah Andrology laboratory, consented for research. Samples were collected after 2–5 days abstinence and subjected to a density gradient (to purify viable, motile, mature sperm) and treated with somatic cell lysis buffer (0.1% SDS, 0.5% Triton X-100 in DEPC H_2O) for 20 min on ice to eliminate white blood cell contamination. Samples were centrifuged at 10,000*g* for 3 min, and the sperm pellet was resuspended in PBS and used immediately for chromatin preparation. Clontech human fibroblast cells (Lonza cc-2251) were cultured (37 °C and 5% CO_2) in DMEM containing 10% FBS and supplemented with penicillin and streptomyocin.

Chromatin immunoprecipitation. Standard ChIP methods were used[35], but we omitted crosslinking and used the following salt concentrations in the numbered buffers[35]: (1) 150 mM NaCl, (2) 250 mM NaCl, (3) 200 mM LiCl, and (4) 150 mM NaCl (the PBS wash). Antibodies used were: anti-H3K27me3 (Upstate 07-449), H3K4me3 (Abcam 8580), H3K4me2 (Abcam 32356), TH2B (Upstate 07-680), H2A.Z (Abcam 4174) and H3K9me3 (Abcam 8898). For each, 4 µl of antibody was coupled to 100 µl of Dynabeads (Invitrogen). After ChIP, samples for sequencing were not amplified, whereas for arrays the DNA was amplified (WGA, Sigma) before hybridization.

Methylation profiling using MeDIP. MeDIP procedures for sperm and primary human fibroblasts (Clonetech) were performed as described previously[30].

Sequencing. Sequencing used the Illumina GAII (Illumina Inc.) with standard protocols. Read numbers are final mapped microsatellite filtered reads (26–36 bases). Nucleosomes from D1: 19,658,110, D2–D4: 18,842,467, D1–4: 25,933,196 with equal contribution from each donor (random sub-sampling). Input, human sperm DNA: 17,991,622, H3K4me3: 13,337,105, H3K27me3:10,344,413, and H2A.Z: 5,449,000. All genomics data sets have been deposited in the Gene Expression Omnibus (GEO) under the SuperSeries GSE15594.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 3 March; accepted 27 May 2009.
Published online 14 June 2009.

1. Ward, W. S. & Coffey, D. S. DNA packaging and organization in mammalian spermatozoa: comparison with somatic cells. *Biol. Reprod.* **44,** 569–574 (1991).
2. Wykes, S. M. & Krawetz, S. A. The structural organization of sperm chromatin. *J. Biol. Chem.* **278,** 29471–29477 (2003).
3. Balhorn, R., Brewer, L. & Corzett, M. DNA condensation by protamine and arginine-rich peptides: analysis of toroid stability using single DNA molecules. *Mol. Reprod. Dev.* **56,** 230–234 (2000).
4. Gatewood, J. M., Cook, G. R., Balhorn, R., Schmid, C. W. & Bradbury, E. M. Isolation of four core histones from human sperm chromatin representing a minor subset of somatic histones. *J. Biol. Chem.* **265,** 20662–20666 (1990).
5. Kimmins, S. & Sassone-Corsi, P. Chromatin remodelling and epigenetic features of germ cells. *Nature* **434,** 583–589 (2005).
6. Nix, D. A., Courdy, S. J. & Boucher, K. M. Empirical methods for controlling false positives and estimating confidence in ChIP-Seq peaks. *BMC Bioinformatics* **9,** 523 (2008).
7. Reik, W., Santos, F. & Dean, W. Mammalian epigenomics: reprogramming the genome for development and therapy. *Theriogenology* **59,** 21–32 (2003).
8. Santos, F., Hendrich, B., Reik, W. & Dean, W. Dynamic reprogramming of DNA methylation in the early mouse embryo. *Dev. Biol.* **241,** 172–182 (2002).
9. Rangasamy, D., Berven, L., Ridgway, P. & Tremethick, D. J. Pericentric heterochromatin becomes enriched with H2A.Z during early mammalian development. *EMBO J.* **22,** 1599–1607 (2003).
10. Bernstein, B. E. *et al.* A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell* **125,** 315–326 (2006).
11. Bernstein, B. E. *et al.* Methylation of histone H3 Lys 4 in coding regions of active genes. *Proc. Natl Acad. Sci. USA* **99,** 8695–8700 (2002).
12. Boyer, L. A. *et al.* Core transcriptional regulatory circuitry in human embryonic stem cells. *Cell* **122,** 947–956 (2005).
13. Muller, J. & Kassis, J. A. Polycomb response elements and targeting of Polycomb group proteins in *Drosophila*. *Curr. Opin. Genet. Dev.* **16,** 476–484 (2006).
14. Schwartz, Y. B. *et al.* Genome-wide analysis of Polycomb targets in *Drosophila melanogaster*. *Nature Genet.* **38,** 700–705 (2006).
15. Tanay, A., O'Donnell, A. H., Damelin, M. & Bestor, T. H. Hyperconserved CpG domains underlie Polycomb-binding sites. *Proc. Natl Acad. Sci. USA* **104,** 5521–5526 (2007).

16. Cao, R. *et al.* Role of histone H3 lysine 27 methylation in Polycomb-group silencing. *Science* **298,** 1039–1043 (2002).
17. Lee, T. I. *et al.* Control of developmental regulators by Polycomb in human embryonic stem cells. *Cell* **125,** 301–313 (2006).
18. Takahashi, K. *et al.* Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell* **131,** 861–872 (2007).
19. Kopp, J. L., Ormsbee, B. D., Desler, M. & Rizzino, A. Small increases in the level of Sox2 trigger the differentiation of mouse embryonic stem cells. *Stem Cells* **26,** 903–911 (2008).
20. Wernig, M. *et al. In vitro* reprogramming of fibroblasts into a pluripotent ES-cell-like state. *Nature* **448,** 318–324 (2007).
21. Mohn, F. *et al.* Lineage-specific polycomb targets and de novo DNA methylation define restriction and potential of neuronal progenitors. *Mol. Cell* **30,** 755–766 (2008).
22. Illingworth, R. *et al.* A novel CpG island set identifies tissue-specific methylation at developmental gene loci. *PLoS Biol.* **6,** e22 (2008).
23. Down, T. A. *et al.* A Bayesian deconvolution strategy for immunoprecipitation-based DNA methylome analysis. *Nature Biotechnol.* **26,** 779–785 (2008).
24. Farthing, C. R. *et al.* Global mapping of DNA methylation in mouse promoters reveals epigenetic reprogramming of pluripotency genes. *PLoS Genet.* **4,** e1000116 (2008).
25. Landgraf, P. *et al.* A mammalian microRNA expression atlas based on small RNA library sequencing. *Cell* **129,** 1401–1414 (2007).
26. Glazov, E. A., McWilliam, S., Barris, W. C. & Dalrymple, B. P. Origin, evolution, and biological role of miRNA cluster in DLK-DIO3 genomic region in placental mammals. *Mol. Biol. Evol.* **25,** 939–948 (2008).
27. Takada, S. *et al. Delta-like* and *Gtl2* are reciprocally expressed, differentially methylated linked imprinted genes on mouse chromosome 12. *Curr. Biol.* **10,** 1135–1138 (2000).
28. da Rocha, S. T., Edwards, C. A., Ito, M., Ogata, T. & Ferguson-Smith, A. C. Genomic imprinting at the mammalian *Dlk1-Dio3* domain. *Trends Genet.* **24,** 306–316 (2008).
29. Li, S. S., Liu, Y. H., Tseng, C. N. & Singh, S. Analysis of gene expression in single human oocytes and preimplantation embryos. *Biochem. Biophys. Res. Commun.* **340,** 48–53 (2006).
30. Weber, M. *et al.* Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. *Nature Genet.* **39,** 457–466 (2007).
31. Fouse, S. D. *et al.* Promoter CpG methylation contributes to ES cell gene regulation in parallel with Oct4/Nanog, PcG complex, and histone H3 K4/K27 trimethylation. *Cell Stem Cell* **2,** 160–169 (2008).
32. Vire, E. *et al.* The Polycomb group protein EZH2 directly controls DNA methylation. *Nature* **439,** 871–874 (2006).
33. Schlesinger, Y. *et al.* Polycomb-mediated methylation on Lys27 of histone H3 pre-marks genes for de novo methylation in cancer. *Nature Genet.* **39,** 232–236 (2007).
34. Ooi, S. K. *et al.* DNMT3L connects unmethylated lysine 4 of histone H3 to *de novo* methylation of DNA. *Nature* **448,** 714–717 (2007).
35. Gordon, M. *et al.* Genome-wide dynamics of SAPHIRE, an essential complex for gene activation and chromatin boundaries. *Mol. Cell. Biol.* **27,** 4058–4069 (2007).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank B. Dalley for microarray and sequencing expertise, B. Schackmann for oligonucleotides, K. Boucher for statistical analysis, J. Wittmeyer for yeast nucleosomes and helpful comments, and T. Parnell for helpful comments. Financial support was from the Department of Urology (genomics and support of S.S.H.), the Howard Hughes Medical Institute (HHMI) (genomics, biologicals and support of J.P. and H.Z.), CA24014 and CA16056 for core facilities, and the Huntsman Cancer Institute (bioinformatics and support of D.A.N.). B.R.C. is an investigator with the HHMI.

Author Contributions B.R.C., D.T.C. and S.S.H. were involved in the overall design. D.T.C. and S.S.H. were responsible for acquisition of samples, clinical logistics, patient consenting and Institutional Review Board documents. B.R.C., S.S.H., D.A.N. and H.Z. designed detailed molecular and genomics approaches. D.A.N. carried out data processing and array analysis. S.S.H. and D.A.N. performed sequencing analysis. S.S.H. carried out experiments and produced the figures. J.P. carried out immunoblotting and bisulphite sequencing. B.R.C. wrote the manuscript.

Author Information The raw unfiltered reads (fastq format) are deposited at the Gene Expression Omnibus (GEO) under the SuperSeries GSE15594, which encompasses the Subseries entries GSE15690 for ChIP-seq data and GSE15701 for ChIP-chip data. Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to D.T.C. (douglas.carrell@hsc.utah.edu) or B.R.C. (brad.cairns@hci.utah.edu).

doi:10.1038/nature08162 nature

METHODS

Partitioning of histone- and protamine-associated DNA. Chromatin was prepared from 40 million sperm as described previously[36] in the absence of crosslinking reagent, treated with sequential and increasing MNase (10–160 U), and centrifuged to sediment protamine-associated DNA, releasing mononucleosomes. The pooled mononucleosomes were used for ChIP, or the DNA was extracted and gel purified (~140–155 bp) for sequencing and array analysis.

ChIP and preparation for genomics methods. All ChIPs for sequencing were performed using the same pool of mononucleosomes from pooled donors. For arrays, a single pool was used from D1. ChIP methods were as described previously[35] but were performed without a crosslinking agent and slight modifications to the salt levels (250 mM NaCl, 200 mM LiCl), and the TE wash was replaced with a 150 mM PBS wash. ChIP methods used anti-H3K27me3 (Upstate 07-449), H3K4me3 (Abcam 8580), H3K4me2 (Abcam 32356), TH2B (Upstate 07-680), or H2A.Z (Abcam 4174) antibodies. For each, 4 µl of antibody was coupled to 100 µl of Dynabeads (Invitrogen). After the ChIP procedure, the DNA was amplified (WGA, Sigma) before hybridization to arrays, whereas samples used for Solexa were not amplified. For sequencing, DNA lengths corresponding to mononucleosomes with adapters (220–280 bp) were gel purified after the addition of the Illumina adaptors. This size selection was also performed for the nucleosomal DNA from pooled donors not subjected to ChIP.

Methylation profiling using MeDIP. This procedure was described previously[30]. In brief, sonicated sperm DNA was obtained from two different donors and sonicated fibroblast DNA was obtained from Clontech primary human fibroblasts (Lonza CC-2251) (4 µg, 300–1,000-bp fragments). Immunoprecipated DNA was washed, subjected to whole genome amplification (Sigma Aldrich). Amplified DNA (6 µg) was labelled with Cy5, and input DNA (6 µg) was labelled with Cy3 (Bio labs) by standard methods. Samples were hybridized to Agilent expanded promoter arrays, treated according to standard Agilent conditions, and scanned in an Agilent scanner.

Computational analytical methods. The software used in this analysis are open source and available from the TIMAT2 (http://timat2.sourceforge.net) and USeq (http://useq.sourceforge.net) project websites. Human annotation and genomic sequence (May 2004, NCBI Build 35, HG17 and March 2006, NCBI Build 36.1, HG18) were obtained from the UCSC Genome Bioinformatic website.

Low-level ChIP-chip analysis. Processing of the Agilent microarray promoter data was performed in three basic steps: data normalization, sliding window summaries, and enriched region identification. For each data set, the median unadjusted signal intensities from the Cy3 and Cy5 channels were extracted. Probes were then mapped to the HG17 or HG18 builds. Biological replicas were quantile normalized and median scaled to 100 (ref. 37). This normalization was applied to the treatment (ChIP samples) and control (whole genomic input DNA for the MeDIP and protamine data sets or DNA derived from mononucleosomes) replicas separately (see later for replica-averaged R^2). Probe level 'Oligo' summaries were calculated by taking the $\log_2$ ratio (mean treatment replicas/mean control replicas). 'Window' level summaries were generated by identifying windows of a particular size (100 bp for data sets derived from mononucleosomes, 675 bp for MeDIP and protamine data sets) containing a minimum number of oligonucleotide start positions (one for the data sets derived from mononucleosomes, three for the MeDIP and protamine data sets), and calculating an all pair (treatment versus control) relative difference pseudo median. This window summary score was assigned to the centre position of the window 'Pse' or represented as heat map 'PseHM' data. Extended regions of high-scoring windows, called 'intervals', were identified by merging windows that exceed a set threshold and are located within 250 bp of one another. Intervals were then ranked by their best window score. Relative difference pseudo median scores were converted to $\log_2$ ratio values.

The average R^2 values for microarray data were as follows: 0.85 for the three D1 MNase replicas; 0.89 for the three Protamine replicas; 0.96 for the two H3C replicas; 0.94 for the two H3K4me2 replicas; 0.93 for the two TH2B replicas; 0.96 for the three H3K4me3 replicas; and 0.93 for the two H3K27me3 replicas. The average MeDIP R^2 values for the three replicas of each donor were as follows: D2 average $R^2 = 0.97$ and D4 = 0.89, and the correlation between D2 versus D4 was 0.87. The average R^2 for the two primary human fibroblast MeDIP replicas was 0.86.

Low-level Chip-seq analysis. The DNA samples derived from mononucleosomes, and the sonicated control input genomic DNA were prepared for sequencing using Illumina's ChIP-seq kit. The 26-bp and 36-bp reads were generated using Illumina's Genome Analyser II and their standard software pipeline. Reads were mapped to the March 2006 NCBI Build 36.1 human genome using the pipeline's eland_extended aligner.

The USeq package[6] was used to identify regions of histone enrichment relative to input control. This entailed selecting reads that mapped with an alignment score ≥13 ($-10\log_{10}(0.05)$), shifting their centre position 73 bp 3′ to accommodate the 146-bp mononucleosome fragment length, and using a sliding window of 300 bp to score each region in the genome for significant histone enrichment. Significance was determined by calculating a binomial P value for each 300-bp window and controlled for multiple testing by applying Storey's q value FDR estimation[38,39].

Read numbers. Note the sperm genome has only 4% of the genome in nucleosomes. For nucleosome enrichment D1 had 19,658,110 reads, and the pool of three additional donors had 18,842,467 reads. The raw correlation for D1 versus the donor pool was $r = 0.7$. For all the analysis containing pool donors (D1, and a pooled sample of three additional individuals D2, D3 and D4) we used 25,933,196 mapped filtered reads with equal contribution from each donor (random subsampling). A total of 17,991,622 reads were generated from control input human sperm DNA, 3,337,105 reads from the H3K4me3 sample, 10,344,413 reads for H3K27me3, and 5,449,000 reads for H2Az. The raw unfiltered reads (fastq format) are deposited at GEO under the superseries GSE15594, which encompasses the Subseries entries GSE15690 for ChIP-seq and GSE15701 for ChIP-chip data.

To assess histone enrichment consistency, the QCSeqs application in the USeq package[6] was used to correlate the read counts between the D1 and pooled sample by calculating a Pearson correlation on the basis of the number of mapped reads falling within 500-bp windowed regions stepped every 250 bp across all chromosomes. Only windows with five or more reads in either of the samples were included in the correlation.

To create lists of candidate histone enriched regions, q-value thresholds of 20 (0.01) and 30 (0.001) ($-10\log_{10}(q\text{ value})$) were selected. Overlapping windows that pass a given threshold were merged and scores from the best window assigned to the enriched region. The normalized window score was then used to rank and sort the regions.

A modification was made to score gene promoters and miRNAs for significant histone enrichment. The first step was to define regions for scoring. For gene promoters, the start of the first exon was used to define its hypothetical promoter by selecting a region 9 kb upstream and 2 kb downstream. For miRNAs, the centre position of each was expanded ±300 bp. These defined regions were scored for significant enrichment using the window statistics above.

High-level ChIP-chip and ChIP-seq analysis. Intersect regions. To identify regions of significant intersection between enriched region lists from various data sets, the USeq IntersectRegions application was used. This application counts the number of intersections between two lists of genomic coordinates that occur within a minimum 'max gap' distance. To estimate confidence in the intersections, a thousand 'random' data sets are generated that were matched to the chromosome and size of the original regions, and randomly picked from the interrogated regions on the array or sequenced regions in the genome. These randomized data sets were used to calculate a P value for the intersection and fold enrichment (fraction real intersection/fraction average random data set intersection) over random. Initial pilots that imposed a fraction GC match when picking random regions showed little difference with non-GC-matched random data sets and were thus subsequently dropped.

Find neighbouring genes (FNG). Genes associating with histones or histone modifications were determined using the FNG application in the USeq package. The gene lists were uploaded in GoMiner (http://discover.nci.nih.gov/gominer/htgm.jsp) to identify over represented Gene Ontology terms.

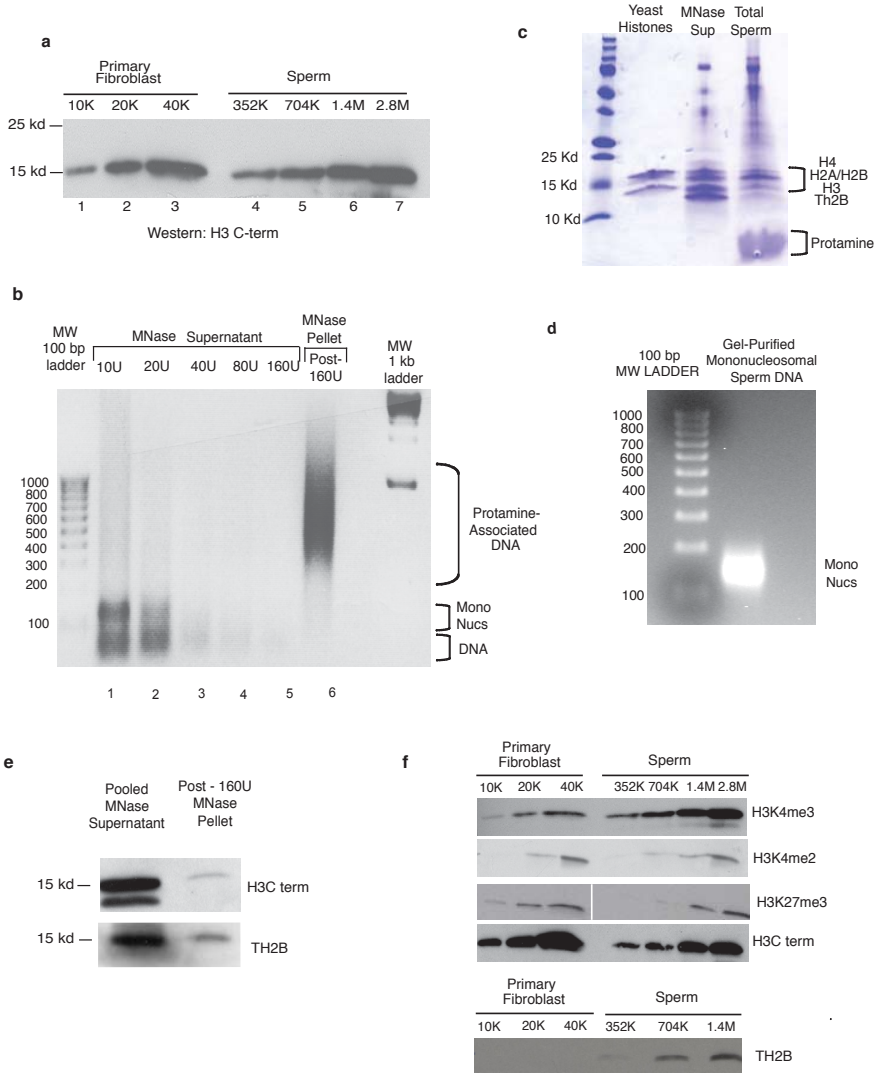
Intersect lists. To determine whether the 4- and 8-cell transcripts identified in early human embryo correlated with any of our histone modifications we used The IntersectLists USeq application which uses random permutation to calculate the significance of intersection between two lists of genes.

Aggregate plots. The USeq AggregatePlots application was used to compare the degree of enrichment and distribution of histone reads surrounding the TSS of developmental and non developmental genes. The gene classes were derived on the basis of Gene Ontology term categories.

36. Zalenskaya, I. A., Bradbury, E. M. & Zalensky, A. O. Chromatin structure of telomere domain in human sperm. *Biochem. Biophys. Res. Commun.* **279,** 213–218 (2000).
37. Bolstad, B. M., Irizarry, R. A., Astrand, M. & Speed, T. P. A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformatics* **19,** 185–193 (2003).
38. Dudoit, S., Gilbert, H. N. & van der Laan, M. J. Resampling-based empirical Bayes multiple testing procedures for controlling generalized tail probability and expected value error rates: focus on the false discovery rate and simulation study. *Biom. J.* **50,** 716–744 (2008).
39. Storey, J. D. & Tibshirani, R. Statistical significance for genomewide studies. *Proc. Natl Acad. Sci. USA* **100,** 9440–9445 (2003).

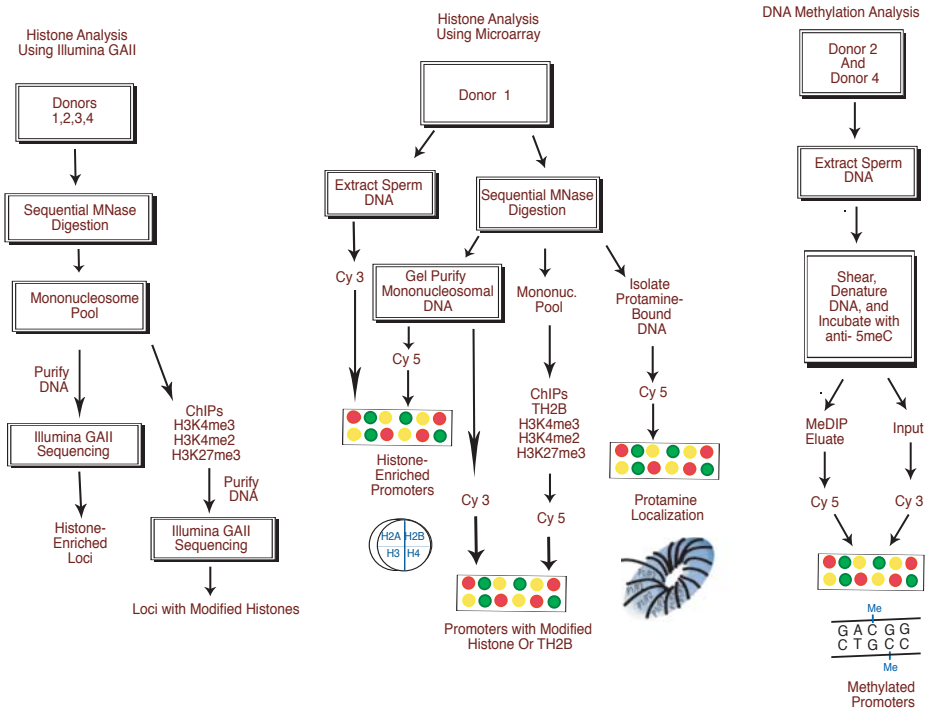
doi: 10.1038/nature08162 nature

SUPPLEMENTARY INFORMATION



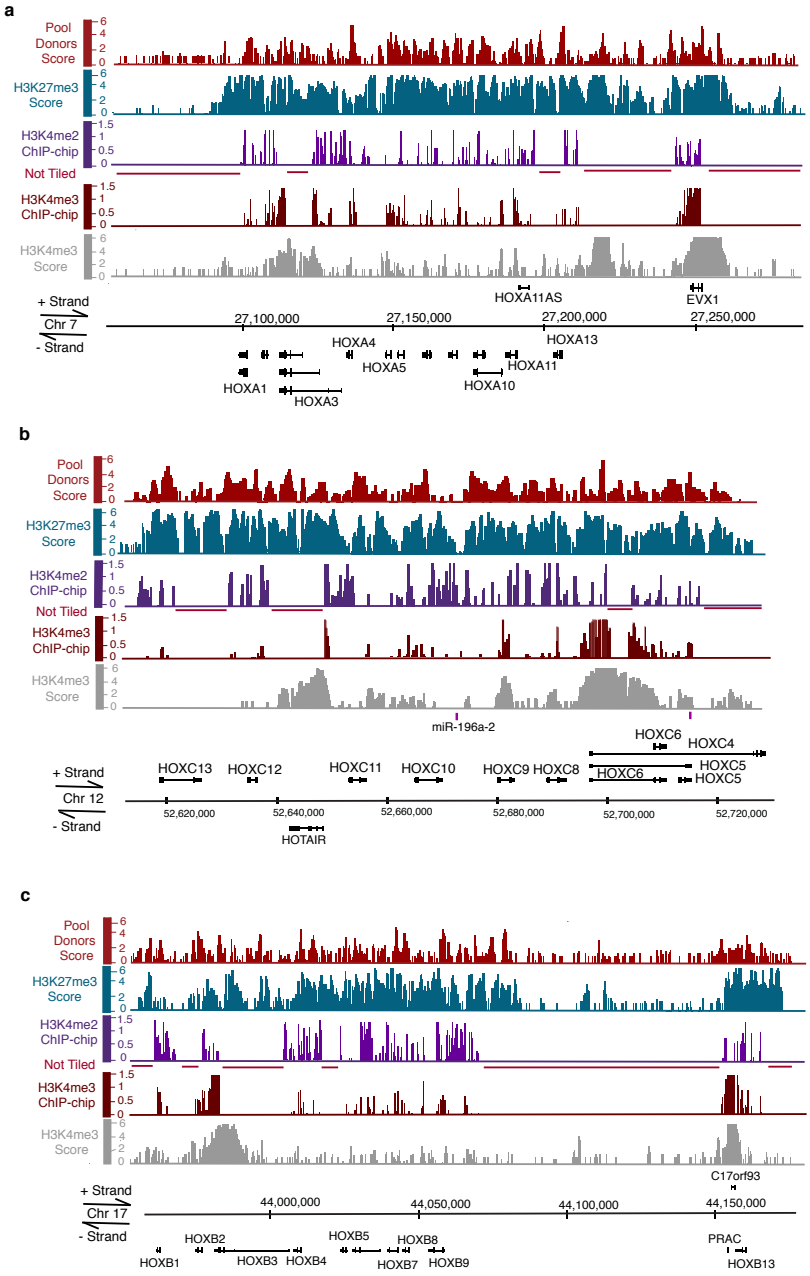
Supplemental Fig. 1: Composition of human sperm chromatin. **a**, Quantifying histone content of primary fibroblast or human sperm cells by immunoblot analysis with the H3C terminus antibody. **b**, Sequential digestion of sperm chromatin with increasing concentrations of micrococal nuclease (MNase) releases mononucleosomes (lanes 1 and 2), whereas protamine-packaged chromatin resists MNase (lane 6). **c**, Characterizing the mononucleosome fraction released into the MNase supernatant pool from panel **b**. **d**, Gel-purified mononucleosomal DNA used for array hybridization or sequencing. **e**, Quantification of the amount of histone released by MNase treatment. Supernatants were pooled. Here, cell equivalents were loaded in each lane; 4% of the total supernatant or protamine pellet. The gel was subjected to immunoblotting and quantified on a Typhoon (Amersham). **f**, Western analysis, involving titrations for bulk levels of H3K4me3, H3K4me2, H3K27me3 in primary fibroblast cells and mature sperm cells. Quanitation by Typhoon (Amersham) reveals that sperm bear ~4% of the histone H3 present in a primary fibroblast.

doi: 10.1038/nature08162



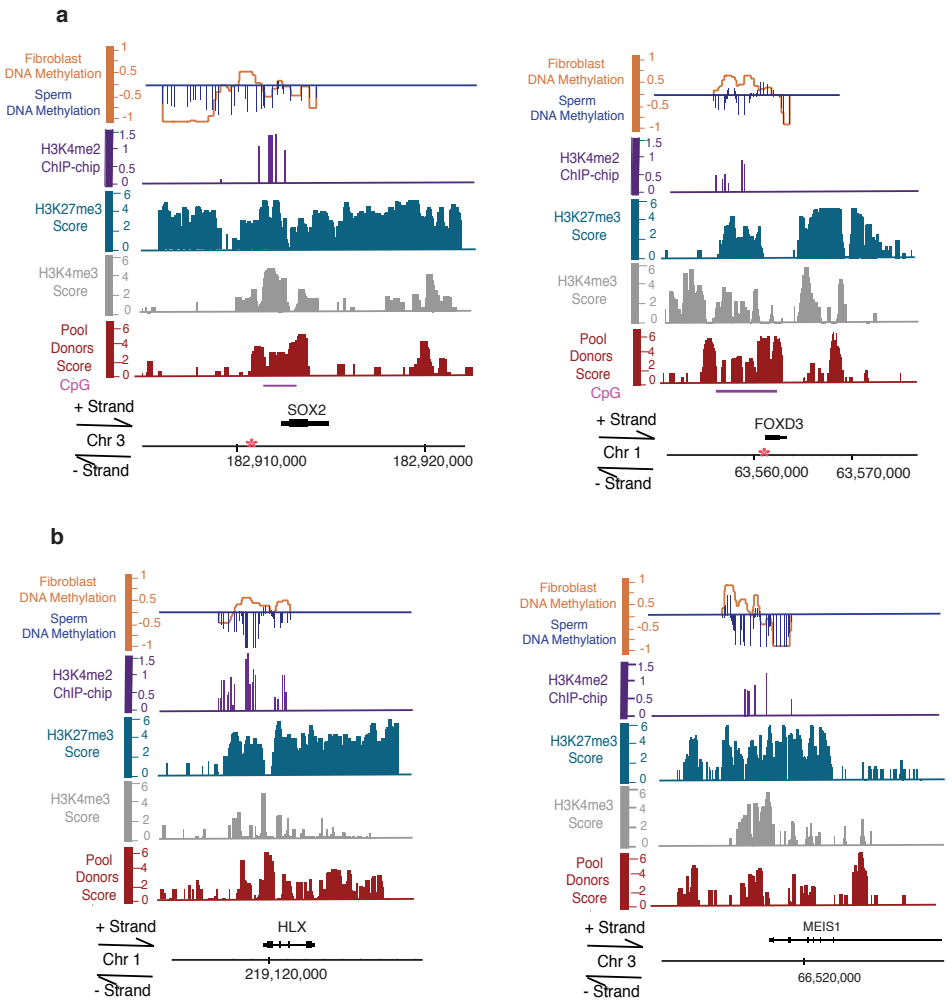
Supplemental Fig. 2: Schematic representation of experimental procedures. Two fertile donors were used for methylation studies, one donor (D1) was used for all histone modifications studied on the arrays. A pool of fertile donors were utilized for mononucleosome localization and characterization and to extend the analysis genome-wide using Illumina GAII.

doi: 10.1038/nature08162



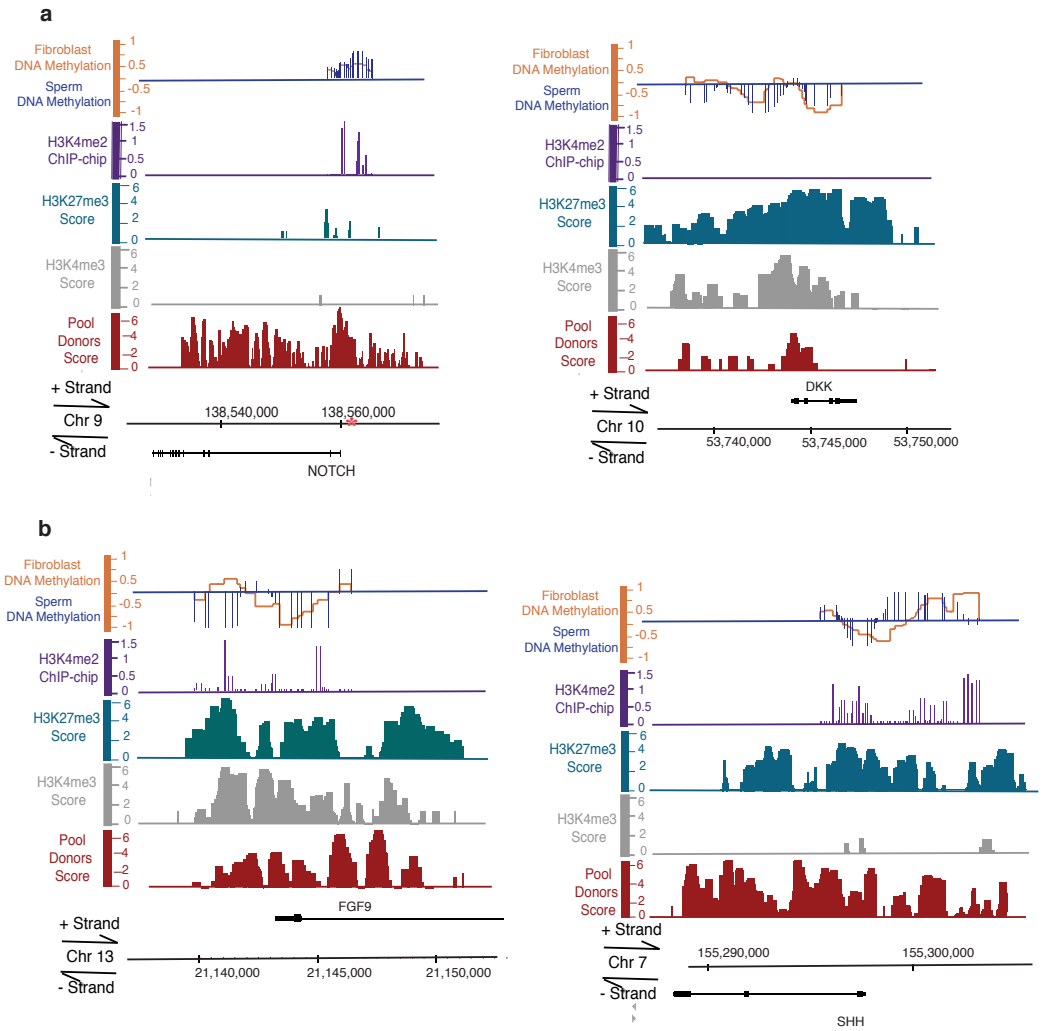
Supplemental Fig. 3: Chromatin attributes of the *HOXA, HOXB,* and *HOXC* loci. Histone enrichment (red bars), or histone modifications (H3K4me3 array results (ruby), H3K4me3 sequencing normalized difference scores (grey), H3K27me3 sequencing normalized difference scores (teal blue) or H3K4me2 (violet)). The y-axis is the signal intensity (log2 for array data, or normalized difference score for Illumina GAII sequencing) and the x-axis is the annotated physical map (HG18). **a**, The *HOXA* locus. **b**, The *HOXC* locus **c,** The *HOXB* locus.

doi: 10.1038/nature08162



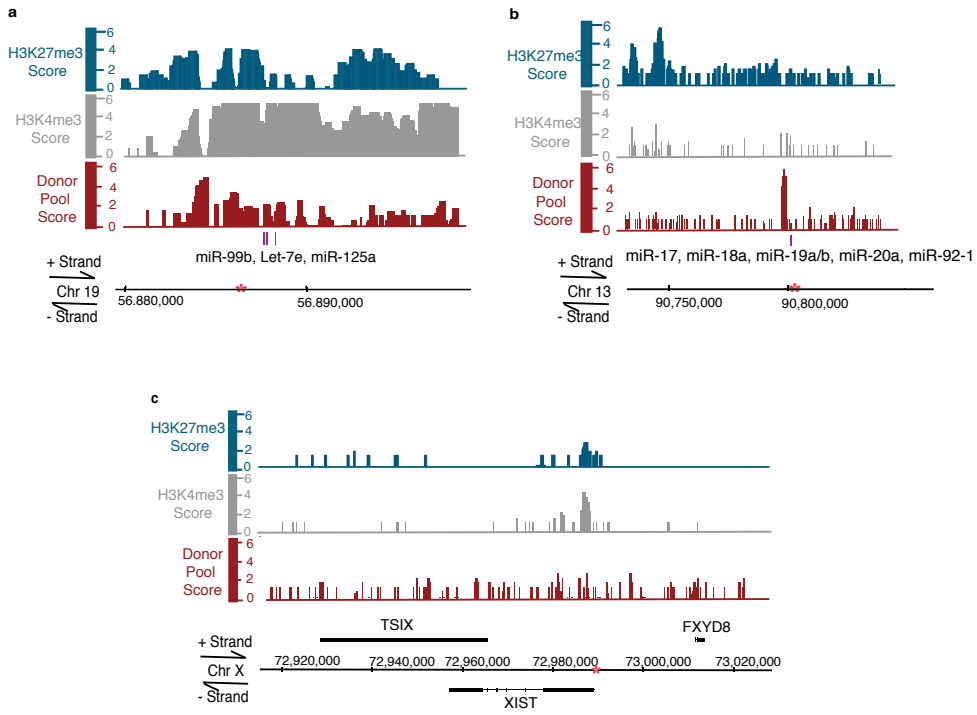
Supplemental Fig. 4: Certain self-renewal genes as well as genes required for embryonic development generally lack DNA methylation and are bivalent. **a**, SOX2 and FOXD3 are member of the pluripotency netwok. *SOX2* is demethylated and characterized by the presence of H3K4me3 and H3K27me3, whereas *FOXD3* is hypermethylated near their transcription start sites. *OCT4* and *NANOG* are also hypermethylated (Supplementary Fig. 10c). **b**, Genes involved in embryonic development are typically DNA hypomethylated, and have high levels of H3K4me2/3 and H3K27me3 around their start sites The red asterisks indicate the region amplified for bisulfite sequencing in Supplementary Fig. 10. The y-axis is the signal intensity (log2 for ChIP-chip arrays, or normalized difference for Illumina GAII sequencing score) and the x-axis is the annotated physical map (HG18).

doi: 10.1038/nature08162



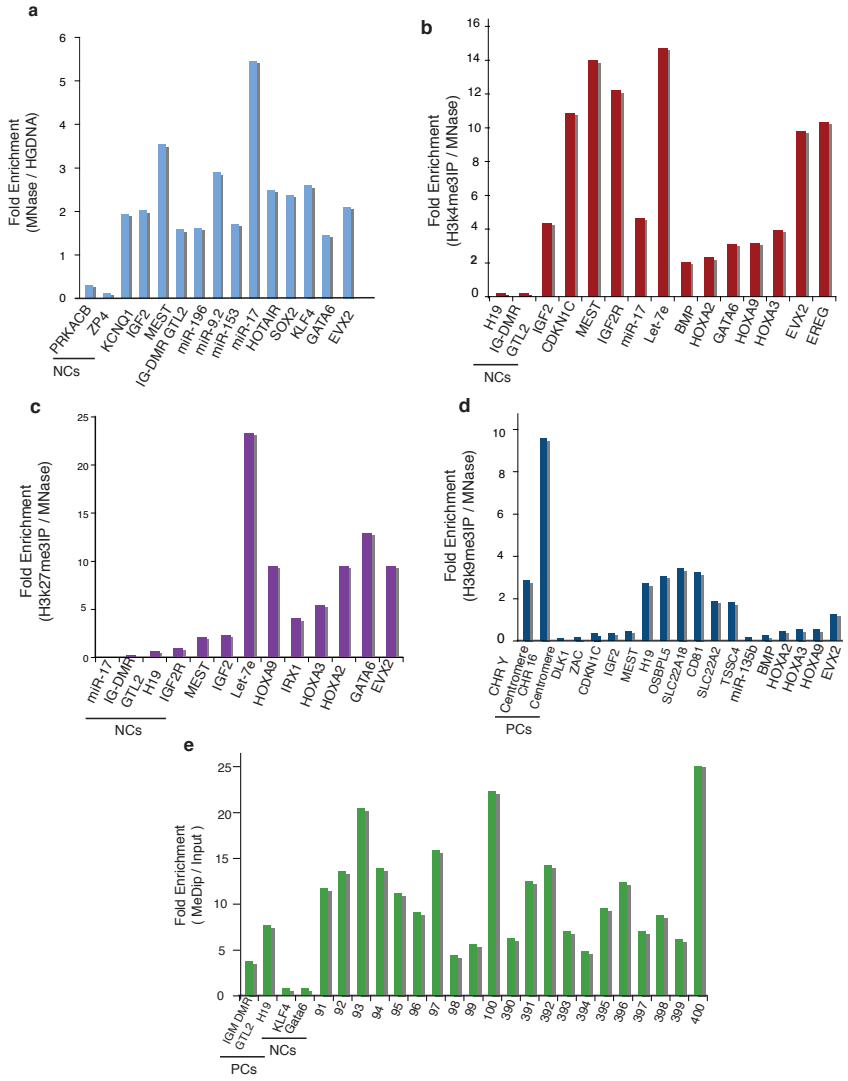
Supplemental Fig 5: Developmental and signaling factors are deficient in DNA methylation, although Notch pathway members are hypermethylated. **a**, Notch signaling pathway members, *DKK1* (hypomethylated) and *NOTCH1* (hypermethylated). **b**, FGF signaling pathway members, *FGF9* (hypomethylated) and regulator *SHH* (hypermehylated). The red asterisks indicates the region amplified for bisulfite sequencing. The y-axis is the signal intensity (log2 for ChIP-chip arrays, or normalized difference for Illumina GAII score) and the x-axis is the annotated physical map (HG18).

doi: 10.1038/nature08162



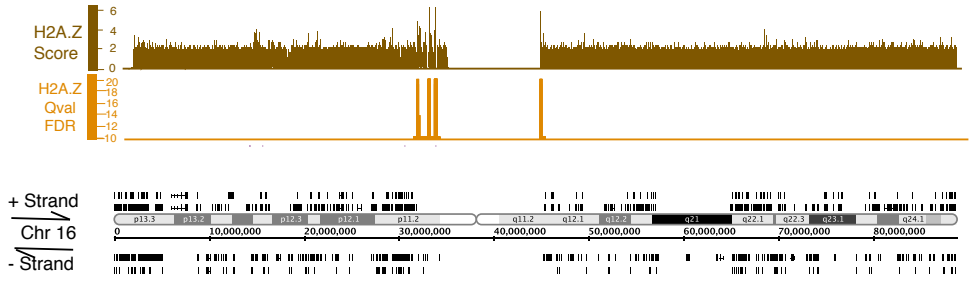
Supplemental Fig6: Histone retention at miRNAs and non-coding RNAs. **a**, A miRNA cluster with high H3K4me3 and H3K27me3. **b,** A miRNA cluster region with high levels of histone in the promoter region of the pri-miRNA, but lacking H3K4me3 and H3K27me3. **c,** The non-coding RNA *XIST* is enriched for H3K4me3 and H3K27me3 at the TSS. The read counts for the *X*-chromosome are half of those on autosomes due to the presence of either *X* or *Y* in sperm. The y-axis is the normalized difference score for sequencing. Asterisks (*) note the locations tested by bisulphite sequencing in Supplementary Fig. 10.

doi: 10.1038/nature08162



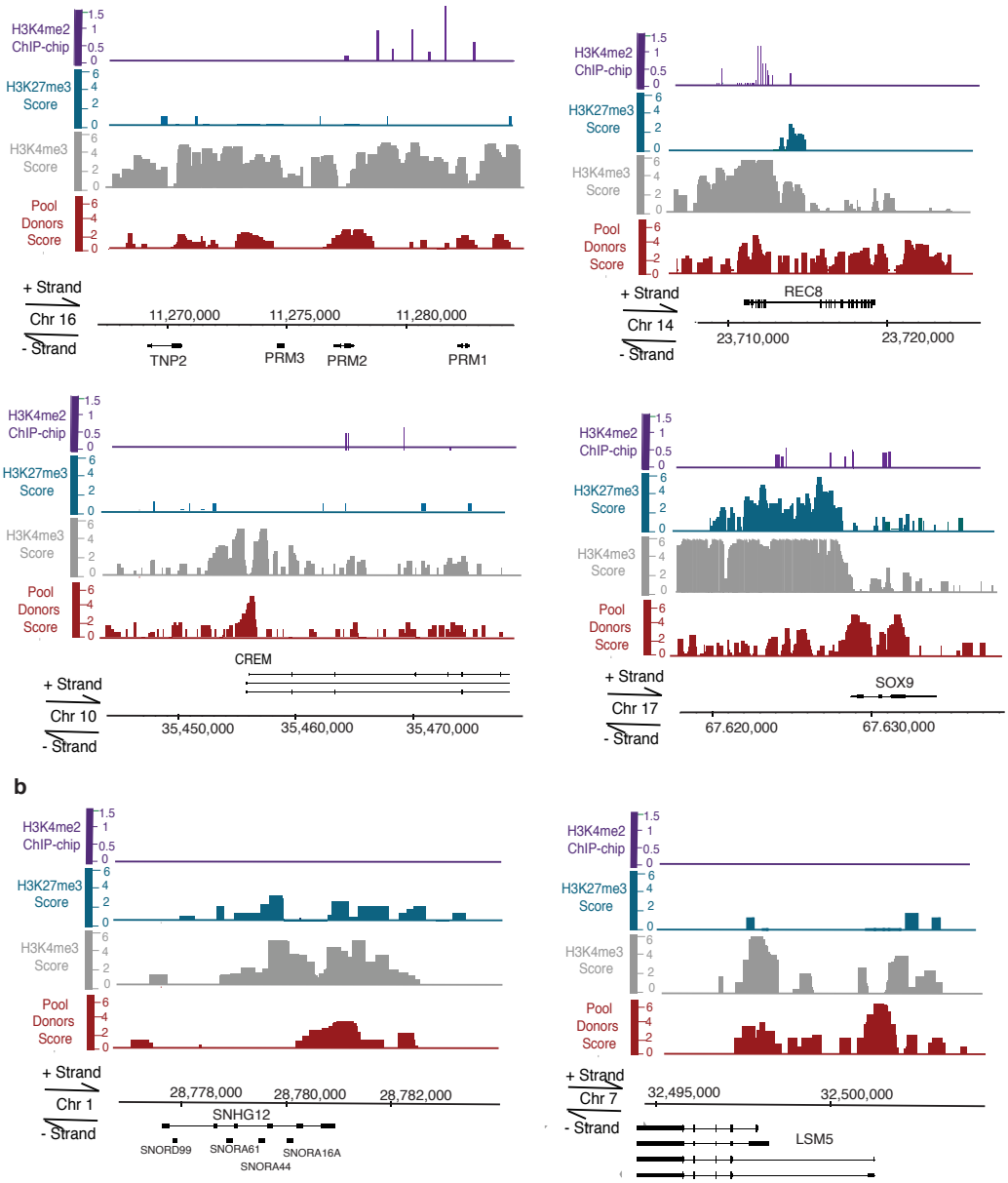
Supplemental Fig 7: qPCR testing of epigenetically modified loci enriched by Illumina GAII sequencing and/or array analysis. **a,** qPCR testing of histone occupancy at both maternally and paternally imprinted genes (*KCNQ1, IGF2, MEST*, and IG-DMR), miRNAs and noncoding RNA (*miR-196*, *miR-9.2*, *miR-153-1*, *miR-17*, and *HOTAIR*), and at a subset of developmental genes. Negative controls (NCs) (*PRKACB* and *ZP4*) are regions that had very low levels of histone by Illumina GAII sequencing and/or array data. Fold enrichment of histone at these promoters was determined by MNase signal divided by the total genomic DNA signal **b**, Fold enrichment of H3K4me3 was determined by normalizing signal from the H3K4me3 IP eluate to the signal from MNase (histone pool). Two maternally-imprinted loci in sperm were used as negative controls. **c**, H3K27me3 and **d**, H3K9me3 enrichment were determined as described above. H3K9me3 positive controls (PCs) were two pericentromeric heterochromatin loci. **e**, qPCR testing of MeDIP data. Enriched loci from MeDIP arrays were binned into the top 100 regions or 400 enriched regions. qPCR of MeDIP eluates were performed for the bottom 10 regions in each of the top 100 and 400 bins. Since all 20 regions enriched for DNA methylation, a cutoff of the top 400 genes (approximately 2-fold) was our stringent cutoff for DNA methylation. qPCR fold enrichment was compared to input (total sheared genomic DNA). Positive controls were two known methylated (imprinted) regions and negative controls were regions that are demethylated in sperm when compared to fibroblast.

doi: 10.1038/nature08162



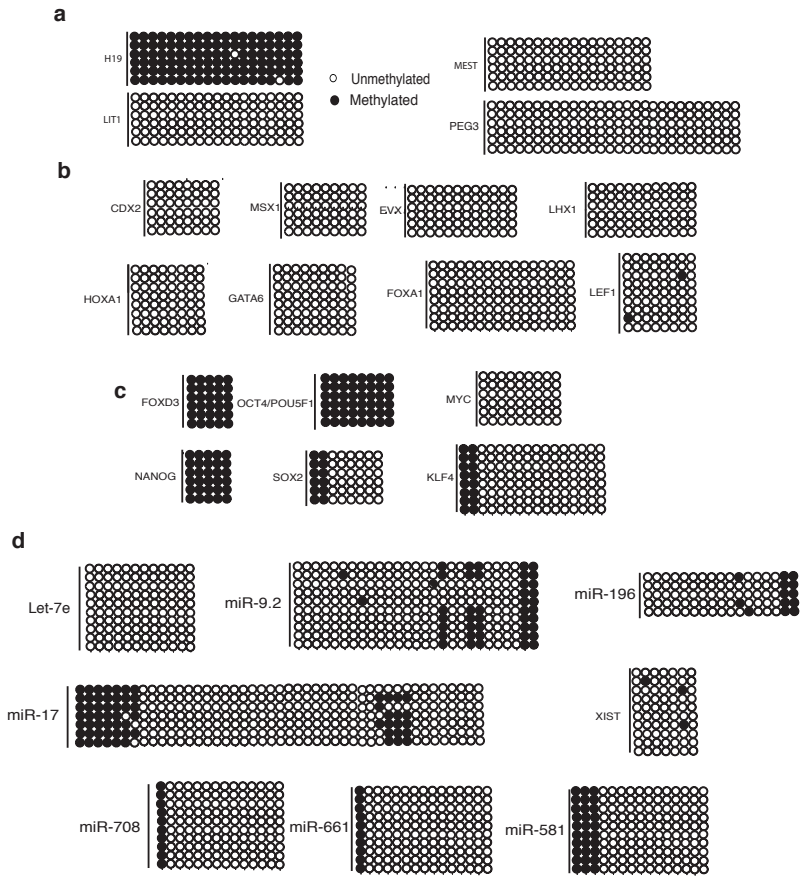
Supplemental Fig 8: H2A.Z localizes to pericentric heterochromatin in the mature human sperm. Brown bars are the normalized difference scores for pooled donor H2A.Z across chromosome 16, and in orange is the FDR. Other chromosomes showed similar peaks flanking the centromere. Pericentric heterochromatin was highly enriched with H2A.Z (FDR <0.05).

doi: 10.1038/nature08162



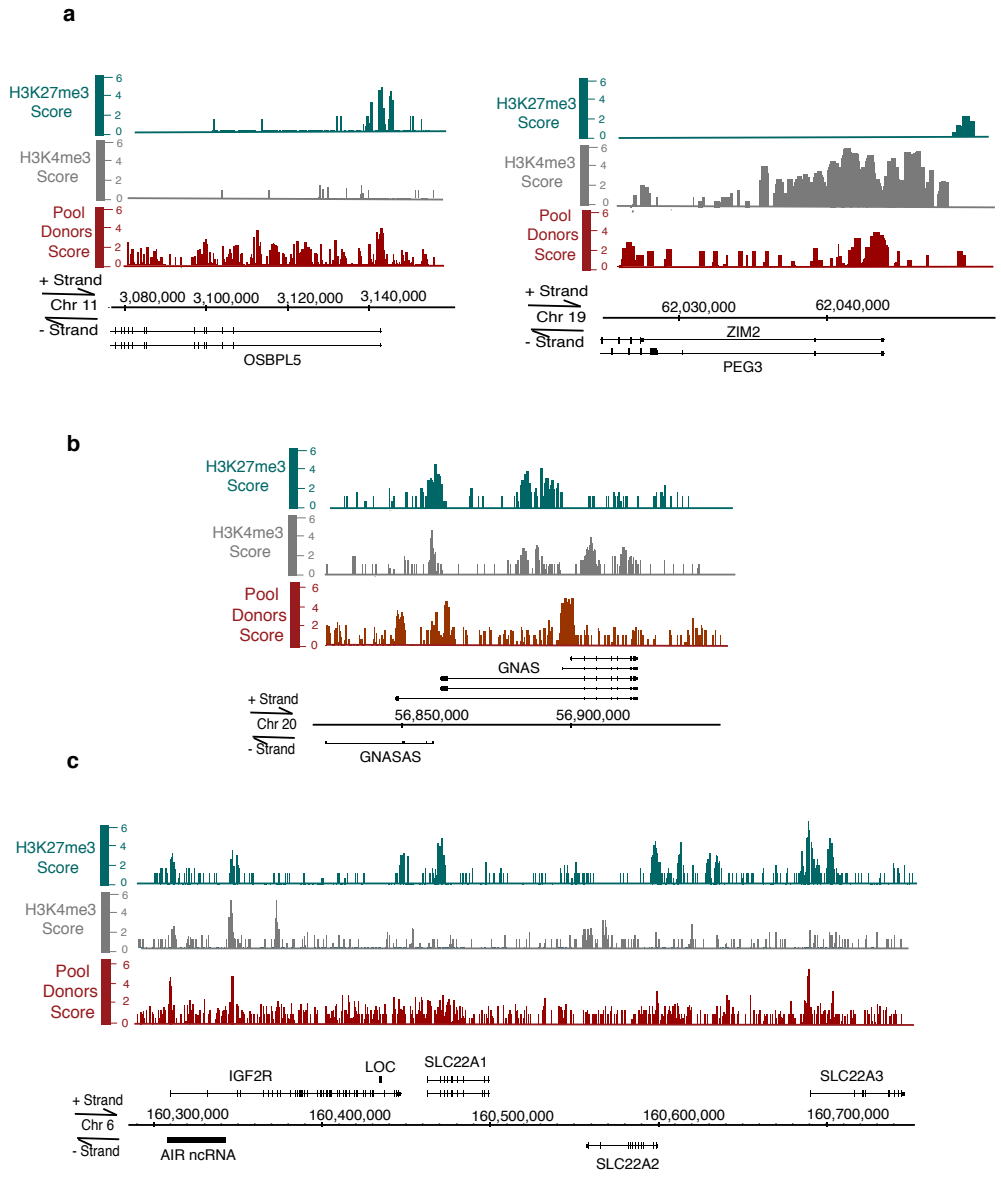
Supplemental Fig 9: Genes required for sperm development generally lack DNA methylation and are bound by H3K4me3. **a,** Four genes expressed at different stages of spermatogenesis remain DNA demethylated and retain H3K4me3 enrichment. **b**, Gene promoters involved in RNA processing, a process utilized intensely during spermiogenesis, are also demethylated and H3K4me3 bound.

doi: 10.1038/nature08162



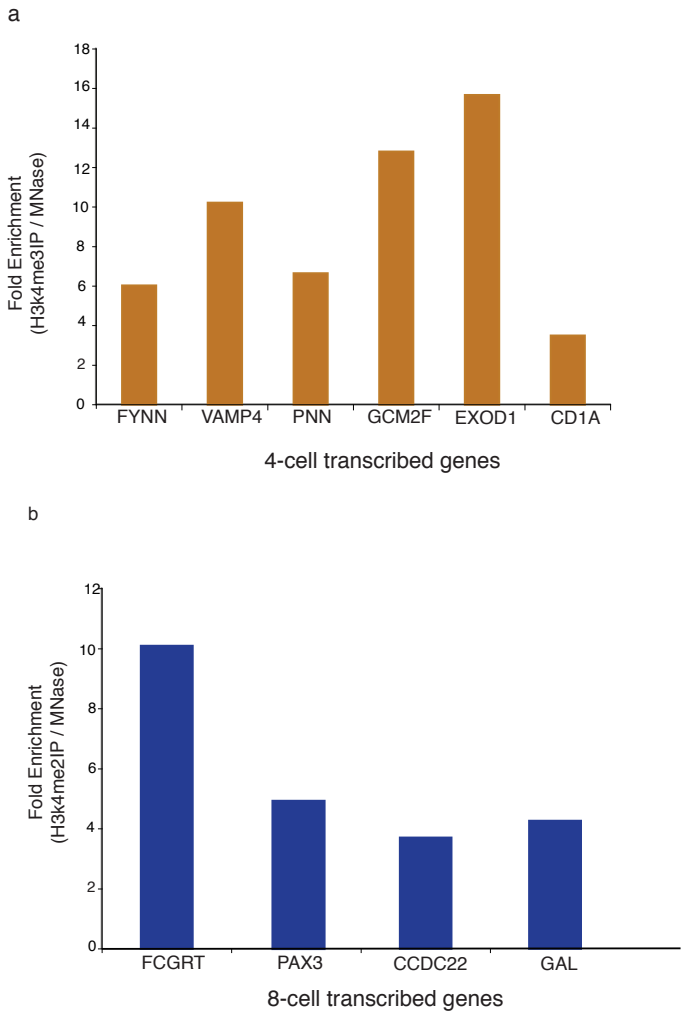
Supplemental Figure 10: DNA hypomethylation at developmental promoters and miRNAs were verified by bisulfite sequencing. **a**, bisulfite sequencing of promoters known to bear (*H19*) or lack (*LIT1, PEG3 and MEST*) paternal methylation in sperm chromatin. CpGs are represented as open dots (if unmethylated) or filled dots (if methylated). **b**, Hypomethylation at developmental transcription factors and **c,** a subset of the pluripotency network promoters. **d**, The TSS of the miRNAs tested were generally hypomethylated.

doi: 10.1038/nature08162



Supplemental Figure 11: H3K4me3 is generally present at paternally-expressed genes and non-coding RNAs. **a**, *OSBPL-5* (a maternally-expressed gene) lacks H3K4me3, whereas *PEG3* (a paternally-expressed gene) has high and broad H3K4me3 . **b,** The TSS of the *AIR* transcript localizes with H3K4me3, whereas promoters silenced by *AIR* (*SLC*s) lack H3K4me3 **c**, Similarly, the *GNASAS* is paternally expressed and has H3K4me3, whereas the remaining promoters are maternally expressed and lack H3K4me3.

doi: 10.1038/nature08162



Supplemental Fig 12: H3K4me2/3 chromatin modifications are correlated with early embryonic genes expression at the 4 and 8 cell stage. **a**, A subset of genes enriched at the 4 cell stage have significant levels of H3K4me3 **b**, whereas genes enriched at the 8-cell stage were associated with high levels of H3K4me2. Fold enrichment for H3K4me3/2 was determined by signal from IP eluate divided the signal derived from the pooled mononucleosomes.

doi: 10.1038/nature08162

Supplemental Table 1: Histone enriched promoters (D1 array)

Go Category	Total Genes	Changed Genes	Enrichment	FDR
Sequence-specific DNA binding	*425*	*32*	*4.601991*	*0*
Transcription factor activity	*755*	*38*	*3.076248*	*0*
Transcription regulator activity	*1090*	*46*	*2.579384*	*0*
Multicellular organismal development	*1620*	*59*	*2.225982*	*0*
DNA binding	*1522*	*54*	*2.168522*	*0*
Regulation of transcription DNA-dependent	*1467*	*51*	*2.124833*	*0*
Transcription DNA-dependent	*1510*	*52*	*2.104801*	*0*
RNA biosynthetic process	*1512*	*52*	*2.102017*	*0*
Regulation of transcription	*1580*	*52*	*2.011551*	*0*
Transcription	*1623*	*53*	*1.995915*	*0*
Developmental process	*1644*	*53*	*1.97042*	*0*
RNA metabolic process	*2265*	*73*	*1.969878*	*0*
Regulation of cellular metabolic process	*1827*	*58*	*1.940324*	*0*
Regulation of metabolic process	*1772*	*55*	*1.897071*	*0*
Regulation of cellular process	*1839*	*57*	*1.894427*	*0*
Regulation of biological process	*2889*	*84*	*1.777119*	*0*
Multicellular organismal process	*3134*	*89*	*1.735704*	*0*
Biological regulation	*2648*	*73*	*1.68496*	*0*
System development	*3396*	*93*	*1.673786*	*0*
Nucleobase nucleoside metabolic process	*1231*	*40*	*1.986034*	*0.0015*
Nucleic acid binding	*2489*	*66*	*1.620704*	*0.001905*
Transcription from RNA polymerase II promoter	*2348*	*63*	*1.639937*	*0.002174*
Anatomical structure development	*460*	*21*	*2.79027*	*0.002273*
Organ development	*1465*	*44*	*1.835692*	*0.0025*
Skeletal development	*869*	*31*	*2.180352*	*0.002692*
Urogenital system development	*174*	*12*	*4.215186*	*0.0028*
Kidney development	*31*	*5*	*9.858096*	*0.004545*
Wnt receptor activity	*29*	*5*	*10.537964*	*0.005*
Growth factor activity	*7*	*3*	*26.194368*	*0.007941*

doi: 10.1038/nature08162

Supplemental Table 2: D1 Histone-enriched loci (Illumina GAII FDR< 0.0001)

Go Category	Total Genes	Changed Genes	Enrichment	FDR
Cell fate commitment	*75*	*60*	*1.59848*	*0*
Sequence-specific DNA binding	*424*	*337*	*1.588112*	*0*
Cellular morphogenesis during differentiation	*125*	*99*	*1.582495*	*0*
Cell projection organization and biogenesis	*169*	*131*	*1.548823*	*0*
Cell part morphogenesis	*169*	*131*	*1.548823*	*0*
Embryonic morphogenesis	*88*	*68*	*1.543986*	*0*
Regionalization	*82*	*63*	*1.535126*	*0*
Neurogenesis	*221*	*168*	*1.518918*	*0*
Wnt receptor signaling pathway	*107*	*80*	*1.493907*	*0*
Regulation of cell differentiation	*119*	*88*	*1.477587*	*0*
Regulation of transcription from RNA polymerase II promoter	*99*	*72*	*1.453164*	*0*
Organ morphogenesis	*304*	*221*	*1.452566*	*0*
Embryonic development	*226*	*164*	*1.449949*	*0*
Regulation of developmental process	*191*	*138*	*1.443653*	*0*
Voltage-gated ion channel activity	*171*	*123*	*1.43723*	*0*
Nervous system development	*604*	*433*	*1.432413*	*0*
Cation channel activity	*228*	*162*	*1.419703*	*0*
Transcription factor activity	*791*	*552*	*1.394376*	*0*
Muscle development	*136*	*94*	*1.38104*	*0*
Central nervous system development	*190*	*129*	*1.356605*	*0*
Skeletal development	*193*	*130*	*1.34587*	*0*
Anatomical structure morphogenesis	*855*	*575*	*1.343751*	*0*
System development	*1396*	*934*	*1.336838*	*0*
Multicellular organismal development	*1868*	*1248*	*1.334919*	*0*
Channel or pore class transporter activity	*363*	*242*	*1.332067*	*0*
Enzyme linked receptor protein signaling pathway	*228*	*152*	*1.332067*	*0*
Positive regulation of transcription DNA-dependent	*180*	*120*	*1.332067*	*0*
Cell morphogenesis	*374*	*249*	*1.330286*	*0*
Cell Differentiation	*1437*	*874*	*1.330286*	*0*
Positive regulation of transcription	*227*	*151*	*1.329133*	*0*
Anatomical structure development	*1679*	*1107*	*1.317389*	*0*
Positive regulation of cell proliferation	*192*	*126*	*1.311253*	*0*
Organ development	*996*	*650*	*1.303981*	*0*
Cell fate commitment	*75*	*60*	*1.59848*	*0*
Sequence-specific DNA binding	*424*	*337*	*1.588112*	*0*
Cellular morphogenesis during differentiation	*125*	*99*	*1.582495*	*0*
Positive regulation of biological process	*850*	*513*	*1.548823*	*0*
51674 localization of cell	*324*	*203*	*1.251896*	*0*
32502 developmental process	*2619*	*1639*	*1.250434*	*0*
06812 cation transport	*432*	*270*	*1.248812*	*0*
15075 ion transporter activity	*622*	*387*	*1.243191*	*0*
42127 regulation of cell proliferation	*383*	*238*	*1.241639*	*0*
65009 regulation of a molecular function	*400*	*246*	*1.228831*	*0*
06366 transcription from RNA polymerase II promoter	*532*	*326*	*1.2244*	*0*
50790 regulation of catalytic activity	*381*	*233*	*1.221935*	*0*
05576 extracellular region	*1056*	*596*	*1.127716*	*0*
06351 transcription DNA-dependent	*1866*	*1050*	*1.124333*	*0*

doi: 10.1038/nature08162

Supplemental Table 2 continued: D1 Histone-enriched loci (Illumina GAII FDR< 0.0001)

Go Category	Total Genes	Changed Genes	Enrichment	FDR
RNA biosynthetic process	*1869*	*1051*	*1.123597*	*0*
Extracellular region	*1056*	*596*	*1.127716*	*0*
Guanyl-nucleotide exchange factor activity	*123*	*84*	*1.364556*	*0.000073*
Blood vessel development	*133*	*90*	*1.352098*	*0.000074*
Ras protein signal transduction	*176*	*115*	*1.305577*	*0.000074*
Negative regulation of developmental process	*65*	*49*	*1.50626*	*0.000075*
Transport	*2094*	*1134*	*1.082066*	*0.000075*
Embryonic development ending in birth or egg hatching	*81*	*59*	*1.455406*	*0.000076*
Extracellular matrix structural constituent	*84*	*61*	*1.451001*	*0.000076*
Transporter activity	*1090*	*611*	*1.120036*	*0.000077*
Cyclic nucleotide metabolic process	*34*	*29*	*1.704262*	*0.000078*
Positive regulation of developmental process	*49*	*39*	*1.590324*	*0.000078*
Vasculature development	*135*	*92*	*1.361668*	*0.000079*
Anion transport	*161*	*107*	*1.32793*	*0.000079*
Extracellular matrix organization and biogenesis	*44*	*36*	*1.634809*	*0.00008*
Heart development	*80*	*59*	*1.473599*	*0.000081*
Extracellular matrix organization and biogenesis	*44*	*36*	*1.634809*	*0.00008*
Heart development	*80*	*59*	*1.473599*	*0.000081*
Voltage-gated potassium channel complex	*80*	*59*	*1.473599*	*0.000081*
Chordate embryonic development	*80*	*59*	*1.473599*	*0.000081*
Developmental maturation	*48*	*38*	*1.581829*	*0.000145*
Kidney development	*29*	*25*	*1.7225*	*0.000201*
Transcriptional activator activity	*243*	*152*	*1.24984*	*0.000203*
Anterior posterior pattern formation	*50*	*39*	*1.558518*	*0.000204*
Cyclic nucleotide biosynthetic process	*26*	*23*	*1.76755*	*0.000205*
Establishment of localization	*2154*	*1162*	*1.077898*	*0.000207*
Extracellular region part	*697*	*400*	*1.146686*	*0.000211*
Anatomical structure formation	*132*	*89*	*1.347204*	*0.000213*
Sensory organ development	*56*	*43*	*1.534255*	*0.000214*
Metanephros development	*23*	*21*	*1.824352*	*0.000216*
Blood vessel morphogenesis	*120*	*81*	*1.348717*	*0.00025*
Ionotropic glutamate receptor activity	*18*	*17*	*1.887094*	*0.000252*
Glutamate-gated ion channel activity	*18*	*17*	*1.887094*	*0.000252*
Muscle contraction	*149*	*98*	*1.314187*	*0.000255*
Brain development	*101*	*70*	*1.384822*	*0.000256*

doi: 10.1038/nature08162

Supplemental Table 3: Donor pool of histone-enriched loci (Illumina GAII FDR< 0.0001)

Go Category	Total Genes	Changed Genes	Enrichment	FDR
RNA polymerase II transcription factor activity	*25*	*22*	*2.187319*	*0*
Cell fate commitment	*69*	*53*	*1.909221*	*0*
Regionalization	*86*	*60*	*1.734133*	*0*
Wnt receptor signaling pathway	*121*	*84*	*1.725534*	*0*
Pattern specification process	*123*	*85*	*1.717684*	*0*
Embryonic morphogenesis	*93*	*64*	*1.710514*	*0*
Sensory organ development	*80*	*55*	*1.708843*	*0*
Negative regulation of cell differentiation	*67*	*46*	*1.706525*	*0*
Cellular morphogenesis during differentiation	*124*	*85*	*1.703832*	*0*
Neurogenesis	*257*	*171*	*1.653836*	*0*
Embryonic development	*93*	*61*	*1.630333*	*0*
Chordate embryonic development	*93*	*61*	*1.630333*	*0*
Brain development	*133*	*87*	*1.625912*	*0*
Sequence-specific DNA binding	*488*	*311*	*1.584054*	*0*
Positive regulation of transcription from RNA polymerase II promoter	*145*	*92*	*1.577064*	*0*
Embryonic development	*221*	*140*	*1.574582*	*0*
Cell projection organization and biogenesis	*193*	*121*	*1.558323*	*0*
Cell part morphogenesis	*193*	*121*	*1.558323*	*0*
Regulation of cell differentiation	*157*	*98*	*1.551515*	*0*
Cell morphogenesis	*256*	*158*	*1.534075*	*0*
Cellular structure morphogenesis	*256*	*158*	*1.534075*	*0*
Central nervous system development	*227*	*140*	*1.532963*	*0*
Nervous system development	*675*	*408*	*1.502401*	*0*
Positive regulation of RNA metabolic process	*229*	*138*	*1.497867*	*0*
Skeletal development	*203*	*121*	*1.481559*	*0*
Vasculature development	*165*	*96*	*1.446162*	*0*
Organ morphogenesis	*355*	*205*	*1.435341*	*0*
Cell migration	*222*	*128*	*1.433133*	*0*
Anatomical structure morphogenesis	*823*	*457*	*1.380212*	*0*
Transcription activator activity	*284*	*157*	*1.374076*	*0*
System development	*1538*	*817*	*1.320369*	*0*
Multicellular organismal development	*2093*	*1104*	*1.31108*	*0*
Positive regulation of cellular process	*952*	*501*	*1.308068*	*0*
Anatomical structure development	*1768*	*930*	*1.307465*	*0*
Cell development	*1089*	*565*	*1.289585*	*0*
Cell differentiation	*1636*	*835*	*1.268623*	*0*
Cellular developmental process	*1636*	*835*	*1.268623*	*0*
Organ development	*1106*	*564*	*1.267516*	*0*
Developmental process	*2848*	*1443*	*1.259377*	*0*
Intracellular signaling cascade	*1291*	*653*	*1.257235*	*0*
Regulation of developmental process	*729*	*367*	*1.251319*	*0*
Regulation of RNA metabolic process	*2115*	*1049*	*1.232806*	*0*
Regulation of transcription DNA-dependent	*2103*	*1043*	*1.232749*	*0*
Regulation of transcription	*2228*	*1104*	*1.231639*	*0*
Regulation of gene expression	*2358*	*1159*	*1.221713*	*0*
Transcription DNA-dependent	*2159*	*1061*	*1.221497*	*0*
RNA biosynthetic process	*2163*	*1061*	*1.219238*	*0*

doi: 10.1038/nature08162

Supplemental Table 3 continued: Donor pool of histone-enriched loci (Illumina GAII FDR< 0.0001)

Go Category	Total Genes	Changed Genes	Enrichment	FDR
Regulation of metabolic process	*2629*	*1285*	*1.214904*	*0*
Transcription	*2315*	*1129*	*1.212195*	*0*
Anatomical structure formation	*152*	*89*	*1.455378*	*0.000072*
Transmembrane receptor protein tyrosine kinase activity	*62*	*43*	*1.723877*	*0.000074*
Small GTPase regulator activity	*201*	*112*	*1.385005*	*0.000136*
Respiratory tube development	*43*	*32*	*1.849741*	*0.000138*
Insulin receptor signaling pathway	*31*	*25*	*2.004508*	*0.000139*
Appendage morphogenesis	*37*	*28*	*1.880987*	*0.000189*
Limb morphogenesis	*37*	*28*	*1.880987*	*0.000189*
Appendage development	*37*	*28*	*1.880987*	*0.000189*
Limb development	*37*	*28*	*1.880987*	*0.000189*
Regulation of anatomical structure morphogenesis	*69*	*45*	*1.621037*	*0.000485*
Transcription corepressor activity	*106*	*64*	*1.500734*	*0.000491*
BMP signaling pathway	*18*	*16*	*2.209413*	*0.000539*
Regulation of neuron differentiation	*26*	*21*	*2.007592*	*0.000595*
Localization of cell	*365*	*184*	*1.25301*	*0.000694*
Protein-tyrosine kinase activity	*159*	*89*	*1.391305*	*0.000706*
Rho protein signal transduction	*101*	*61*	*1.501198*	*0.00071*
Small conjugating protein ligase activity	*137*	*78*	*1.415153*	*0.000769*
Forebrain development	*46*	*32*	*1.729106*	*0.000773*
Voltage-gated cation channel activity	*141*	*80*	*1.410264*	*0.000778*
Blood vessel morphogenesis	*145*	*82*	*1.405644*	*0.000787*
Tube development	*114*	*67*	*1.460829*	*0.000795*
Cartilage development	*35*	*26*	*1.846438*	*0.0008*
Regulation of cellular component organization and biogenesis	*241*	*127*	*1.309834*	*0.00082*
Mesoderm formation	*14*	*13*	*2.308048*	*0.000851*
Heart development	*93*	*56*	*1.496699*	*0.000947*
Regulation of neurogenesis	*41*	*29*	*1.7581*	*0.000952*
Negative regulation of developmental process	*314*	*159*	*1.258627*	*0.000964*
Regulation of cell proliferation	*456*	*223*	*1.215541*	*0.000984*
Voltage-gated ion channel activity	*189*	*102*	*1.34143*	*0.00099*
Voltage-gated channel activity	*189*	*102*	*1.34143*	*0.00099*
Actin filament-based process	*206*	*109*	*1.315191*	*0.001133*
Regulation of anatomical structure morphogenesis	*69*	*45*	*1.621037*	*0.000485*
Transcription corepressor activity	*106*	*64*	*1.500734*	*0.000491*
BMP signaling pathway	*18*	*16*	*2.209413*	*0.000539*
Regulation of neuron differentiation	*26*	*21*	*2.007592*	*0.000595*
Localization of cell	*365*	*184*	*1.25301*	*0.000694*
Protein-tyrosine kinase activity	*159*	*89*	*1.391305*	*0.000706*
Rho protein signal transduction	*101*	*61*	*1.501198*	*0.00071*
Small conjugating protein ligase activity	*137*	*78*	*1.415153*	*0.000769*
Forebrain development	*46*	*32*	*1.729106*	*0.000773*
Voltage-gated cation channel activity	*141*	*80*	*1.410264*	*0.000778*
Blood vessel morphogenesis	*145*	*82*	*1.405644*	*0.000787*
Tube development	*114*	*67*	*1.460829*	*0.000795*
Cartilage development	*35*	*26*	*1.846438*	*0.0008*
Regulation of cellular component organization and biogenesis	*241*	*127*	*1.309834*	*0.00082*

doi: 10.1038/nature08162

Supplemental Table 4: TH2B Enriched Promoters (D1 array)

GO CATEGORY	TOTAL GENES	CHANGED GENES	ENRICHMENT	FDR
Beta DNA polymerase activity	*3*	*3*	*22.225524*	*0.018333*
Multidrug transport	*3*	*3*	*22.225524*	*0.018333*
Cation transport	*380*	*35*	*2.047088*	*0.023333*
Metal ion transport	*310*	*30*	*2.150857*	*0.0275*
Voltage-gated potassium channel	*72*	*11*	*3.395566*	*0.0325*
Potassium ion transport	*141*	*17*	*2.679673*	*0.03375*
Alpha-type channel activity	*333*	*32*	*2.135786*	*0.035*
Voltage-gated ion channel activity	*161*	*18*	*2.484841*	*0.035455*
Potassium ion binding	*106*	*14*	*2.935447*	*0.036*
Transporter activity	*1067*	*73*	*1.520584*	*0.037143*
Adenylate cyclase activity	*14*	*5*	*7.937687*	*0.04*
Channel or pore class transporter activity	*338*	*33*	*2.169948*	*0.06*

doi: 10.1038/nature08162

Supplemental Table 5: H3K4me2 enriched promoters (D1 array)

GO CATEGORY	TOTAL GENES	CHANGED GENES	ENRICHMENT	FDR
Multicellular organismal development	*1620*	*148*	*1.394279*	*0.005*
Developmental process	*2265*	*197*	*1.327398*	*0.01*
Sequence-specific DNA binding	*425*	*50*	*1.795495*	*0.026667*
Anatomical structure development	*1465*	*132*	*1.375116*	*0.03*
System development	*1231*	*113*	*1.400953*	*0.031429*
Cell-cell signaling	*525*	*57*	*1.656985*	*0.035*
Organ development	*869*	*84*	*1.47524*	*0.035*
Menstrual cycle	*30*	*9*	*4.578511*	*0.04*
Multicellular organism reproduction	*45*	*10*	*3.39149*	*0.065294*
Reproductive process in a multicellular organism	*45*	*10*	*3.39149*	*0.065294*
Multicellular organismal process	*2648*	*212*	*1.221859*	*0.067333*

doi: 10.1038/nature08162

Supplemental Table 6: H3K4me3 enriched promoters (D1 array)

GO CATEGORY	TOTAL GENES	CHANGED GENES	ENRICHMENT	FDR
Nuclear pore	*44*	*15*	*3.620699*	*0*
mRNA metabolic process	*198*	*44*	*2.36016*	*0*
mRNA processing	*165*	*35*	*2.25288*	*0*
Chromosome	*204*	*40*	*2.082494*	*0*
RNA processing	*266*	*51*	*2.036303*	*0*
Nuclear part	*596*	*109*	*1.94238*	*0*
RNA binding	*481*	*78*	*1.722279*	*0*
Cell cycle	*606*	*98*	*1.717542*	*0*
Cell cycle process	*530*	*82*	*1.643205*	*0*
RNA metabolic process	*1827*	*256*	*1.488179*	*0*
Transcription DNA-dependent	*1510*	*199*	*1.399684*	*0*
RNA biosynthetic process	*1512*	*199*	*1.397833*	*0*
Regulation of transcription DNA-dependent	*1467*	*193*	*1.397272*	*0*
Transcription	*1644*	*216*	*1.395423*	*0*
DNA binding	*1522*	*199*	*1.388648*	*0*
Regulation of transcription	*1580*	*206*	*1.384727*	*0*
Regulation of nucleobase nucleoside nucleotide and nucleic acid metabolic process	*1623*	*211*	*1.380759*	*0*
Regulation of cellular metabolic process	*1772*	*229*	*1.372542*	*0*
Regulation of metabolic process	*1839*	*237*	*1.368739*	*0*
Regulation of cellular process	*2889*	*341*	*1.253605*	*0*
Regulation of biological process	*3134*	*362*	*1.226771*	*0*
RNA splicing	*137*	*30*	*2.325705*	*0.000227*
Macromolecule localization	*548*	*83*	*1.608612*	*0.000233*
Intracellular transport	*494*	*75*	*1.612457*	*0.0004*
Cellular protein metabolic process	*2294*	*271*	*1.254671*	*0.000417*
RNA localization	*36*	*13*	*3.835259*	*0.000426*
Ligase activity	*238*	*42*	*1.874244*	*0.000727*
Establishment of cellular localization	*596*	*86*	*1.53252*	*0.000741*
Specific RNA polymerase II transcription factor activity	*29*	*11*	*4.028548*	*0.000755*
Translation initiation factor activity	*46*	*14*	*3.232393*	*0.000833*
Spliceosome	*88*	*21*	*2.53449*	*0.000847*
Nucleic acid transport	*35*	*12*	*3.641389*	*0.000862*
RNA transport	*35*	*12*	*3.641389*	*0.000862*
Establishment of RNA localization	*35*	*12*	*3.641389*	*0.000862*
Ribonucleoprotein complex	*328*	*53*	*1.716153*	*0.001475*
Nuclear membrane part	*54*	*15*	*2.9502*	*0.001791*
Pore complex	*54*	*15*	*2.9502*	*0.001791*
Tricarboxylic acid cycle	*22*	*9*	*4.344839*	*0.001846*
Acetyl-CoA catabolic process	*22*	*9*	*4.344839*	*0.001846*
Cellular localization	*611*	*86*	*1.494897*	*0.001905*
Ubiquitin cycle	*267*	*45*	*1.790009*	*0.001935*
Translation regulator activity	*99*	*21*	*2.25288*	*0.004405*
Spermatogenesis	*141*	*27*	*2.033755*	*0.004444*
Male gamete generation	*141*	*27*	*2.033755*	*0.004444*

doi: 10.1038/nature08162

Supplemental Table 6 continued: H3K4me3 enriched promoters (D1)

GO CATEGORY	TOTAL GENES	CHANGED GENES	ENRICHMENT	FDR
Transcription regulator activity	*1090*	*136*	*1.325154*	*0.004458*
RNA export from nucleus	*20*	*8*	*4.248287*	*0.004512*
Translation factor activity nucleic acid binding	*86*	*19*	*2.346438*	*0.004828*
Protein transport	*460*	*66*	*1.523842*	*0.004884*
Microtubule-based process	*136*	*26*	*2.030431*	*0.005114*
Protein modification process	*1218*	*149*	*1.29925*	*0.006517*
Nuclear chromosome	*55*	*14*	*2.703456*	*0.007444*
Acetyl-CoA metabolic process	*27*	*9*	*3.540239*	*0.00828*
Transcription from RNA polymerase II promoter	*460*	*65*	*1.500754*	*0.00837*
Nucleobase nucleoside nucleotide and nucleic acid transport	*44*	*12*	*2.89656*	*0.008404*
Organelle organization and biogenesis	*711*	*93*	*1.389208*	*0.008438*
Microtubule cytoskeleton organization and biogenesis	*57*	*14*	*2.608598*	*0.008454*
Sexual reproduction	*218*	*36*	*1.75388*	*0.008526*
Meiotic recombination	*18*	*7*	*4.130279*	*0.010918*
Tricarboxylic acid cycle intermediate metabolic process	*23*	*8*	*3.694163*	*0.01101*
Nuclear export	*29*	*9*	*3.296085*	*0.018762*
Cofactor catabolic process	*29*	*9*	*3.296085*	*0.018762*
Gamete generation	*184*	*31*	*1.78936*	*0.018835*
Protein complex	*1361*	*161*	*1.256382*	*0.01902*
Intracellular protein transport	*289*	*44*	*1.616995*	*0.019208*
Endomembrane system	*331*	*49*	*1.572251*	*0.0194*

doi: 10.1038/nature08162

Supplemental Table 7: H3K4me3-enriched loci as determined from donor pool (Illumina GAII FDR< 0.001)

GO CATEGORY	TOTAL GENES	CHANGED GENES	ENRICHMENT	FDR
RNA splicing	*64*	*46*	*1.622*	*0*
spliceosome	*119*	*85*	*1.612*	*0*
ATP-dependent helicase activity	*88*	*62*	*1.59*	*0*
mRNA processing	*235*	*156*	*1.48*	*0*
Protein folding	*151*	*99*	*1.48*	*0*
Helicase activity	*129*	*84*	*1.4699*	*0*
ribonucleoprotein complex biogenesis and assembly	*186*	*134*	*1.468*	*0*
mRNA metabolic process	*272*	*177*	*1.463*	*0*
RNA processing	*404*	*253*	*1.413636*	*0*
Ribonucleoprotein complex	*400*	*250*	*1.41*	*0*
nucleolus	*158*	*97*	*1.385*	*0*
Microtubule-based process	*190*	*116*	*1.378*	*0*
Ligase activity	*338*	*204*	*1.362*	*0*
Translation	*351*	*210*	*1.350*	*0*
Mitotic cell cycle	*295*	*168*	*1.285*	*0*
Cell cycle phase	*323*	*183*	*1.285*	*0*
Nucleoplasm	*442*	*250*	*1.279302*	*0*
Nucleoplasm part	*381*	*215*	*1.272843*	*0*
Cell cycle process	*395*	*253*	*1.257553*	*0*
Transcription factor binding	*390*	*214*	*1.238*	*0*
RNA metabolic process	*2624*	*1411*	*1.213*	*0*
Transcription from RNA polymerase II promoter	*610*	*326*	*1.206*	*0*
Nucleobase nucleoside nucleotide metabolic process	*3279*	*1734*	*1.19*	*0*
DNA binding	*2080*	*1099*	*1.199734*	*0*
Cell cycle	*691*	*364*	*1.18*	*0*
Gene expression	*3028*	*1878*	*1.18*	*0*
Transcription	*2315*	*1207*	*1.173*	*0*
Transcription DNA-dependent	*2159*	*1121*	*1.172*	*0*
RNA biosynthetic process	*2163*	*1222*	*1.171*	*0*
Regulation of gene expression	*2358*	*1223*	*1.1698*	*0*
Transcription regulator activity	*1309*	*678*	*1.169*	*0*
Regulation of transcription DNA-dependent	*2103*	*1089*	*1.1689*	*0*
Regulation of transcription	*2228*	*1146*	*1.161*	*0*
Post-translational protein modification	*1346*	*777*	*1.133388*	*0*
Ribonucleotide binding	*1537*	*879*	*1.12284*	*0*

doi: 10.1038/nature08162

Supplemental Table 7 continued: H3K4me3-enriched loci as determined from donor pool (Illumina GAII FDR< 0.001)

GO CATEGORY	Total Genes	Changed Genes	Enrichment	FDR
M phase	*261*	*148*	*1.28*	*0.000098*
Mitochondrion	*807*	*413*	*1.15*	*0.00099*
Ribosome biogenesis and assembly	*86*	*57*	*1.49*	*0.0001*
Regulation of cell cycle	*272*	*154*	*1.278*	*0.000102*
Ubiquitin-dependent protein catabolic process	*164*	*98*	*1.3489*	*0.000185*
RNA helicase activity	*28*	*24*	*1.85*	*0.000187*
Protein RNA complex Assembly	*105*	*67*	*1.440*	*0.000189*
Spindle	*70*	*48*	*1.547*	*0.000192*
Spermatogenesis	*202*	*106*	*1.296*	*0.00082*
Male gamete generation	*202*	*106*	*1.296*	*0.00082*
Response to DNA damage stimulus	*278*	*152*	*1.234*	*0.003*
Mitosis	*198*	*112*	*1.276*	*0.003*
Flagellum	*30*	*23*	*1.73*	*0.00331*
Regulation of translation	*99*	*61*	*1.347*	*0.00349*
Centrosome	*124*	*74*	*1.347*	*0.00349*
Gamete generation	*247*	*136*	*1.24*	*0.00353*
Regulation of RNA cellular biosynthetic process	*128*	*76*	*1.3403*	*0.00357*
Negative regulation of cell cycle	*138*	*81*	*1.323143*	*0.00038*
mRNA splice site selection	*13*	*12*	*2.08*	*0.00397*
rRNA processing	*61*	*40*	*1.48839*	*0.00515*
nuclear chromosome part	*61*	*40*	*1.488*	*0.00515*
Translation initiation factor activity	*58*	*44*	*1.489455*	*0.0056*
Negative regulation of cellular process	*1023*	*579*	*1.111234*	*0.00057*
Chromosome organization and biogenesis	*345*	*211*	*1.200787*	*0.000606*
Regulation of protein metabolic process	*301*	*184*	*1.200201*	*0.001384*
RNA splicing via transesterification reactions	*64*	*47*	*1.441852*	*0.001392*
RNA splicing via transesterification reactions with bulged adenosine as nucleophile	*64*	*47*	*1.441852*	*0.001392*
Nuclear mRNA splicing via spliceosome	*64*	*47*	*1.441852*	*0.001392*
rRNA metabolic process	*64*	*47*	*1.441852*	*0.001392*
Establishment of cellular localization	*766*	*439*	*1.125223*	*0.0012*
Transcription factor complex	*161*	*105*	*1.28046*	*0.001208*
Establishment of protein localization	*674*	*389*	*1.133163*	*0.00125*
Regulation of cyclin-dependent protein kinase activity	*48*	*37*	*1.513433*	*0.001258*
Protein tyrosine phosphatase activity	*97*	*67*	*1.356144*	*0.001438*
Regulation of translation	*99*	*68*	*1.348579*	*0.001657*
Interphase of mitotic cell cycle	*84*	*59*	*1.379036*	*0.001667*
G1 S transition of mitotic cell cycle	*33*	*27*	*1.606396*	*0.001677*
Nucleolar part	*39*	*31*	*1.56063*	*0.001718*
Embryonic developmemt	*220*	*120*	*1.225*	*0.01*

doi: 10.1038/nature08162

Supplemental Table 8: Donor pool H3K27me3-enriched loci (Illumina GAII FDR <0.0001)

GO CATEGORY	TOTAL GENES	CHANGED GENES	ENRICHMENT	FDR
Wnt receptor signaling	*20*	*18*	*2.706*	*0*
Embryonic organ development	*20*	*18*	*2.706*	*0*
Transmembrane receptor protein	*18*	*16*	*2.706*	*0*
Inner ear morphogenesis	*27*	*24*	*2.619*	*0*
Mesenchymal cell development and differentiation	*23*	*19*	*2.484*	*0*
Cell fate commitment	*69*	*54*	*2.353*	*0*
Embryonic morphogenesis	*93*	*71*	*2.295*	*0*
Lung development	*42*	*31*	*2.219*	*0*
Cyclic nucleotide metabolic process	*37*	*27*	*2.194*	*0*
Appendage morphogenesis	*37*	*27*	*2.1943*	*0*
Limb morphogenesis	*37*	*27*	*2.1943*	*0*
Appendage development	*37*	*27*	*2.1943*	*0*
Limb development	*37*	*27*	*2.1943*	*0*
Sensory organ development	*80*	*58*	*2.1800*	*0*
Potassium ion binding	*123*	*89*	*2.171*	*0*
Regionalization	*86*	*62*	*2.16*	*0*
Anterior posterior pattern formation	*54*	*38*	*2.116*	*0*
Axonogenesis	*112*	*77*	*2.06*	*0*
Pattern specification process	*123*	*84*	*2.0535*	*0*
Regulation of anatomical structure morphogenesis	*69*	*47*	*2.048*	*0*
Neuron differentiation	*206*	*139*	*2.029*	*0*
Forebrain development	*46*	*31*	*2.026*	*0*
Developmental maturation	*52*	*35*	*2.02*	*0*
Neuron morphogenesis during differentiation	*118*	*79*	*2.013*	*0*
Skeletal development	*203*	*133*	*1.970*	*0*
Neurite development	*133*	*87*	*1.966*	*0*
Neurogenesis	*265*	*165*	*1.930*	*0*
Cell migration	*222*	*142*	*1.9217*	*0*
Brian development	*133*	*85*	*1.921*	*0*
Embryonic development	*221*	*40*	*1.904*	*0*
Sequence specific DNA binding	*488*	*309*	*1.904*	*0*
Tube Development	*114*	*70*	*1.86*	*0*
Vasculature development	*165*	*101*	*1.846*	*0*
Organ morphogenesis	*335*	*215*	*1.821*	*0*
Blood vessel development	*162*	*98*	*1.819*	*0*
Central nervous system development	*227*	*137*	*1.814*	*0*
Heart development	*93*	*56*	*1.8106*	*0*
Anatomical structure formation	*152*	*91*	*1.8002*	*0*
Bone remodeling	*96*	*57*	*1.785*	*0*
Chordate embryonic development	*93*	*55*	*1.778*	*0*

doi: 10.1038/nature08162

Supplemental Table 8: Donor pool H3K27me3-enriched loci (Illumina GAII FDR <0.0001)

GO CATEGORY	TOTAL GENES	CHANGED GENES	ENRICHMENT	FDR
System process	*1264*	*504*	*1.1989*	*0*
Ligand gated ion channel	*97*	*54*	*1.674*	*0*
Embryonic limb morphogenesis	*33*	*24*	*2.186*	*0.000044*
Embryonic appendage morphogenesis	*33*	*24*	*2.186*	*0.00044*
Neural crest cell development and differentiation	*14*	*13*	*2.792*	*0.00082*
Metanephros development	*23*	*18*	*2.3533*	*0.000114*
Voltage-gated calcium channel complex	*21*	*17*	*2.434*	*0.000115*
Eye morphogenesis	*21*	*17*	*2.434*	*0.000115*
Eye development	*42*	*28*	*2.004*	*0.000116*
Transcription	*2315*	*570*	*1.183421*	*0.000153*
Dorsal ventral pattern formation	*28*	*18*	*3.089797*	*0.000154*
Endoderm development	*9*	*9*	*4.806351*	*0.000155*
Negative regulation of cell differentiation	*67*	*32*	*2.295571*	*0.000155*
Developmental maturation	*52*	*27*	*2.495605*	*0.000156*
Ligand-gated ion channel activity	*97*	*42*	*2.081101*	*0.000158*
Morphogenesis of an epithelium	*63*	*29*	*2.212447*	*0.00018*
Neuron fate commitment	*14*	*11*	*3.776419*	*0.000181*
Regulation of heart contraction	*42*	*22*	*2.517613*	*0.000182*
Tube morphogenesis	*82*	*35*	*2.051491*	*0.000183*
Tissue remodeling	*105*	*42*	*1.92254*	*0.000183*
Positive regulation of transcription DNA-dependent	*227*	*76*	*1.609175*	*0.000184*
Somitogenesis	*16*	*12*	*3.604763*	*0.000185*
Biological process	*12711*	*2729*	*1.031904*	*0.000186*
Growth factor activity	*164*	*59*	*1.729114*	*0.000187*

doi: 10.1038/nature08162

Supplemental Table 9: Loci enriched for H3K4me3 and H3k27me3 derived from donor pool sequencing data (Illumina GAII FDR< 0.0001)

GO CATEGORY	TOTAL GENES	CHANGED GENES	ENRICHMENT	FDR
Cell fate determination	*27*	*14*	*4.239878*	*0*
Endocrine system development	*33*	*17*	*4.212346*	*0*
Cell fate commitment	*69*	*34*	*4.0292*	*0*
Neuron migration	*36*	*16*	*3.634181*	*0*
Embryonic morphogenesis	*93*	*41*	*3.604873*	*0*
Appendage morphogenesis	*37*	*16*	*3.53596*	*0*
Limb morphogenesis	*37*	*16*	*3.53596*	*0*
Appendage development	*37*	*16*	*3.53596*	*0*
Limb development	*37*	*16*	*3.53596*	*0*
Forebrain development	*46*	*19*	*3.377418*	*0*
Sensory organ development	*80*	*30*	*3.06634*	*0*
Anterior posterior pattern formation	*54*	*20*	*3.028484*	*0*
Brain development	*133*	*47*	*2.889584*	*0*
Regionalization	*86*	*30*	*2.852409*	*0*
Heart development	*93*	*32*	*2.813559*	*0*
Embryonic development	*221*	*74*	*2.737969*	*0*
Pattern specification process	*123*	*41*	*2.725636*	*0*
Homophilic cell adhesion	*133*	*43*	*2.643662*	*0*
Sequence-specific DNA binding	*488*	*155*	*2.597173*	*0*
Central nervous system development	*227*	*69*	*2.485491*	*0*
Chordate embryonic development	*93*	*28*	*2.461864*	*0*
Neurogenesis	*257*	*73*	*2.322623*	*0*
Tube development	*114*	*32*	*2.295272*	*0*
Skeletal development	*203*	*56*	*2.255698*	*0*
Organ morphogenesis	*355*	*97*	*2.234253*	*0*
Positive regulation of transcription from RNA polymerase II promoter	*145*	*39*	*2.199306*	*0*
Circulatory system process	*157*	*42*	*2.187453*	*0*
Blood circulation	*157*	*42*	*2.187453*	*0*
Regulation of cell differentiation	*157*	*42*	*2.187453*	*0*
Muscle development	*166*	*42*	*2.068856*	*0*
Transcription factor activity	*881*	*221*	*2.051188*	*0*
Neuron development	*152*	*38*	*2.044227*	*0*
Nervous system development	*675*	*168*	*2.035141*	*0*
Vasculature development	*165*	*41*	*2.031837*	*0*
Anatomical structure morphogenesis	*823*	*186*	*1.848001*	*0*
Organ development	*1106*	*249*	*1.840913*	*0*
Positive regulation of transcription DNA-dependent	*227*	*51*	*1.837102*	*0*
System development	*1538*	*340*	*1.807639*	*0*
Positive regulation of transcription	*278*	*60*	*1.7648*	*0*
Transcription regulator activity	*1309*	*281*	*1.755318*	*0*
Multicellular organismal development	*2093*	*446*	*1.742427*	*0*
Anatomical structure development	*1768*	*376*	*1.73898*	*0*
Positive regulation of metabolic process	*408*	*84*	*1.683481*	*0*

doi: 10.1038/nature08162

Supplemental Table 9 continued: Loci enriched for H3K4me3 and H3k27me3 derived from donor pool sequencing data (Illumina GAII FDR< 0.0001)

GO CATEGORY	TOTAL GENES	CHANGED GENES	ENRICHMENT	FDR
Positive regulation of cellular metabolic process	*396*	*81*	*1.672549*	*0*
Regulation of transcription from RNA polymerase II promoter	*424*	*87*	*1.677809*	*0*
Biological adhesion	*683*	*139*	*1.664114*	*0*
Cell-cell signaling	*611*	*119*	*1.592556*	*0*
Developmental process	*2848*	*537*	*1.541783*	*0*
Transcription from RNA polymerase II promoter	*610*	*113*	*1.514738*	*0*
Cell differentiation	*1636*	*292*	*1.459448*	*0*
Cellular developmental process	*1636*	*292*	*1.459448*	*0*
Multicellular organismal process	*3267*	*567*	*1.419133*	*0*
Positive regulation of cellular process	*952*	*164*	*1.408627*	*0*
Cell development	*1089*	*186*	*1.396607*	*0*
Positive regulation of biological process	*1046*	*177*	*1.383664*	*0*
Negative regulation of cellular process	*1023*	*171*	*1.366814*	*0*
DNA binding	*2080*	*347*	*1.364128*	*0*
Regulation of transcription	*2228*	*368*	*1.350584*	*0*
Regulation of RNA metabolic process	*2115*	*348*	*1.34542*	*0*
Regulation of transcription DNA-dependent	*2103*	*346*	*1.345321*	*0*
Regulation of nucleobase nucleoside nucleotide and nucleic acid metabolic process	*2282*	*374*	*1.340124*	*0*
Transcription DNA-dependent	*2159*	*351*	*1.329363*	*0*
RNA biosynthetic process	*2163*	*351*	*1.326904*	*0*
Regulation of gene expression	*2358*	*382*	*1.324673*	*0*
Transcription	*2315*	*373*	*1.317489*	*0*
Biological regulation	*4522*	*682*	*1.233227*	*0*
Regulation of biological process	*4060*	*605*	*1.21848*	*0*
Cell communication	*3573*	*524*	*1.199188*	*0*
Positive regulation of nucleobase nucleoside nucleotide and nucleic acid metabolic process	*291*	*61*	*1.714059*	*0.000085*
Negative regulation of biological process	*1089*	*179*	*1.344046*	*0.000086*
Embryonic limb morphogenesis	*33*	*14*	*3.468991*	*0.000088*
Embryonic appendage morphogenesis	*33*	*14*	*3.468991*	*0.000088*
Positive regulation of RNA metabolic process	*229*	*51*	*1.821058*	*0.000089*
Substrate specific channel activity	*365*	*73*	*1.635381*	*0.00009*
Signal transduction	*3247*	*466*	*1.173526*	*0.000164*
Blood vessel development	*162*	*39*	*1.968515*	*0.000165*
Neurotransmitter binding	*101*	*28*	*2.266865*	*0.000167*
Positive regulation of heart contraction	*5*	*5*	*8.176907*	*0.000244*
Morphogenesis of an epithelium	*63*	*20*	*2.595843*	*0.000317*
Regulation of developmental process	*729*	*124*	*1.390859*	*0.000347*
Cellular morphogenesis during differentiation	*124*	*31*	*2.044227*	*0.00035*
Anatomical structure formation	*152*	*36*	*1.936636*	*0.000362*

doi: 10.1038/nature08162

Supplemental Table 10: Promoters deficient in DNA methylation (D2 and D4 array)

GO CATEGORY	TOTAL GENES	CHANGED GENES	ENRICHMENT	FDR
Embryonic development	*199*	*22*	*3.061998*	*0*
Multicellular organismal development	*1620*	*102*	*1.743896*	*0*
System development	*1231*	*83*	*1.867478*	*0*
Nucleus	*2828*	*153*	*1.498468*	*0*
RNA biosynthetic process	*1512*	*95*	*1.740232*	*0*
Transcription	*1644*	*202*	*1.70159*	*0*
Transcription regulator activity	*1090*	*75*	*1.905768*	*0*
Anatomical structure development	*1465*	*92*	*1.739344*	*0*
Regulation of transcription	*1580*	*97*	*1.700396*	*0*
RNA metabolic process	*1827*	*108*	*1.637271*	*0*
Nucleobase nucleoside nucleotide and nucleic acid metabolic process	*2489*	*136*	*1.513385*	*0*
Regulation of cellular metabolic process	*1772*	*105*	*1.641198*	*0*
Nucleic acid binding	*2348*	*130*	*1.533489*	*0*
Regulation of transcription DNA-dependent	*1467*	*91*	*1.718093*	*0*
DNA binding	*1522*	*93*	*1.692402*	*0*
Regulation of metabolic process	*1839*	*106*	*1.596465*	*0*
Organ development	*869*	*61*	*1.94422*	*0*
Biopolymer metabolic process	*3392*	*170*	*1.388125*	*0*
Developmental process	*2265*	*123*	*1.504085*	*0*
Transcription factor activity	*755*	*54*	*1.980989*	*0*
Transcription from RNA polymerase II promoter	*460*	*38*	*2.288027*	*0*
Regulation of transcription from RNA polymerase II promoter	*297*	*26*	*2.424668*	*0.000588*
Female pronucleus	*3*	*3*	*27.697168*	*0.001081*
Nervous system development	*553*	*39*	*1.953326*	*0.001053*
Central nervous system development	*179*	*18*	*2.78519*	*0.00125*
Dorsal ventral pattern formation	*22*	*6*	*7.553773*	*0.001463*
Positive regulation of nucleobase nucleoside nucleotide and nucleic acid metabolic process	*199*	*19*	*2.644453*	*0.002381*
Gamete generation	*184*	*18*	*2.709506*	*0.002558*
Anatomical structure formation	*122*	*14*	*3.178364*	*0.002727*
Anatomical structure morphogenesis	*730*	*46*	*1.745301*	*0.003556*
Notch signaling pathway	*34*	*7*	*5.702358*	*0.003478*
Pronucleus	*4*	*3*	*20.772876*	*0.004681*
M phase	*175*	*17*	*2.690582*	*0.005417*
Multicellular organismal process	*2648*	*127*	*1.328376*	*0.0054*
Regionalization	*72*	*10*	*3.846829*	*0.00549*
Cell cycle phase	*214*	*19*	*2.459094*	*0.005385*
Negative regulation of cellular process	*776*	*47*	*1.677535*	*0.006038*
Sequence-specific DNA binding	*425*	*30*	*1.955094*	*0.006667*
Negative regulation of cellular metabolic process	*256*	*21*	*2.272033*	*0.007636*
Cell cycle process	*530*	*35*	*1.829058*	*0.007544*
Negative regulation of biological process	*807*	*48*	*1.647415*	*0.007414*
Chromosome	*204*	*18*	*2.443868*	*0.007288*
Brain development	*93*	*11*	*3.276009*	*0.011*
Positive regulation of transcription	*192*	*17*	*2.452353*	*0.011148*

doi: 10.1038/nature08162

Supplemental Table 10 continued : Promoters deficient in DNA methylation (D1 and D2 array)

GO CATEGORY	TOTAL GENES	CHANGED GENES	ENRICHMENT	FDR
Positive regulation of cellular metabolic process	*264*	*21*	*2.203184*	*0.010968*
Gastrulation	*31*	*6*	*5.360742*	*0.013492*
Positive regulation of transcription DNA-dependent	*147*	*14*	*2.637826*	*0.019692*
Meiosis	*44*	*7*	*4.406368*	*0.020149*
M phase of meiotic cell cycle	*44*	*7*	*4.406368*	*0.020149*
Sexual reproduction	*218*	*18*	*2.286922*	*0.021618*
Meiotic cell cycle	*45*	*7*	*4.308448*	*0.022754*
Mitosis	*135*	*13*	*2.667135*	*0.023714*
Cellular protein complex disassembly	*14*	*4*	*7.913477*	*0.026197*
Positive regulation of metabolic process	*280*	*21*	*2.077288*	*0.025833*
Male pronucleus	*2*	*2*	*27.697168*	*0.042405*
Regulation of translational elongation	*2*	*2*	*27.697168*	*0.042405*
Heart development	*75*	*9*	*3.32366*	*0.042683*
Heart morphogenesis	*7*	*3*	*11.870215*	*0.045833*
Vasculature development	*122*	*12*	*2.724312*	*0.045412*
Forebrain development	*25*	*5*	*5.539434*	*0.046292*
Spermatogenesis	*141*	*13*	*2.55364*	*0.045495*
Male gamete generation	*141*	*13*	*2.55364*	*0.045495*

doi: 10.1038/nature08162

Supplemental Table 11: Promoters that share histone enrichment and DNA hypomethylation (array)

GO CATEGORY	TOTAL GENES	CHANGED GENES	ENRICHMENT	FDR
Sequence-specific DNA binding	*425*	*35*	*5.037969*	*0*
Developmental process	*2265*	*37*	*2.186039*	*0*
Multicellular organismal development	*1620*	*30*	*2.478168*	*0*
DNA binding	*1522*	*27*	*2.373961*	*0*
Anatomical structure development	*1465*	*26*	*2.374981*	*0.002*
Transcription factor activity	*755*	*17*	*3.013189*	*0.003333*
RNA metabolic process	*1827*	*29*	*2.124144*	*0.002857*
Nucleic acid binding	*2348*	*34*	*1.937784*	*0.0025*
Regulation of transcription	*1467*	*25*	*2.280522*	*0.003333*
Neural tube patterning	*2*	*2*	*133.821053*	*0.006*
System development	*1231*	*22*	*2.391603*	*0.007273*
Transcription DNA-dependent	*1510*	*25*	*2.21558*	*0.0075*
RNA biosynthetic process	*1512*	*25*	*2.21265*	*0.006923*
Transcription regulator activity	*1090*	*20*	*2.455432*	*0.009286*
Transcription from RNA polymerase II promoter	*460*	*12*	*3.490984*	*0.008667*
Regulation of transcription	*1580*	*25*	*2.117422*	*0.0225*
Heart development	*75*	*5*	*8.921404*	*0.022353*
Regulation of nucleobase nucleoside nucleotide and nucleic acid metabolic process	*1623*	*25*	*2.061322*	*0.022778*
Regulation of metabolic process	*1839*	*27*	*1.964746*	*0.0235*
Skeletal development	*174*	*7*	*5.383606*	*0.023333*
Transcription	*1644*	*25*	*2.034992*	*0.022273*
Nucleobase and nucleic acid metabolic process	*2489*	*33*	*1.774245*	*0.032609*
Regulation of cellular metabolic process	*1772*	*26*	*1.963514*	*0.035*
Regulation of bone remodeling	*21*	*6*	*19.117293*	*0.0368*
Cell-cell signaling	*525*	*12*	*3.058767*	*0.035385*
Multicellular organismal process	*2648*	*34*	*1.718246*	*0.034074*
Voltage-gated potassium channel activity	*93*	*5*	*7.19468*	*0.040345*
Regulation of biological process	*3134*	*38*	*1.622591*	*0.039*
Biological regulation	*3396*	*40*	*1.57622*	*0.04*
Nervous system development	*553*	*12*	*2.903893*	*0.046875*
Alpha-type channel activity	*333*	*9*	*3.616785*	*0.045758*
Channel or pore class transporter activity	*338*	*9*	*3.563282*	*0.046176*
Anatomical structure morphogenesis	*730*	*14*	*2.566431*	*0.045143*
Positive regulation of cell differentiation	*28*	*3*	*14.33797*	*0.058056*
Regulation of cellular process	*2889*	*35*	*1.621231*	*0.057568*
Cellular morphogenesis during differentiation	*108*	*5*	*6.195419*	*0.056053*
Positive regulation of cellular process	*671*	*13*	*2.592658*	*0.056154*
Cell development	*859*	*15*	*2.336805*	*0.065854*
Potassium channel activity	*118*	*5*	*5.670384*	*0.067674*
Organ development	*869*	*15*	*2.309915*	*0.066136*

doi: 10.1038/nature08162

Supplemental Table 12: Sperm DNA demethylation extends beyond CpGs

GO CATEGORY	TOTAL GENES	CHANGED GENES	ENRICHMENT	FDR
Nucleus	*2665*	*103*	*1.668257*	*0*
Transcription DNA-dependent	*1398*	*64*	*1.97604*	*0*
RNA biosynthetic process	*1399*	*64*	*1.974627*	*0*
Regulation of transcription DNA-dependent	*1353*	*62*	*1.977957*	*0*
Regulation of RNA metabolic process	*1365*	*62*	*1.960568*	*0*
Transcription regulator activity	*996*	*50*	*2.166874*	*0*
Transcription	*1518*	*65*	*1.848266*	*0*
Regulation of transcription	*1457*	*63*	*1.866396*	*0*
RNA metabolic process	*1698*	*70*	*1.77944*	*0*
Regulation of nucleobase nucleoside nucleotide and nucleic acid metabolic process	*1500*	*64*	*1.841669*	*0*
Regulation of gene expression	*1540*	*65*	*1.821862*	*0*
Regulation of cellular process	*2683*	*97*	*1.560537*	*0*
Regulation of cellular metabolic process	*1708*	*69*	*1.74375*	*0*
Regulation of biological process	*2873*	*101*	*1.51743*	*0*
Intracellular	*6181*	*180*	*1.257004*	*0*
Gene expression	*2015*	*77*	*1.649448*	*0*
DNA binding	*1398*	*59*	*1.821662*	*0*
Biopolymer metabolic process	*3194*	*108*	*1.459526*	*0*
Regulation of metabolic process	*1757*	*69*	*1.695119*	*0*
Intracellular part	*5850*	*171*	*1.26172*	*0*
Nucleobase nucleoside nucleotide and nucleic acid metabolic process	*2249*	*82*	*1.573792*	*0*
Embryonic development	*183*	*15*	*3.538043*	*0.00037*
Transcription factor activity	*688*	*34*	*2.133111*	*0.000385*
Biological regulation	*3250*	*107*	*1.421096*	*0.0004*
Nucleic acid binding	*2071*	*76*	*1.584004*	*0.000417*
Embryonic morphogenesis	*78*	*10*	*5.533862*	*0.000435*
Macromolecule metabolic process	*4200*	*131*	*1.34631*	*0.000455*
Positive regulation of transcription	*214*	*16*	*3.227224*	*0.00069*
Anatomical structure morphogenesis	*643*	*32*	*2.148137*	*0.000714*
Positive regulation of nucleobase nucleoside nucleotide and nucleic acid metabolic process	*225*	*16*	*3.069449*	*0.000857*
Positive regulation of metabolic process	*324*	*20*	*2.664452*	*0.000882*
Cellular component organization and biogenesis	*1639*	*62*	*1.63281*	*0.000909*
Membrane-bounded organelle	*4209*	*129*	*1.32292*	*0.000938*
Intracellular membrane-bounded organelle	*4207*	*129*	*1.323549*	*0.000968*
Positive regulation of cellular metabolic process	*312*	*20*	*2.766931*	*0.001*
Positive regulation of RNA metabolic process	*171*	*13*	*3.281483*	*0.003864*
Primary metabolic process	*4979*	*145*	*1.257039*	*0.003902*
RNA polymerase II transcription factor activity	*170*	*13*	*3.300786*	*0.003953*

doi: 10.1038/nature08162

Supplemental Table 12 continued: Sperm DNA demethylation extends beyond CpGs

GO CATEGORY	TOTAL GENES	CHANGED GENES	ENRICHMENT	FDR
Positive regulation of transcription DNA-dependent	*170*	*13*	*3.300786*	*0.003953*
Embryonic development ending in birth or egg hatching	*83*	*9*	*4.680447*	*0.004*
Chordate embryonic development	*83*	*9*	*4.680447*	*0.004*
Transcription from RNA polymerase II promoter	*450*	*24*	*2.302087*	*0.004211*
Positive regulation of transcription from RNA polymerase II promoter	*109*	*10*	*3.960011*	*0.004222*
Pattern specification process	*102*	*10*	*4.231777*	*0.004324*
Anatomical structure development	*1378*	*52*	*1.628835*	*0.004681*

doi: 10.1038/nature08162

Supplemental Table 13: Gene promoters occupied by Suz12 in ES cell are DNA demethylated and histone bound in sperm

GO CATEGORY	TOTAL GENES	CHANGED GENES	ENRICHMENT	FDR
Sequence-specific DNA binding	*425*	*51*	*10.377959*	*0*
Transcription factor activity	*755*	*65*	*7.445556*	*0*
Transcription regulator activity	*1090*	*69*	*5.471537*	*0*
Multicellular organismal development	*1620*	*41*	*4.407517*	*0*
DNA binding	*1522*	*72*	*4.348046*	*0*
Regulation of transcription DNA-dependent	*1467*	*72*	*4.112349*	*0*
Transcription DNA-dependent	*1510*	*37*	*4.267268*	*0*
RNA biosynthetic process	*1512*	*37*	*4.261624*	*0*
developmental process	*2265*	*84*	*3.38306*	*0*
Regulation of transcription	*1580*	*72*	*4.078212*	*0*
RNA metabolic process	*1827*	*39*	*3.717502*	*0*
Regulation of nucleobase nucleoside	*1623*	*37*	*3.970163*	*0*
Nucleotide and nucleic acid metabolic process				
Regulation of metabolic process	*1839*	*39*	*3.693245*	*0*
Transcription	*1644*	*72*	*3.78945*	*0*
Regulation of cellular metabolic process	*1772*	*38*	*3.734608*	*0*
Nucleic acid binding	*2348*	*42*	*3.115132*	*0*
Regulation of biological process	*3134*	*48*	*2.667273*	*0*
Regulation of cellular process	*2889*	*45*	*2.712627*	*0*
Multicellular organismal process	*2648*	*43*	*2.827976*	*0*
Biological regulation	*3396*	*48*	*2.461494*	*0*
Nucleobase nucleoside nucleotide and nucleic acid metabolic process	*2489*	*40*	*2.798725*	*0*
System development	*1231*	*58*	*4.07*	*0*
Anatomical structure development	*1465*	*36*	*3.776226*	*0*
Organ development	*869*	*44*	*4.408878*	*0*
Nucleus	*2828*	*39*	*2.401654*	*0*
Nervous system development	*553*	*16*	*5.038718*	*0*
Transcription from RNA polymerase II promoter	*460*	*13*	*4.92165*	*0*
Anatomical structure morphogenesis	*730*	*36*	*4.264915*	*0*
Cellular metabolic process	*5390*	*50*	*1.615498*	*0*
Skeletal development	*174*	*8*	*8.006928*	*0*
Primary metabolic process	*5420*	*50*	*1.606556*	*0*
Lung development	*32*	*4*	*21.768836*	*0.00027*
Neural tube patterning	*2*	*2*	*174.150685*	*0.000263*
Respiratory tube development	*33*	*4*	*21.109174*	*0.000256*
Cellular process	*8815*	*65*	*1.284151*	*0.000476*
Central nervous system development	*179*	*7*	*6.810362*	*0.000465*
Positive regulation of transcription from RNA polymerase II promoter	*80*	*5*	*10.884418*	*0.000455*
Cell differentiation	*1210*	*38*	*2.590671*	*0.001957*
Cellular developmental process	*1210*	*18*	*2.590671*	*0.001957*
Brain development	*93*	*11*	*10.2294*	*0.00383*
Positive regulation of transcription DNA-dependent	*147*	*6*	*7.108191*	*0.00551*
Metabolic process	*6020*	*50*	*1.446434*	*0.0054*

doi: 10.1038/nature08162

Supplemental Table 13: Gene promoters occupied by Suz12 in ES cell are DNA demethylated and histone bound in sperm

GO CATEGORY	TOTAL GENES	CHANGED GENES	ENRICHMENT	FDR
Regulation of transcription from RNA polymerase II promoter	*297*	*8*	*4.690928*	*0.006923*
Neuron fate specification	*5*	*2*	*69.660274*	*0.009811*
Pattern specification process	*111*	*15*	*7.844625*	*0.012321*
Kidney development	*29*	*3*	*18.015588*	*0.01614*
Tube development	*70*	*4*	*9.951468*	*0.019483*
Urogenital system development	*31*	*3*	*16.853292*	*0.020339*
Positive regulation of cellular metabolic process	*264*	*7*	*4.617632*	*0.020667*
positive regulation of transcription	*192*	*6*	*5.442209*	*0.020984*
Cell fate commitment	*74*	*7*	*8.13551*	*0.021452*
Embryonic limb morphogenesis	*34*	*3*	*15.366237*	*0.023231*
Embryonic appendage morphogenesis	*34*	*3*	*15.366237*	*0.023231*
Organ morphogenesis	*274*	*7*	*4.449105*	*0.022879*
Embryonic development	*199*	*36*	*15.250774*	*0.022206*
Positive regulation of nucleobase nucleoside Nunucleotide and nucleic acid metabolic process	*199*	*6*	*5.250774*	*0.022206*
Cell development	*859*	*24*	*2.41575*	*0.022899*
Positive regulation of metabolic process	*280*	*7*	*4.353767*	*0.022571*
Appendage morphogenesis	*36*	*4*	*14.512557*	*0.022055*
Limb morphogenesis	*36*	*4*	*14.512557*	*0.022055*

doi: 10.1038/nature08162

Supplemental Table 14: Promoters that acquire methylation in fibroblasts compared to sperm

GO CATEGORY	TOTAL GENES	CHANGED GENES	ENRICHMENT	FDR
Embryonic development	*199*	*43*	*2.393575*	*0*
Anatomical structure morphogenesis	*730*	*105*	*1.593302*	*0*
Regulation of transcription DNA-dependent	*1467*	*205*	*1.547945*	*0*
Transcription DNA-dependent	*1510*	*208*	*1.525872*	*0*
RNA biosynthetic process	*1512*	*208*	*1.523854*	*0*
Regulation of transcription	*1580*	*216*	*1.514357*	*0*
Transcription	*1644*	*220*	*1.482356*	*0*
RNA metabolic process	*1827*	*241*	*1.461202*	*0*
Multicellular organismal development	*1620*	*201*	*1.374399*	*0*
Tube development	*70*	*20*	*3.164927*	*0.000526*
Negative regulation of cell differentiation	*49*	*15*	*3.390993*	*0.001034*
Transcription from RNA polymerase II promoter	*460*	*69*	*1.661587*	*0.001071*
Negative regulation of developmental process	*60*	*17*	*3.138553*	*0.001111*
Cellular component organization and biogenesis	*1763*	*209*	*1.313184*	*0.00125*
Developmental process	*2265*	*259*	*1.266669*	*0.001304*
System development	*1231*	*155*	*1.394779*	*0.001364*
Anatomical structure development	*1465*	*180*	*1.361027*	*0.001429*
Regulation of cell differentiation	*105*	*23*	*2.426444*	*0.001935*
Embryonic morphogenesis	*78*	*19*	*2.698303*	*0.002*
Organ morphogenesis	*274*	*45*	*1.819255*	*0.0025*
Nervous system development	*553*	*77*	*1.542401*	*0.003529*
Lung development	*32*	*11*	*3.807803*	*0.003636*
Synapse organization and biogenesis	*23*	*9*	*4.334574*	*0.003889*
Respiratory tube development	*33*	*11*	*3.692415*	*0.004*
Nucleosome assembly	*41*	*12*	*3.24212*	*0.005405*
Chromosome organization and biogenesis	*236*	*38*	*1.783624*	*0.005417*
Tube morphogenesis	*43*	*12*	*3.091324*	*0.005532*
Calcium-dependent cell-cell adhesion	*20*	*8*	*4.430898*	*0.005641*
Formation of primary germ layer	*21*	*8*	*4.219903*	*0.005652*
Gastrulation	*31*	*10*	*3.573305*	*0.005682*
Regulation of developmental process	*162*	*29*	*1.982963*	*0.005714*
Organ development	*869*	*108*	*1.376689*	*0.005778*
Branching morphogenesis of a tube	*30*	*10*	*3.692415*	*0.005789*
Chromosome organization and biogenesis	*226*	*37*	*1.813531*	*0.005814*
Chromatin assembly or disassembly	*87*	*19*	*2.419168*	*0.005854*
Macromolecular complex assembly	*359*	*53*	*1.635359*	*0.006*
Pattern specification process	*111*	*22*	*2.19549*	*0.006327*
Morphogenesis of a branching structure	*32*	*10*	*3.461639*	*0.0064*
Mesoderm morphogenesis	*22*	*8*	*4.028089*	*0.008039*
Protein-DNA complex assembly	*86*	*18*	*2.318493*	*0.008393*
Cellular component assembly	*389*	*55*	*1.566191*	*0.008462*
Positive regulation of cell differentiation	*28*	*9*	*3.560543*	*0.008491*
Regionalization	*72*	*16*	*2.46161*	*0.008545*

doi: 10.1038/nature08162

Supplemental Table 14: Promoters that acquire methylation in fibroblasts compared to sperm

GO CATEGORY	TOTAL GENES	CHANGED GENES	ENRICHMENT	FDR
Chromatin assembly	*52*	*13*	*2.769311*	*0.008704*
Synaptogenesis	*18*	*7*	*4.307817*	*0.009298*
DNA packaging	*180*	*30*	*1.846207*	*0.009483*
Embryonic arm morphogenesis	*3*	*3*	*11.077244*	*0.012623*
Arm morphogenesis	*3*	*3*	*11.077244*	*0.012623*
Positive regulation of osteoblast differentiation	*3*	*3*	*11.077244*	*0.012623*
Heart development	*75*	*16*	*2.363145*	*0.015625*
Response to hypoxia	*19*	*7*	*4.08109*	*0.016*
Chordate embryonic development	*69*	*15*	*2.408097*	*0.016866*
Anatomical structure formation	*122*	*22*	*1.997536*	*0.018169*
Mesoderm formation	*20*	*7*	*3.877035*	*0.018378*
Embryonic development ending in birth or egg hatching	*70*	*15*	*2.373695*	*0.018429*
Mitotic sister chromatid segregation	*15*	*6*	*4.430898*	*0.018933*
Sensory organ development	*57*	*13*	*2.526389*	*0.019211*
Mesoderm development	*44*	*11*	*2.769311*	*0.019221*
Cell cycle phase	*214*	*33*	*1.708173*	*0.019367*
Cell differentiation	*1210*	*138*	*1.263355*	*0.019634*
Cell cycle	*606*	*76*	*1.389225*	*0.020814*
Sister chromatid segregation	*16*	*6*	*4.153967*	*0.025057*
Cell fate determination	*27*	*8*	*3.282146*	*0.025455*
Protein catabolic process	*185*	*29*	*1.736433*	*0.025495*
Regulation of transcription from RNA polymerase II promoter	*297*	*42*	*1.566479*	*0.025556*
Cell fate commitment	*74*	*15*	*2.245387*	*0.02573*
Macromolecule catabolic process	*326*	*45*	*1.529067*	*0.025895*
Embryonic limb morphogenesis	*34*	*9*	*2.932212*	*0.025957*
Embryonic appendage morphogenesis	*34*	*9*	*2.932212*	*0.025957*
Dorsal ventral pattern formation	*22*	*7*	*3.524578*	*0.025978*
Regulation of gliogenesis	*4*	*3*	*8.307933*	*0.033469*
RNA interference	*4*	*3*	*8.307933*	*0.033469*
Regulation of glial cell differentiation	*4*	*3*	*8.307933*	*0.033469*
Tissue morphogenesis	*55*	*12*	*2.416853*	*0.0372*
Mesodermal cell fate commitment	*8*	*4*	*5.538622*	*0.04181*
Mitotic chromosome condensation	*8*	*4*	*5.538622*	*0.04181*
Pancreas development	*8*	*4*	*5.538622*	*0.04181*
Mesodermal cell differentiation	*8*	*4*	*5.538622*	*0.04181*
Appendage morphogenesis	*36*	*9*	*2.769311*	*0.042364*
Cell cycle process	*530*	*66*	*1.37943*	*0.044123*
Chromosome condensation	*13*	*5*	*4.260479*	*0.044554*
Positive regulation of developmental process	*43*	*10*	*2.576103*	*0.044696*
Anterior posterior pattern formation	*44*	*10*	*2.517556*	*0.048103*
Sex differentiation	*66*	*13*	*2.181881*	*0.050168*
Embryonic pattern specification	*25*	*7*	*3.101628*	*0.050339*

CHAPTER 6

ALTERATIONS IN SPERM DNA METHYLATION PATTERNS AT IMPRINTED LOCI IN TWO CLASSES OF INFERTILITY

Published in *Fertility and Sterility*

Authors

Saher Sue Hammoud

Jahnvi Purwar

Christian Pflueger

Bradley R. Cairns

Douglas T. Carrell

Alterations in sperm DNA methylation patterns at imprinted loci in two classes of infertility

Saher Sue Hammoud, B.S.,[a,b] *Jahnvi Purwar, B.S.,*[d] *Christian Pflueger, M.S.,*[d] *Bradley R. Cairns, Ph.D.,*[d,e] *and Douglas T. Carrell, Ph.D.*[a,b,c]

[a]Andrology and IVF Laboratories, Division of Urology, Department of Surgery, [b]Department of Physiology, and [c]Department of Obstetrics and Gynecology, School of Medicine, and [d]Department of Oncological Sciences, Huntsman Cancer Institute, University of Utah; and [e]Howard Hughes Medical Institute, University of Utah School of Medicine, Salt Lake City, Utah

Objective: To evaluate the associations between proper protamine incorporation and DNA methylation at imprinted loci.
Design: Experimental research study.
Setting: Research laboratory.
Patient(s): Three populations were tested—abnormal protamine patients, oligozoospermic patients, and fertile donors.
Intervention(s): The CpG methylation patterns were examined at seven imprinted loci sequenced: *LIT1, MEST, SNRPN, PLAGL1, PEG3, H19,* and *IGF2.*
Main Outcome Measure(s): The DNA methylation patterns were analyzed using bisulfite sequencing. The percentage of methylation was compared between fertile and infertile patients displaying abnormal protamination.
Result(s): At six of the seven imprinted genes, the overall DNA methylation patterns at their respective differentially methylated regions were significantly altered in both infertile patient populations. When comparing the severity of methylation alterations among infertile patients, the oligozoospermic patients were significantly affected at mesoderm-specific transcript (*MEST),* whereas abnormal protamine patients were affected at *KCNQ1,* overlapping transcript 1 *(LIT1),* and at small nuclear ribonucleoprotein polypeptide N *(SNRPN).*
Conclusion(s): Patients with male factor infertility had significantly increased methylation alteration at six of seven imprinted loci tested, with differences in significance observed between oligozoospermic and abnormal protamine patients. This could suggest that risk of transmission of epigenetic alterations may be different with diagnoses. However, this study does not provide a causal link for epigenetic inheritance of imprinting diseases, but does show significant association between male factor infertility and alterations in sperm DNA methylation at imprinted loci. (Fertil Steril® 2010;94:1728–33. ©2010 by American Society for Reproductive Medicine.)

Key Words: Imprinting, Beckwith-Wiedemann syndrome and epigenetic alterations, Angelman syndrome, chromatin, assisted reproductive technology, IVF, ICSI, oligozoospermic, protamines

Genomic imprinting is established and inherited during gametogenesis and preimplantation to ensure parent-of-origin monoallelic gene expression (1, 2). The mechanism by which either one of the two alleles are differentially expressed is not completely understood; however, it is known that the majority of imprinted genes are clustered and are predominately regulated by imprinting control regions (ICRs) (3, 4). At present, approximately 80 imprinted genes have been identified, many of which are implicated in tumorigenesis, fetal growth regulation, and embryonic development (5–8). Pathological perturbation in the methylation imprints during gametogenesis or development can give rise to growth-related syndromes and is frequently observed in cancer (9–20).

After fertilization, both parental genomes are globally demethylated through active or passive demethylation mechanisms, whereas the methylation patterns at imprinted genes are maintained and only erased and re-established in the primordial germ cell. The presence of abnormal methylation patterns residing in gametes raises concerns, as these may be inherited and maintained in the embryo. Meta-analysis showed that children born from assisted reproductive technology (ART) have a fourfold increased incidence of Beckwith-Weidemann syndrome compared with children conceived naturally (21–24). In addition, imprinting syndromes such as Angelman, Prader-Willi, and Silver-Russell have been associated with ART, although no strong correlations were established. Currently, it is unclear whether imprinting abnormalities arise from the ART procedure itself or from pre-existing methylation aberrations in the gametes of infertile patients (25–27).

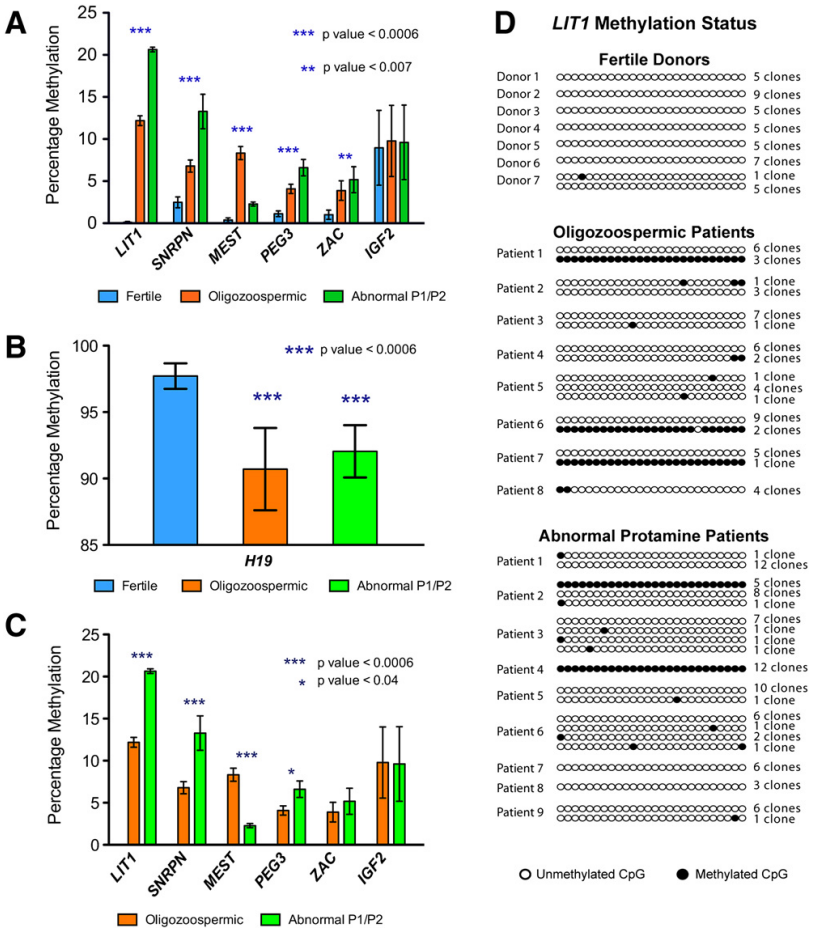
Recent studies have shown that epigenetic abnormalities are common in the sperm of severely oligozoospermic patients, favoring the latter hypothesis (26, 27). Whether epigenetic alterations at imprinted loci of infertile men are limited to oligozoospermic patients or whether epigenetic alterations extend beyond oligozoospermic patients is unknown. In this study we examine methylation changes in patients with an alternative cause for their male factor infertility—patients with abnormal sperm protamine replacement of histones. Protamines 1 and 2 are sperm-specific nuclear proteins that are incorporated into the DNA in a 1:1 ratio and ensure chromatin

Received April 8, 2009; revised August 5, 2009; accepted September 8, 2009; published online November 1, 2009.
S.S.H. has nothing to disclose. J.P. has nothing to disclose. C.P. has nothing to disclose. B.R.C. has nothing to disclose. D.T.C. has nothing to disclose.
Saher Sue Hammoud and Jahnvi Purwar contributed equally.
Reprint requests: Douglas T. Carrell, Ph.D., University of Utah IVF and Andrology Laboratories, 675 Arapeen Drive, Suite 205, Salt Lake City, UT 84108 (FAX: 801-581-6127; E-mail: douglas.carrell@hsc.utah.edu).

0015-0282/$36.00
doi:10.1016/j.fertnstert.2009.09.010

FIGURE 1

The overall methylation patterns at both paternally and maternally imprinted genes were altered in the sperm of infertile patients. (**A,B,C**) The mean percentage of methylation with standard error. $P < .05$ is significant. (**A**) The percentage of methylated CpGs at normally paternally demethylated loci. (**B**) The percentage of demethylation at a paternally methylated DMR of *H19*. (**C**) Comparing methylation changes between the two infertile patient populations. (**D**) Methylation status at the differentially methylated region of *LIT1* for fertile donors, oligozoospermic patients, and abnormal protamine patients.



Hammoud. Imprinting abnormalities in infertile men. Fertil Steril 2010.

condensation. The average P1:P2 ratio in fertile men is ~1, whereas in some infertile patients this ratio is significantly altered (28, 29) and consequently associated with severe sperm defects that can usually be addressed through ART (30, 31). It has been proposed that chromatin packaging may have a role in properly establishing and maintaining methylation patterns, hence, hypothetically, patients with abnormal protamine ratios may be at an increased risk of conceiving an ART offspring with imprinting disease (32, 33). This study evaluates the relationship between protamine ratios and methylation patterns at seven imprinted loci in the sperm of abnormal protamine patients or oligozoospermic patients. We reveal significant changes in the overall DNA methylation patterns at six of these loci, with varying impact on methylation patterns within each class of infertility: oligozoospermic or abnormal protamine levels (p-value < 0.05, Figure 1). These data suggest that aberrant imprinting patterns are observed in patients with abnormal protamine ratios, and that the abnormal patterns may vary among different pathologies, providing a spectrum of risks for transmitting epigenetic abnormalities to the embryo.

MATERIALS AND METHODS

Patient Population

Of the seven tested imprinted loci, six are paternally demethylated and expressed: KCNQ1 overlapping transcript 1 (*LIT1*), insulin-like growth factor 2 (*IGF2*), paternally expressed gene 3 (*PEG3*), pleiomorphic adenoma gene-like 1 (*PLAGL1* also known as *ZAC*), small nuclear ribonucleoprotein

polypeptide N *(SNRPN),* and mesoderm-specific trancript (*MEST*), and one is maternally expressed and is normally DNA methylated in sperm (*H19*). For each locus 10 oligozoospermic (sperm count $\leq 10 \times 10^6$/mL), 10 abnormal protamine replacement patients (average sperm count of $73 \times 10^6 \pm 60$ SD/mL), and 5 known fertile donors were evaluated. For *LIT1* only, eight oligozoospermic patients and nine abnormal protamine patients were evaluated.

Sample Collection and Bisulfite Treatment

Institutional Review Board (IRB) approval was obtained before initiation of this study. Frozen sperm DNA samples were treated with sodium bisulfite to convert unmethylated cytosines to uracil and leaving methylated cytosines unchanged, as previously described by Clark et al. (34). DNA was purified using Qiagen DNeasy clean up kit (Qiagen, Valencia, CA) and eluted twice, each time with 100 μL of elution buffer. The purified DNA was desulfonated by the addition of 20 μL NaOH and incubated at 37°C for 15 minutes. After incubation, 22 μL of 4 M NaOAC, glycogen, and two volumes of ethanol were added to precipitate the DNA overnight at -20°C. Precipitated DNA was washed twice with 70% ethanol and eluted in 30 μL of elution buffer.

PCR Amplification of Bisulfite Converted DNA

Primer sequences and temperatures for *SNRPN, PEG3, ZAC, MEST, LIT1, H19 ICR*, and *IGF2* are available upon request (35, 36). The polymerase chain reaction (PCR) reactions were performed in 50-μL volume reactions containing 5 μL of 10 × PCR buffer–$MgCl_2$ (Invitrogen, Carlsbad, CA), 5 μL of 10 × Enhancer Buffer (Invitrogen), 1.5 μL of $MgCl_2$, 1 μL of 10 mM dNTPs, 0.5 μL of *Taq* (Invitrogen), 2.5 μL of each forward and reverse primer (10 μM stock), and 30 μL of water. The PCR results were analyzed on a 1% agarose gel, and gel purified if multiple products were detected.

TOPO TA Cloning and Sequencing

The PCR products were cloned into a TOPO 2.1 pCR vectors (Invitrogen) and plated onto KAN-X-GAL plates for blue-white screening. Positive colonies were reinoculated into LB-KAN (50 μg/mL), cultured overnight, and plasmids were purified using the Qiagen 96-well clean-up kit. To address sperm sample heterogeneity five or more clones/alleles were sequenced per patient for each of the imprinted loci (sequencing done at Genewiz San Diego Laboratory).

Data Visualization and Analysis

The CG/TG-analyzer, a Perl program, was used to examine the methylation status of a bisulfite-converted sequence and provides an output in the form of 1s and 0s, where 1s represent methylated cytosines and 0s represent unmethylated cytosines (thymine). The CpG positions were defined in a multifasta file, text-based file containing multiple DNA or protein sequences, which includes the CpG position number flanked by four nucleotides on each side. The output was used to calculate the percentage of CpG methylation (program is be available upon request). To compare the overall methylation profile in infertile patients versus fertile donors (Fig. 1), the Wilcoxon-Mann-Whitney test was used. This test is a nonparametric significance test for assessing whether two independent samples of observations came from the same distribution. To determine significance between fertile donors and oligozoospermic patients or fertile and abnormal protamine patients the percentage of methylated CpGs represented in columns 2 and 3 (in Tables 1, 2, and 3) were compared as independent sample populations. A P value $< .05$ was considered significant. The χ^2 analysis was used to compare the percentage of methylated CpGs in the abnormal protamine or oligozoospermic patients with known fertile donors.

RESULTS

Six imprinted genes, that are normally paternally demethylated, were examined: *LIT1, SNRPN, MEST, ZAC, PEG3,* and *IGF2.* Here, all except *IGF2,* showed significant hypermethylation in oligozoospermic and abnormal protamine patients compared with fertile donors (Fig. 1A). Furthermore, the differentially methylated region (DMR) of *H19* (a paternally methylated locus) was

TABLE 1

The percentage of methylated CpGs in the DMR of *LIT1* of oligozoospermic and abnormal protamine patients.

CpG	Abnormal P1/P2 (n = 9)	Oligozoospermic (n = 8)	Fertile donors (n = 7)	Fertile vs. abnormal	Fertile vs. oligozoospermic
CpG 1	25.882	18.181	0	0.0003	0.0035
CpG 2	20	18.181	0	0.0021	0.0035
CpG 3	20	10.909	0	0.0021	0.0271
CpG 4	20	10.909	2.38	0.0066	0.17
CpG 5	21.176	10.909	0	0.0015	0.0271
CpG 6	20	10.909	0	0.0021	0.0271
CpG 7	21.176	10.909	0	0.0015	0.0271
CpG 8	20	10.909	0	0.0021	0.0271
CpG 9	20	10.909	0	0.0021	0.0271
CpG 10	20	10.909	0	0.0021	0.0271
CpG 11	21.176	12.277	0	0.0015	0.0186
CpG 12	20	10.909	0	0.0021	0.0271
CpG 13	21.176	10.909	0	0.0015	0.0271
CpG 14	20	14.454	0	0.0021	0.0101
CpG 15	20	10.909	0	0.0021	0.0271
CpG 16	20	7.272	0	0.0021	0.0742
CpG 17	20	10.909	0	0.0021	0.0271
CpG 18	21.176	12.272	0	0.0015	0.0093
CpG 19	20	10.909	0	0.0021	0.0271
CpG 20	20	10.909	0	0.0021	0.0271
CpG 21	21.176	16.363	0	0.0015	0.0059
CpG 22	21.176	16.363	0	0.0015	0.0059

Note: DMR = differentially methylated region.

Hammoud. Imprinting abnormalities in infertile men. Fertil Steril 2010.

TABLE 2

The percentage of methylated CpG in the DMR of *SNRPN*.

CpGs	Abnormal P1/P2 (n = 11)	Oligozoospermic (n = 13)	Fertile donors (n = 5)	Fertile vs. abnormal	Fertile vs. oligozoospermic
CpG 1	4.3	4.0	0	0.152	0.169
CpG 2	5.8	5.0	0	0.09	0.123
CpG 3	5.8	5.0	0	0.09	0.123
CpG 4	5.8	5.0	0	0.09	0.123
CpG 5	5.8	5.0	0	0.09	0.123
CpG 6	10.6	5.0	0	0.026	0.123
CpG 7	14.8	8.0	4.3	0.08	0.413
CpG 8	8.7	5.0	0	0.04	0.123
CpG 9	13.0	5.0	4.3	0.10	0.864
CpG 10	23.1	16	6.5	0.05	0.114
CpG 11	10.1	6.0	0	0.026	0.09
CpG 12	11.6	6.1	8.7	0.618	0.526
CpG 13	15.9	8.0	6.5	0.1	0.753
CpG 14	47.8	10	2.2	0.0001	0.09
CpG 15	11.6	6.1	0	0.017	0.08
CpG 16	5.8	4.0	2.2	0.351	0.566
CpG 17	11.6	13	2.2	0.065	0.039
CpG 18	15.9	12.2	6.5	0.130	0.295
CpG 19	15.9	5.1	6.7	0.140	0.705
CpG 20	17.4	5.1	0	0.003	0.119
CpG 21	17.6	4.1	2.2	0.011	0.560

Note: DMR = differentially methylated region.

Hammoud. Imprinting abnormalities in infertile men. Fertil Steril 2010.

significantly hypomethylated in both infertile classes ($P<.006$ for all except *ZAC*, $P<.002$) (Fig. 1B). Thus, these infertile patients show methylation alterations at six of seven loci tested. However, when comparing overall methylation changes between the two infertile populations, abnormal protamine patients show more extensive hypermethylation at the DMRs of *LIT1* and *SNRPN* in comparison with oligozoospermic patients. In contrast, hypermethylation at *MEST* is significantly higher in oligozoospermic patients (p-value < 0.006, Fig. 1C).

Notably, in both patient populations, the locus that displays the highest number of affected CpGs is *LIT1*. In the DMR of *LIT1*, the percentage of methylated CpGs ranged from 7%–18% or 20%–25% for oligozoospermic or abnormal protamine patients, respectively (Table 1). In contrast, for fertile donors, virtually all CpGs were demethylated. The percentages of methylated CpGs in oligozoospermic and abnormal protamine patients were statistically significant when compared with fertile donors (p-value < 0.05, Table 1). To address the uniformity of methylation changes at *LIT1* in individual sperm from a single patient, we sequenced multiple alleles (5–12) from each patient, and found striking heterogeneity. In three of the eight oligozoospermic patients, *LIT1* was completely methylated in 20%–30% of the alleles, whereas in the other five patients, only sporadic increases were observed (Fig. 1D). Similarly, in the abnormal protamine category one patient always displayed complete methylation, a second displayed methylation on 50% of his alleles, and the remainder (seven) displayed little or no increase.

Consistent with the findings reported previously, the DMR of *SNRPN* was also susceptible to acquiring methylation in infertile men. Abnormal protamine patients had a significant increase in CpG methylation (methylation at individual CpGs typically ranged from 4%–20%) (p-value < 0.05 Table 2). Alterations were also observed in oligozoospermic patients (range of methylation, 4%–8%), but the increase lacked statistical significance (Table 2). At *SNRPN*, alterations in methylation were common (observed at a majority of the alleles) but typically involved only a moderate number of CpGs acquiring methylation. However, in both patient categories, a small number of patients displayed complete methylation at 10% of the alleles tested.

Methylation levels in the DMR of *MEST* (for each CpG) ranged from 7%–19% or 1%–3% in oligozoospermic or abnormal protamine patients, respectively (Table 3). The changes in methylation at many of the CpGs in oligozoospermic patients were near the range of statistical significance ($P=.07$; Table 3). In addition, 3 of 10 oligozoospermic patients had 12%–33% of their alleles completely methylated, whereas the remaining 7 patients displayed very little change. Likewise, in the abnormal protamine class, one patient had 14% of his alleles completely methylated and in the remaining nine patients, there was virtually no change observed. In contrast, very few individual CpGs were significantly ($P<.05$) affected in *PEG3, ZAC*, *IGF2* promoter 3, and *H19* in infertile patients (data not shown).

DISCUSSION

In this study we evaluated the methylation status of seven imprinted loci in two patient populations: oligozoospermic and abnormal protamine ratio patients. The overall methylation patterns in sperm of infertile patients were significantly altered at all imprinted loci (except *IGF2*) when compared with fertile donors. However, when comparing the two infertile patient populations, oligozoospermic patients were hypermethylated at *MEST*, an imprinted gene associated with Silver-Russell syndrome, whereas abnormal protamine patients had significant changes at *LIT1* and *SNRPN* (Figure 1), genes that may be associated with cases of transient neonatal diabetes mellitus and Angelman syndrome. These data suggest that risk of transmission of epigenetic alterations may be different with diagnoses.

TABLE 3

The percentage of methylated CpGs at the DMR of *MEST* in oligozoospermic and abnormal protamine patients.

CpG	Abnormal P1/P2 (n = 10)	Oligozoospermic (n = 10)	Fertile donors (n = 5)	Fertile vs. abnormal	Fertile vs. oligozoospermic
CpG 1	1.785	14.28	0	0.2346	0.0167
CpG 2	1.785	19.04	0	0.2346	0.0063
CpG 3	3.571	7.1428	0	0.1515	0.070
CpG 4	3.571	7.142	3.4	0.483	0.250
CpG 5	1.785	7.1428	0	0.2346	0.070
CpG 6	1.785	9.5238	0	0.2346	0.436
CpG 7	3.571	7.1428	0	0.1515	0.070
CpG 8	1.785	7.1428	0	0.2346	0.070
CpG 9	1.785	7.1428	0	0.2346	0.070
CpG 10	1.785	7.1428	0	0.2346	0.070
CpG 11	1.785	7.1428	0	0.2346	0.070
CpG 12	3.571	7.1428	0	0.1515	0.070
CpG 13	1.785	9.523	3.4	0.642	0.1604
CpG 14	3.571	4.7619	0	0.1515	0.1167
CpG 15	1.785	7.1428	0	0.2346	0.070
CpG 16	3.571	7.1428	0	0.1515	0.070
CpG 17	1.785	7.1428	0	0.2346	0.070
CpG 18	0	7.1428	0	NA	0.070

Note: DMR = differentially methylated region.

Hammoud. Imprinting abnormalities in infertile men. Fertil Steril 2010.

Our data evaluate and demonstrate abnormal imprinting in a different class of abnormal spermatogenesis, abnormal replacement of nuclear proteins by protamine 1 and protamine 2. It was our hypothesis that abnormal chromatin packaging may be associated with methylation defects, which is supported by the data presented from this study. These data, along with previously published data from oligozoospermic patients, reveal that alteration in DNA methylation patterns are common at a handful of imprinted loci tested, suggesting that imprinting abnormalities may reside in the sperm of infertile patients (25–27), but whether these alterations can be inherited is uncertain. Remarkably, when examining normally demethylated DMRs, the alleles of infertile patients are often either unaffected or entirely methylated, suggesting a bistable status, and a susceptibility to complete methylation. Clearly, complete methylation of a normally unmethylated locus may lead to an imprinting disorder in the embryo if proper imprint reestablishment mechanisms are not implemented. Also of note are the small differences in the degree of methylation within some genes and alleles. It is important to determine whether this abnormal methylation has reached a threshold level that might lead to complete methylation in the embryo (at a certain unknown probability) and confer disease, or whether there is a gradual continuum with a threshold for disease.

Whether imprinting diseases in ART offspring arise as a result of abnormal methylation of gametes, or acquire methylation changes during in vitro culture, or both, is still unknown. Current human data suggest that methylation alteration at imprinted loci may reside in gametes and may be inherited by the embryo. Supporting evidence comes from two reports showing that a gain in methylation on the paternal alleles of *LIT1* or *MEST* in sperm is maintained in the baby and associated with transient neonatal diabetes (37) or Silver-Russell syndrome (38). The findings suggest that paternal imprints in sperm may be needed for a healthy and uncomplicated pregnancy. The need to study sperm from fathers of children with imprinting diseases is imperative.

This study does not report a causal link between abnormal methylation of imprinted genes and disease. The relative risk of the defects reported in our study to patients is unknown. However, we demonstrate a link between abnormal spermatogenesis and abnormal methylation of genes associated with rare imprinting diseases previously reported to have elevated incidences in ART offspring (21–24). This suggests that such a link may be strengthened in infertile men with known abnormalities in chromatin packaging. Characterizing these epigenetic alterations in the sperm of infertile men may help predict the likelihood of IVF success rate.

REFERENCES

1. Morison IM, Paton CJ, Cleverley SD. The imprinted gene and parent-of-origin effect database. Nucleic Acids Res 2001;29:275–6.
2. Ferguson-Smith AC, Surani MA. Imprinting and the epigenetic asymmetry between parental genomes. Science 2001;293:1086–9.
3. Verona RI, Mann MR, Bartolomei MS. Genomic imprinting: intricacies of epigenetic regulation in clusters. Annu Rev Cell Dev Biol 2003;19:237–59.
4. Spahn L, Barlow DP. An ICE pattern crystallizes. Nat Genet 2003;35:11–2.
5. Morgan HD, Santos F, Green K, Dean W, Reik W. Epigenetic reprogramming in mammals. Hum Mol Genet 2005. 14 Spec No 1: R47–R58.
6. Jaenisch R, Bird A. Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. Nat Genet 2003;33. Suppl: 245–54.
7. Royo H, Bortolin ML, Seitz H, Cavaille J. Small noncoding RNAs and genomic imprinting. Cytogenet Genome Res 2006;113:99–108.
8. Seitz H, Royo H, Bortolin ML, Lin SP, Ferguson-Smith AC, Cavaille J. A large imprinted microRNA gene cluster at the mouse Dlk1-Gtl2 domain. Genome Res 2004;14:1741–8.
9. Bjornsson HT, Brown LJ, Fallin MD, Rongione MA, Bibikova M, Wickham E, et al. Epigenetic specificity of loss of imprinting of the IGF2 gene in Wilms tumors. J Natl Cancer Inst 2007;99:1270–3.
10. Chao W, D'Amore PA. IGF2: epigenetic regulation and role in development and disease. Cytokine Growth Factor Rev 2008;19:111–20.

11. Cui H. Loss of imprinting of IGF2 as an epigenetic marker for the risk of human cancer. Dis Markers 2007;23:105–12.
12. Hemberger M. Epigenetic landscape required for placental development. Cell Mol Life Sci 2007;64: 2422–36.
13. Henckel A, Feil R. Differential epigenetic marking on imprinted genes and consequences in human diseases. Med Sci (Paris) 2008;24:747–52.
14. Lemeta S, Jarmalaite S, Pylkkanen L, Bohling T, Husgafvel-Pursiainen K. Preferential loss of the nonimprinted allele for the ZAC1 tumor suppressor gene in human capillary hemangioblastoma. J Neuropathol Exp Neurol 2007;66:860–7.
15. Murrell A, Ito Y, Verde G, Huddleston J, Woodfine K, Silengo MC, et al. Distinct methylation changes at the IGF2-H19 locus in congenital growth disorders and cancer. PLoS ONE 2008;3:e1849.
16. Nakano S, Murakami K, Meguro M, Soejima H, Higashimoto K, Urano T, et al. Expression profile of LIT1/KCNQ1OT1 and epigenetic status at the KvDMR1 in colorectal cancers. Cancer Sci 2006;97:1147–54.
17. Nowaczyk MJ, Carter MT, Xu J, Huggins M, Raca G, Das S, et al. Paternal deletion 6q24.3: a new congenital anomaly syndrome associated with intrauterine growth failure, early developmental delay and characteristic facial appearance. Am J Med Genet A 2008;146:354–60.
18. Shin JY, Fitzpatrick GV, Higgins MJ. Two distinct mechanisms of silencing by the KvDMR1 imprinting control region. EMBO J 2008;27:168–78.
19. Valleley EM, Cordery SF, Bonthron DT. Tissue-specific imprinting of the ZAC/PLAGL1 tumour suppressor gene results from variable utilization of monoallelic and biallelic promoters. Hum Mol Genet 2007;16:972–81.
20. Yoshimizu T, Miroglio A, Ripoche MA, Gabory A, Vernucci M, Riccio A, et al. The H19 locus acts in vivo as a tumor suppressor. Proc Natl Acad Sci U S A 2008;105:12417–22.
21. Cox GF, Burger J, Lip V, Mau UA, Sperling K, Wu BL, et al. Intracytoplasmic sperm injection may increase the risk of imprinting defects. Am J Hum Genet 2002;71:162–4.
22. Niemitz EL, DeBaun MR, Fallon J, Murakami K, Kugoh H, Oshimura M, et al. Microdeletion of LIT1 in familial Beckwith-Wiedemann syndrome. Am J Hum Genet 2004;75:844–9.
23. DeBaun MR, Niemitz EL, Feinberg AP. Association of in vitro fertilization with Beckwith-Wiedemann syndrome and epigenetic alterations of LIT1 and H19. Am J Hum Genet 2003;72:156–60.
24. Gosden R, Trasler J, Lucifero D, Faddy M. Rare congenital disorders, imprinted genes, and assisted reproductive technology. Lancet 2003;361: 1975–7.
25. Kobayashi H, Sato A, Otsu E, Hiura H, Tomatsu C, Utsunomiya T, et al. Aberrant DNA methylation of imprinted loci in sperm from oligospermic patients. Hum Mol Genet 2007;16:2542–51.
26. Marques CJ, Carvalho F, Sousa M, Barros A. Genomic imprinting in disruptive spermatogenesis. Lancet 2004;363:1700–2.
27. Marques CJ, Costa P, Vaz B, Carvalho F, Fernandes S, Barros A, et al. Abnormal methylation of imprinted genes in human sperm is associated with oligozoospermia. Mol Hum Reprod 2008;14:67–74.
28. Aoki VW, Emery BR, Liu L, Carrell DT. Protamine levels vary between individual sperm cells of infertile human males and correlate with viability and DNA integrity. J Androl 2006;27:890–8.
29. Balhorn R, Reed S, Tanphaichitr N. Aberrant protamine 1/protamine 2 ratios in sperm of infertile human males. Experientia 1988;44:52–5.
30. Aoki VW, Moskovtsev SI, Willis J, Liu L, Mullen JB, Carrell DT. DNA integrity is compromised in protamine-deficient human sperm. J Androl 2005;26: 741–8.
31. Aoki VW, Liu L, Jones KP, Hatasaka HH, Gibson M, Peterson CM, et al. Sperm protamine 1/protamine 2 ratios are related to in vitro fertilization pregnancy rates and predictive of fertilization ability. Fertil Steril 2006;86:1408–15.
32. Paldi A. Genomic imprinting: could the chromatin structure be the driving force? Curr Top Dev Biol 2003;53:115–38.
33. Rousseaux S, Caron C, Govin J, Lestrat C, Faure AK, Khochbin S. Establishment of male-specific epigenetic information. Gene 2005;345:139–53.
34. Clark SJ, Harrison J, Paul CL, Frommer M. High sensitivity mapping of methylated cytosines. Nucleic Acids Res 1994;22:2990–7.
35. El-Maarri O, Seoud M, Coullin P, Herbiniaux U, Oldenburg J, Rouleau G, et al. Maternal alleles acquiring paternal methylation patterns in biparental complete hydatidiform moles. Hum Mol Genet 2003;12:1405–13.
36. Kamikihara T, Arima T, Kato K, Matsuda T, Kato H, Douchi T, et al. Epigenetic silencing of the imprinted gene ZAC by DNA methylation is an early event in the progression of human ovarian cancer. Int J Cancer 2005;115:690–700.
37. Arima T, Kamikihara T, Hayashida T, Kato K, Inoue T, Shirayoshi Y, et al. ZAC, LIT1 (KCNQ1OT1) and p57KIP2 (CDKN1C) are in an imprinted gene network that may play a role in Beckwith-Wiedemann syndrome. Nucleic Acids Res 2005;33:2650–60.
38. Kagami M, Nagai T, Fukami M, Yamazawa K, Ogata T. Silver-Russell syndrome in a girl born after in vitro fertilization: partial hypermethylation at the differentially methylated region of PEG1/MEST. J Assist Reprod Genet 2007;24: 131–6.

CHAPTER 7

GENOMEWIDE ANALYSIS IDENTIFIES CHANGES IN HISTONE RETENTION AND EPIGENETIC MODIFICATIONS AT DEVELOPMENTAL AND IMPRINTED GENE LOCI IN THE SPERM OF INFERTILE MEN

Submitted for Consideration of Publication in *Human Reproduction*

Authors

Saher Sue Hammoud

David A. Nix

Ahmad O. Hammoud

Mark Gibson MD

Bradley R. Cairns PhD

Douglas T. Carrell

Abstract

Background

The sperm chromatin of fertile men retains a small number of nucleosomes that are enriched at developmental gene promoters and imprinted gene loci. This unique chromatin packaging at certain gene promoters provides these genomic loci the ability to convey instructive epigenetic information to the zygote, potentially expanding the role and significance of the sperm epigenome in embryogenesis. In this study, we hypothesize that changes in chromatin packaging may be associated with poor reproductive outcome.

Methods

Seven patients with reproductive dysfunction were recruited: three had unexplained poor embryogenesis during IVF and four were diagnosed with male infertility and previously shown to have altered protamination. Genome-wide analysis of the location of histones and histone modifications was analyzed by isolation and purification of DNA bound to histones and protamines. The histone bound fraction of DNA was analyzed using high throughput sequencing, both initially and following chromatin immunoprecipation. The protamine-bound fraction was hybridized to Agilent arrays. DNA methylation was examined using bisulfite sequencing.

Results

Unlike fertile men, six of seven infertile men had non-programmatic (randomly distributed) histone retention genomewide. Interestingly, in contrast to the total histone pool, the localization of H3K4me or H3K27me was mostly similar in the gametes of

infertile men compared to fertile men. However, there was a reduction in the amount of H3K4me or H3K27me retained at developmental transcription factors and certain imprinted genes. Finally, the methylation status of candidate developmental promoters and imprinted loci, were altered in a subset of the infertile men.

Conclusion

This initial genome-wide analysis of epigenetic markings in the sperm of infertile men demonstrates differences in composition and epigenetic markings compared to fertile men, especially at certain imprinted and developmental loci. Although no single locus displays a complete change in chromatin packaging or DNA modification, the data suggest that moderate changes throughout the genome exist and may have a cumulative detrimental affect on fecundity.

Introduction

The incidence of infertility is rising and estimated to affect one of six couples. As a result, the use of assisted reproductive technologies (ART) is rising and it is now estimated that ART accounts for 1-2% of all births in developed countries.[1] Although ART has become a widely accepted and implemented therapy for many forms of infertility, there have been concerns about safety of ART for resulting offspring. Recently, a series of reports raised the concern that ART is associated with an increased incidence of major congenital malformations, chromosomal aberrations, miscarriage rates, intrauterine growth restriction, early childhood cancers, and imprinting disorders (such as Angelmann's syndrome, Beckwith-Weidmann's syndrome, and

Retinoblastoma).[2,3,4] These reports suggest that many of the effects of ART on offspring are genetic or epigenetic in nature.[5,6]

The precedence for epigenetic abnormalities in ART stems from earlier work on animal models which suggests an increased incidence of imprinting errors in offspring conceived by ART compared to natural conception.[7] This rise in imprinting errors was attributed to embryo or gamete manipulation, in-vitro culture conditions, hormonal stimulation, or ovulation induction.[8-12] However, a currently accepted alternative notion is that the increased incidence of imprinting disorders may be due to facilitation of conception using gametes of infertile couples that may have elevated risk of epigenomic errors. This view aligns with the limited number of reports showing abnormal methylation patterns at imprinted loci in the gametes of infertile men.[13-19]

Recently, the understanding of the germline epigenome has expanded. We, and others, have shown that the sperm genome is hypomethylated with respect to a differentiated somatic cell,[20,21] and that the small percentage of nucleosomes retained in the mature sperm of human and mouse are enriched at developmental promoters (i.e. HOX clusters), miRNA genes, and imprinted loci.[21-23] Furthermore, the promoters of genes involved in spermatogenesis, cell cycle, and cell metabolism were associated with H3 Lysine 4 methylation (H3K4me3: an activating histone modification), while lacking H3 Lysine 27 methylation (H3K27me3: a repressive histone modification), in keeping with their activation during gametogenesis. In contrast, many of the promoters for genes encoding transcripton factors important for embryonic development and morphogenesis bear two marks with antagonistic roles: H3K4me3 and H3K27me3 - termed ‘bivalent’ chromatin. Here, large regions of H3K27me3 often overlap with smaller regions of

H3K4me3. This bivalent chromatin observed in sperm was first described in embryonic stem cell (ESC), and is believed to help poise genes for either activation or repression later in development.[24] These findings suggest that the sperm genome may be packaged and poised for two programs: the first is a reminiscent gametogenesis program (active chromatin marks), and the second is a future embryonic program (bivalent domains).

The conservation in epigenomic packaging in both the mouse and humans suggests a potential role for paternal histones in early developing embryos.[21-23] It is understood that some cases of IVF failure appear to be due to paternal factors, and one current focus of research is to determine whether abnormal epigenetics in gametes could account for failure of embryogenesis[25,26] and subsequent infertility. We hypothesized that the retained nucleosomes in the male gamete exhibit a functional role in early embryogenesis and that changes in histone retention and histone modifications may have important clinical ramifications on embryo outcome and/or fertility potential. Therefore, we studied the epigenome of two classes of infertility, men with abnormal semen parameters (also previously shown to have altered P1/P2 ratios) and men with unexplained, poor embryogenesis during IVF therapy. This is the first report to show alterations genome-wide in the epigenomic landscape of infertile men.

Materials and Methods

Study Participants

Sperm samples were obtained from seven infertility patients and four donors of known fertility attending the University of Utah Andrology and IVF Laboratories. All patients were consented under an IRB approved protocol. The first four patients (patients

1-4) had presumed male factor infertility who had an abnormal semen analysis and an abnormal protamine 1-to-protamine 2 (P1/P2) ratio (2 had lower than normal P1/P2 ratio (<0.8) and 2 with an elevated P1/P2 ratio (>1.2)). Two of the patients had previously undergone IVF without obtaining a pregnancy, although morphologically normal embryos were transferred. The other two patients had undergone intrauterine inseminations without obtaining a pregnancy.

The remaining three patients had unexplained embryo arrest during IVF. Two of the poor embryo patients had good quality embryos on day 2 of embryo culture, but all embryos arrested on day 3. The third patient had pronuclear arrest of all embryos. Electron microscopy revealed normal sperm centrosome morphology in this patient. Sperm aneuploidy testing for five chromosomes was also performed on this patient and the aneuploidy rate was not elevated.

Semen Sample Collection and Clinical Parameters

Semen samples were collected after 2-5 days of abstinence and subjected to somatic cell lysis (0.1% SDS, 0.5% Triton-X in DEPC H_2O) for 20 minutes on ice and washed with 1X PBS to eliminate white blood cell contamination. Semen parameters were measured following WHO criteria.[27] Sperm count, motility, morphology, DNA damage, P1/P2 ratio, and percentage of histone retention are summarized in Table 7.1. DNA damage was assessed using the TUNNEL assay[28,29] and P1/P2 ratio were determined as previously described.[30]

Partitioning of histone and protamine associated DNA was performed as previously described.[21] Briefly, 10-20 million sperm cells were treated with an increasing

concentration of micrococcal nuclease (MNase 10-240U). The histone fraction was either used directly for chromatin immunoprecipitation (ChIP) or the DNA was purified and mononucleosomal length fragments (~140-155 bp) were gel purified and subjected to Illumina GAIIx, ~10-20 million reads were mapped for each sample. The protamine associated DNA was hybridized to the agilent expanded promoter arrays, a two-slide set (244k X2).[21]

Chromatin IP and Preparation for Genomics Methods

ChIP and sequencing methods were as described previously.[21] ChIPs used the following antibodies: anti-H3K27me3 (Upstate 07-449) and H3K4me3 (Abcam 8580). For each modification ~10-20 million reads were aligned.

Methylation Profiling using Bisulfite Sequencing: Sperm DNA was extracted as previously described.[19] 1ug of genomic DNA was converted using the epitect bisulfite kit (Qiagen, Valencia, CA). PCR products were cloned into TOPO2.1 vector (Invitrogen), and bacterial cells were grown on kanamyocin (50ug/ml) and X-gal (25ug/ml) plates overnight. Colony PCR was used to screen 10-20 colonies, however, 8-10 positive clones were submitted to sequencing.

Computational Data Analysis

Low and high level Agilent array and Illumina GAIIx sequencing data analysis were performed as previously described.[21] Both Timat2 and Useq analysis packages are available at the SourceForge projects (http://useq.sourceforge.net/, http://timat2.sourceforge.net/). For the array data, the relative difference pseudo median

scores were converted to log2 ratio values and a Log2 value of 1 (2-fold enrichment) was the selected threshold for all analysis. For the sequencing data, 36 bp reads were generated using Illumina's Genome Analyser IIx and processed according to their standard software pipeline. Reads were mapped to the March 2006 NCBI Build 36.1 human genome and analyzed using the USeq ChIP-Seq application where a q-value threshold of 20 (FDR< 0.01), Log2 ratio 1 was used to select regions for analysis.

Results

Histone Localization is Altered in the Gametes of Infertile Men

To better understand the relationship between the sperm epigenome and infertility we localized nucleosomes in the gametes of seven men with infertility. All sperm samples were subjected to sequential increasing amounts of micrococcal nuclease digestion, and mononucleosomal length fragments were gel purified and submitted to high-throughput sequencing to determine histone localization. Interestingly, unlike the consistent and repeated pattern of histone localization noted in fertile men, all four men with abnormal semen parameters/protamine ratios and two of the three men from couples with embryo failure couples (6 of 7 infertile) had dispersed histone localization patterns genome-wide suggesting that histone retention in infertile men maybe random (Figure 7.1). For fertile donors and all infertile patients, we depict both sequencing read scores and the results of a statistical windowing program that quantifies the significance of read/histone enrichment (Qval, see methods). This distribution of nucleosomes is illustrated across the HOXD cluster (Figure 7.1); however, the observation of random histone localization for six of the seven patients extended genomewide.

An altered histone localization profile in infertile men might be related to the percentage of the genome packaged in histone, as the failure to efficiently remove histone from the bulk genome will obscure the observation of relative retention patterns at developmental loci. Previous work from this laboratory showed that the histone retention in normal subjects was estimated to be in the range of 3-5% (Table 7.1). In contrast to fertile men, three of the four men with abnormal semen parameters/abnormal protamine ratios, and all of the men from couples with embryonic failure, had elevated histone retention in their bulk genome. Histone levels in infertile men were variable and ranged from 5-32% (Table 7.1). This supports the notion that greater histone retention overall obscures observation of specific enrichment at developmental loci. However, one of the embryo failure subjects (patient 7), despite a markedly elevated histone retention (16%: Figure 7.1, Table 7.1) retained relative histone enrichment at expected promoters. Furthermore, a patient with an abnormal protamine ratio (patient 2), although retaining a normal range of histone (5%: Table 7.1, Figure 7.1), did not display a normal retention profile. Therefore, measuring bulk nucleosome retention is insufficient to conclude nucleosomes are properly placed – one must determine the localization.

The loss of apparent histone enrichment at particular promoters may be due to excess histone across the genome, or instead due to the loss of histone from promoters at developmental and gametogenesis genes. To test this, we examined the level and placement of histone modifications, reasoning that if these attributes are retained, then properly marked histones remain. Here, we performed chromatin immunoprecipitation (ChIP) experiments for two histone modifications, H3K4me3 and H3K27me3 in our seven study subjects. The ChIP data was normalized to each subject's mononucleosome

input and a 300 bp window was used to identify enriched regions. We found that H3K4me and H3K27me methylation patterns in infertile men were generally similar to that of fertile donors (Figures 7.2, 7.3 and Table 7.2). Therefore, it seems likely that the abnormal histone localization pattern in infertile men is because of an incomplete replacement of histone by protamine in the genome as whole, resulting in an increased background of abnormally retained histone that obscures the pattern of appropriately retained histones seen in normal sperm.

The Sperm Epigenome of Infertile Men

This genome-wide similarity in H3K4me3 and H3K27me3 datasets observed in fertile subjects and in subjects with evidence for reproductive dysfunction (Figure 7.2 and 7.3) was substantiated using gene ontology analysis (Figure 7.4) and Pearson correlations (Table 7.2). Similar to fertile controls, H3K4me gene ontology analysis revealed that H3K4me localization in subjects with sperm/protamine abnormalities or embryonic failure was enriched at the promoters of cell cycle genes, cellular metabolic processes, RNA processing and splicing, spermatogenesis gene promoters, and developmental transcription factors (FDR 40 = < 1/1000). In contrast to H3K4me localization, H3K27me localization in control and abnormal subjects was similar, showing enrichment only at developmental gene promoters (FDR 40 = < 1/1000). These results imply that as in normal fertile men, sperm from men with reproductive disorders exhibit H3K4me enrichment at the gametogenesis program and H3K27me represses loci in the future embryonic program. The similarity in modification datasets for the two groups of subjects was also confirmed by the Pearson correlation coefficients on raw datasets

(Summarized in Table 7.2). The correlation coefficients (r) for many of the datasets were estimated 0.7 or greater. These findings confirm that modification profiles for the fertile subjects and subjects with reproductive dysfunction are highly correlated, but not completely so, such that differences exist between individuals and classes (Summarized in Table 7.2). An illustrative gene specific difference for H3K4me comparing normal and abnormal subjects can be observed at imprinted loci (Figure 7.5). In fertile subjects, nearly all paternally expressed imprinted genes were enriched with H3K4me, while maternally imprinted loci lacked H3K4me3. In contrast, among the seven subjects with evidence for reproductive dysfunction, three (patients 3, 5, 6) showed absence of H3K4me3 at the promotors of 3 paternally expressed imprinted loci (PEG10, SGCE, and NAP1L5), and 6 showed a gain of H3K4me3 at a maternally imprinted gene (ZNF264) (Figure 7.5).

Although the localization of modified nucleosomes was similar to controls in all seven subjects with evidence for reproductive abnormalities, a differential analysis was performed to identify loci that bear differences in the amount of H3K4me/H3K27me enrichment (read count at promoters). The analysis outputs 1) enriched regions: which are regions enriched in the patient and reduced in controls and/or 2) reduced regions: which are regions deficient in patients and elevated in controls. These regions are subsequently submitted for gene ontology analysis to identify altered programs or pathways in infertile men. Interestingly, the top 300 genes with reduced H3K4me3 were predominately genes of bivalent promoters involved in embryonic processes such as multicellular oragnismal development, system development, organ development, pattern

specification and not the promoters of cell cycle or spermatogenesis (FDR 40 < 1/1000). Similarily, the top 300 genes with reduced H3K27me3 promoters in the subjects were developmental promoters. This latter finding is expected because H3K27me3 is enriched only at developmental promoters in patients and controls (FDR 40 < 1/1000).

Changes in Chromatin Modifications Do Not Generally Result in Dramatic Changes in DNA Methylation

Recently, structural and in vitro data showed that H3K4 methylation deters DNA methylation by DNMT3A2 and DNMT3L in mice.[31] To examine whether a change in chromatin modifications might result in a change in DNA methylation, bisulfite sequencing of several candidate developmental loci that lacked or had significantly reduced levels of H3K4me3 were assessed in three infertile men (patients 1, 4, 6). As seen in fertile control subjects, the developmental promoters tested in the three infertile men were generally unmethylated with very few changes in methylation observed relative to controls (Figure 7.6). Expanding the DNA methylation analysis to imprinted loci and including all subjects with evidence for reproductive dysfunction revealed more pronounced changes at the imprinted loci of patients 2 and 3. However, the remainder of the subjects showed very limited changes in methylation (Figure 7.7). The gain of DNA methylation did not always correlate with a loss of K4me. Subjects 2 and 3 had reduced or lost K4me at PEG10 whereas 30-70% of their alleles gained methylation (Figures 7.5 and 7.7). In contrast, patients 5 and 6 lost K4me at PEG10 but did not acquire methylation (Figures 7.5 and 7.7). Therefore, although we could not always correlate a loss of H3K4me with a gain in DNA methylation, infertile men appear more susceptible

to changes in DNA methylation independent of the changes in chromatin in the mature sperm.

Changes in Protamine Localization Are Observed in the Gametes of Infertile Patients

Protamines were localized in three infertile men with an abnormal P1/P2 ratio. The data presented are the averages from two biological replicas. The correlations between the replicas were greater than 0.8. Similar to fertile men, protamines were present throughout in the genome as shown in Figure 7.8a. Gene ontology analysis did not enrich for a particular cellular process. However, the correlation and hierarchal clustering results show that patients clustered away from donors (Figure 7.8b), suggesting potential differences in protamine levels as well as localization genome-wide in the gametes of infertile men.

Discussion

This paper provides several lines of evidence that the genome packaging and epigenomic modifications are altered in the gametes of some infertile men. A few reports have shown that the methylation at several imprinted loci were altered in the gametes of infertile men.[13-15,17,21,32,33] We add that the alteration in the epigenome extends beyond DNA methylation at imprinted loci and entails changes in histone modifications and localization genome-wide at a large number of loci.

Altered histone localization patterns identified in the sperm of infertile men generally reflect a more random pattern of nucleosome retention rather than the localized

retention seen in all fertile donors. This incomplete removal of nucleosomes might be attributed to problems with the chromatin remodeling machinery or improper histone acetyl transferase (HAT) activity during the histone-to-protamine exchange, however, the effect of improper histone retention in early embryos is currently not well understood.

While nucleosomes were randomly retained genome-wide in infertile men, the localization of modified nucleosomes was similar to fertile controls suggesting that the establishment of epigenetic marks in the spermatogonial stem cell remain largely intact in infertile men. However, subtle but potentially important differences were identified. For example developmental promoters enriched with H3K4 and H3K27me in known fertile men were lost or significantly reduced in most infertile men. The most striking observation is the loss of H3K4me3 from three paternally imprinted loci, a mark normally highly correlated with hypomethylation, in keeping with structural and functional studies of DNMT3L and DNMT3A in mice.[31] However, the loss or reduction of K4me3 in the mature sperm correlated with a gain in DNA methylation in only one of the seven patients. This shows that the loss of H3K4me3 is likely insufficient to cause DNA methylation, suggesting that additional mechanisms are likely involved in preventing DNA methylation at these imprinted loci.

The clinical ramifications of an altered H3K4me/H3K27me bound promoters in the gametes of infertile men is not yet known. However, clinical case studies have established that imprinting diseases in particular are not binary (presence or absence), but instead can result from moderate changes in DNA methylation status, e.g., in Beckwith-weidmann's syndrome a 20% reduction in the number of methylated CpGs in a paternally-methylated DMR results in disease onset. Therefore, the fidelity of the

epigenome may need to be quite high to ensure embryo viability and health. Future studies that determine the relevance of chromatin marks in the germ-line are essential.

Although the number of patients assessed in this study is small (necessarily so due to the extreme costs associated with genome-wide analysis) the results are intriguing. Previously we, and others, have shown that all fertile men tested programmatically retain nucleosomes genomewide. In contrast to fertile controls, six of seven infertile men had abnormal histone retention patterns, suggesting that histone to protamine exchange is defective in many infertile men. Furthermore, a subset of the patients displayed abnormalities in histone modifications (H3K4me and H3K27me) and DNA methylation at imprinted and developmental loci, which was absent in the fertile control pooled sample. This study demonstrates differences in the chromatin composition and epigenetic marking in the sperm of infertile men, especially at imprinted and developmental loci. Although no single locus displays a complete change in chromatin packaging or DNA modification, these moderate changes at many loci may have a cumulative detrimental affect on fecundity. This is aligned with current notions, since infertility is a multifactorial disease, and like other complex diseases, numerous etiologies likely contribute to male infertility or poor embryo outcome, with epigenetics being just one potential contributor.

In summary, this is the first report examining the sperm epigenome (histone localization and histone modifications) genome-wide in men with abnormal spermatogenesis or repeated poor embryo outcome. The risks associated with germ-line epigenetic abnormalities for the embryo is unclear, but we know subtle differences in chromatin packaging were observed in the gametes of infertile men, and the penetrance

of epigenetic abnormalities according to the literature is low, therefore it is not entirely surprising that the frequency of epigenetic anomalies from ART is rare. The low disease penetrance can be attributed to several selected defense mechanism acquired through evolution and natural selection to ensure the formation of a viable conceptus.

References

1. Manipalviratn S, DeCherney A, Segars J. Imprinting disorders and assisted reproductive technology. Fertil Steril 2009;91(2):305-15.

2. Hansen M, Kurinczuk JJ, Bower C, Webb S. The risk of major birth defects after intracytoplasmic sperm injection and in vitro fertilization. N Engl J Med 2002;346(10):725-30.

3. Reefhuis J, Honein MA, Schieve LA, Correa A, Hobbs CA, Rasmussen SA. Assisted reproductive technology and major structural birth defects in the United States. Hum Reprod 2009;24(2):360-6.

4. Rebar RW, DeCherney AH. Assisted reproductive technology in the United States. N Engl J Med 2004;350(16):1603-4.

5. Nelissen EC, van Montfoort AP, Dumoulin JC, Evers JL. Epigenetics and the placenta. Hum Reprod Update 2010.

6. Bourque DK, Avila L, Penaherrera M, von Dadelszen P, Robinson WP. Decreased placental methylation at the H19/IGF2 imprinting control region is associated with normotensive intrauterine growth restriction but not preeclampsia. Placenta 2010;31(3):197-202.

7. Maher ER. Imprinting and assisted reproductive technology. Human Molecular Genetics 2005;14 Spec No 1:R133-8.

8. Mann MR, Lee SS, Doherty AS, Verona RI, Nolen LD, Schultz RM, Bartolomei MS. Selective loss of imprinting in the placenta following preimplantation development in culture. Development 2004;131(15):3727-35.

9. Ludwig M, Katalinic A, Gross S, Sutcliffe A, Varon R, Horsthemke B. Increased prevalence of imprinting defects in patients with Angelman syndrome born to subfertile couples. J Med Genet 2005;42(4):289-91.

10. Khosla S, Dean W, Brown D, Reik W, Feil R. Culture of preimplantation mouse embryos affects fetal development and the expression of imprinted genes. Biol Reprod 2001;64(3):918-26.

11. DeBaun MR, Niemitz EL, Feinberg AP. Association of in vitro fertilization with Beckwith-Wiedemann syndrome and epigenetic alterations of LIT1 and H19. Am J Hum Genet 2003;72(1):156-60.

12. Cox GF, Burger J, Lip V, Mau UA, Sperling K, Wu BL, Horsthemke B. Intracytoplasmic sperm injection may increase the risk of imprinting defects. Am J Hum Genet 2002;71(1):162-4.

13. Marques CJ, Carvalho F, Sousa M, Barros A. Genomic imprinting in disruptive spermatogenesis. Lancet 2004;363(9422):1700-2.

14. Marques CJ, Costa P, Vaz B, Carvalho F, Fernandes S, Barros A, Sousa M. Abnormal methylation of imprinted genes in human sperm is associated with oligozoospermia. Mol Hum Reprod 2008;14(2):67-74.

15. Marques CJ, Francisco T, Sousa S, Carvalho F, Barros A, Sousa M. Methylation defects of imprinted genes in human testicular spermatozoa. Fertil Steril 2010;94(2):585-94.

16. Kobayashi H, Hiura H, John RM, Sato A, Otsu E, Kobayashi N, Suzuki R, Suzuki F, Hayashi C, Utsunomiya T and others. DNA methylation errors at imprinted loci after assisted conception originate in the parental sperm. Eur J Hum Genet 2009;17(12):1582-91.

17. Kobayashi H, Sato A, Otsu E, Hiura H, Tomatsu C, Utsunomiya T, Sasaki H, Yaegashi N, Arima T. Aberrant DNA methylation of imprinted loci in sperm from oligospermic patients. Human Molecular Genetics 2007;16(21):2542-51.

18. Kagami M, Nagai T, Fukami M, Yamazawa K, Ogata T. Silver-Russell syndrome in a girl born after in vitro fertilization: partial hypermethylation at the differentially methylated region of PEG1/MEST. J Assist Reprod Genet 2007;24(4):131-6.

19. 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 2009.

20. Weber M, Hellmann I, Stadler MB, Ramos L, Paabo S, Rebhan M, Schubeler D. Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. Nature Genetics 2007;39(4):457-66.

21. 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(7254):473-8.

22. Brykczynska U, Hisano M, Erkek S, Ramos L, Oakeley EJ, Roloff TC, Beisel C, Schubeler D, Stadler MB, Peters AH. Repressive and active histone methylation mark distinct promoters in human and mouse spermatozoa. Nat Struct Mol Biol 2010;17(6):679-87.

23. Arpanahi A, Brinkworth M, Iles D, Krawetz SA, Paradowska A, Platts AE, Saida M, Steger K, Tedder P, Miller D. Endonuclease-sensitive regions of human spermatozoal chromatin are highly enriched in promoter and CTCF binding sequences. Genome Res 2009;19(8):1338-49.

24. Bernstein BE, Mikkelsen TS, Xie X, Kamal M, Huebert DJ, Cuff J, Fry B, Meissner A, Wernig M, Plath K and others. A bivalent chromatin structure marks key developmental genes in embryonic stem cells. Cell 2006;125(2):315-26.

25. Tesarik J. Paternal effects on cell division in the human preimplantation embryo. Reprod Biomed Online 2005;10(3):370-5.

26. Carrell DT. Contributions of spermatozoa to embryogenesis: assays to evaluate their genetic and epigenetic fitness. Reprod Biomed Online 2008;16(4):474-84.

27. World Health Organization. WHO laboratory manual for the examination of human semen and sperm-cervical mucus interaction. Cambridge, UK ; New York, NY: Published on behalf of the World Health Organization Cambridge University Press; 1999. x, 128 p. p.

28. Zini A, Boman JM, Belzile E, Ciampi A. Sperm DNA damage is associated with an increased risk of pregnancy loss after IVF and ICSI: systematic review and meta-analysis. Hum Reprod 2008;23(12):2663-8.

29. Aoki VW, Moskovtsev SI, Willis J, Liu L, Mullen JB, Carrell DT. DNA integrity is compromised in protamine-deficient human sperm. J Androl 2005;26(6):741-8.

30. Carrell DT, Liu L. Altered protamine 2 expression is uncommon in donors of known fertility, but common among men with poor fertilizing capacity, and may reflect other abnormalities of spermiogenesis. J Androl 2001;22(4):604-10.

31. Ooi SK, Qiu C, Bernstein E, Li K, Jia D, Yang Z, Erdjument-Bromage H, Tempst P, Lin SP, Allis CD and others. DNMT3L connects unmethylated lysine 4 of histone H3 to de novo methylation of DNA. Nature 2007;448(7154):714-7.

32. Poplinski A, Tuttelmann F, Kanber D, Horsthemke B, Gromoll J. Idiopathic male infertility is strongly associated with aberrant methylation of MEST and IGF2/H19 ICR1. Int J Androl 2010;33(4):642-9.

33. Boissonnas CC, Abdalaoui HE, Haelewyn V, Fauque P, Dupont JM, Gut I, Vaiman D, Jouannet P, Tost J, Jammes H. Specific epigenetic alterations of IGF2-H19 locus in spermatozoa from infertile men. Eur J Hum Genet 2010;18(1):73-80.

Table 7.1. Description of Patient Population and Summary of Semen Parameters

Patient Code	Female Age	Male Age	Progressive Motility	Morphology	Sperm Count	% TUNNEL Positive	P1/P2 Ratio	%Histone Retention	Number Of IUI	Number Of IVF	Embryo Quality
Donor Pool	NA	26-35	62	49	132	5	0.923	3-5%	NA	NA	Many Pregnancies
Patient 1	31	31	28	14	119	52	1.73	25%	0	2	Day3 embryos-No pregnancy twice
Patient 2	33	33	37	15	75	40	0.732	5%	2	0	No IVF Data
Patient 3	29	32	23	10	10.7	2	1.539	13%	0	1	Day No pregnancy
Patient 4	30	30	37	15	75.5	30	0.739	8%	3	0	No IVF Data
Patient 5	27	29	7	11	2.45	42	0.8	32%	3	1	Pronuclear arrest (twice)
Patient 6	26	29	42	35	49	19	0.96	16%	1	3	All embryos arrested on day3 (twice)
Patient 7	27	32	50	39	36	15	0.72	16.6%	0	2	All embryos arrested day3 (twice)

Table 7.2. Correlation of Histone Modification Between Fertile and Infertile Men

Comparison	r
Donor H3K4me3 vs. Patient 1 H3K4me3	0.786
Donor H3K4me3 vs. Patient 2 H3K4me3	0.826
Donor H3K4me3 vs. Patient 3 H3K4me3	0.788
Donor H3K4me3 vs. Patient 4H3K4me3	0.868
Donor H3K4me3 vs. Patient 5 H3K4me3	0.698
Donor H3K4me3 vs. Patient 6 H3K4me3	0.749
Donor H3K4me3 vs. Patient 7 H3K4me3	0.749
Donor H3K27me3 vs Patient 1 H3K27me3	0.807
Donor H3K27me3 vs. Patient 2 H3K27me3	0.771
Donor H3K27me3 vs. Patient 3 H3K27me3	0.818
Donor H3K27me3 vs. Patient 4 H3K27me3	0.787
Donor H3K27me3 vs. Patient 5 H3K27me3	0.666
Donor H3K27me3 vs. Patient 6 H3K27me3	0.775
Donor H3K27me3 vs. Patient 7 H3K27me3	0.785

Figure 7.1. High throughput sequencing of mononucleosomes reveals relatively random histone retention at the HOX loci in most infertile patients. Mapped sequencing reads were scored for enrichment (score) and for significance. Regions significantly enriched for histone relative to the input control (sheared total sperm DNA) were identified using a 300-bp window metric. For display, we depict score and false discovery rate (FDR) window scores (-10Log(FDR)). Note a FDR of 20 is equal to <0.01 and FDR 40 < 0.001.

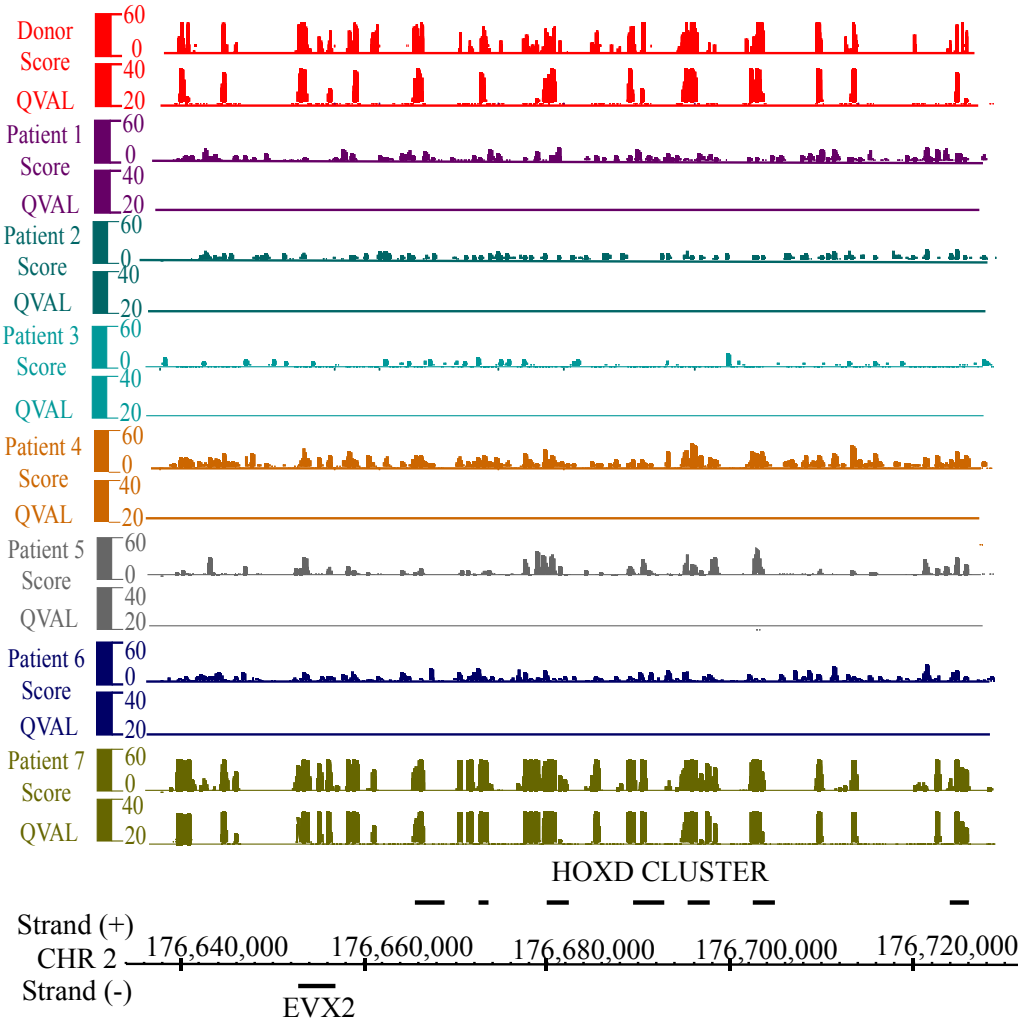

Donor Score
60
0
QVAL
40
20
Patient 1 Score
Patient 2 Score
Patient 3 Score
Patient 4 Score
Patient 5 Score
Patient 6 Score
Patient 7 Score
HOXD CLUSTER
Strand (+)
CHR 2
Strand (-)
176,640,000
176,660,000
176,680,000
176,700,000
176,720,000
EVX2

Figure 7.2. H3K4me3 localization patterns are generally normal in the gametes of infertile men. This figure illustrates H3K4me3 retention aligned to the physical map of the HOXD locus. The y-axis depicts Storey q-value false discovery rate (-10Log(FDR)). Note a FDR of 20 is equal to <0.01 and FDR 40< 0.001.

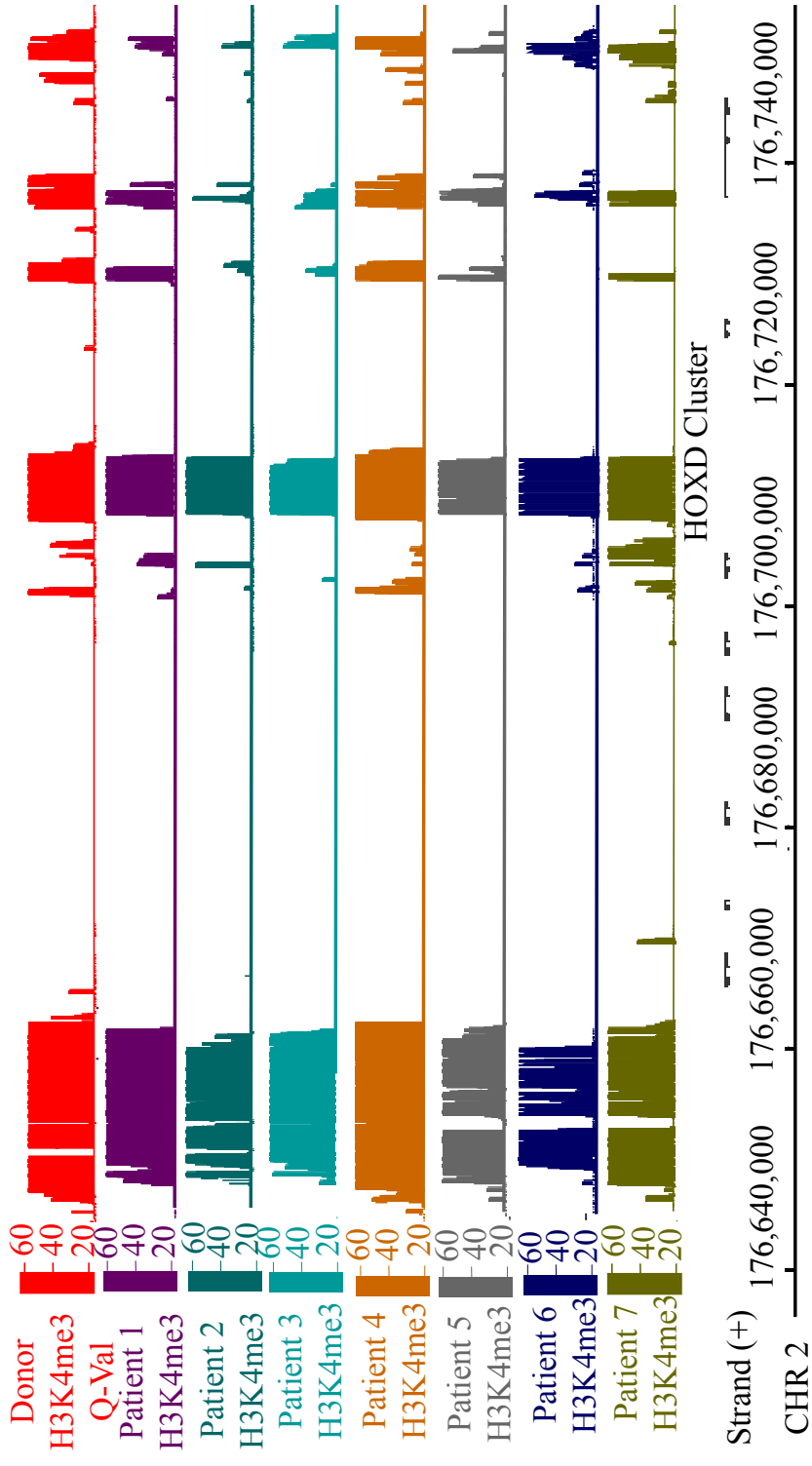

Donor
H3K4me3
Q-Val
Patient 1
H3K4me3
Patient 2
H3K4me3
Patient 3
H3K4me3
Patient 4
H3K4me3
Patient 5
H3K4me3
Patient 6
H3K4me3
Patient 7
H3K4me3
60
40
20
Strand (+)
176,640,000
176,660,000
176,680,000
176,700,000
176,720,000
176,740,000
CHR 2
HOXD Cluster
Strand (-)
EVX2

Figure 7.3. H3K27me3 localization patterns are generally normal in the gametes of infertile men. This figure illustrates H3K27me3 retention aligned to the physical map of the HOXD locus. The y-axis depicts Storey q-value false discovery rate (-10Log(FDR)). Note a FDR of 20 is equal to <0.01 and FDR 40< 0.001.

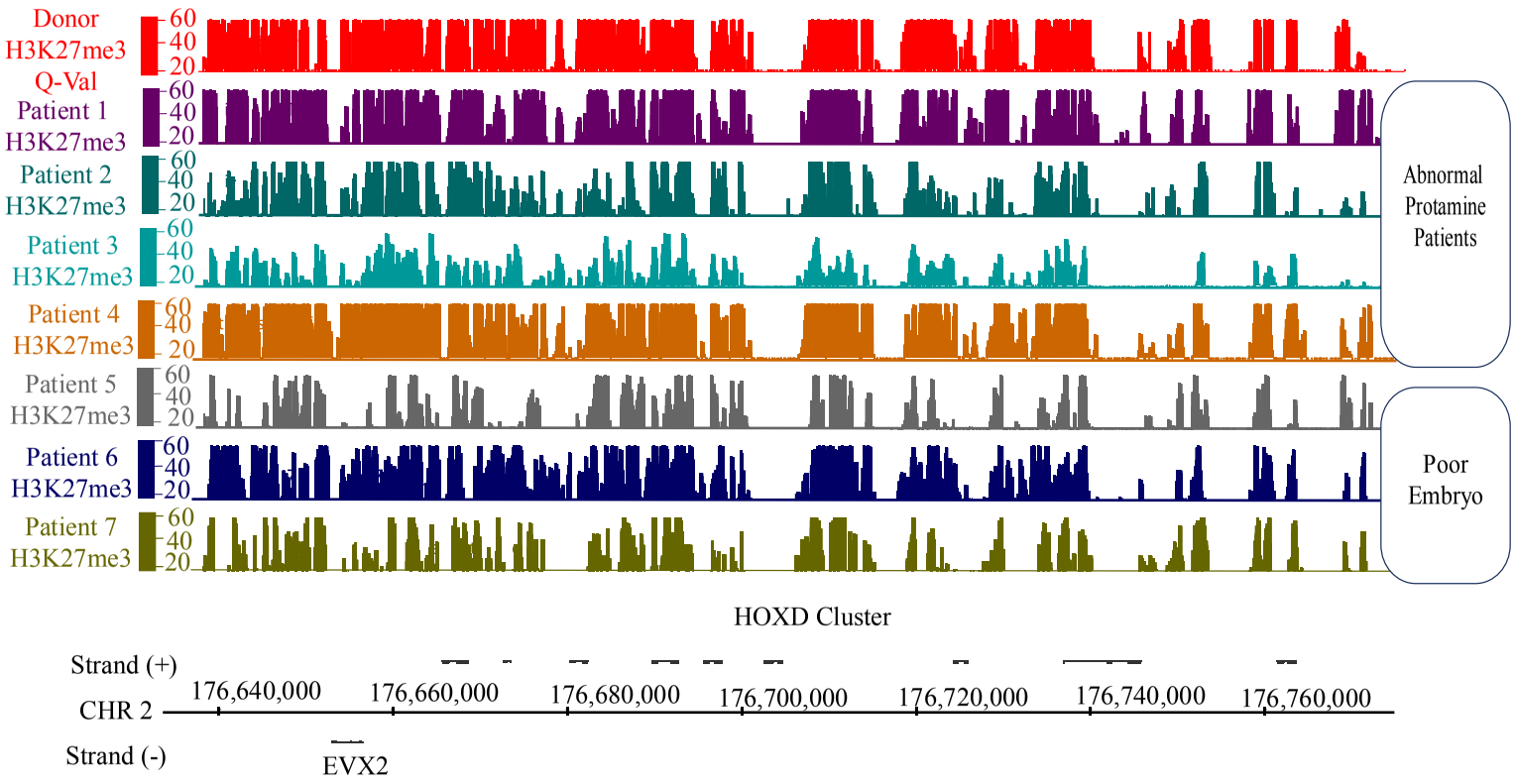

Donor
H3K27me3
Q-Val
Patient 1
H3K27me3
Patient 2
H3K27me3
Patient 3
H3K27me3
Patient 4
H3K27me3
Patient 5
H3K27me3
Patient 6
H3K27me3
Patient 7
H3K27me3
60
40
20
Abnormal
Protamine
Patients
Poor
Embryo
HOXD Cluster
Strand (+)
CHR 2
176,640,000
176,660,000
176,680,000
176,700,000
176,720,000
176,740,000
176,760,000
Strand (-)
EVX2

Figure 7.4. Gene Ontology Analysis reveals similar K4 and K27me localization patterns in fertile and infertile men. GoMiner was used to identify enriched categories, and all categories displayed have an FDR < 0.001. The top five general categories are listed, after omitting nearly identical/redundant classes.


Patient 4 H3k4me3
RNA Splicing
Cell cycle Processes
RNA metabolic process
Spermatogenesis
Transcription factor activity
Patient 7 H3k4me3
RNA metabolic Process
Transcription
RNA processing
Mitosis/Meiosis
Gamete Generation
Patient 3 H3k4me3
RNA metabolic Process
Transcription
Regulation of transcription
RNA processing
Spermatogenesis
Patient 1 H3k4me3
RNA metabolic Process
Transcription
RNA binding/Splicing
Mitosis
Spermatogenesis
Patient 2 H3k4me3
RNA metabolic Process
RNA binding/Splicing
Cell Cycle
Transcription
Spermatogenesis
Patient 5 H3k4me3
RNA Splicing
Cell cycle Processes
RNA metabolic process
Spermatogenesis
Transcription factor activity
Patient 4H3K27me3
Anatomical Structure Dev.
System Dev
Multicellular Organismal Process
Organ Development
Cell Differentiation
Cell-cell signaling
Patient 3 H3K27me3
Anatomical Structure Dev.
System Dev
Multicellular Organismal Process
Organ Development
Cell Differentiation
Sequence Specific DNA Binding
Patient 5 H3K27me3
Anatomical Structure Dev.
System Dev
Multicellular Organismal Process
Organ Development
Cell Differentiation
Sequence Specific DNA Binding
Patient 2 H3K27me3
Anatomical Structure Dev.
System Dev
Multicellular Organismal Process
Developmental Process
Cell Differentiation
Organ Morphogenesis
Patient 6 H3K27me3
Anatomical Structure Dev.
System Dev
Multicellular Organismal Process
Developmental Process
Cell Differentiation
Organ Morphogenesis
Patient 7 H3K27me3
Anatomical Structure Dev.
System Dev
Multicellular Organismal Process
Organ Development
Cell Differentiation
Cell-cell signaling

Figure 7.5. A few maternally or paternally imprinted loci are improperly modified in the gametes of infertile men. (A) Depicts two paternally imprinted genes PEG10 and SGCE that lack or have reduced levels of H3K4me3 in patients (B) Represents a maternally imprinted gene (ZNF264) that lacks K4me in fertile donors but acquires K4me in infertile men. (C) and (D) represent two properly marked materally imprinted (H19) or paternally imprinted gene (PEG3).

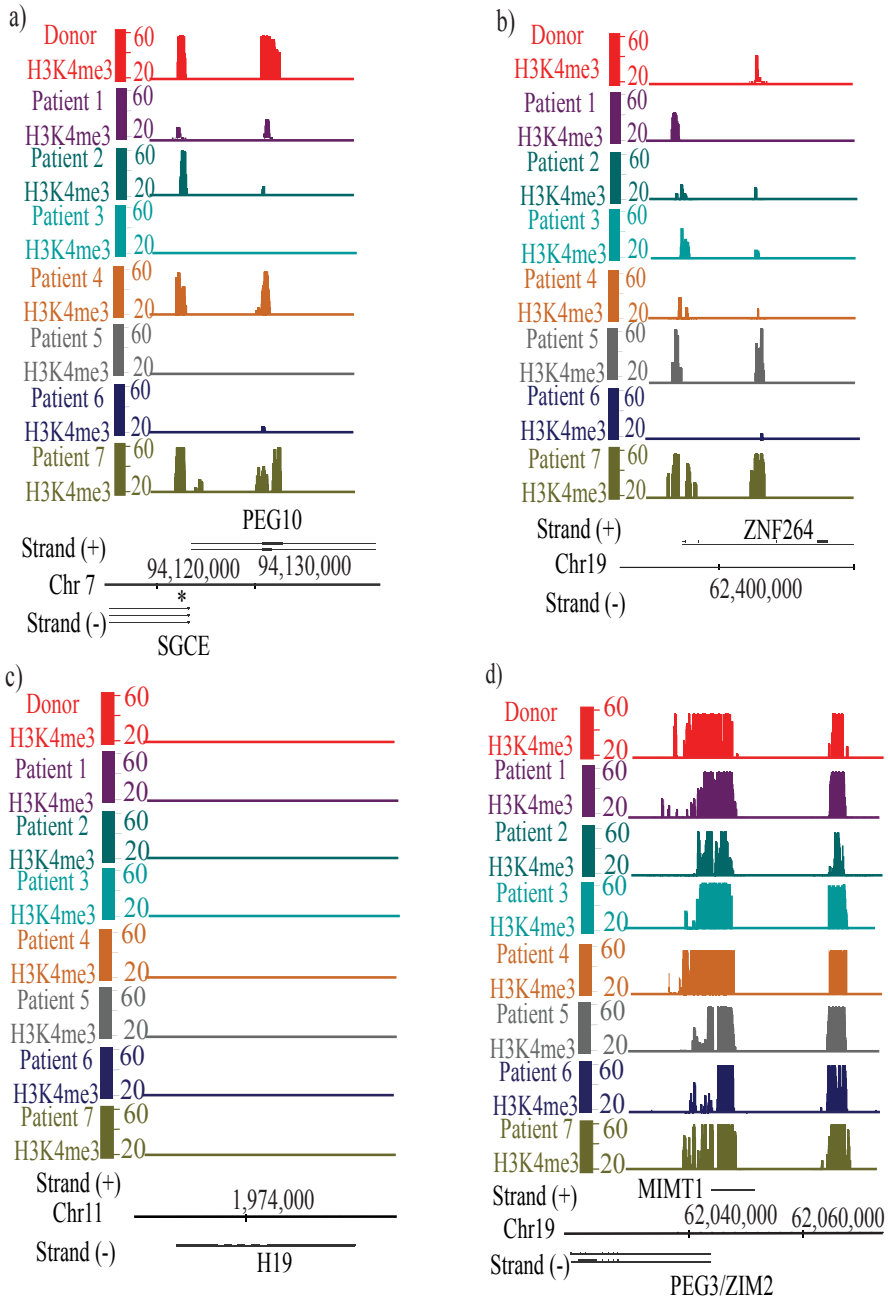

a)
Donor H3K4me3
Patient 1 H3K4me3
Patient 2 H3K4me3
Patient 3 H3K4me3
Patient 4 H3K4me3
Patient 5 H3K4me3
Patient 6 H3K4me3
Patient 7 H3K4me3
60
20
PEG10
Strand (+)
Chr 7
94,120,000
94,130,000
Strand (-)
SGCE
b)
ZNF264
Chr19
62,400,000
c)
1,974,000
Chr11
H19
d)
MIMT1
62,040,000
62,060,000
PEG3/ZIM2

Figure 7.6. DNA hypomethylation is largely maintained at developmental promoters of infertile men. Bisulfite sequencing of developmental promoters that lack or have reduced levels of K4me. CpGs are represented as open dots (if unmethylated) or filled dots (if methylated). Lines represent the number of alleles sequenced per patient.

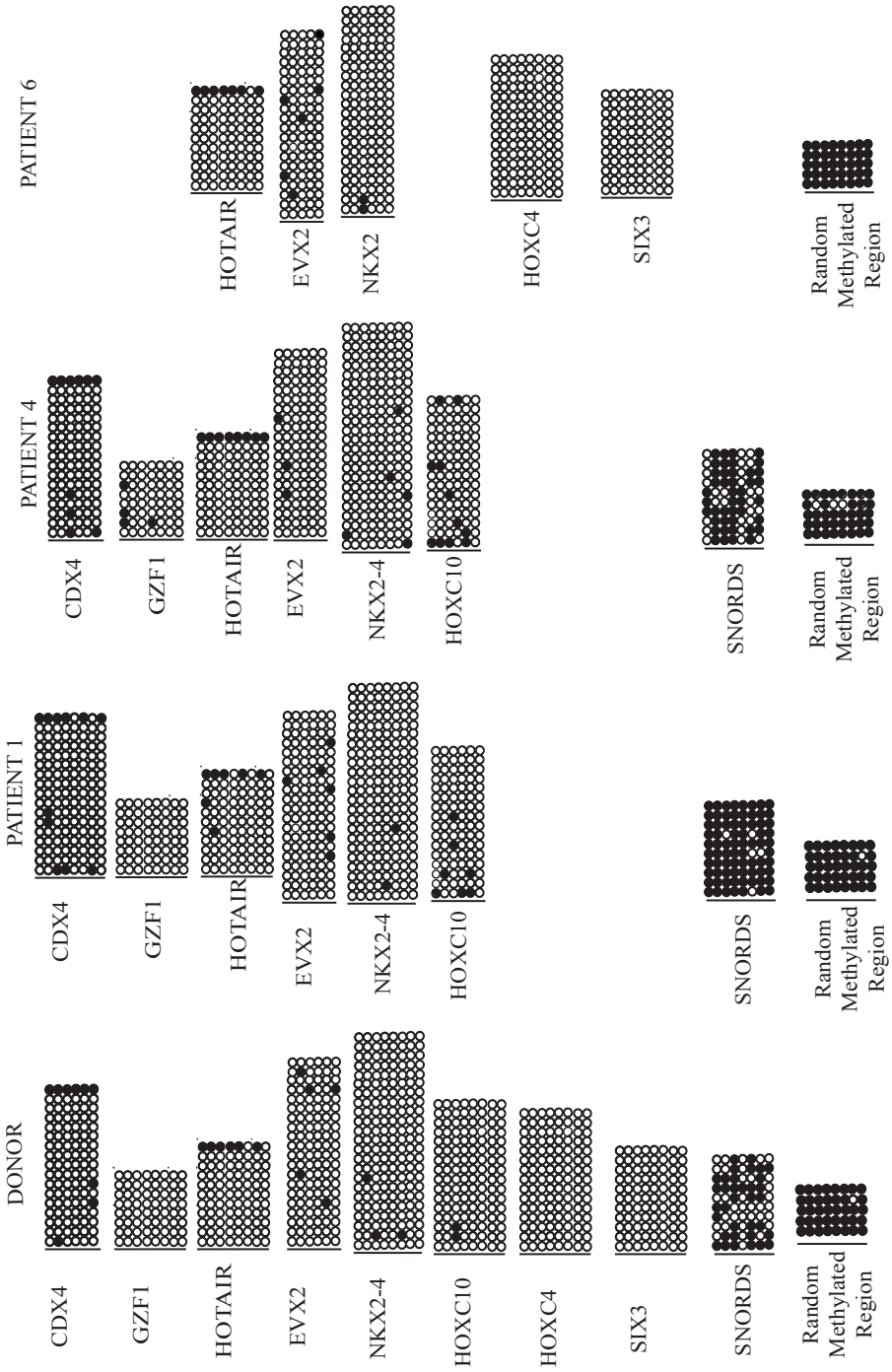
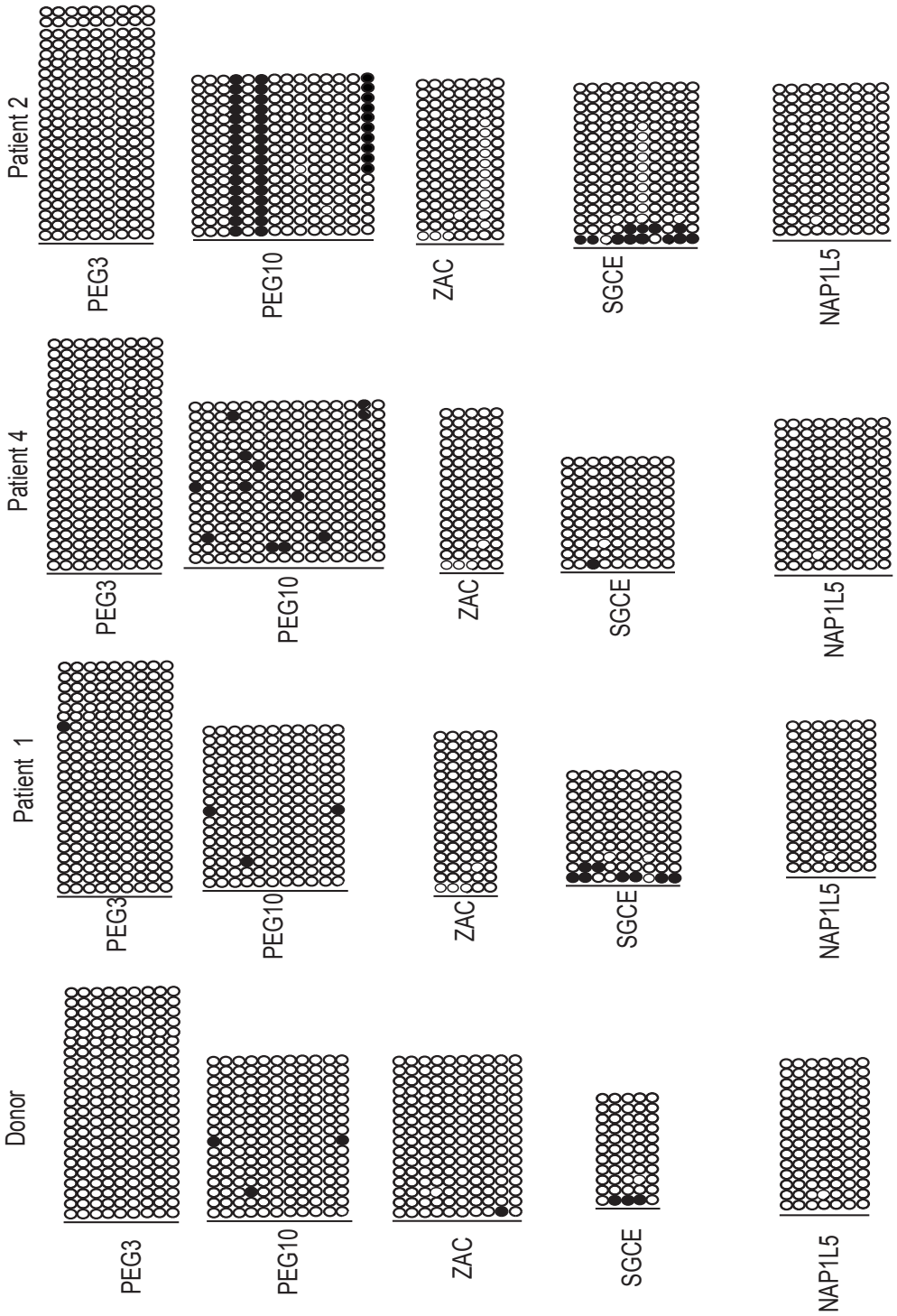

DONOR
CDX4
GZF1
HOTAIR
EVX2
NKX2-4
HOXC10
HOXC4
SIX3
SNORDS
Random Methylated Region
PATIENT 1
CDX4
GZF1
HOTAIR
EVX2
NKX2-4
HOXC10
SNORDS
Random Methylated Region
PATIENT 4
CDX4
GZF1
HOTAIR
EVX2
NKX2-4
HOXC10
SNORDS
Random Methylated Region
PATIENT 6
HOTAIR
EVX2
NKX2
HOXC4
SIX3
Random Methylated Region

Figure 7.7. DNA methylation patterns are altered at a subset of paternally imprinted loci in the gametes of infertile men. CpGs are represented as open dots (if unmethylated) or filled dots (if methylated).


Donor
PEG3
PEG10
ZAC
SGCE
NAP1L5
Patient 1
PEG3
PEG10
ZAC
SGCE
NAP1L5
Patient 4
PEG3
PEG10
ZAC
SGCE
NAP1L5
Patient 2
PEG3
PEG10
ZAC
SGCE
NAP1L5

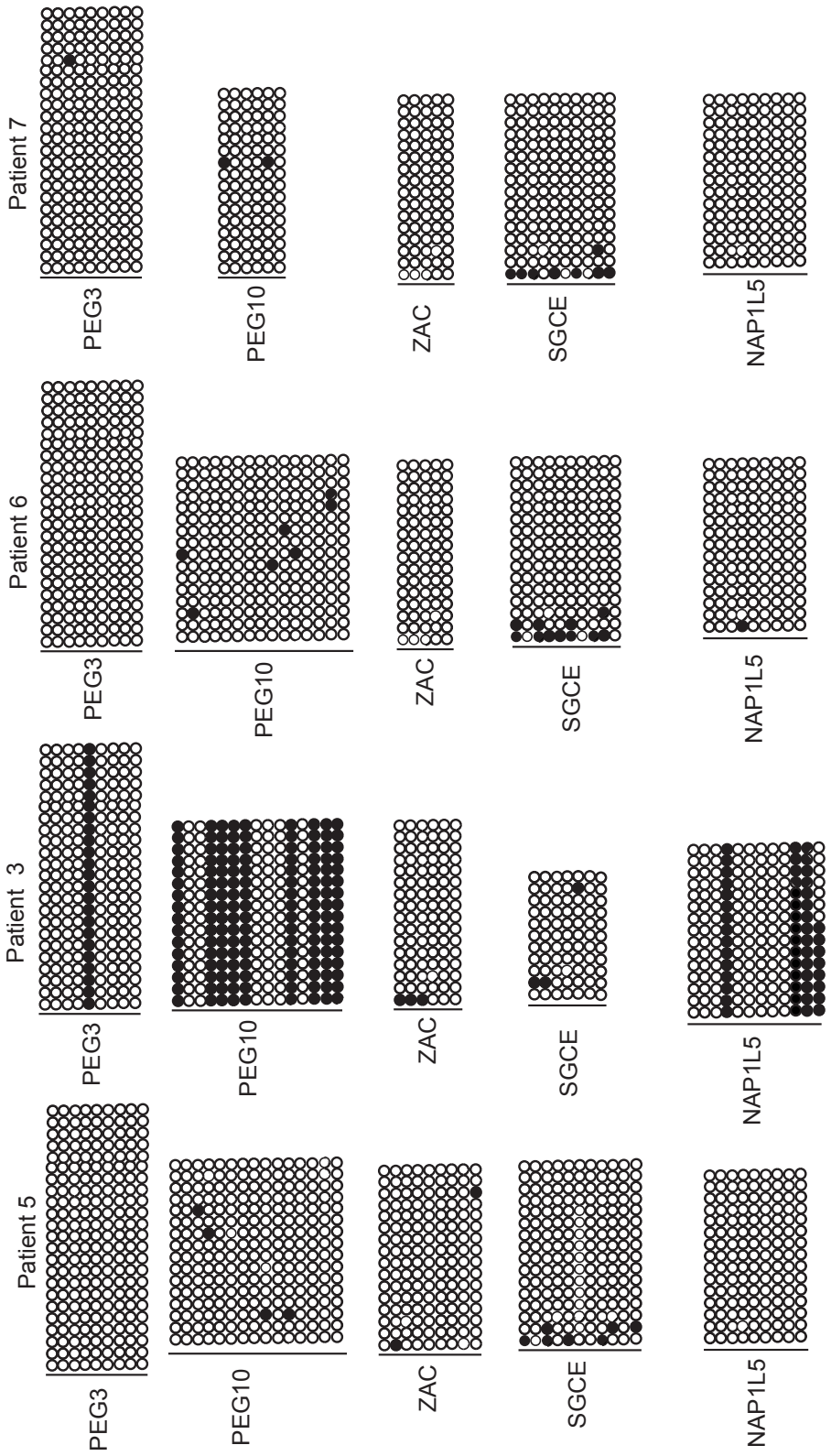
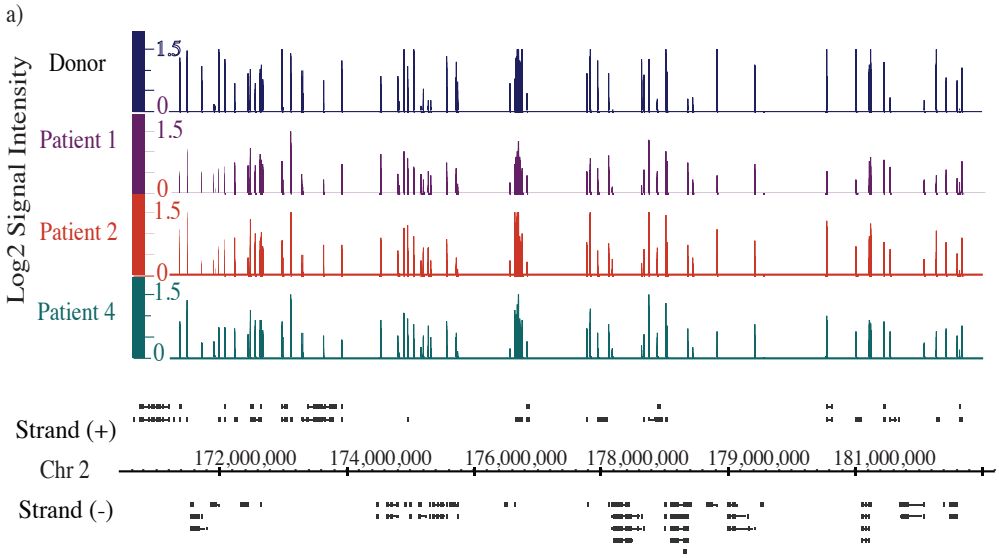
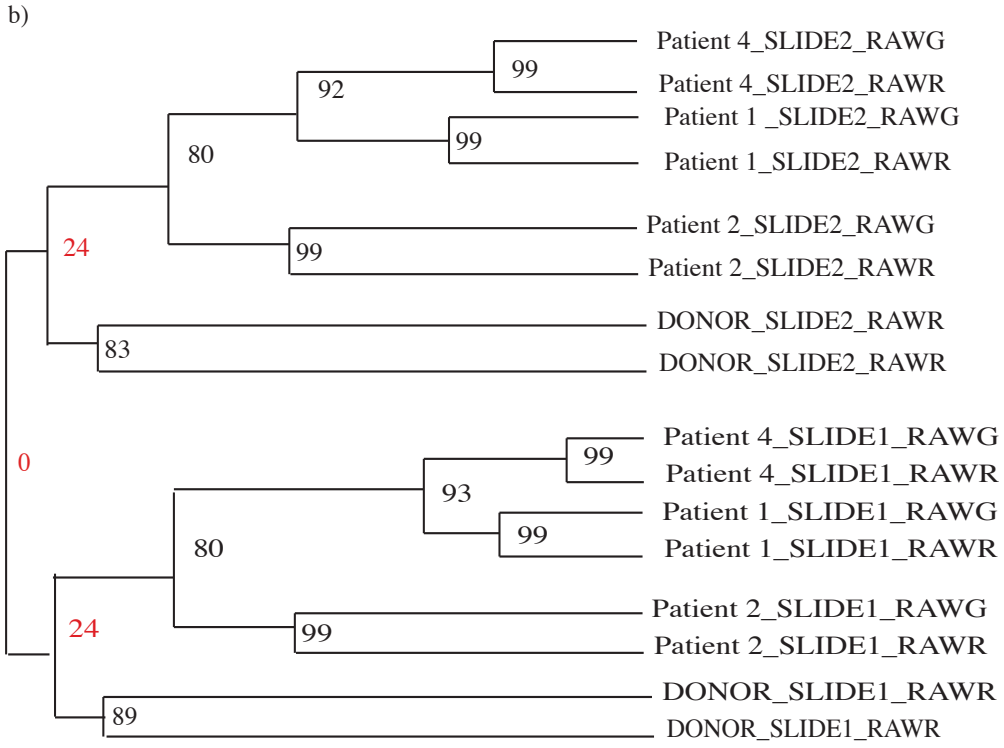

Patient 5
PEG3
PEG10
ZAC
SGCE
NAP1L5
Patient 3
PEG3
PEG10
ZAC
SGCE
NAP1L5
Patient 6
PEG3
PEG10
ZAC
SGCE
NAP1L5
Patient 7
PEG3
PEG10
ZAC
SGCE
NAP1L5

Figure 7.8. Genome-wide protamine localization and Pearson correlations suggest differences in protamine localization. a. Protamine localization in fertile and infertile men. The y-axis is the signal intensity (log2 for ChIP-chip arrays) and the x-axis is the annotated physical map (HG18). B. Hierarchal clustering of raw datasets shows that patients cluster away from controls.


a)
Log2 Signal Intensity
Donor
Patient 1
Patient 2
Patient 4
1.5
0
Strand (+)
Chr 2
Strand (-)
172,000,000
174,000,000
176,000,000
178,000,000
179,000,000
181,000,000


b)
Patient 4_SLIDE2_RAWG
Patient 4_SLIDE2_RAWR
Patient 1 _SLIDE2_RAWG
Patient 1_SLIDE2_RAWR
Patient 2_SLIDE2_RAWG
Patient 2_SLIDE2_RAWR
DONOR_SLIDE2_RAWR
DONOR_SLIDE2_RAWR
Patient 4_SLIDE1_RAWG
Patient 4_SLIDE1_RAWR
Patient 1_SLIDE1_RAWG
Patient 1_SLIDE1_RAWR
Patient 2_SLIDE1_RAWG
Patient 2_SLIDE1_RAWR
DONOR_SLIDE1_RAWR
DONOR_SLIDE1_RAWR
99
92
80
24
0
83
93
89

CHAPTER 8

SUMMARY AND DISCUSSION

Genetic Mutations in the Protamine Cluster and its Role in Infertility

The hypothesis underlying this work is that the genetic and epigenetic states of the mature sperm have important roles in the establishment of proper germ cell identity and function, and in early developmental capacity. We reason that perturbations to the genome or epigenome of the male germ cell may underlie a subset of idiopathic cases of male infertility.

Initially, we investigated genetic causes of male infertility. Abnormal expression of protamine genes is associated with male infertility and reduced reproductive efficiency. Work from our lab and others has shown that abnormal protamine expression could not be explained by protamine gene body mutations.[1-6] As an alternative we tested whether the abnormal protamine ratio in sperm may be due to mutations in the 5' and 3' untranslated regions of the protamine genes, which are known to contain regulatory elements important for modulating gene transcription or gene translation. We identified a number of genetic variants in the protamine gene UTRs, however, these variants were present in similar frequencies in both patients and controls (Chapter 3). In a subsequent study performed by an independent group, a significant 5'UTR SNP that specifically associated with abnormal protamine was identified in patients with less than 9% normal sperm head morphology.[7] The differences reported between the two studies may be due to differences in patient selection criteria. Our patients were selected according to the P1/P2 ratio alone with no preset threshold for semen parameter such as morphology or motility. However, we have previously shown that patients with abnormal protamine ratio typically have poor semen parameters. An alternative explanation for the genetic differences reported in the two studies may be due to the differences in patient ethnicities.

Previous works has revealed that certain polymorphisms are more prevalent in some ethnic groups than in others.[8]

Although many of the SNPs reported to date in the protamine gene or UTRs alone do not confer susceptibility to protamine abnormality, one can speculate that if the protamine polymorphisms segregate with other polymorphisms located in nearby genes, then a particular haplotype may contribute to an altered P1/P2 ratio in patients. This aligns with the current understanding that many common diseases in humans are not caused by a single genetic variant within a single gene but are influenced by complex interactions among multiple genes as well as environmental and lifestyle factors. [9] The candidate gene approach applied for decades to identify genetic variants contributing to disease states has been slow, inefficient, and arduous. Whereas, the introduction of genome-wide association studies, in which millions of SNPs are assayed in large patient populations, have provided a new tool for investigating the genetic complexity of multifactorial diseases.[10-14] A low but significant number of SNPs have been associated with a number of diseases including infertility, however, the limitation has been the small effect conferred by common SNP variants to a disease state with a typical odd ratio of 1.1 -1.2.[9] A low odd ratio may seem clinically irrelevant for estimating the risk a particular haplotype poses, but can help identify biological pathways that may be used as potential drug targets to treat or impede disease progression.

Epigenetics and its Role in Infertility

Currently, the only treatment option for severe male factor infertility is intracytoplasmic sperm injection (ICSI), due in large part to the lack of understanding of

the underlying mechanism contributing to infertility. Although ICSI has revolutionized the treatment of male infertility and provided a quick fix for the couple undergoing infertility treatment, this treatment strategy may place parents at risk of transmitting genetic and epigenetic abnormalities to the progeny. The long-term implications and safety of assisted reproductive technologies (ART), specifically intracytoplasmic sperm injection (ICSI), have not been defined since the first ICSI child was conceived in the early '90s.[15]

With the growing understanding of the role and significance of epigenetics in cellular processes, developmental pathways, and disease states (i.e., complex and monogenic diseases such as cancer and imprinting diseases), we and other raised the question of whether epigenetic abnormalities in gametes may be associated with infertility or poor reproductive outcome. To achieve a better understanding of the role of epigenetics in infertility, we first investigated the epigenetic state of fertile men. This work revealed many new and interesting concepts, which revised our understanding of the potential paternal contribution to zygotic development. Our data suggest that the sperm epigenome retains a substantial amount of epigenetic information, although the vast majority of the genome is packaged with protamines. However, we find that the small amounts of nucleosomes retained were not randomly distributed, but enriched at many developmental promoters, miRNA genes and imprinted loci. These initial findings suggest that histone retention might be programmatic.

To infer a developmental role to the retained nucleosomes, canonical nucleosomes alone will be insufficient to differentiate retained vs newly incorporated

maternal histones shortly after fertilization, but rather these retained nucleosomes should be secondarily marked by methylation or acetylation, etc. Genome-wide maps for a number of modifications reveal that the sperm genome is in fact poised for two programs: the first is a gametogenesis program which is highly enriched with H3K4me3 (a mark of gene activation) and the second is a developmental program that correlates with H3K27me3 (a repressive mark) or the presence of two contrasting marks H3K4me and H3K27me- reminiscent of bivalent domains in embryonic stem cells. Furthermore, the majority of developmental transcription and differentiation factors were found to be hypomethylated,[16-18] but acquired methylation in a differentiated tissue (Chapter 5).

Retaining a stem-like state in a mature fully differentiated germ cell suggests either these marks have an instructive role in the developing embryo or are residual marks from the spermatogonial stem cell and a property of the germline. In short, direct evidence for either hypothesis is lacking. Prior to our work, it was long thought that the potential contribution from the paternal genome to the zygote was limited due to changes in chromatin composition during spermatogenesis, which resulted in the vast replacement of histones with protamines,[19,20] proteins not known to propagate information via modifications. One hypothesis is that the male genome does not convey gene packaging information (beyond the methylated imprinted genes), as it can simply be 'reprogrammed' and repackaged by the egg following fertilization to achieve totipotency. This notion has predominated given the active DNA demethylation observed in mice and humans shortly upon fertilization.[21,22] Although demethylation of a portion of the genome clearly occurs, this has only been studied in 'bulk' and we currently lack an understanding of which genes in the genome are susceptible or resistant to DNA

demethylation, beyond the known resistant imprinted genes. A full understanding of the methylation state of the gametes before and after fertilization will be an important step forward.

An alternative perspective is that the sperm chromatin is indeed poised and that the small amounts of histones retained in the paternal genome may be instructive. The conservation in chromatin packaging and DNA hypomethylation at many developmental loci in human, mouse, and zebrafish sperm genomes makes the genome-wide demethylation of the paternal genome all the more curious, if needed (in part) for totipotency, and raises many questions regarding the extent of embryonic reprogramming.[17,23,24]

Although a definitive answer for genome-wide reprogramming at fertilization in mammals is unknown, in early zebrafish embryos it was found that H3K4me and H3K27me marks were erased and later re-established after fertilization.[25] This could imply that the zebrafish zygotic genome may undergo complete erasure with epigenetic reestablishment occurring from a clean slate. Alternatively, germ cells may retain cellular memory but the levels of H3K4me and H3K27me may be extremely low and beyond antibody detection limits. In contrast to the finding in zebrafish, Brykczynska et al. (2010), mentioned in his discussion that paternal histones in both mouse and human embryos were retained in the embryo (unpublished data from Antione Peter's Lab).[24] Whether this discrepancy between human/mouse vs. zebrafish is true is uncertain.

However, an indirect approach to assess the potential significance of the sperm epigenome to development and fertility and overcome the known limitation of human genetic experiments will be to establish a direct correlation between alterations in

chromatin composition, or epigenetic state, to cases of male infertility or poor reproductive outcome. Animal studies [26] and human meta-analysis [27-30] have shown that ART increases the risk of imprinting disorders, but whether this is due to epigenetic alterations in the gametes or derived from in-vitro culture and manipulation is unclear. It was recently shown that embryo culture conditions induced persistent epigenetic changes at the agouti locus in mice. This is the first study to establish a strong correlation between preimplantation embryo growth conditions and persistent epigenetic changes.[31] Furthermore, we and others have concurrently showed that the epigenetic changes at maternally and paternally imprinted loci reside in the gametes of infertile men (Chapter 6), but whether these epimutations in sperm can be inherited is unknown.[32-38] Remarkably, when we were examining normally demethylated differentially methylated regions (DMRs), the alleles of infertile patients are often either unaffected, have a moderate number of CpGs affected, or entirely methylated. Clearly, complete methylation of a normally unmethylated locus may lead to an imprinting disorder in the embryo if proper imprint reestablishment mechanisms are not implemented.[39,40] Also the small number of CGs affected within some genes and alleles raise concerns regarding whether the few abnormally methylated Cs have reached a critical threshold that results in the recruitment of methyl binding proteins and DNA methyltransferases and lead to complete methylation in the embryo (at a certain unknown probability) and confer disease. Therefore, these data suggest that epigenetic changes, specifically DNA methylation in gamates, if not corrected may account for the increased incidence of imprinting disorders in offspring conceived by ART.

To determine whether epigenetic changes in gametes of infertile men were restricted to imprinted loci or extended genome-wide we characterized the sperm epigenome of two patient populations: infertile men with poor semen parameters and abnormal protamine ratios and the second population were recurrent poor reproductive outcome patients. Unlike fertile men, infertile men have nonprogrammatic histone retention genomewide, and the abnormal histone localization pattern is a result of an incomplete replacement of histone by protamine in the genome as whole. This results in an increased background of abnormally retained histone that obscures the pattern of appropriately retained histones seen in normal sperm. This incomplete removal of nucleosomes might be attributed to improper histone acetyl transferase activity in the round spermatid or due to the inefficiency in chromatin remodeling machinery.

These findings raise concerns regarding the clinical ramifications of an altered histone retention profile in the sperm of infertile men. Although the perdurance of sperm nucleosomes in the embryo has not been established, we speculate that if histones are replaced following fertilization, then nonprogrammatic histone retention may not necessarily affect the embryo other than potentially increasing levels of sperm DNA damage, which may be corrected by the oocyte's DNA repair machinery.[41] In contrast, if nucleosomes are retained in the male pronucleus but active genome-wide reprogramming in the male pronucleus requires naked DNA, then this non-programmatic histone retention may affect the efficiency and extent of reprogramming of the paternal pronucleus. Several studies, including recent nuclear transfer studies of either gynogenetic or androgenetic embryos, have suggested that the active demethylation of the paternal genome takes place during the remodeling of sperm chromatin following

fertilization but prior to pronuclear formation.[42,43] During this period of protein exchange the paternal DNA is presumably loosely packaged or possibly naked,[44] therefore, providing a unique opportunity for active DNA demethylation.

While nucleosomes were randomly retained genome-wide in infertile men, the localization of modified nucleosomes was similar to fertile controls suggesting that the establishment of epigenetic marks in the spermatogonial stem cell remains largely intact in infertile men, however, H3K4me and H3K27me levels were significantly reduced or absent at a number of developmental promoters and imprinted loci. The loss or reduction in H3K4me at the developmental and imprinted loci did not always correlate with a gain in DNA methylation in sperm. However, the loss of H3K4me may predispose these regions to acquiring DNA methylation in the early embryo since DNMT3A/B/3L specifically interact with the amino terminus of H3 and this interaction was strongly inhibited by H3K4me.[45]

Conclusion and Future Direction

Together this body of work provides a better understanding of the sperm genetic and epigenetic states and their potential role in spermatogenesis and fertility. Most importantly, we have revealed new features of germline chromatin, which has significantly expanded our understanding of the potential role of the sperm epigenome. Furthermore, we show for the first time genomewide changes in chromatin packaging and poising are common in the gametes of infertile men.

Given that in males the spermatogenic process is continuous throughout life, it is essential to elucidate how genetic and epigenetic processes are influenced by aging and

environmental cues. Recently, a sizeable body of evidence suggested that epigenetic alterations acquired with age might also be inherited by the offspring.[46-48] In humans, advanced paternal age (APA) is associated with an increased risk of neuro developmental and neuro psychiatric disorders such as schizophrenia, autism, bipolar disorder, dyslexia, reduced intelligence, and neural tube defects.[46-48] Therefore, evaluating semen samples collected at different time-points or environmental exposures will be important to identify whether methylation changes occur genomewide or at a subset of genes promoters (ex; promoters with or without repeat elements).

While it has long been known that the maternal contribution to the embryo is extensive, little is known about the genome status and the role of genome packaging (chromatin), modification (methylation) and expression (RNA) in the creation of a totipotent egg or embryo. Like sperm, a multitude of contributing mechanisms can be envisioned for the oocyte including the loading of maternal RNAs (coding and noncoding) that promote totipotency, the loading and function of key transcription factor proteins (including pluripotency/self renewal factors), and chromatin structures that enable (or prevent) the expression of particular embryonic developmental regulators.

A full understanding of both maternal and paternal chromatin status and the extent of embryonic reprogramming in early embryos will clarify the role and significance of germ cell chromatin to development, fertility, and transgenerational inheritance. These questions have recently become feasible as a result of the recent advances in sophisticated genomic methodologies such as high through-put genome-wide sequencing and tools for methylation analysis at single base pair resolution.

References

1. Schlicker M, Schnulle V, Schneppel L, Vorob'ev VI, Engel W. Disturbances of nuclear condensation in human spermatozoa: search for mutations in the genes for protamine 1, protamine 2 and transition protein 1. Hum Reprod 1994;9(12):2313-7.

2. Queralt R, Oliva R. Identification of conserved potential regulatory sequences of the protamine-encoding P1 genes from ten different mammals. Gene 1993;133(2):197-204.

3. de Yebra L, Ballesca JL, Vanrell JA, Bassas L, Oliva R. Complete selective absence of protamine P2 in humans. J Biol Chem 1993;268(14):10553-7.

4. Tanaka H, Miyagawa Y, Tsujimura A, Matsumiya K, Okuyama A, Nishimune Y. Single nucleotide polymorphisms in the protamine-1 and -2 genes of fertile and infertile human male populations. Mol Hum Reprod 2003;9(2):69-73.

5. Kichine E, Msaidie S, Bokilo AD, Ducourneau A, Navarro A, Levy N, Terriou P, Collignon P, Boetsch G, Chiaroni J and others. Low-frequency protamine 1 gene transversions c.102G->T and c.-107G->C do not correlate with male infertility. J Med Genet 2008;45(4):255-6.

6. Aoki VW, Christensen GL, Atkins JF, Carrell DT. Identification of novel polymorphisms in the nuclear protein genes and their relationship with human sperm protamine deficiency and severe male infertility. Fertil Steril 2006;86(5):1416-22.

7. Gazquez C, Oriola J, de Mateo S, Vidal-Taboada JM, Ballesca JL, Oliva R. A common protamine 1 promoter polymorphism (-190 C->A) correlates with abnormal sperm morphology and increased protamine P1/P2 ratio in infertile patients. J Androl 2008;29(5):540-8.

8. Spielman RS, Bastone LA, Burdick JT, Morley M, Ewens WJ, Cheung VG. Common genetic variants account for differences in gene expression among ethnic groups. Nat Genet 2007;39(2):226-31.

9. Eichler EE, Flint J, Gibson G, Kong A, Leal SM, Moore JH, Nadeau JH. Missing heritability and strategies for finding the underlying causes of complex disease. Nat Rev Genet 2010;11(6):446-50.

10. Aston KI, Krausz C, Laface I, Ruiz-Castane E, Carrell DT. Evaluation of 172 candidate polymorphisms for association with oligozoospermia or azoospermia in a large cohort of men of European descent. Hum Reprod 2010;25(6):1383-97.

11. Aston KI, Carrell DT. Genome-wide study of single-nucleotide polymorphisms associated with azoospermia and severe oligozoospermia. J Androl 2009;30(6):711-25.

12. Frazer KA, Ballinger DG, Cox DR, Hinds DA, Stuve LL, Gibbs RA, Belmont JW, Boudreau A, Hardenbol P, Leal SM and others. A second generation human haplotype map of over 3.1 million SNPs. Nature 2007;449(7164):851-61.

13. Hardy J, Singleton A. Genomewide association studies and human disease. N Engl J Med 2009;360(17):1759-68.

14. Franke A, McGovern DP, Barrett JC, Wang K, Radford-Smith GL, Ahmad T, Lees CW, Balschun T, Lee J, Roberts R and others. Genome-wide meta-analysis increases to 71 the number of confirmed Crohn's disease susceptibility loci. Nat Genet 2010;42(12):1118-25.

15. Palermo G, Joris H, Devroey P, Van Steirteghem AC. Pregnancies after intracytoplasmic injection of single spermatozoon into an oocyte. Lancet 1992;340(8810):17-8.

16. Down TA, Rakyan VK, Turner DJ, Flicek P, Li H, Kulesha E, Graf S, Johnson N, Herrero J, Tomazou EM and others. A Bayesian deconvolution strategy for immunoprecipitation-based DNA methylome analysis. Nat Biotechnol 2008;26(7):779-85.

17. 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(7254):473-8.

18. Weber M, Hellmann I, Stadler MB, Ramos L, Paabo S, Rebhan M, Schubeler D. Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. Nature Genetics 2007;39(4):457-66.

19. Wykes SM, Krawetz SA. Conservation of the PRM1 --> PRM2 --> TNP2 domain. DNA Seq 2003;14(5):359-67.

20. Ward WSC, D.S. . DNA packaging and organization in mammalian spermatozoa: comparison with somatic cells. . Biol Reprod 1991;44:569-74.

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

22. Mayer W, Niveleau A, Walter J, Fundele R, Haaf T. Demethylation of the zygotic paternal genome. Nature 2000;403(6769):501-2.

23. Arpanahi A, Brinkworth M, Iles D, Krawetz SA, Paradowska A, Platts AE, Saida M, Steger K, Tedder P, Miller D. Endonuclease-sensitive regions of human spermatozoal chromatin are highly enriched in promoter and CTCF binding sequences. Genome Res 2009;19(8):1338-49.

24. Brykczynska U, Hisano M, Erkek S, Ramos L, Oakeley EJ, Roloff TC, Beisel C, Schubeler D, Stadler MB, Peters AH. Repressive and active histone methylation mark distinct promoters in human and mouse spermatozoa. Nat Struct Mol Biol 2010;17(6):679-87.

25. Vastenhouw NL, Zhang Y, Woods IG, Imam F, Regev A, Liu XS, Rinn J, Schier AF. Chromatin signature of embryonic pluripotency is established during genome activation. Nature 2010;464(7290):922-6.

26. Young LE, Sinclair KD, Wilmut I. Large offspring syndrome in cattle and sheep. Rev Reprod 1998;3(3):155-63.

27. Gosden R, Trasler J, Lucifero D, Faddy M. Rare congenital disorders, imprinted genes, and assisted reproductive technology. Lancet 2003;361(9373):1975-7.

28. DeBaun MR, Niemitz EL, Feinberg AP. Association of in vitro fertilization with Beckwith-Wiedemann syndrome and epigenetic alterations of LIT1 and H19. Am J Hum Genet 2003;72(1):156-60.

29. Niemitz EL, DeBaun MR, Fallon J, Murakami K, Kugoh H, Oshimura M, Feinberg AP. Microdeletion of LIT1 in familial Beckwith-Wiedemann syndrome. Am J Hum Genet 2004;75(5):844-9.

30. Cox GF, Burger J, Lip V, Mau UA, Sperling K, Wu BL, Horsthemke B. Intracytoplasmic sperm injection may increase the risk of imprinting defects. Am J Hum Genet 2002;71(1):162-4.

31. Morgan HD, Jin XL, Li A, Whitelaw E, O'Neill C. The culture of zygotes to the blastocyst stage changes the postnatal expression of an epigentically labile allele, agouti viable yellow, in mice. Biol Reprod 2008;79(4):618-23.

32. Kagami M, Nagai T, Fukami M, Yamazawa K, Ogata T. Silver-Russell syndrome in a girl born after in vitro fertilization: partial hypermethylation at the differentially methylated region of PEG1/MEST. J Assist Reprod Genet 2007;24(4):131-6.

33. 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 2009;94(5):1728-1733.

34. Marques CJ, Francisco T, Sousa S, Carvalho F, Barros A, Sousa M. Methylation defects of imprinted genes in human testicular spermatozoa. Fertil Steril 2010;94(2):585-94.

35. Marques CJ, Carvalho F, Sousa M, Barros A. Genomic imprinting in disruptive spermatogenesis. Lancet 2004;363(9422):1700-2.

36. Marques CJ, Costa P, Vaz B, Carvalho F, Fernandes S, Barros A, Sousa M. Abnormal methylation of imprinted genes in human sperm is associated with oligozoospermia. Mol Hum Reprod 2008;14(2):67-74.

37. Kobayashi H, Hiura H, John RM, Sato A, Otsu E, Kobayashi N, Suzuki R, Suzuki F, Hayashi C, Utsunomiya T and others. DNA methylation errors at imprinted loci after assisted conception originate in the parental sperm. Eur J Hum Genet 2009;17(12):1582-91.

38. Kobayashi H, Sato A, Otsu E, Hiura H, Tomatsu C, Utsunomiya T, Sasaki H, Yaegashi N, Arima T. Aberrant DNA methylation of imprinted loci in sperm from oligospermic patients. Human Molecular Genetics 2007;16(21):2542-51.

39. Reik W, Santos F, Dean W. Mammalian epigenomics: reprogramming the genome for development and therapy. Theriogenology 2003;59(1):21-32.

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

41. Menezo Y, Dale B, Cohen M. DNA damage and repair in human oocytes and embryos: a review. Zygote 2010;18(4):357-65.

42. Barton SC, Arney KL, Shi W, Niveleau A, Fundele R, Surani MA, Haaf T. Genome-wide methylation patterns in normal and uniparental early mouse embryos. Human Molecular Genetics 2001;10(26):2983-7.

43. Nonchev S, Tsanev R. Protamine-histone replacement and DNA replication in the male mouse pronucleus. Mol Reprod Dev 1990;25(1):72-6.

44. Rodman TC, Pruslin FH, Hoffmann HP, Allfrey VG. Turnover of basic chromosomal proteins in fertilized eggs: a cytoimmunochemical study of events in vivo. J Cell Biol 1981;90(2):351-61.

45. Ooi SK, Qiu C, Bernstein E, Li K, Jia D, Yang Z, Erdjument-Bromage H, Tempst P, Lin SP, Allis CD and others. DNMT3L connects unmethylated lysine 4 of histone H3 to de novo methylation of DNA. Nature 2007;448(7154):714-7.

46. Frans EM, Sandin S, Reichenberg A, Lichtenstein P, Langstrom N, Hultman CM. Advancing paternal age and bipolar disorder. Arch Gen Psychiatry 2008;65(9):1034-40.

47. Garcia-Palomares S, Pertusa JF, Minarro J, Garcia-Perez MA, Hermenegildo C, Rausell F, Cano A, Tarin JJ. Long-term effects of delayed fatherhood in mice on postnatal development and behavioral traits of offspring. Biol Reprod 2009;80(2):337-42.

48. Garcia-Palomares S, Navarro S, Pertusa JF, Hermenegildo C, Garcia-Perez MA, Rausell F, Cano A, Tarin JJ. Delayed fatherhood in mice decreases reproductive fitness and longevity of offspring. Biol Reprod 2009;80(2):343-9.

APPENDIX

SEQUENCE ALTERATIONS IN THE YBX2 GENE ARE ASSOCIATED WITH MALE FACTOR INFERTILITY

Fertility and Sterility, 2009, Vol 91(4),
pages 1090-1095.

Sequence alterations in the YBX2 gene are associated with male factor infertility

Sue Hammoud, B.S.,[a,b] *Benjamin R. Emery, B.S.,*[a,b] *Diane Dunn, M.S.,*[c] *Robert B. Weiss, Ph.D.,*[c] *and Douglas T. Carrell, Ph.D.*[a,b,d]

[a] Andrology and In Vitro Fertilization Laboratories, Division of Urology, Department of Surgery; [b] Department of Physiology; [c] Department of Human Genetics; and [d] Department of Obstetrics and Gynecology, School of Medicine, University of Utah, Salt Lake City, Utah

Objective: To investigate YBX2 gene alterations in men with severe defects in spermatogenesis, including azoospermia or severe oligozoospermia, and protamine deregulation. MSY2 has been identified as a central component in the regulation of spermatogenesis in mice, but the potential role of its human orthologue, YBX2 or "Contrin," in human infertility is not known.
Design: A prospective cohort study.
Setting: University infertility clinic and associated research laboratory.
Patient(s): A total of 288 men were evaluated. Diagnoses were made of complete azoospermia, severe oligozoospermia, and protamine deregulation, or men were of known paternity.
Intervention(s): Deoxyribonucleic acid (from peripheral blood) and semen samples were collected and analyzed for gene mutations and semen parameters respectively.
Main Outcome Measure(s): YBX2 gene alterations.
Result(s): YBX2 sequence analysis revealed 15 polymorphic sites, of which seven polymorphisms were present at a statistically higher frequency in one or both of the patient populations than in controls. Of these seven, two resulted in an amino acid substitution in the highly conserved cold shock domain and one resulted in a highly significant synonymous change in exon 8 of infertile patients. The frequency of single nucleotide polymorphisms was significantly elevated in patients with infertility, particularly in men with abnormal protamine expression.
Conclusion(s): These data indicate a significant association between gene alterations in the YBX2 gene and abnormal spermatogenesis in humans, including a potential role in altering protamine expression, and implicate YBX2 gene alterations as a potential cause of male factor infertility. (Fertil Steril® 2009;91:1090–5. ©2009 by American Society for Reproductive Medicine.)

Key Words: YBX2, Contrin, SNPs, spermatogenesis, protamine, translation regulation, transcription factor

One of the distinctive characteristics of chromatin remodeling during spermatogenesis is the sequential transition of the histone-bound genome in primary spermatocytes to a protamine-bound genome in the elongating spermatid (1–4). This transition requires temporal uncoupling of DNA transcription and messenger RNA (mRNA) translation in the developing spermatid (5–7) and results in a highly condensed and likely transcriptionally silent chromatin structure (2). Therefore, it is believed that the regulation of protamine incorporation is related strongly to the temporal uncoupling of transcription and translation in the developing sperm (8).

The mechanism by which spermatogenic translational repression operates is not entirely understood. However, it is clear that messenger ribonucleoproteins (mRNPs), complexes of RNA-binding proteins (RBP) and mRNAs, are present in the round and elongating spermatid stages (9). The phosphorylated RBPs bind mRNAs with little sequence specificity (10) and associate with mRNPs to inhibit translation (11). Modifications to the mRNP complexes release translatable mRNAs to reinitiate protein synthesis.

A number of RNA-binding proteins in mice, and their human homologues, have been identified in germ cells (12–14). One family of RBPs is the Y-box protein family, which is known to include the mouse variants: MSY1, MSY2, and MSY4 (15–18). The mouse Y-box protein MSY2 is one of the most abundant DNA/RNA-binding proteins, constituting 0.7% of the total protein in the spermatid (19). In the mouse testis, MSY2 expression is highest in the round spermatid, where translational repression dominates, suggesting a role for MSY2 as an mRNA-stabilizing protein (20).

MSY2 may also function as a transcription factor in mammalian male germ cells because it binds to the Y-box sequence in the mouse protamine 2 promoter (21) and stimulates transcription from the mouse protamine 2 promoter in in vitro assays (22). It has also been postulated that MSY2 acts as

Received October 3, 2007; revised January 2, 2008; accepted January 3, 2008; published online March 12, 2008.
Reprint requests: Douglas T. Carrell, Ph.D., University of Utah IVF and Andrology Laboratories, 675 Arapeen Drive, Suite 205, Salt Lake City, UT 84108. (FAX: 1-801-581-6127; E-mail: douglas.carrell@hsc.utah.edu).

0015-0282/09/$36.00
doi:10.1016/j.fertnstert.2008.01.009

a nuclear shuttle protein, binding specific transcripts and removing them from the spermatocyte nucleus (19, 23, 24). These data suggest that MSY2 functions as a transcription factor promoting the transcription of testis-specific genes and also functions as a structural component of mRNPs in the cytoplasm of maturing germ cells (14, 25).

Gene-targeting studies have demonstrated that MSY2 knockout mice appear phenotypically normal but males are sterile (23). Spermatogenesis is disrupted in the postmeiotic null germ cells with many amorphous and multinucleated spermatids in the testis, but no spermatozoa are present in the epididymis (24). The efficiency of nuclear histones' replacement with protamines in sperm from these transgenic mice has not been reported, but it is likely that it has been affected. This conclusion is supported by the analysis of testicular sperm morphology, which indicates a failure of proper elongation of sperm, a process tightly coupled to protamine incorporation.

Recent studies demonstrate that a human homologue of MSY2 is expressed in human testis and is known as YBX2 or "Contrin" (26). To determine whether mutations in the YBX2 gene are associated with specific subtypes of male infertility, we evaluated the genomic sequence of YBX2 in patients with azoospermia, oligozoospermia, and protamine deficiency in comparison with men of known paternity.

MATERIALS AND METHODS

After obtaining University of Utah Institutional Review Board approval for all aspects of the study, 288 patients were recruited for YBX2 gene sequencing. These patients were classified into three groups: group 1 contained 96 men of known paternity from the Utah Genetic Reference Project (UGRP), group 2 contained 47 men with azoospermia and 49 men with severe oligozoospermia ($<5 \times 10^6$ sperm/mL), and group 3 included 96 infertile men with an abnormal protamine 1 to protamine 2 ratio (P1/P2). The above samples have been screened previously for alterations in the protamine and transition protein gene regions (27). Patients with abnormal protamine expression were selected on the basis of previous work identifying YBX2 as an important component of spermatogenic gene expression (16, 19). To increase the chance of identifying a mutation in the YBX2 gene responsible for male infertility and to decrease possible confounding factors, patients were excluded if they had any suspected or known causes of male factor infertility. This criterion excluded patients with a known Y-chromosome microdeletion, cystic fibrosis, varicocele, Klinefelter's syndrome, or prior exposure to chemotherapeutics or radiation. Deoxyribonucleic acid was extracted from peripheral lymphocytes with use of the Puregene DNA extraction kit (Gentra Systems, Minneapolis, MN).

Evaluation of Sperm P1/P2 Ratio

Study participants with an abnormal protamine ratio were identified with use of nuclear protein extraction, gel electrophoresis, and densitometry analysis. Sperm nuclear proteins were extracted from cryopreserved semen aliquots as previously described (28). Gel electrophoresis reagents were obtained from Bio-Rad Laboratories (Hercules, CA). All other reagents were obtained from Sigma Chemical Company (St. Louis, MO).

Screening for Mutations

Primer sets were designed and optimized to amplify the exonic regions of YBX2 with their bordering intronic sequences by standard polymerase chain reaction (PCR) techniques. Primary PCR products were cleaned with use of ExoSAP-IT (USB, Cleveland, OH) and sequenced in the forward and reverse direction.

Samples were sequenced with use of ABI 3700 capillary sequencer. Sequence traces were assembled with use of the Phrap software (available from http://www.phrap.org) and analyzed for significant changes with use of Phred and Consed. Phred assigns a quantitative value to quality of each sequenced base. This base quality provides a probabilistic estimate of the correctness of the base call.

Potential mutations were identified with use of Consed, which has the ability to search for high-quality base discrepancies in the assembled sequences. Visual analysis was also used to confirm identified polymorphisms and potential mutations. The frequencies of novel single nucleotide polymorphisms (SNPs) were compared by χ^2 analysis with Fisher correction when necessary and use of SPSS (SPSS Inc., Chicago, IL) and Intercooled Stata (StataCorp, College Station, TX).

RESULTS

Direct sequencing of the YBX2 gene in the control and study populations revealed 13 SNPs, one dinucleotide deletion, and a single trinucleotide deletion (Table 1). These polymorphic sites were found in exonic and neighboring intronic regions. The frequencies of the heterozygous and homozygous mutations are reported in Table 1. Gene alterations were considered significant at a P value $\leq .05$.

Among these gene alterations, four are exonic polymorphisms that resulted in either an amino acid change or amino acid deletion. Three of these SNPs localized to the highly conserved region of exon 1 and one to exon 5. Two of the three SNPs occurring in exon 1 were found at a significantly increased frequency in either one or both of the study population groups (oligozoospermic/azoospermic or abnormal P1/P2) when compared with controls (UGRP patients), whereas the Ser53-Pro amino acid change in exon 1 was present at similar frequencies in all populations tested. The Val9-Gly change in exon 1, which occurs at a significantly higher incidence in men with protamine abnormalities than in UGRP controls, has been reported previously as a miscellaneous difference between a reported mRNA sequence and the human genome (build 36) (29). However, the previously unreported Pro80-Gln alteration in exon 1 is not found in any of the sequenced UGRP controls but is prevalent in both patient groups (oligozoospermic/azoospermic or abnormal P1/P2)

TABLE 1

Description of identified gene polymorphisms in the human YBX2 gene.

Gene alteration	Region	No. of patients scored for control/ azoospermia / abnormal P1/P2	UGRP (control) genotype allelic frequencies	Azoospermic/ oligozoospermic genotype allelic frequencies	Abnormal P1/ P2 genotype allelic frequencies	Amino acid change	Control vs. patients with azoospermia (*P* value)	Control vs. abnormal P1/P2 (*P* value)
gggG/Atcg	5′UTR	48/48/80	GG: 1 GA: 0 AA: 0	GG: 1 GA: 0 AA: 0	GG: 0.987 GA: 0.013 AA: 0	N/A	NS	NS
cggT/Gggc	Exon 1	44/54/84	GG: 0.063 GT: 0.238 TT: 0.698	GG: 0.055 GT: 0.205 TT: 0.740	GG: 0.149 GT: 0.414 TT: 0.437	**Val-Gly**	NS	**.002**
ggcC/Tccc	Exon 1	40/42/85	CC: 0.289 CT: 0.409 TT: 0.303	CC: 0.349 CT: 0.422 TT: 0.229	CC: 0.188 CT: 0.582 TT: 0.229	**Ser-Pro**	NS	NS
cccC/Aggc	Exon 1	48/40/80	CA: 0 CC: 1	CA: 0.062 CC: 0.938	CA: 0.037 CC: 0.963	**Pro-Gln**	**.019**	**.08**
ttt-/Tctt	Intron 1	48/48/80	–/–: 1 –/T: 0	–/–: 0.99 –/T: 0.01	–/–: 1 –/T: 0	N/A	NS	NS
ggaC/Tagg	Exon 5	88/88/95	CC: 1 CT: 0	CC: 1 CT: 0	CC: 0.989 CT: 0.011	No change	NS	NS
tctTCT/-acc	Exon 5	88/88/95	–/TCT: 0 TCT/TCT: 1	–/TCT: 0 TCT/TCT: 1	–/TCT: 0.011 TCT/TCT: 0.989	**Phe deletion**	NS	NS
ccgA/Tcgg	Exon 5	88/88/95	AA: 1 AT: 0 TT: 0	AA: 1 AT: 0 TT: 0	AA: 0.989 AT: 0.011 TT: 0	No change	NS	NS
tccC/Tctt	Intron 6	70/45/45	CC: 1 CT: 0 TT: 0	CC: 1 CT: 0 TT: 0	CC: 0.901 CT: 0.041 TT: 0.058	N/A	NS	**.007**
agcC/Tttg	Intron 6	61/76/57	CT: 0.328 CC: 0.475 TT: 0.197	CT: 0.368 CC: 0.539 TT: 0.092	CT: 0.018 CC: 0.825 TT: 0.158	N/A	NS	**.0001**
tgt-/Tcct	Intron 6	49/73/60	–/–: 0.429 –/T: 0.429 T/T: 0.143	–/–: 0.548 –/T: 0.356 T/T: 0.096	–/–: 0.339 –/T: 0.603 T/T: 0.058	N/A	NS	NS
tgaCC/-aac	Intron 7	48/48/80	–/CC: 0 CC/CC: 1	–/CC: 0 CC/CC: 1	–/CC: 0.022 CC/CC: 0.978	N/A	NS	.003

Hammoud. Male factor infertility and the YBX2 gene. Fertil Steril 2009.

TABLE 1

Continued.

Gene alteration	Region	No. of patients scored for control/ azoospermia / abnormal P1/P2	UGRP (control) genotype allelic frequencies	Azoospermic/ oligozoospermic genotype allelic frequencies	Abnormal P1/ P2 genotype allelic frequencies	Amino acid change	Control vs. patients with azoospermia (*P* value)	Control vs. abnormal P1/P2 (*P* value)
aaaC/Acaa	Intron 7	41/41/52	AA: 0 AC: 0 CC: 1	AA: 0 AC: 0 CC: 1	AA: 0.288 AC: 0.288 CC: 0.423	N/A	NS	**.0001**
tcaA/Gagg	Exon 8	49/83/52	AA: 1 AG: 0 GG: 0	AA: 0.434 AG: 0.361 GG: 0.205	AA: 0.404 AG: 0.308 GG: 0.288	No change	**.0001**	**.0001**
ctcC/Tccc	Exon 9	63/79/63	CC: 1 CT: 0	CC: 0.975 CT: 0.025	CC: 1 CT: 0	N/A	NS	NS

Note: Gene alterations are reported as major allele/minor allele; frequencies are divided into homozygous and heterozygous changes. N/A= not applicable; NS = not significant. Bold values highlight significant *P* values for SNPs. Bold characters emphasize amino acid changes or deletions.

Hammoud. Male factor infertility and the YBX2 gene. Fertil Steril 2009.

at an allelic frequency of 0.062 (P=.019) for the men with azoospermia or oligozoospermia and 0.037 (P=.08) in the abnormal protamine group. The Phe229 deletion (the deletion does not cause any frameshift events or early termination sites) in exon 5 was present only in the patients with abnormal protamine but absent in the remaining two populations (oligozoospermic/azoospermic and UGRP). However, the prevalence of this deletion was very low and did not achieve significance.

In addition, three synonymous SNPs were identified in exons 5 and 8. The SNP in exon 8 is drastically increased in infertile patients (those with oligozoospermia or azoospermia or abnormal P1/P2) and absent from the control population (UGRP). Single nucleotide polymorphisms identified in the 5′ UTR and in the noncoding region of exon 9 are limited to infertile patients but are present at insignificant frequencies (Table 1). The remaining six SNPs were localized to intron/exon borders; four are significantly increased in the study populations (Table 1).

When calculating the overall allelic frequency of gene alterations in the three different patient populations, we note that the frequency is significantly higher in the infertile patients (oligozoospermic/azoospermic or abnormal P1/P2) when compared with controls. However, when we calculated the allelic frequency of polymorphisms that resulted in amino acid changes, the frequency was not different for patients with azoospermia or oligozoospermia when compared with controls but was significantly higher in patients with abnormal protamine expression. The corresponding frequencies and *P* value results are summarized in Table 2.

DISCUSSION

This is the first report to establish a potential link between genomic alterations of a key spermatogenesis regulator (YBX2) and male factor infertility. The subgroups of infertility patients tested are men with azoospermia or severe oligozoospermia and men with protamine deregulation. Variations in the number of reads sequenced for each gene alteration are dependent on Guanine and Cytosine (GC) content, the number of repetitive nucleotides, and DNA fidelity. Among the identified novel gene polymorphisms reported here are several alterations that correlate with sperm production and/or sperm function, in that there is an overall increase in the frequency of men carrying any one of the identified polymorphisms in the azoospermic/oligospermic or abnormal protamine groups (16.50% and 25.20%, respectively) when compared with the UGRP (fertile control) population (11.60%) (Table 2). However, these data indicate that gene alterations of the YBX2 gene and bordering intronic areas are far more common in these two subgroups of infertile men than is seen in the general population. Thus our findings suggest that the regulatory role of YBX2 in regard to male factor fertility and specifically to patients with abnormal protamine expression is paramount, just as has been shown in the mouse model with the MSY2 homologue (19, 23, 24, 30).

The gene alterations identified in exon 1 become more interesting when one considers that this is a highly conserved

TABLE 2

Significant categories of gene alterations by subject groups and comparisons between infertile men and controls.

	Total No. of alleles observed	No. of unique alterations	UGRP (control) (%)	Azoospermic/ oligozoospermic[a] (%)	Abnormal P1/P2[a] (%)	Azoospermic/ oligozoospermic vs. control (*P* value)
Total gene alterations	2812	15	11.60	16.50	25.20	<.001
Gene alterations with statistical difference between controls and study groups	1150	7	11.43	21.50	28.50	<.0001
Gene alterations that cause aa changes	925	4	21.10	24.00	40.5	NS
Gene alterations with statistical difference between controls and study groups and cause aa changes	489	2	12.00	15.50	31.00	NS

Note: Gene alterations are reported as a percentage of the total number of alleles analyzed carrying any one of the alterations from the category given in the left-hand column. NS = not significant; aa = amino acid.

[a] $P<.0001$, abnormal vs. control.

Hammoud. Male factor infertility and the YBX2 gene. Fertil Steril 2009.

cold shock domain among mammalian species and contains a putative casein kinase 2 phosphorylation site (26). The Val9-Gly substitution is in close proximity to this putative region, whereas the remaining polymorphisms in this exon are in highly conserved regions from mouse to human (26).

It is intriguing that four of the nine polymorphisms identified in the nearby intronic and noncoding regions occur at significantly higher frequencies in the infertile groups than in the UGRP controls. These are all novel gene alterations and are not overlapping with genes in the plus or minus direction of the YBX2 gene. Further analysis of sequence alterations in intronic and noncoding regions is required to determine whether these modifications potentially may alter mRNA splicing or create de novo splice sites that may alter spermatogenesis.

The significance of the findings reported here is threefold. First, this is the first report of an infertility gene that has been characterized with several significant and potentially meaningful alterations that are likely associated with spermatogenesis. Second, these data clearly indicate that infertility patients have a significantly increased frequency of gene alterations compared with the control patients (UGRP) in the YBX2 gene. Third, many of the identified alterations are novel. Currently this proposition is limited to the populations of men with decreased or few sperm in the ejaculate and men with an abnormal pattern of protamine expression; however, a strong relationship between YBX2 gene fidelity and male factor infertility looks promising. Currently there is no proved causal association between YBX2 alteration and protamine deregulation, but it may be one of the potential causes. However, it has been hypothesized that the manifestation of protamine deregulation in men is an indication of a more global spermatogenic defect (8). Future studies will examine the relationship of YBX2 regulatory mechanisms, sperm concentration, and protamine deregulation.

REFERENCES

1. Oliva R, Bazett-Jones D, Mezquita C, Dixon GH. Factors affecting nucleosome disassembly by protamines in vitro. Histone hyperacetylation and chromatin structure, time dependence, and the size of the sperm nuclear proteins. J Biol Chem 1987;262:17016–25.
2. Steger K. Transcriptional and translational regulation of gene expression in haploid spermatids. Anat Embryol (Berl) 1999;199:471–87.
3. Hecht NB. Regulation of 'haploid expressed genes' in male germ cells. J Reprod Fertil 1990;88:679–93.
4. Dadoune JP. The nuclear status of human sperm cells. Micron 1995;26: 323–45.
5. Steger K, Klonisch T, Gavenis K, Drabent B, Doenecke D, Bergmann M. Expression of mRNA and protein of nucleoproteins during human spermiogenesis. Mol Hum Reprod 1998;4:939–45.
6. Steger K, Klonisch T, Gavenis K, Behr R, Schaller V, Drabent B, et al. Round spermatids show normal testis-specific H1t but reduced cAMP-responsive element modulator and transition protein 1 expression in men with round-spermatid maturation arrest. J Androl 1999;20:747–54.
7. Steger K, Fink L, Klonisch T, Bohle RM, Bergmann M. Protamine-1 and -2 mRNA in round spermatids is associated with RNA-binding proteins. Histochem Cell Biol 2002;117:227–34.
8. Carrell DT, Emery BR, Hammoud S. Altered protamine expression and diminished spermatogenesis: what is the link? Hum Reprod Update 2007;13:313–27.
9. Kleene KC. Patterns of translational regulation in the mammalian testis. Mol Reprod Dev 1996;43:268–81.
10. Marello K, LaRovere J, Sommerville J. Binding of Xenopus oocyte masking proteins to mRNA sequences. Nucleic Acids Res 1992;20: 5593–600.
11. Richter JD, Smith LD. Reversible inhibition of translation by Xenopus oocyte-specific proteins. Nature 1984;309:378–80.
12. Cooke HJ, Elliott DJ. RNA-binding proteins and human male infertility. Trends Genet 1997;13:87–9.
13. Wolffe AP, Tafuri S, Ranjan M, Familari M. The Y-box factors: a family of nucleic acid binding proteins conserved from Escherichia coli to man. New Biol 1992;4:290–8.
14. Wolffe AP. Structural and functional properties of the evolutionarily ancient Y-box family of nucleic acid binding proteins. Bioessays 1994;16: 245–51.
15. Davies HG, Giorgini F, Fajardo MA, Braun RE. A sequence-specific RNA binding complex expressed in murine germ cells contains MSY2 and MSY4. Dev Biol 2000;221:87–100.
16. Gu W, Tekur S, Reinbold R, Eppig JJ, Choi YC, Zheng JZ, et al. Mammalian male and female germ cells express a germ cell–specific Y-box protein, MSY2. Biol Reprod 1998;59:1266–74.
17. Tafuri SR, Familari M, Wolffe AP. A mouse Y box protein, MSY1, is associated with paternal mRNA in spermatocytes. J Biol Chem 1993;268: 12213–20.
18. Tafuri SR, Wolffe AP. Selective recruitment of masked maternal mRNA from messenger ribonucleoprotein particles containing FRGY2 (mRNP4). J Biol Chem 1993;268:24255–61.
19. Yang J, Medvedev S, Reddi PP, Schultz RM, Hecht NB. The DNA/RNA-binding protein MSY2 marks specific transcripts for cytoplasmic storage in mouse male germ cells. Proc Natl Acad Sci U S A 2005;102: 1513–8.
20. Oko R, Korley R, Murray MT, Hecht NB, Hermo L. Germ cell–specific DNA and RNA binding proteins p48/52 are expressed at specific stages of male germ cell development and are present in the chromatoid body. Mol Reprod Dev 1996;44:1–13.
21. Nikolajczyk BS, Murray MT, Hecht NB. A mouse homologue of the Xenopus germ cell–specific ribonucleic acid/deoxyribonucleic acid–binding proteins p54/p56 interacts with the protamine 2 promoter. Biol Reprod 1995;52:524–30.
22. Yiu GK, Murray MT, Hecht NB. Deoxyribonucleic acid–protein interactions associated with transcriptional initiation of the mouse testis–specific cytochrome c gene. Biol Reprod 1997;56:1439–49.
23. Yang J, Medvedev S, Yu J, Tang LC, Agno JE, Matzuk MM, et al. Absence of the DNA-/RNA-binding protein MSY2 results in male and female infertility. Proc Natl Acad Sci U S A 2005;102:5755–60.
24. Yang J, Morales CR, Medvedev S, Schultz RM, Hecht NB. In the absence of the mouse DNA/RNA-binding protein MSY2, messenger RNA instability leads to spermatogenic arrest. Biol Reprod 2007;76: 48–54.
25. Ladomery M, Sommerville J. A role for Y-box proteins in cell proliferation. Bioessays 1995;17:9–11.
26. Tekur S, Pawlak A, Guellaen G, Hecht NB. Contrin, the human homologue of a germ-cell Y-box–binding protein: cloning, expression, and chromosomal localization. J Androl 1999;20:135–44.
27. Aoki VW, Christensen GL, Atkins JF, Carrell DT. Identification of novel polymorphisms in the nuclear protein genes and their relationship with human sperm protamine deficiency and severe male infertility. Fertil Steril 2006;86:1416–22.
28. Carrell D, Liu L. Altered protamine 2 expression is uncommon in donors of known fertility, but common among men with poor fertilizing capacity, and may reflect other abnormalities of spermiogenesis. J Androl 2001;22:604–10.
29. Strausberg RL, Feingold EA, Grouse LH, Derge JG, Klausner RD, Collins FS, et al. Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences. Proc Natl Acad Sci U S A 2002;99:16899–903.
30. Yang J, Medvedev S, Yu J, Schultz RM, Hecht NB. Deletion of the DNA/RNA-binding protein MSY2 leads to post-meiotic arrest. Mol Cell Endocrinol 2006;250:20–4.