DNA METHYLATION IN MALE GERM CELLS: THE ACQUISITION AND MAINTENANCE OF UNIQUE GENOME-WIDE PATTERNS

by

Christopher Charles Oakes

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of Doctor of Philosophy

February 2007

Department of Pharmacology and Therapeutics
McGill University
Montreal, Quebec, Canada

Your file Votre référence
ISBN: 978-0-494-32374-8
Our file Notre référence
ISBN: 978-0-494-32374-8

This thesis is dedicated to...

...My parents who have always stood by me and have inspired me to do my best - without their relentless support, none of this would have been possible; to my brother and sister, the best siblings one could have; and to my grandparents, especially in memory of my grandfather, who was eternally supportive of my education.

...And to my wife, whose patience, encouragement and love were essential elements in the completion of this work. I could not have done it without you.

"We shall not cease from exploration, and the end of all our exploring will be to arrive where we started and know the place for the first time."

T. S. Eliot

ABSTRACT

The development of healthy gametes is paramount to the health of progeny and to the survival of a species. Epigenetic information contained within gametic DNA in the form of DNA methylation is essential for germ cell and embryo development. DNA methylation is a genome-wide phenomenon involved in the control of gene expression and chromosome structure and stability. During germ line development, patterns of DNA methylation are established in a sex- and sequence-specific manner. The primary goal of the work presented in this thesis is to gain an understanding of the nature of the genome-wide pattern of DNA methylation in germ cells and to study its progression during germ cell development. The complexity of male germ cell development has been well studied in mice and thus makes an excellent system in which to study germ cell DNA methylation. Firstly, genome-wide patterns of DNA methylation in adult male germ cells were determined using a variety of techniques. Results from these studies demonstrate that the DNA methylation pattern in male germ cells is highly distinct from that of somatic cells. The reorganization of the germ cell pattern is associated with chromosomal features such as the chromosomal banding pattern and regional GC content. Secondly, by examining purified populations of male germ cells, we have determined that patterns of DNA methylation are being acquired during spermatogenesis. *De novo* methylation and demethylation events occur in a sequence-specific manner prior to the meiotic phase of germ cell development. Finally, the stability of these patterns was studied by perturbing DNA methyltransferase activity. The study of germ cells lacking a functional *Dnmt3L* gene demonstrates that the abnormalities displayed in these cells are associated with a failure to acquire normal levels of DNA methylation. In addition, the treatment of mice with the hypomethylating agent, 5-aza-2'-deoxycytidine, results in adverse effects on sperm function and is associated with sequence-specific hypomethylation. Collectively, these studies have uncovered several novel aspects of DNA methylation in male germ cells and contribute to our understanding of the role(s) for epigenetic phenomena in the development and maintenance of healthy gametes.

RÉSUMÉ

La formation de gamètes sains est cruciale au bon développement de la progéniture et à la survie de l'espèce. L'information épigénétique contenue dans l'ADN des gamètes, sous la forme d'ADN méthylé, est essentielle au développement des cellules germinales et de l'embryon. La méthylation de l'ADN est un phénomène qui s'étend à l'ensemble du génome et qui est impliqué dans le contrôle de l'expression des gènes tout en contribuant à la stabilité et à la structure des chromosomes. Les patrons de méthylation sont établis durant le développement des cellules germinales en fonction du genre de l'organisme et de la nature des séquences étudiées. Le but principal des travaux présentés dans cette thèse était d'approfondir notre compréhension de la nature des patrons de méthylation de l'ADN de la cellule germinale ainsi que d'étudier comment ces patrons sont modifiés durant le développement du gamète chez la souris. Comme le développement des cellules germinales mâles a été bien étudié, nous avons choisi d'utiliser ce système pour nos études. Premièrement, nous avons déterminé les patrons génomiques de méthylation de l'ADN du spermatozoïde en utilisant diverses techniques. Ces études ont démontré que le patron de méthylation des cellules germinales mâles est hautement distinct de celui des cellules somatiques. La réorganisation du patron gamétique est associée avec des caractéristiques chromosomiques telles que le patron des bandes chromosomiques ainsi que le contenu régional en GC. Deuxièmement, en examinant diverses populations de cellules germinales purifiées, nous avons déterminé que les patrons de méthylation sont acquis au cours de la spermatogenèse. Des évènements de méthylation *de novo* et de déméthylation se produisent avant la méiose. Finalement, nous avons étudié la stabilité de ces patrons en perturbant l'activité des ADN méthyltransférases, les enzymes responsables de l'établissement des patrons de méthylation. L'étude de cellules germinales ne comprenant pas de gène *Dnmt3L* fonctionnel a démontré que le phénotype adverse présenté par ces cellules était du à leur incapacité à acquérir des taux de méthylation normaux. Aussi, le traitement de souris avec l'agent hypométhylant 5-aza-2'-deoxycytidine a résulté en des conséquences adverses

sur la fonction du spermatozoïde et fut associé avec l'hypométhylation de séquences spécifiques. Ensemble, ces études mettent en lumière plusieurs nouveaux aspects de la méthylation de l'ADN chez le gamète mâle et contribuent à notre compréhension du rôle des phénomènes épigénétiques dans le développement de gamètes sains.

TABLE OF CONTENTS

LIST OF FIGURES

Chapter 1:

Chapter 2:

Chapter 3:

Chapter 6:

LIST OF TABLES

Chapter 5:

FORMAT OF THE THESIS

This thesis comprises six chapters, four chapters of which in the form in which they were submitted or will be submitted for publication. Chapter I, is an introduction that includes background material relevant to this thesis. Chapter II has been published in the journal, *Epigenetics* (1:146-152, 2006). Chapter III has been published in the journal, *Proceedings of the National Academy of Sciences* (104(1):228-33, 2007). Chapter IV is currently under consideration for publication in the journal, *Developmental Biology.* Chapter V has been submitted to *The Journal of Pharmacology and Experimental Therapeutics.* Connecting texts are provided in accordance with Section C of the Guidelines for Submitting a Doctoral or Master's Thesis as a manuscript based thesis.

CONTRIBUTION OF AUTHORS

For chapters II, III and IV, the candidate designed the studies and performed experimental assays and data analysis. The three manuscripts were written by the candidate with assistance from the co-authors. Chapter V is a collaborative effort equally performed by the candidate and Dr. Tamara Kelly. All work was conducted under the supervision of Dr. Jacquetta Trasler and Dr. Bernard Robaire.

In chapter II, the DNA methylation assay, quantitative analysis of DNA methylation by real-time PCR (qAMP), was conceived and developed by the candidate. Tissue collection, DNA extraction, DNA methylation analysis by restriction landmark genomic scanning (RLGS) and qAMP and all data analysis were performed by the candidate. Some of the qAMP data was generated by Liyuan Deng. qAMP primer design was performed by Dr. Sophie La Salle. The bisulfite sequencing analysis of the *U2af1-rs1* gene was performed with the assistance of Nicole Darricarrere.

In chapter III, tissue collection, DNA extraction, RLGS, RLGS spot identification, qAMP and all data analysis were performed by the candidate. Some of the primers for the qAMP assay were designed by Dr. Sophie La Salle. Mouse crosses and isolation of spermatogonia were carried out by Dr. Sophie La Salle. *Dnmt3L*-deficient mice were obtained from Tim Bestor (Columbia University, NY). *Oct4*/GFP mice were obtained from Hans Scholer (Max Planck Institute, Göttingen, Germany).

In chapter IV, tissue collection, DNA extraction, RLGS, RLGS spot identification, qAMP, Southern blots and all data analysis were performed by the candidate. Testicular germ cells were isolated by Dr. Sophie La Salle. The IAP probe was obtained from T. Bestor.

Chapter V was a collaborative effort of the candidate and Dr. Tamara Kelly. Animal treatments, sperm analysis and measurement of embryonic development were performed by Dr. Tamara Kelly. DNA extraction, Southern

blots, RLGS and qAMP assays were performed by the candidate. The manuscript was co-written by Dr. Tamara Kelly and the candidate.

ABBREVIATIONS AND NOMENCLATURE

Ag	type A spermatogonia
CASA	computer-assisted sperm analysis
C_t	threshold cycle value
ΔC_t	change in threshold cycle value
DM	differentially methylated
DMR	differentially methylated region
DNMT	DNA methyltransferase
dpc	day post-coitum
dpp	day post-partum
ERVK	endogenous retrovirus-K
ES cell	embryonic stem cell
ETn	early transposon
FACS	fluorescence activated cell sorting
GFP	green fluorescent protein
IAP	intracisternal A particle
ICF	immunodeficiency, centromeric region instability and facial anomalies syndrome
L/Z	leptotene/zygotene spermatocytes
LINE	long interspersed transposable element
LTR	long terminal repeat
MaLR	mammalian retrotransposon-like
MT	mouse transcript
Pa	pachytene spermatocytes
p-Ag	primitive type-A spermatogonia
PGC	primordial germ cell
EP	early pachytene spermatocytes
qAMP	quantitative analysis of DMA methylation using real-time
PCR	polymerase chain reaction
qPCR	quantitative or Real-Time PCR, RT-PCR

Rd	round spermatids
RLGS	restriction landmark genomic scanning
RNAi	RNA interference
RLTR	retrotransposon long terminal repeat
SAH	S-adenosylhomocysteine
SAM	S-adenosylmethionine
SD	standard deviation
SEM	standard error of the mean
SINE	short interspersed transposable element
TE	transposable element

NOMENCLATURE

The nomenclature used throughout this thesis is based on the conventions described by Maltais *et al.*, (2002), Roberts *et al.*, (2003) and Wain *et al.*, (2002). Mouse genes are written in lowercase italics (e.g. *Dnmt3a)*, transcripts are in lowercase letters (e.g. Dnmt3a), while proteins are in uppercase letters (e.g. DNMT3a); human genes, transcripts and proteins are distinguished by the use of uppercase letters. Only genes are italicized. A lower case suffix is used when referring to the murine protein (e.g. DNMT3a). The only exception to these conventions is for *Dnmt3L,* as the "L" will always be written using an uppercase letter because the use of a lowercase "l" can be confused with the number "one".

ACKNOWLEDGEMENTS

It seems like a lifetime ago when I think back to the time when I started on the path of becoming a scientist. Although, part of the beauty of this path is that it is a never ending journey, I honestly cannot say that there wasn't a single step that I have taken thus far where there wasn't someone helping me along the way.

The mistakes, however, are all my own.

There are many people to whom I owe many thanks. I hope one day that I can reciprocate their generosity.

First and foremost, I would like to thank my supervisors, **Dr. Jacquetta Trasler** and **Dr. Bernard Robaire**. Thank you both for being so supportive of my ideas and enthusiasm. I really feel that the freedom that you have given to me over the years to pursue many different aspects of my projects has given me a wealth of knowledge and experience. The lessons learned from turning failure into success has given me the confidence one needs for future endeavors and taught me more about being a scientist than anything else. You have both been excellent mentors, and not just about how to succeed in science.

I am very grateful to my collaborators, **Dr. Christoph Plass** and **Dr. Dominic Smiraglia**. Thank you, Christoph, for welcoming me into your lab and for your support over the years. I only hope that I may one day assist a young graduate student the way that you did for me (and continue to do so). Dominic, I have to thank you for many things: your assistance, guidance, patience and for being a good friend. It's nice to share ideas with someone who is as equally enthusiastic about science as yourself and to have someone you can always count on.

To the **Department of Pharmacology** and **The Montreal Children's Hospital Research Institute**: thank you **Helene**, **Tina** and **Pam** for being very supportive over the years; Helene you were a wonderful help to me over the years. Thanks to **Eric Simard** and **Margaret Holding** for keeping everything running smoothly and to the Institute for funding my endavours. I would also like to thank **Liyuan Deng** for her assistance and excellent expertise. Your much appreciated help came at critical times.

To **Sophie La Salle,** my 'male partner', I can honestly say that I couldn't have done it without you. Thanks for always offering to help with so many aspects of my projects. You were a great listener when I needed to talk out my ideas and many of your ideas contributed greatly to my work. We made a pretty good team - I hope we can work together again some day.

To the members of the Trasler lab, past and present, you all helped me along the way, most importantly in making an environment that was fun to be a part of. I would like to thank **Tonia Doerksen** for getting me started and for putting up with a very enthusiastic and naïve young student; **Tamara Kelly** for your hard work and collaboration; **Diana Lucifero** for demonstrating how a Ph.D should be done; **Marc Toppings** and the Friday Thompson house crew for the excellent, pertinent debates about the ethical aspects of science; **Oana Neaga** for your brilliantly easy-going personality that was always a joy to be around; the next generation: **Donovan Chan, Amanda Fortier, Flavia Lopes, Kirsten Niles, Wells Cushnie and Nicole Darricarrere** you made my final year or so in the lab an equally enjoyable time - I think I gained more knowledge than I passed on; and finally, a huge thanks to **Josee Martel** who manages to balance it all, thank you so much for always dropping what you were doing so that 'Mr. Rush' always kept on schedule.

To the 1st floor Pharmacology crew, there are so many of you and you all deserve my thanks. In particular, I would like to thank **Sara Ali-Khan** for having the most

wonderful personality and truly being an inspiration; **Rob Vinson** for being a person of diverse talent and good friend; and to **Kate Jervis** for being a very special person - your encouragement has always meant a lot.

CHAPTER I

INTRODUCTION

Germ cells provide the essential link between generations. They are the only cells that naturally carry out the process of transmitting genetic information, making them of paramount importance to the survival and evolution of a species. Germ cells have highly specialized characteristics that allow for this process to reliably take place; yet, from two single cells, the myriad of cell-types that make up an individual are derived. It is in this sense that germ cells embody the potential for cellular totipotency and a species' immortality.

The birth of a healthy child is the successful result of an exceedingly complex set of genetically-driven interactions that occur in precise time and space within the developing embryo. As the knowledge of the genes involved in this process move forward, it has become clear that there are components to the genetic material that are contributed by both parents that go beyond the underlying genetic sequence. This emerging component is termed 'epigenetic' and can be considered as the next frontier in genetics. Over the past decade, an appreciation has been growing for a vast role for epigenetics in the control of many aspects of development and disease. It has been widely accepted that, unlike genetic information, epigenetic information is highly dynamic in defined periods of our lifecycle. Specifically, it is within male and female germ cells where a critical period of epigenetic reprogramming takes place, establishing epigenetic marks that can persist for the life of an individual supporting both normal development and contributing to overall health. It is the exploration of the nature of the epigenetic program in male germ cells, its acquisition and stability during adult life, with which the theme of this thesis is concerned.

1 Statement of Investigation

The mechanisms that underlie epigenetic phenomena relate to how local chromatin structure influences the expression of genes. These mechanisms include an array of covalent modifications to core histone proteins, interactions with small RNA molecules, and DNA methylation. Recent work has shown that these phenomena work closely together to stabilize chromatin states. Disruption of various aspects of these mechanisms has been shown to be related to the

abnormal gene expression patterns seen in many types of cancer and other diseases. To date, DNA methylation remains the best studied of these mechanisms.

DNA methylation is present in many types of plants and vertebrate animals (Tweedie et al., 1997; Vanyushin, 2006). It is present in a non-random fashion in the genomes of mammals. The mouse is an excellent model to understand various aspects of the biology of DNA methylation and has been intensively studied. Mice possess homologs of all of the known human DNA methyltransferase enzymes. The organization and overall amount of DNA methylation is vastly similar between mice and humans. Recently, it was concluded that syntenic chromosomal regions retain a pattern of DNA methylation that is highly concordant between the two species (Eckhardt et al., 2006). Furthermore, in the mouse, it is possible to extend our knowledge with interventional studies which procure important information through genetic ablation and drug treatment approaches that are not possible in humans.

The findings of many of these studies point to a critical aspect that sets epigenetics apart from classical genetics – reversibility. The labile nature of DNA methylation versus genetic sequence information makes it a fascinating candidate for acquired diseases and raises hope for pharmacological and other approaches for prevention and intervention. There are time points in development where DNA methylation patterns are reset, giving all progeny a "clean epigenetic slate" to begin life. One of these reprogramming events happens in early embryonic development where somatic patterns are set, the other occurs in germ cells (Figure 1.1). Continual reprogramming with each generation raises the possibility of the occurrence of 'epimutations' in these critical developmental widows. Furthermore, some acquired epigenetic states can escape reprogramming of DNA methylation patterns and be passed on to progeny, in what is called 'epigenetic inheritance' (Morgan et al., 1999; Rakyan et al., 2003). Clearly, a fundamental understanding of DNA methylation in germ cells is paramount to the understanding of these very important processes.

Previous work in looking at DNA methylation in germ cells, particularly the male, has hinted that DNA methylation patterns are unique in these cells. Due to the larger number of germ cells produced in the male than in the female, their greater accessibility and their continual renewal, most germ cell studies focus on spermatozoa and testicular tissues. Various techniques have been used for these studies; however, no study has provided a non-biased genome-wide picture of the status of methylation in germ cells. I have employed a technique called restriction landmark genomic scanning (RLGS) to accomplish this. Chapter II describes the development and evaluation of a technique that works in concert with RLGS. This technique was a necessary component to achieve this goal of a comprehensive analysis of DNA methylation patterns in male germ cells, the focus of chapter III. As mentioned earlier, previous studies have found that, during a defined window in pre-natal development, patterns of DNA methylation are reset; however, it is not known if patterns continue to be acquired outside of this window. Data presented in chapter IV addresses the question of whether or not patterns of DNA methylation continue to be acquired in post-natal life, in a particular period of gametogenesis called spermatogenesis. After defining the nature of DNA methylation during spermatogenesis, chapter V explores the relationship between DNA methyltransferases and DNA methylation in this period by perturbing their function by a pharmacological intervention.

This introductory chapter will provide an overview of the aspects of male germ cell biology and DNA methylation that are relevant to the work presented in this thesis. As the work presented in this thesis relates solely to male germ cells, less emphasis will be placed on female germ cells. This chapter will begin with the origins and development of male germ cells in both pre- and post-natal developmental windows. Next, I will discuss aspects of DNA methylation in mammals, and, finally, what is known about DNA methylation in male germ cells and approaches that can be taken to study DNA methylation in these cells.

1.1 Gametogenesis

1.1.1 Embryonic Origins of Germ Cells

The germ line is defined as the continuous lineage of cells that perpetuate the lifecycle of a species. Other than a brief period after fertilization and persisting into the early stages of embryonic development, germ cells carry the germ line in mammals. In the mouse, identifiable germ cells first appear around 6.5 dpc (days *post-coitum*) as a cluster of a few cells that arises out of the embryonic endoderm and are termed primordial germ cells (PGCs) (Figure 1.1). These cells migrate out of this compartment, through the developing embryo and arrive a few days later to colonize the genital ridges around 10.5 dpc. These cells will mitotically divide during the period of migration and after colonization of the genital ridges, creating a population of approximately 25,000 germ cells by 13.5 dpc (Donovan et al., 1987; Tam and Snow, 1981).

1.1.2 Female Germ Cell Development

The appearance and behavior of PGCs from their appearance to the colonization of the genital ridge are indistinguishable between male and female (McLaren, 1995). The gonads acquire a sexually-dimorphic appearance around 12.5 dpc as germ cells are organized into cord structures in the male but not in the female (Hogan et al., 1986). Early female fetal germ cells, termed 'oogonia', stop mitotic divisions and enter meiosis as early as 14 dpc (Wasserman and Albertini, 1994). Pairing and recombination of homologous chromosomes occur until the time of birth where oogonia are found to be arrested in meiosis (Bakken and McClanahan, 1978). These cells, now called non-growing oocytes, will remain arrested until puberty (which occurs shortly after birth in mice) where some are recruited into the oocyte growth phase by hormonal signals. During this period, many non-growing oocytes will degenerate by atresia. In the oocyte growth period, oocytes will increase their diameter from 10-20 μm to 80 μm and undergo other ultra-structural changes. The completion of oocyte development

prior to the ovulation of a fully-grown oocyte occurs in the meiotic maturation period where meiosis resumes and is completed upon fertilization.

1.1.3 Male Germ Cell Development

1.1.3.1 Male Pre-Natal Germ Cell Development

There is a striking dimorphism of germ cell development between male and female. After the initial spermatogenic cord formation begins to occur around 12.5 dpc specifically in the male, male PGCs continue to divide mitotically until being arrested around 14.5 dpc, but do not enter meiosis (Nagano et al., 2000; Vergouwen et al., 1991). PGCs will resume their mitotic activity immediately after birth (Coffigny et al., 1999). Over the next few days of post-natal life, the male germ cells will migrate towards the basement membrane of the spermatogenic cords. These cells will give rise to both the first spermatogonia and a population of self-renewing spermatogonial stem cells that will continually produce germ cells for the life of the animal (Orwig et al., 2002).

1.1.3.2 Male Post-natal Germ Cell Development

Much literature has been composed that describes many of the intricate details of the male reproductive system in the adult. Much work has been done to characterize the kinetics, regulation, molecular mechanisms and pathology of spermatogenesis (Clermont, 1972; Cooke and Saunders, 2002; Dadoune et al., 2004; de Rooij and Russell, 2000; Ewing et al., 1980; Grootegoed et al., 2000; Hecht, 1998; Russell et al., 1990). Therefore, only an overview of the specific aspects that are relevant to this thesis will be discussed.

Testis Structure

The testis of the adult animal is organized as an array of seminiferous tubules encapsulated by the testicular tunica (reviewed by Russell et al. 1990) (Figure 1.2). Intervening between the tubules are small interstitial spaces comprised of somatic cells, such as the testosterone-producing Leydig cells, as

well as myoid, epithelial and blood-borne cells. Germ cells are housed within the seminiferous tubules that are surrounded by the basal lamina. Spermatogenesis occurs within the tubule in a highly organized histological structure called the seminiferous epithelium (Figure 1.2). Spermatogonial stem cells occupy the basal compartment of the tubule and germ cells of increasing maturity are found as they migrate towards the tubular lumen. The only somatic cell type found within the tubules is the Sertoli cells. These non-proliferative cells nurse the developing germ cells by providing factors necessary for and removing cellular waste and debris created by developing germ cells (Griswold, 1995; Mruk and Cheng, 2004; Sharpe, 1994). The nuclei of Sertoli cells are distributed along the basal lamina of the tubule and the cytoplasmic projections of each cell contact 30-50 germ cells throughout the tubule and at various stages of development (Siu and Cheng, 2004). Tight junctions between adjacent Sertoli cells form the blood-testis barrier separating the basal and adluminal compartments of the seminiferous epithelium.

The timing of the phases of spermatogenesis is highly precise and constant. These kinetics result in a predictable architecture of the seminiferous epithelium where specific cell types within each of the phases of spermatogenesis are found grouped together. These groups are defined as particular stages of spermatogenesis (reviewed by Russell et al., 1990). In the mouse, cross sections of seminiferous tubules can be grouped into 12 stages that are arbitrarily defined by cellular features that are identifiable by light microscopy. This allows for precise histological identification of all germ cells within the various phases of spermatogenesis.

Spermatogenesis

Spermatogenesis is a very complex process whereby haploid spermatozoa are produced from diploid spermatogonial stem cells. In a continuous process that lasts for the lifespan of the male, spermatogenesis achieves the processes of genetic recombination, a dramatic morphological remodeling of the nucleus and cell body, while managing to produce

approximately ten million sperm per gram of testis tissue a day (Amann and Howards, 1980). This is achieved by a highly regulated process that can be broken down into three distinct phases: the proliferative phase, the meiotic phase, and spermiogenesis (Figure 1.3). These phases are also referred to as pre-meiotic or spermatogonial, meiotic or spermatocyte and post-meiotic or spermatid spermatogenic phases. Each phase involves distinct processes and germ cell types (Bellvé et al., 1977b).

In the proliferative phase, diploid spermatogonial stem cells mitotically divide to produce the first undifferentiated spermatogonia. Like the stem cells of other tissues, one daughter cell is destined for differentiation, the other for self-renewal. Although differing theories of the kinetics of spermatogonial stem cell renewal have been proposed (Clermont and Bustos-Obregon, 1968; Huckins, 1971; Oakberg, 1971), both proposed mechanisms involve the existence of stem cells. Cells that are not destined for self-renewal are termed primitive type A spermatogonia and continue to divide several times which amplifies the numbers of germ cells before they begin to differentiate. All germ cells that arise from a single stem cell are connected throughout the duration of spermatogenesis by cytoplasmic bridges that arise from incomplete cytokinesis. Type A spermatogonia divide and differentiate into intermediate spermatogonia, then into type B spermatogonia before exiting the proliferative phase. Commitment to the differentiation into intermediate and type B spermatogonia is associated with visible changes to the nuclear architecture, most notably is the accumulation of heterochromatin in the periphery of the nuclear envelope.

In the meiotic phase, spermatocytes (germ cells of the meiotic phase) will become tetraploid, synapse homologous chromosomes, undergo recombination and reduce their genetic content to haploid. This phase begins with one final mitotic cell division that generates preleptotene spermatocytes from type-B spermatogonia. Preleptotene cells will endo-reduplicate their DNA (DNA replication without cell division) to form the tetraploid leptotene spermatocytes. It is in this stage and the next stage, zygonema, where homologous chromosomes are paired in preparation for recombination. In the lengthy pachytene stage that

follows, a phase that requires approximately a week to complete, homologous chromosomes undergo genetic recombination. The first reductive division occurs to produce secondary (diploid) spermatocytes from diplotene spermatocytes, which quickly divide in the second meiotic reduction to produce haploid spermatids.

Spermiogenesis is the spermatogenic phase where spermatids undergo dramatic morphological changes to both the cytoplasm and nucleus of the cell. Spermatids can be grouped generally by their morphology into round spermatids followed by the elongating spermatid stage (Figure 1.3). Nuclear changes include extensive chromatin remodeling that replaces histones with transition proteins and eventually protamines. (Dadoune et al., 2004; Sassone-Corsi, 2002). Protamines are small basic proteins that coil DNA into a compact structure and condense the chromatin by disulfide bonding. Cytoplasmic changes include the formation of an acrosome, elongation and eventual removal of most of the cytoplasm and the development of a motile flagellum. Cytoplasmic bridges are pinched off in the final stages of spermiogenesis before spermatozoa are released from the testicular epithelium in a process called 'spermiation'.

At which point spermatogenesis is complete; however, additional processes are further required to produce spermatozoa that are motile and are capable of natural fertilization. Spermatozoa that are liberated into the testicular lumen are transported into the caput of the epididymis where the final nuclear compaction and disulfide linkages between protamines are completed. A small amount of additional cytoplasm is shed from the spermatozoa during transit from the caput to the cauda of the epididymis that is important for their motility, and extensive changes occur in the lipid and protein composition of the sperm membrane. Finally, mature spermatozoa are stored in the cauda of the epididymis before being mixed into semen and ejaculated.

During spermatogenesis, several processes necessitate highly ordered, germ cell-specific chromatin states, both during meiotic stages and spermiogenesis. The accomplishment of this is made possible by the

developmentally regulated expression of a large array of germ cell-specific genes (Shima et al., 2004). Although the mechanisms that underlie the establishment of germ cell-specific chromatin states and how the genes that control these processes are regulated are unclear, a large body of evidence from other systems implicates the phenomenon of DNA methylation as being intricately involved in the regulation of chromatin states that can influence both gene expression and chromatin structure.

1.2 DNA Methylation

DNA methylation is an important component of a general mechanism that allows for additional information to be encoded into chromatin in addition to the underlying genetic sequence. These modifications are termed 'epigenetic' and involve an extensive array of modifications to histone proteins (Jenuwein and Allis, 2001) and small, non-coding RNA molecules (Bayne and Allshire, 2005). As all the data in presented in this thesis relates only to DNA methylation, these other related phenomena will not be discussed further.

Although the vast majority of DNA of all organisms is made up of the four native deoxynucleotides (adenine, cytosine, guanine, and thymine), many organisms modify a small proportion of these molecules. In mammalian DNA, the primary modification occurs to cytosine, via the transfer of a methyl group from S-adenosylmethionine (SAM) to the 5^{th} carbon of the cytosine ring to form 5-methylcytosine (5mC) (Figure 1.4). 5mC accounts for approximately 5% of the total cytosine content of DNA, or 1-2% of the overall DNA content in most tissues. In mammals and other organisms, 5mC is found in palindromic 5'-CG-3' sequences within DNA (termed CpG dinucleotides). Approximately 60-80% of CpG dinucleotides in the mammalian genome are methylated (Bestor et al., 1984; Bird, 1980).

A key feature of DNA methylation is that information can be either stably inherited from mother to daughter cells or can be erased and re-established without altering the underlying sequence. It has been well demonstrated that 5mC can be added and removed in specific periods of genomic reprogramming

(Reik et al., 2001). DNA methyltransferase enzymes (DNMTs) methylate cytosine residues, while, although less extensively described, demethylation can occur via a growing number of potential mechanisms that include direct removal of the methyl group (Bhattacharya et al., 1999) and base excision and repair by glycosylase and DNA repair enzymes (Jost and Jost, 1995; Walsh and Xu, 2006).

1.2.1 Biological Roles of DNA methylation

Most roles that have been assigned to 5mC involve the creation of an additional layer of information in the genome. This information is interpreted as a signal to format the surrounding sequence as condensed heterochromatin versus less-condensed euchromatin (Vermaak et al., 2003). DNA methylation is a key component of an array of different phenomena that work together to establish and maintain a heterochromatic state.

DNA methylation has been theorized to behave both on a genome-wide scale (Bestor and Tycko, 1996), as well as on a gene-specific level (Ballestar and Wolffe, 2001). On a genome-wide scale, 5mC promotes chromatin stability and is theorized to constrain the effective size of the genome through the selective exposure of regulatory sequences versus non-regulatory sequences. Generally, it has been shown that the presence of 5mC, particularly in 5' regions, by promoting the formation of local heterochromatin, will down-regulate transcriptional activity.

1.2.1.1 Regulation of Gene Expression

There are several proposed mechanisms that link DNA methylation to the modulation of gene expression via direct interactions with proteins that modulate chromatin structure (Klose and Bird, 2006). On a local level, small changes to the methylation of DNA in key positions (mainly thought to be 5' regions of genes) have been shown to change the ability of transcription-factors to associate with DNA. The simplest mechanism involves a direct physical change that results from the addition of 5mC to a particular sequence. DNA methylation can alter the conformation of DNA (McIntosh et al., 1983) and may result in changed

associations with various regulatory proteins. DNA methylation of key residues also physically blocks binding of proteins that promote transcription (De Smet et al., 1996; Griswold and Kim, 2001; Santoro and Grummt, 2001). More generally, it is believed that DNA methylation plays a role in the regulation of genes by mainly influencing regional chromatin states. There are several members of a methyl-CpG-binding domain (MBD) protein family that promote repressive states in chromatin by specifically recognizing the methylated state of DNA. These include Kaiso, MeCP2 and four MBD proteins (MBD1-4). Commonly, these proteins promote transcriptional repression by complexing proteins that retain histone deacetylase (HDAC) activity that changes the potential of DNA-histone association by modifying the histone's electrical charge. DNA methyltransferase enzymes are able to interact with HDAC proteins and promote repression through this association (Burgers et al., 2002).

A considerable amount of information has been produced by investigating aberrant DNA methylation in disease, mainly cancer (Feinberg, 2004; Robertson, 2005). Many studies have focused on the relationship between gene-specific hypermethylation in tumors and cancer cell lines and loss of expression. Many of these repressed genes that are associated with hypermethylation can be reactivated via the use of 5-azacytidine, a drug that promotes demethylation in replicating cells. These studies provide a considerable amount of information that supports the causative association between hypermethylation and gene repression.

1.2.1.2 Imprinting and X Chromosome Inactivation

Regions of the genome that display allele-specific expression are often differentially methylated (Reik and Walter, 2001). These regions are termed 'imprinted regions' and/or 'imprinted genes'. In most cases, differentially methylated states between alleles observed in offspring are acquired during germ cell development in either parent in a sex-specific manner. Several regions of the genome are imprinted; however, most marks are acquired in the female germ line as opposed to the male. The reason(s) for this dichotomy are unclear; however,

the existence of imprinted regions has been ascribed to the desire of either sex to control the growth potential of the offspring (Moore, 2001). The female takes a more pro-active stance because offspring overgrowth will have a greater effect on her than on the father.

Another role of DNA methylation involves the inactivation of X chromosome in females. X-Linked gene dosage compensation in females is achieved by repression of the genes on the X chromosome during development (Sado and Ferguson-Smith, 2005). It has been shown that many genes on the inactive X chromosome are methylated (Park and Chapman, 1994). In addition, methylation contributes to the regulation of the master switch gene that controls the inactivation of the X chromosome, *Tsix* (Sado et al., 2004).

1.2.1.3 Relationship to Genome Organization

Analysis of the mouse and human genomic sequences reveal that there are approximately 20 and 27 million CpG dinucleotides in the two species, respectively (Fazzari and Greally, 2004). Each CpG dinucleotide is a potential target for DNA methylation. Total genomic sequence can be divided into two broad classes: those sequences that belong to interspersed transposable elements (TEs) and those that do not. In the mouse, approximately 40% of the total number of CpG dinucleotides is composed of TE-derived DNA, which include short and long interspersed transposable elements (SINEs and LINEs, respectively), and many types of long terminal repeat (LTR) retroviral elements. (Figure 1.5). The remaining approximately 60% of the total number of CpGs are found in unique sequence which include non-repetitive exonic, intronic and intergenic sequences. CpG dinucleotides are divided between these genomic compartments in proportion to the amount of underlying sequence. Also, the proportion of sequence in each compartment and the CpG distribution is similar between mice and humans, with somewhat more CpGs found in TEs in humans.

One very interesting feature of the genomic sequence of mammals and other species are CpG islands (CGIs). While the genomes of mammals are overall relatively devoid of CpG dinucleotides, short sequences of high CpG

density are found throughout their genomes. Originally defined based mainly on CpG and G+C density (Gardiner-Garden and Frommer, 1987), definition of higher stringency (Takai and Jones, 2002) has been more widely-adopted for CGIs (G+C ≥ 55%, CpG obs/exp ≥ 0.6, length ≥ 500bp). CGIs account for less than 1% of the overall genomic sequence; however, account for 6-7% of all CpGs (Figure 1.5). CGIs are of special interest because they are commonly associated with the 5' regions of genes and their methylation status has been shown generally to be inversely correlated with gene expression levels. Greater than half of all mammalian genes are associated with CGIs (Ioshikhes and Zhang, 2000), many of which perform house-keeping functions. Another interesting feature of CGIs is that they are generally found to be unmethylated in all tissues. A recent study in humans found that only 9% of assayed CGI were methylated in any of the 12 different tissues studied, and all methylated CGIs were of lower CpG density (Eckhardt et al., 2006). The unmethylated nature of CGIs and their high frequency of association with the 5' regulatory regions of housekeeping genes suggest that the function of CGIs is to provide a constitutively open chromatin state that is permissive for the expression of highly expressed genes.

Unlike CGIs, it is generally believed that non-CGI sequences are highly methylated in mammals. High levels of methylation in TEs are thought to promote overall chromosome stability by repressing their transcription (Yoder et al., 1997b). Also, stable heterochromatic states may mask homologous regions on non-homologous chromosomes that are represented by TE sequences and repress somatic recombination events. It has been shown that in cancer cells, genomic DNA is hypomethylated and genomic instability is a hallmark of the genomes of many tumors and cancer cell lines (Ehrlich, 2006). Current genomic sequence builds do not include centromeric and telomeric regions of chromosomes, which are mostly composed of tandemly repeated satellite and telomeric repeat sequences. Studies using Southern blot analysis have demonstrated that these regions are highly methylated in somatic tissues (Gaillard et al., 1981; Sanford et al., 1984). Non-repetitive, non-CGI sequences are thought to contain high levels of DNA methylation as well; however, they

have been the subject of much less direct interrogation. The purpose(s) of methylation in the non-CGI unique sequence compartment is unknown, despite the majority of the genome belonging to this compartment.

DNA methylation functions in genome organization by promoting the formation of heterochromatin. The chromosomal landscape of higher eukaryotes is heterogeneous with respect to heterochromatin versus euchromatin. Heterogeneity within chromosomes was first visualized by Caspersson (Caspersson et al., 1968), which was later improved by using the Giemsa dye (Dev et al., 1972). At the light microscopy level, the staining of metaphase chromosomes readily reveals a chromosome-specific pattern of alternating dark (Giemsa- or G-) and light (reverse- or R-) bands. G- and R-bands generally define chromosomal regions of heterochromatin versus euchromatin, respectively. These regions are associated with differences in sequence features. G-bands contain fewer genes, longer introns, fewer CGIs, replicate later and have a lower rate of meiotic recombination versus R-bands (Fazzari and Greally, 2004). While TEs and unique sequence are defined as occupying genomic 'compartments', the interspersed nature of TEs causes these compartments to be scattered throughout the genome. However, the distribution of specific sequences is not random and is associated with higher order ultra-structural features of chromosomes. The proportion of TEs to unique sequence and the distribution of various TEs have been shown to be related to the distribution of heterochromatin. G-bands also contain proportionally more total repetitive sequence, including satellite repeats, more LINEs, yet, surprisingly, fewer SINEs. Another sequence-based feature of the chromosomal banding pattern is that G-bands contain below-average and R-bands above-average G+C nucleotide content (Saccone et al., 1993). Defined regions of chromosomes contain high and low GC content (usually several Mb in size) and are defined as isochores (Bernardi et al., 1985). Recently, it has been concluded that this is the defining feature of the banding pattern (Costantini et al., 2006), demonstrating a clear connection between heterochromatin to chromosomal regions of low GC content.

The basis for the linear heterogeneity of GC content is thought to be the result of the association of DNA methylation to heterochromatin. 5-Methylcytosine is relatively unstable compared to cytosine and the other nucleotides and will spontaneously deaminate to form thymine (Shen et al., 1994). Improperly repaired G nucleotides in T/G mismatches would result in the original 5mC/G being replaced with a T/A base pair. This existence of a 5mC-dependent mutational mechanism operating in the germ line of a species over many generations would cause any methylated sequence to gradually lose GC content. The existence of CGIs is explained by this mechanism and due to their property of being consistently unmethylated, thus resistant to 5mC-dependent mutation. This mechanism has also been proposed as a key aspect of genome defense, as targeting 5mC to TEs would not only cause transcriptional inactivation, but mutations would render them nonfunctional over time.

The relationship between heterochromatin and DNA methylation combined with the fact that there are specific sequence features associated with heterochromatin, opens the door to a multitude of possible functions for DNA methylation. As DNA methylation promotes heterochromatin formation, functions may include genomic defense against parasitic elements, long-range interactions influencing gene expression and regulation of mitosis (Talbert and Henikoff, 2006). It has been postulated that by supporting the maintenance of heterochromatin, DNA methylation allows regions containing regulatory sequences and genes to be preferentially available (Bestor and Tycko, 1996). A role for DNA methylation in this process is supported by the fact that $5^{m}C$ is a universal feature of large-genome eukaryotes that contribute to the C-value paradox (Thomas, Jr., 1971); defined as genome size scaling independent of gene number, while many eukaryotes with genome sizes <5 x 10(8) bp do not methylate their DNA (Kidwell, 2002). As a more detailed knowledge of the functions of these sequences evolves, additional roles of DNA methylation will undoubtedly be uncovered.

1.2.2 DNA Methyltransferases

The flexibility of DNA methylation (stability vs. reversibility) is facilitated by different activities attributed to various members of the DNMT family of enzymes. During the S-phase of cell cycle, patterns of DNA methylation are faithfully copied from the existing DNA strand to the newly synthesized strand in an activity termed 'maintenance methylation'. The palindromic nature of the CpG dinucleotide sequence allows for maintenance methylation activity to result in the duplication of the genomic methylation pattern. The other activity attributed to DNMT proteins is the establishment of newly methylated cytosines, termed '*de novo* methylation'. CpG dinucleotides are thought to be primary targets of *de novo* methylation; however, non-CpG methylation may also occur during periods of DNA methylation establishment (Imamura et al., 2005; Ramsahoye et al., 2000).

In mammals, there are three subfamilies of DNMTs that are classified according to sequence similarities (Goll and Bestor, 2005) (Figure 1.6). The primary and initially characterized DNA methyltransferase is DNMT1 (Bestor et al., 1988). Subsequent studies showed that DNMT1 has a greater preference for hemimethylated DNA (Hsieh, 2005; Li et al., 1992; Yoder et al., 1997a) causing it to be assigned the function of maintenance methylation. Further strengthening this classification, DNMT1 is upregulated during the S-phase of the cell cycle (Szyf et al., 1991), unlike other DNMTs (Robertson et al., 2000), and associates with DNA replication machinery (Rountree et al., 2000). However, as DNMT1 has also been shown to have *de novo* activity *in vitro* (Okano et al., 1999), it remains unclear if DNMT1 performs this function *in vivo*.

The essential nature of DNA methylation was first demonstrated by the targeted disruption of DNMT1 (Lei et al., 1996; Li et al., 1992). Disruption of catalytic activity of the protein results in mid-gestational embryonic lethality and a decrease in overall methylation levels to ~5% of normal. Many phenomena associated with DNA methylation were shown to be disrupted, including gene imprinting (Li et al., 1993), X-chromosome inactivation (Panning and Jaenisch, 1996), and transposon silencing (Walsh et al., 1998).

Although DNMT2 is the most conserved DNMT protein between organisms, disruption of this gene results in no decrease of DNA methylation in mouse cells (Okano et al., 1998). Recently, it has been shown that DNMT2 functions as an RNA methyltransferase in mammals (Goll et al., 2006), consistent with the fact that DNMT2 is found in lower organisms that lack detectible quantities of DNA methylation.

The DNMT3 subfamily consists of two catalytically-active members, DNMT3a and DNMT3b, and a homologous protein, DNMT3L (DNA methyltransferase 3-like), which lacks catalytic activity. DNMT3a and DNMT3b are encoded by separate genes and are postulated to function primarily as *de novo* methyltransferases (Okano et al., 1999). Regulation of these genes is complex, as evidenced by the large number of alternative splicing variants that are produced (Aoki et al., 2001; Chen et al., 2002; Ishida et al., 2003; Okano et al., 1999; Weisenberger et al., 2002; Weisenberger et al., 2004). Of interest, a catalytically-functional splice variant of DNMT3a, DNMT3a2 is the predominant protein in germ cells (Chen et al., 2002).

Targeted disruption of both of DNMT3a and DNMT3b further underscores the essential nature of DNMT activity and DNA methylation in normal development. DNMT3a-deficient mice die 3-4 weeks after birth despite having normal levels of global methylation (Okano et al., 1999). Targeted deletion of DNMT3b results in embryonic lethality occurring in mid gestation and causes demethylation of centromeric satellite sequences (Okano et al., 1999). Mice engineered to have a double deletion of DNMT3a/3b die earlier during embryonic development and have globally hypomethylated genomes (Okano et al., 1999). Similar effects of perturbing DNA methylation have been found in studies on embryonic stem cells. In these studies, targeted deletion of DNMT3a or DNMT3b alone has little or no effect on genome-wide methylation, whereas double deletion results in extensive hypomethylation at multiple genomic loci, similar to levels seen in DNMT1-deficient cells (Hattori et al., 2004). DNMT3b is the only DNMT to be associated with a human disease. Various mutations in this gene cause ICF (immunodeficiency, centromeric instability and facial anomalies)

syndrome, a disease associated with cytogenic abnormalities resulting from demethylation of centromeric regions of certain chromosomes (Xu et al., 1999). Together, these studies indicate that DNMT3a and DNMT3b perform shared and separate functions.

Although lacking some amino acids that confer catalytic activity, DNMT3L is similar to other DNMT3 family members in both its N- and C-terminal domains (Aapola et al., 2000). DNMT3L is also of interest because its expression is predominant in the germ cells of both sexes (Aapola et al., 2001). It is important for the establishment of DNA methylation in both male and female germ cells (Bourc'his et al., 2001; Bourc'his and Bestor, 2004; Hata et al., 2002; Webster et al., 2005). Studies showing interactions between DNMT3L and DNMT3a and/or DNMT3b suggests that DNTM3L may function as a cofactor and stimulate *de novo* methylase activity (Chedin et al., 2002; Hata et al., 2002; Nimura et al., 2006; Suetake et al., 2004; Suetake et al., 2006). Consistent with its selective expression in germ cells, targeted disruption of this gene results in no observable somatic phenotype, but mice of both sexes are infertile (Bourc'his et al., 2001; Hata et al., 2002; Webster et al., 2005). (A more comprehensive review of the germ cell-specific effects can be found in the subsequent section of the introduction, DNA Methyltransferases in Male Germ Cells)

1.3 DNA Methylation in Germ Cells

1.3.1 Tissue-Specific DNA Methylation

The reversible nature of DNA methylation led to the postulation that DNA methylation could be involved in the establishment of tissue-specific expression patterns (Holliday and Pugh, 1975; Riggs, 1975). Different tissues acquire differential gene expression patterns although they arose from the same precursor cells during development. The role of DNA methylation as a director of this process is controversial (Bestor and Tycko, 1996), as direct evidence demonstrating causality is elusive. Studies have shown that as early as post-

implantation embryonic development, tissue-specific methylation has occurred, evidenced by all tissues of extra-embryonic origin containing methylation patterns distinct to that of the embryo proper (Chapman et al., 1984; Rossant et al., 1986). More recently, studies have demonstrated that different tissues from within the same adult individual possess differential methylation patterns. Restriction landmark genomic scanning (RLGS), a technique that can visualize several thousand unique loci located throughout the genome simultaneously, has been used to demonstrate that all tissues studied possess unique patterns of methylation to varying degrees (Shiota et al., 2002), and the tissue-specific methylation status of some gene-specific loci correlate with levels of tissue-specific expression (Song et al., 2005). Recently, a study of several tissues in humans has reported similar results, where differences in methylation range between 5-20% between various cell types and tissues (Eckhardt et al., 2006).

1.3.2 Male Germ Cell-Specific DNA Methylation

Although genome-wide patterns of DNA methylation are established *de novo* in early embryos, germ cells undergo a second period of reprogramming that occurs later in embryonic development (Reik et al., 2001). Data from studies on imprinted genes, repeat sequences and germ-cell specific genes (Hajkova et al., 2002; Lees-Murdock et al., 2003; Maatouk et al., 2006) demonstrate that before and upon arrival to the genital ridges around 10.5-11.5 dpc, migrating PGCs possess a pattern of DNA methylation that is somatic in nature. As PGCs colonize the genital ridges, patterns of DNA methylation in PGCs are erased. Germ cell patterns are re-established at a later stage in development in a sex-specific manner. Due to the small numbers of germ cells at these stages, the full extent and sequence specificity of the reprogramming is largely unknown. In the case of imprinted genes, the sequences and the timing of the acquisition of DNA methylation are sex-specific. In female PGCs, DNA methylation of maternally-methylated imprinted DMRs is acquired after birth during the oocyte growth phase (Lucifero et al., 2004). In males, PGCs begin to re-acquire DNA methylation between 15.5-17.5 dpc, a period much earlier than in the female

(Hajkova et al., 2002). Repeat sequences such as the intracisternal A particle (IAP), long interspersed transposable elements (LINEs) and satellite sequences show a similar timing of erasure and establishment (Hajkova et al., 2002; Lees-Murdock et al., 2003; Walsh et al., 1998). However, unlike imprinted genes, methylation is not fully erased at repeat sequences. The timing of the acquisition of DNA methylation coincides with high levels of the germ cell-specific protein *Dnmt3L* in both sexes (Bourc'his et al., 2001; La Salle et al., 2004), as well as a sharp increase in the antigenicity of germ cells for an antibody directed against 5mC in the male (Coffigny et al., 1999).

There are several possibilities that have been suggested to explain the necessity of germ cell-specific reprogramming, which include the regulation of germ cell-specific gene expression, the setting of genomic imprints and the establishment of chromatin structure. Although the primary purpose of germ cell reprogramming is unclear, there are some studies that hint that the pattern of DNA methylation that is acquired as a result of this reprogramming event is distinct from other cells.

1.3.2.1 Methylation and Testis-Specific Gene Expression

The germ cells of the testis display the most unique tissue-specific global transcriptional profile of any tissue or cell type examined in both mouse and human (Shima et al., 2004; Su et al., 2004). The high number of genes expressed in the testis is presumably a requirement of the specialized processes of meiosis and spermiogenesis. A few of these genes have been shown to display testis-specific methylation states (MacLean and Wilkinson, 2005). Included in these are the transition protein genes, *Tp1* & *Tp2*, and the protamine genes, *Prm1* & *Prm2*, that are expressed during spermatid development to facilitate nuclear remodeling (Trasler et al., 1990). *Pdha2* and *Pgk2* are autosomal homologs of the X-linked genes whose activation is required when XY body formation silences the X chromosome during spermatocyte development (Ariel et al., 1991; Iannello et al., 1997). Most testis-specific genes that display a testis-specific methylation state are hypomethylated in their 5' regions, reflecting

their activated status. However, the *Prm1*, *Prm2* and *Tp2* exist as a single gene cluster and, curiously, this region is specifically hypermethylated in the testis despite being the exclusive tissue of expression (Trasler et al., 1990). Another gene, the human cyclin A gene, *CCNA1*, is methylated in germ cells of the testis and hypomethylated in other tissues, even though the testis is its preferential site of expression (Muller-Tidow et al., 2001). Although the number of testis-specific genes that have been shown to be associated with altered methylation pales in comparison to the total number of genes with unique levels of expression in the testis, it is likely that methylation plays a role in the control of these few examples. Why the numerous other testis-specific genes do not show this association and how some testis-specific genes are highly expressed despite being hypermethylated specifically in the testis is unclear and casts doubt on the hypothesis that gene regulation is the primary purpose for testis-specific DNA methylation.

1.3.2.2 Genomic Imprinting in Male Germ Cells

A second function of germ cell reprogramming is the setting down of genomic imprints. Imprinted genes are methylated on either the maternal or paternal allele, resulting in an overall 50% level of methylation in adult somatic tissues. Depending on the parental nature of the imprint, these loci will be either fully methylated or unmethylated in the germ cells of the testis resulting in a methylation state distinct to somatic tissue. There is an extreme dichotomy between the proportion of imprinted regions known to be methylated on paternal versus maternal alleles (Bourc'his and Bestor, 2006). Only three regions have been described to be methylated in sperm versus oocytes. These three paternally-methylated DMRs are found in the *H19-Igf2* region (Tremblay et al., 1995), the *Dlk1-Gtl2* region (Takada et al., 2002) and upstream of the *Rasgrf1* gene (Yoon et al., 2002). The regions marked by paternally-methylated imprints are different from those that are associated with testis-specific genes. Paternally-methylated DMRs tend to be found in intergenic regions that are several kilobases away from the genes that they control, whereas altered methylation

found in testis-specific genes are directly associated. In addition, the methylation of paternally-imprinted DMRs seems to function by blocking access to distant enhancer sequences instead of directly influencing the function of proximal promoter regions (Bourc'his and Bestor, 2006). Although methylation occurs mostly on maternal alleles for imprinted genes, correct levels of expression of imprinted genes in offspring depends on the cumulative methylation state of both parental alleles. Abnormal expression of gametic imprints has been associated with several human disease syndromes, including Prader-Willi, Beckwith-Wiedemann and Angelman syndromes (Robertson, 2005).

1.3.2.3 Repetitive Elements

It has been shown that gametic reprogramming results in a difference in the methylation status of some repetitive sequences between somatic and germ cells. Centromeric repeat sequences, particularly the major and minor satellite sequences, are distinctly hypomethylated in the germ cells of both sexes in humans and mice (Gama-Sosa et al., 1983; Sanford et al., 1984). This is in contrast to findings on other repeats of mainly interspersed transposable element origin. IAP, LINE and major urinary protein (MUP) sequences are methylated in mature sperm (Sanford et al., 1984; Walsh et al., 1998). However, in humans and primates, Alu repeat sequences are hypomethylated in sperm relative to somatic tissues (Hellmann-Blumberg et al., 1993; Kochanek et al., 1993; Liu et al., 1994). Methylation of centromeric repeats are thought to provide mainly a structural function to chromosomes (O'Neill et al., 1998; Viegas-Pequignot and Dutrillaux, 1976; Xu et al., 1999); although the purpose of their hypomethylation in germ cells is not clear, it may represent a modified chromosome structure requirement of germ cell-specific processes, such as meiosis.

A more complete picture of DNA methylation patterns in male germ cells has not been produced. Many repetitive sequences and most non-CGI unique sequences, despite composing a majority of the entire genome, have not been interrogated. Due to the involvement of DNA methylation in the development of male germ cells, it has been postulated that DNA methylation may play a role in

male infertility although very little work has been done to support this hypothesis (Cisneros, 2004). The purpose(s) of the germ cell reprogramming and DNA methylation in germ cells will become clearer once a comprehensive germ cell pattern of DNA methylation is established.

1.3.3 Acquisition of DNA Methylation during Spermatogenesis

Although it has been established that the acquisition of DNA methylation begins during fetal germ cell development, few studies have investigated the changes in DNA methylation during spermatogenesis. Spermatogenesis begins a few days after birth in the mouse, when germ cells resume the cell cycle to produce a pool of undifferentiated type A spermatogonia (Reviewed by Russell et al., 1990). Around 8 dpp, germ cells differentiate into type B spermatogonia and enter meiosis around 10 dpp. For testis-specific genes, the hypomethylated state is fully established prior to spermatogenesis (Kafri et al., 1992; Maatouk et al., 2006). The methylation status of *Pgk2*, has been well studied in various stages of spermatogenesis and is found to remain unmethylated through meiotic and spermiogenic phases (Geyer et al., 2004). Repetitive elements, such as IAP and LINE-1, are fully methylated in spermatozoa and have acquired this state by 17.5 days of gestation (Lees-Murdock et al., 2003; Walsh et al., 1998). Although the major and minor satellite centromeric repeat sequences are partially methylated in germ cells, their status remains unchanged during spermatogenesis (Sanford et al., 1984). Interestingly, studies on imprinted genes have shown that complete acquisition of methylation does not occur until after birth. *H19*, a paternally-methylated imprinted gene, initially acquires methylation before birth; however, complete levels of DNA methylation are not achieved until the pachytene spermatocyte phase of spermatogenesis (Davis et al., 1999). The late acquisition was shown to occur specifically on the allele of maternal origin, indicating the involvement of other epigenetic factors in genomic reprogramming. The two other known paternally methylated imprinted genes, *Rasgrf1* and *Gtl2*, similarly acquire most of their DNA methylation in the pre-natal window, but have yet to acquire the levels found in spermatozoa (Li et al., 2004). Although limited

data point to the acquisition of DNA methylation patterns beyond the fetal development window, a comprehensive study of the timing and sequence-types that may be involved has not been done.

1.3.4 DNA Methyltransferases in Male Germ Cells

The importance of DNA methylation in male germ cells has been highlighted by studies investigating the roles for the DNMT family of enzymes. All known members of the DNMT family are expressed in male germ cells and are expressed in a developmentally regulated fashion to varying degrees, (La Salle et al., 2004; La Salle and Trasler, 2006a; Shima et al., 2004). In addition, some are expressed as alternative transcripts that are only found in germ cells (La Salle and Trasler, 2006b). For example, the expression of Dnmt1 in male pachytene spermatocytes is marked by the inclusion of an alternative 5' exon that prevents translation of the transcript despite an enrichment of the protein in leptotene/zygotene spermatocytes (Jue et al., 1995).

Knock-out studies have highlighted a role for several Dnmts in various stages of germ cell development (Bourc'his et al., 2001; Bourc'his and Bestor, 2004; Kaneda et al., 2004). The germ cell-specific gene, *Dnmt3L*, although lacking catalytic activity, is highly expressed in PGCs beginning at 15.5-18.5 dpc (Bourc'his et al., 2001) and not in somatic tissues (Aapola et al., 2001). Males lacking a functional *Dnmt3L* gene are infertile as a result of a complete lack of germ cells in the adult testes (Bourc'his et al., 2001). Their infertility is a result of a structural failure of meiotic chromosomes to properly synapse which presumably leads to apoptosis via the activation of the pachytene checkpoint (Bourc'his and Bestor, 2004). Intriguingly, a failure to establish normal genomic levels of DNA methylation in spermatocytes isolated from *Dnmt3L*$^{(-/-)}$ animals is observed, including several types of interspersed repeat sequences, such as LINE-1 and IAP. Various paternally-methylated imprinted DMRs are undermethylated including *H19* (Bourc'his and Bestor, 2004; Kaneda et al., 2004; Webster et al., 2005) and *Rasgrf1* (Webster et al., 2005). Further studies have revealed that the failure to acquire DNA methylation extends to the imprinted

intergenic DMR, *Dlk1-Gtl2*, and to widespread non-repetitive loci (La Salle et al., 2007).

Other DNMT enzymes are ubiquitously expressed, thus the dissection of a germ cell-specific effect in the whole-animal knock out models is difficult. Although the germ cell-specific conditional targeted disruption of *Dnmt1* has not been done, *Dnmt3a* and *Dnmt3b* have been conditionally disrupted in germ cells. The conditional disruption of the *Dnmt3a* gene in germ cells leads to a similar phenotype of infertility due to a complete loss of germ cells in the adult testis (Kaneda et al., 2004). Spermatogonia isolated from *Dnmt3a*$^{(-/-)}$ animals show defects in methylation at imprinted genes but not at repeat sequences (Kaneda et al., 2004). Conditional disruption of the *Dnmt3b* gene had no detectible effect on germ cell viability, emphasizing that the roles for these proteins are distinct. The developmental regulation and requirement of various DNMTs during spermatogenesis highlights the importance of DNA methylation for the proper development of fertile male germ cells.

1.3.5 Effect of 5-Aza-2'Deoxycytidine in Male Germ Cells

The disruption of spermatogenesis and lack of fertile germ cells in DNMT knock-out animal models makes the study of germ cells in later stages of development impossible and negates the ability to measure fertility effects. An alternate strategy to perturb DNMT function is the use of the cytidine analog, 5-aza-2'-deoxycytidine (5-azaCdR). 5-azaCdR is currently used as an anti-cancer agent for myelodysplastic syndromes and other types of cancer, due to its ability to cause demethylation and cytotoxicity. Upon interaction with 5-azaCdR that has been incorporated into DNA, DNMT enzymes become irreversibly bound to DNA as a covalent adduct (Gabbara and Bhagwat, 1995; Santi et al., 1984; Taylor and Jones, 1982). Hypomethylation occurs during subsequent rounds of DNA replication due to the depletion of the cellular pool of DNMTs. Adducts are cytotoxic and induce apoptosis in a p53-dependent manner (Juttermann et al., 1994; Schneider-Stock et al., 2005). Notably, decreases in methylation can occur at non-cytotoxic concentrations of 5-azaCdR that do not significantly impair

DNA synthesis (Glazer and Knode, 1984; Mondal and Heidelberger, 1980). Use of 5-azaCdR offers an advantage as this agent inhibits both known (Weisenberger et al., 2004) and potentially unknown methyltransferases.

5-AzaCdR has been used in a series of experiments that further highlights the importance of DNA methylation in male germ cells. Male mice treated with the drug during the entire developmental window of spermatogenesis display considerable adverse effects to their reproductive physiology (Kelly et al., 2003). Male rats treated with an analogous drug, 5-azacytidine, show similar effects (Doerksen et al., 2000; Doerksen and Trasler, 1996). Effects include abnormalities in testicular histology leading to significant reductions in testis weight and reduced sperm counts. Litters sired by treated males are smaller, owing to an increase in pre-implantation embryo loss. These reproductive effects occur at doses that do not affect the overall health of the animals. Interestingly, mice that are heterozygous for a loss-of-function mutation in the catalytic domain of the DNMT1 protein (Lei et al., 1996) are partially protected from the adverse effects of the drug. In both rodent studies, abnormalities were associated with dose-dependent global reductions in DNA methylation. These studies further support the importance of DNA methylation to the proper development of male germ cells; however, effects caused by the cytotoxic properties of the drugs cannot be excluded.

1.4 Approaches to the Study of DNA Methylation

There are many techniques that have been developed for the study of DNA methylation. A summary and comparison of several commonly used methods and the methods used in this thesis is displayed in table 1.1. Until recently, the study of DNA methylation has been largely restricted to investigating individual sequences. Frequently used techniques such as bisulfite sequencing and Southern blotting allow for the analysis of methylation of either unique-copy sequences or a particular repeat. While these approaches have generated useful and informative data, the size and complexity of mammalian genomes makes any study based on a focused number of loci limited in its overall interpretative power.

Global techniques also are available that measure total cytosine content (generally done using high-performance liquid chromatography) or overall methylation at CCGG sites (done using thin-layer chromatography). These techniques are limited because they have low sensitivity and fail to identify sequences affected. Non-biased, genome-wide approaches that allow for the investigation of various sequence types simultaneously have recently been developed. These techniques allow for DNA methylation to be analyzed in more depth and to determine the specificity of alterations in DNA methylation between samples.

1.4.1 Restriction Landmark Genomic Scanning

Restriction landmark genomic scanning (RLGS) is a technique that allows for the investigation of a relatively large number of loci simultaneously. The technique was pioneered by Hayashizaki *et al.* (Hayashizaki et al., 1993) and is in current use in laboratories around the world. Using RLGS, the methylation status of approximately 3000 individual CpG sites in unique-copy and low-copy number sequences across the genome are revealed by digestion of genomic DNA using the methylation sensitive restriction enzyme, NotI (Figure 1.7). DNA fragment ends are then radiolabeled and separated by 2-dimensional electrophoresis producing an array of spots, with each spot corresponding to an individual CpG site located randomly in the genome. The relative intensity of each spot is inversely proportional to the percentage of methylation at that locus. RLGS has been very useful in the study of cancer (Rush and Plass, 2002) and other diseases related to DNA methylation (Kondo et al., 2000).

1.4.1.1 Virtual RLGS

An early drawback of the RLGS technique was the difficulty involved in identifying the genomic loci associated with spots of interest. This was initially overcome by the production of RLGS spot cloning libraries; a bacterial plasmid library containing individual RLGS DNA fragments that could be used to identify RLGS spots (Smiraglia et al., 1999; Yu et al., 2004). Although effective, spot

identification using the RLGS library is fairly laborious and inefficient. The usefulness of the RLGS technique has been increased substantially as a result of the recent development of a virtual RLGS (vRLGS) resource. In combination with the sequencing of the human and mouse genomes, vRLGS produces a computer-generated spot pattern that closely resembles the actual RLGS profile of a given species. Total genomic sequence is digested '*in silico*' and carefully formulated migration algorithms predict spot positions in two-dimensions (Figure 1.8) (D. Smiraglia, unpublished). This allows for hastened spot identification by identifying candidate loci (virtual spots that are in close proximity to the spot of interest on overlapping virtual and actual profiles). The correct candidate locus is identified by obtaining plasmid or bacterial artificial chromosome (BAC) clones that contain each candidate locus and mixing clone DNA with genomic DNA as background. Virtual RLGS analysis also allows for determination of the subset of sequences that are visible on actual RLGS profiles. Due to the high GC content of the NotI recognition sequence (GCGGCCGC), approximately 70% of the loci visualized by RLGS correspond to CpG islands, despite CGIs encompassing less than 1% of the genomic sequence (Fazzari and Greally, 2004). Because CGIs are mostly associated with genes, the total dataset generated by RLGS is gene-centric. Approximately 15% of the analyzable loci are found in interspersed transposable elements. Specifically, 240 and 60 NotI sites are found in the internal regions and LTRs of the IAP and early transposon (ETn) interspersed repeats, respectively. Due to their repetitive nature, their 'migration' produces a highly discernable pattern on vRLGS gels. However, they are rarely observed on actual RLGS profiles due to their constitutively hypermethylated nature. The remaining 15% of the loci analyzed result from NotI sites that are found randomly in unique sequences, mostly in intronic and intergenic sequences.

1.4.2 Tiling Arrays

An emerging approach to investigate genome-wide patterns of DNA methylation is the use of tiling array technology. Tiling arrays are constructed to contain thousands to millions of individual DNA fragments (features), which are

attached to a solid support. Tiling arrays can be designed to analyze parts of the genome (i.e. individual chromosomes) and/or specific classes of sequences (i.e. promoters). Individual array features can be larger (~1 kb) (Lippman et al., 2004) and result in low resolution analysis or can be designed using oligonucleoties of 20-50 nucleotides in length (high resolution) (Schumacher et al., 2006). DNA is prepared by either immunoprecipitation using an antibody to 5mC or digestion using methylation sensitive restriction enzymes followed by dual fluorescent labeling. Although the promise of this technology is high, it is currently expensive to design and manufacture the arrays, making analysis of numerous samples/conditions not feasible. Data that are generated using this technique are non-quantitative making the investigation of partial - yet potentially biologically important changes in DNA methylation that may occur at a particular locus difficult to interpret. In addition, because the length of individual tiles on an array is relatively short, interpretation of interspersed repetitive loci is problematic due to cross hybridization. This leaves approximately 40% of the total sequence of the mammalian genome uninvestigated using this method. Alternatively, because RLGS fragment sizes are relatively large, the methylation status of a single interspersed repetitive element located in a unique genomic position can be interrogated.

1.4.3 Quantitative Analysis of Unique Loci

The investigation of DNA methylation using RLGS generates information on multiple single CpGs located in numerous loci across the genome. Confirmation of the RLGS findings and further investigation of the methylation status of additional CpGs at these and other loci presents a formidable challenge. A suitable technique for this purpose would have to satisfy two major requirements: firstly, it would need to be flexible and rapid enough to handle numerous individual genomic regions, and, secondly, be quantitative so that partial changes in DNA methylation could be reproducibly detected.

There are many techniques that have been developed that can assess methylation with single CpG resolution (for review see (Tollefsbol, 2004). This

was initially done by digesting DNA using methylation-sensitive restriction enzymes followed by sequence detection using Southern blotting. Since then, several additional techniques have been developed; however, most do not produce a quantitative assessment. Of the methods that are quantitative, the bisulfite sequencing technique is the most widely used. Bisulfite sequencing generates methylation information by initially treating the DNA with sodium bisulfite which selectively converts non-methylated cytosines to uracil while leaving methylated cytosine resides unaffected (Clark et al., 1994). Following conversion, CpGs of interest are flanked by primers and PCR amplified. DNA fragments are sub-cloned and sequenced. Although this technique is unmatched by all other techniques because all CpGs in a small region are analyzed, the primary drawback is the length of time required for bisulfite conversion and subsequent sub-cloning and sequencing steps. There have recently been techniques that are quantitative and avoid lengthy cloning steps including QMSP (Lo et al., 1999), QAMA (Zeschnigk et al., 2004), Bio-CORE (Brena et al., 2006); however, these techniques still involve the use of sodium bisulfite. Further drawbacks of the use of bisulfite include high rates of sample loss (Grunau et al., 2001), instability of bisulfite-treated DNA, PCR amplification bias and artifacts that result from incomplete conversion of cytosines to uracil.

Chapter II of this thesis describes in detail the development and evaluation of an approach that avoids the drawbacks of the use of bisulfite, while taking advantage of the speed and accuracy of real-time PCR technology.

1.5 Rationale for Thesis Studies

As our understanding grows about the nature of genes and how their behavior is related to the chromosomal environment in which they function, so is our appreciation of the influence of epigenetics and DNA methylation. The links between abnormalities in DNA methylation and broad disease states, such as cancer, and natural phenomena, such as aging, are undeniable. The fact that DNA methylation represents an adaptable, essential piece of our genetic makeup that is acquired in every individual of each generation speaks to its importance. It

is within our germ cells that the epigenetic patterns are established that support healthy embryonic development and the health of our offspring. Alterations in germ cell acquired patterns of methylation are most likely the cause of imprinting-related disease syndromes and may be related to infertility. Furthermore, genes that are normally restricted to germ cells are commonly found erroneously in tumors and associated with abnormal states of DNA methylation (Simpson et al., 2005). A clear appreciation of the genome-wide nature of the germ cell epigenetic program, its acquisition, and its stability is paramount to the understanding of these very important processes and how they may contribute to human disease.

As discussed previously, there are several lines of evidence that demonstrate that DNA methylation and germ cell function/viability are linked. The important members of the DNMT family of enzymes are highly regulated during pre-natal development and spermatogenesis. Germ cell-specific disruption of DNMTs results in cell death and alterations in DNA methylation in germ cells. Treatment of mice with 5-aza-2'-deoxycytidine results in decreased spermatogenesis and abnormal embryo outcomes.

Due to the strong association between DNA methylation and several aspects of germ cell biology, it is very likely that distinct properties of DNA methylation are playing key roles in male germ cells. The purpose of this dissertation is to advance our current understanding of DNA methylation in male germ cells. To address this, the following questions were asked:

1. What is the pattern of DNA methylation of male germ cells of the adult? How does this pattern differ from that of other cells?

2. Are patterns of DNA methylation acquired during spermatogenesis?

3. Is it possible to perturb patterns of DNA methylation in male germ cells by altering DNMT function during spermatogenesis? How does this effect germ cell quality?

To ask these questions, we employed the RLGS technique, due to its ability to determine methylation at multiple sequences on a genome-wide scale. In order to fully employ the RLGS technique to it fullest extent, it was necessary to develop a technique that could be used in concert with RLGS to verify and expand upon the RLGS findings. This newly developed technique, termed quantitative analysis of DNA methylation using real-time PCR (qAMP), was extensively tested in its ability to accurately assess site specific levels of methylation determined by RLGS. The technique was further evaluated against other sites of known methylation, including imprinted genes, and by using bisulfite sequencing, the widely-accepted gold standard of DNA methylation analysis.

The use of RLGS and qAMP in combination with other standard techniques for DNA methylation analysis has allowed for the studies presented in this thesis to assess DNA methylation in male germ cells in a scope and level of detail not accomplished previously. The results of these studies have important implications for understanding the roles of DNA methylation in germ cells and many other biological systems.

Table 1.1: A comparison of methods for DNA methylation analysis:

Method	Bisulfite-Based?	Scope	Accuracy	Sensitivity	Processing Time, Cost
Southern Blotting	N	genome-wide	quantitative	low	5-6 days, low
HPLC	N	genome-wide	quantitative	low	1 day, low
Tiling Arrays	Y/N	genome-wide/ loci-specific	semi-quantitative	low	1-2 days, Very high
RLGS	N	genome-wide/ loci-specific	quantitative	low	7-14 days, high
Bisulfite Sequencing	Y	loci-specific	semi-quantitative	high	5-7 days, high
MSP	Y	loci-specific	non-quantitative	high	2-4 days, low
qAMP	N	loci-specific	quantitative	medium-high	1 day, medium

HPLC, high-performance liquid chromatography; RLGS, restriction landmark genomic scanning; qAMP, quantitative analysis of DNA methylation by real-time PCR; MSP, methylation-specific PCR

Figure 1.1: Diagram of epigenetic reprogramming events that occur in the male germ line. At fertilization and during the preimplantation stages of embryonic development, gametic patterns of DNA methylation are erased. At ~3.5 dpc, the inner cell mass cells of the blastocyst, cells that are the progenitor cells of all of the embryonic cell lineages, establish a new pattern of DNA methylation. This pattern is the foundation for all of the somatic cell lineages. Germ cells are specified ~6.5-7.5 dpc from the proximal epiblast and migrate through the developing embryo to arrive at the genital ridges at ~10.5-11.5 dpc. Germ cells arrive at the genital ridges with a somatic pattern of DNA methylation. Around this time, the somatic pattern is erased, and, in the male, a new germ cell pattern is established beginning ~13.5 dpc which continues to be established during fetal development.

Epigenetic programming in the male germ line:

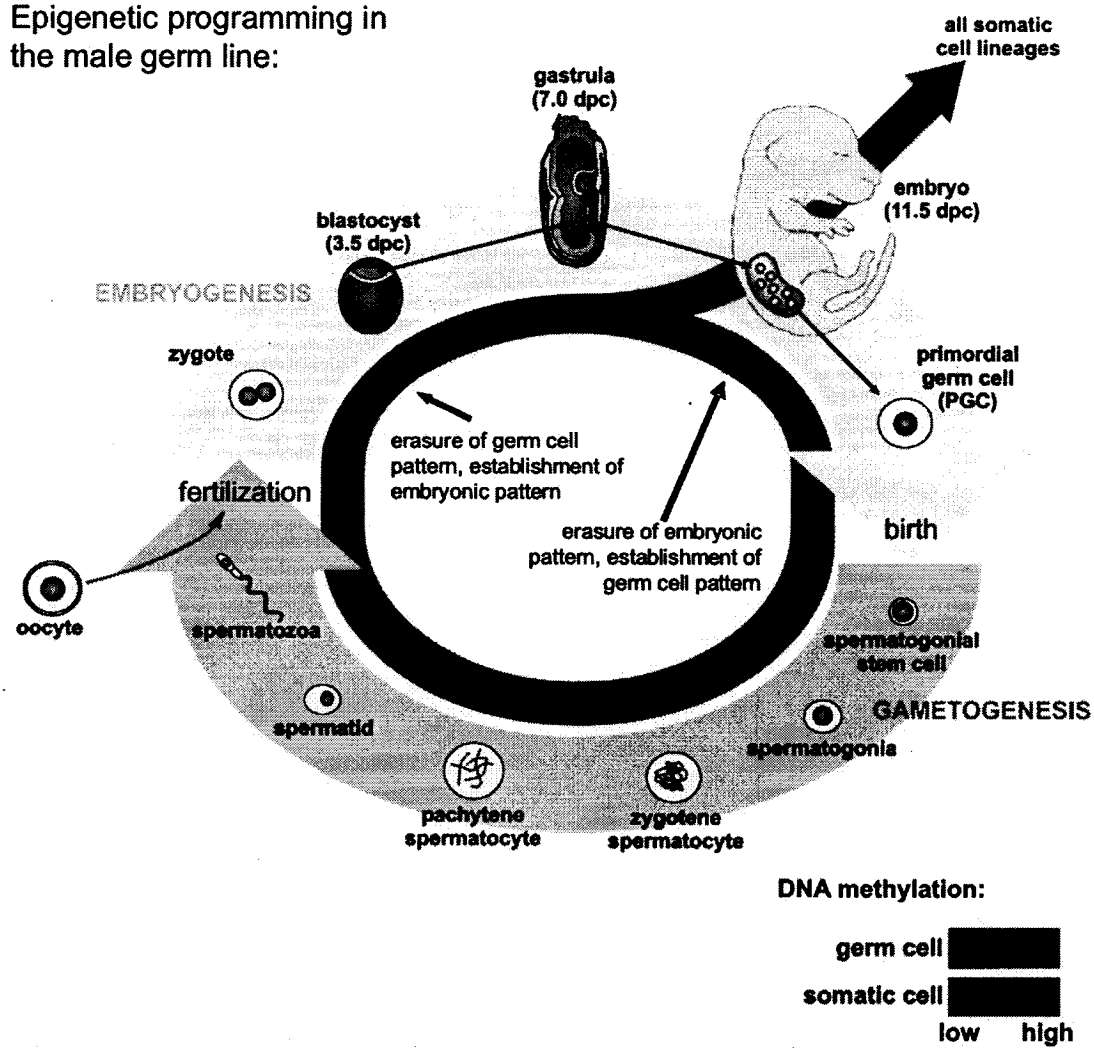


Figure 1.2: Schematic representation of seminiferous tubules within the testis. Germ cell populations are shown in the context of their localization in the tubule, with immature spermatogonia lying along the basement membrane and spermatids near the lumen. The basal and adluminal compartments of the testis are also shown. Adapted from (Gilbert, 2000) and reprinted with permission from Sinauer Associates, Inc.

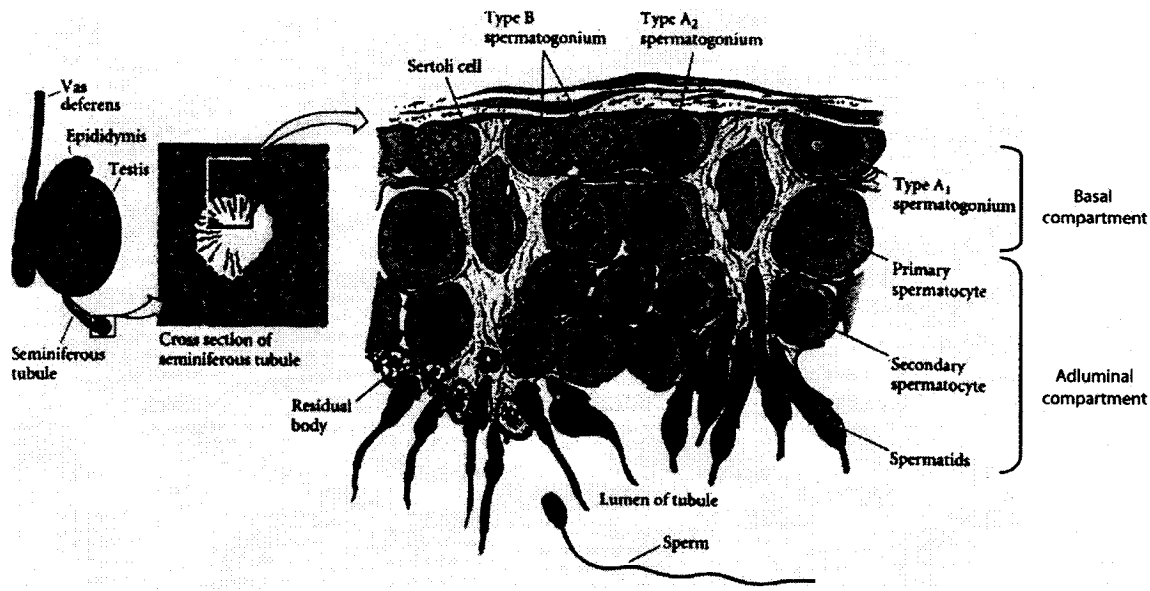


Figure 1.3: Schematic representation of spermatogenesis. A) The development of mature spermatozoa results from germ cells passing through three phases: proliferation, meiosis and spermiogenesis. Mitotic diploid spermatogonia divide and differentiate before entry into meiosis. Tetraploid spermatocytes pair homologous chromosomes, undergo recombination and divide twice without DNA replication to yield haploid spermatids. Spermatids undergo the process of spermiogenesis which results in the formation of mature spermatozoa.

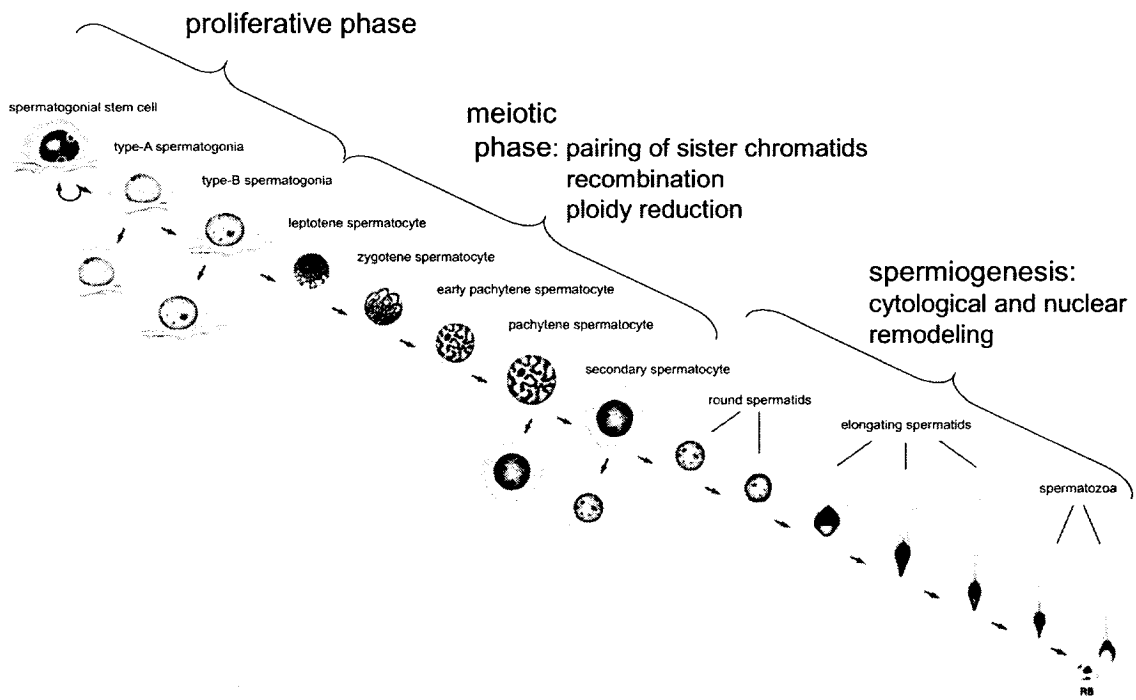


Figure 1.4: Schematic representation of the conversion of cytosine to 5-methylcytosine by DNA methyltransferases. DNMTs catalyze the transfer of a methyl group from the cofactor S-adenosylmethionine (SAM) to the position 5' of the cytosine ring, creating 5-methylcytosine (5mC) and S-adenosylhomocysteine (SAH).

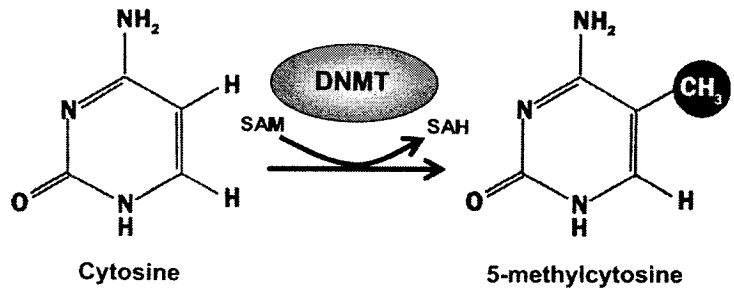


Figure 1.5: The distribution of CpG dinucleotides in the human and mouse genomes. The proportion of CpG dinucleotides located in unique sequence (right half) and transposable element (left half) compartments of the genomes of mouse and human are shown. The proportion of total CpGs found in CpG islands is shown in green. CpG islands are exclusively made up of non-transposable element sequence. Data obtained from Fazzari and Greally (2004).

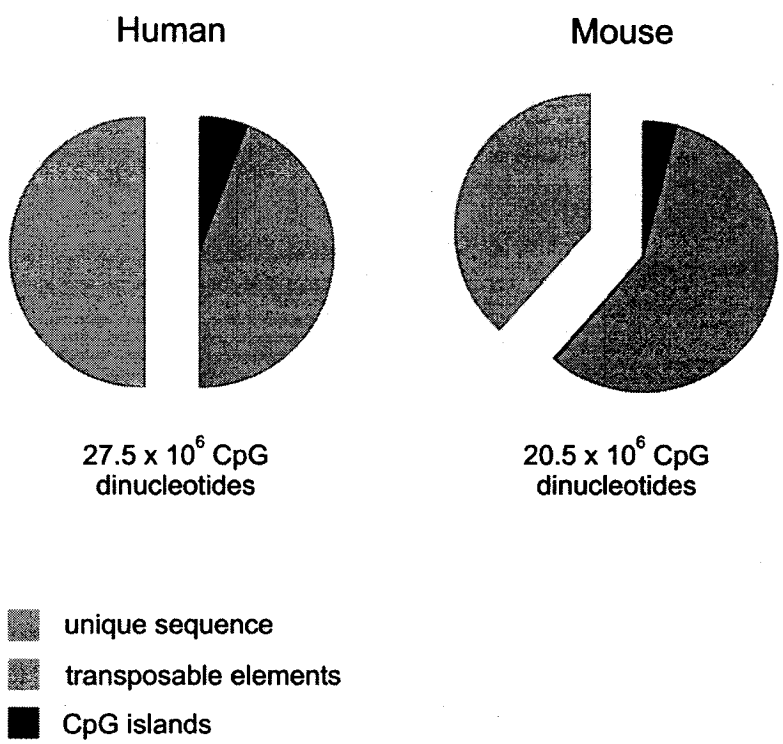


Figure 1.6: Organization of known mammalian DNA methyltransferases. Specific motifs are represented by black boxes; five of the important amino acid motifs involved in catalytic activity are illustrated to demonstrate homology in the catalytic domain (adapted from Bestor (2000) *Hum Mol Gen.* 9:2395). The primary methylation activity ascribed to each protein is indicated (Goll and Bestor, 2005).

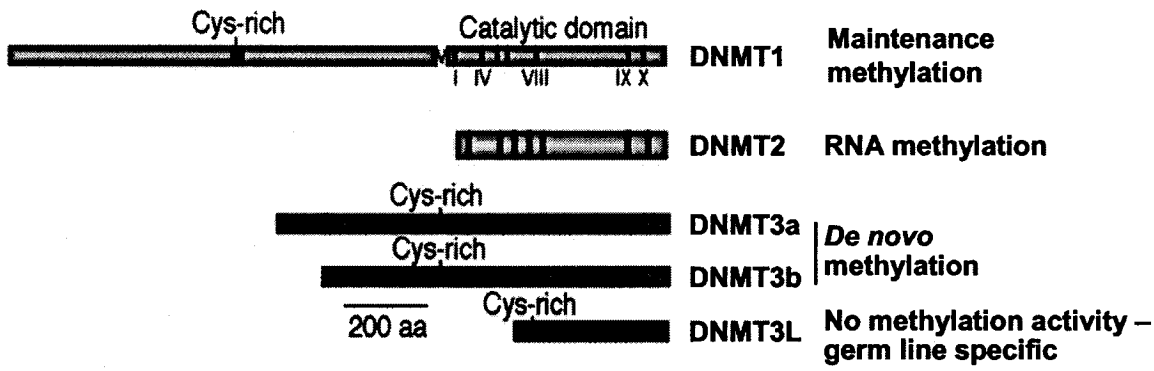


Figure 1.7: Schematic diagram of the RLGS method. NotI sites in genomic DNA are cleaved only if they are unmethylated. Both 5' and 3' ends are labeled by a fill-in reaction with radioactive nucleotides. NotI-NotI DNA fragments are reduced to smaller sizes with the EcoRV restriction enzyme which is not sensitive to DNA methylation. These DNA fragments are run in an agarose tube gel (1^{st} dimension electrophoresis). The gel is extracted from the tubular sheath and while the fragments remain within the gel, they are digested *in situ* with another methylation insensitive enzyme, HinFI. The agarose tube gel is then connected to a polyacrylamide gel and the fragments are run in the second dimension. Spots visible on RLGS profiles represent hypomethylated loci, whereas spots that are missing in comparison to others are hypermethylated.

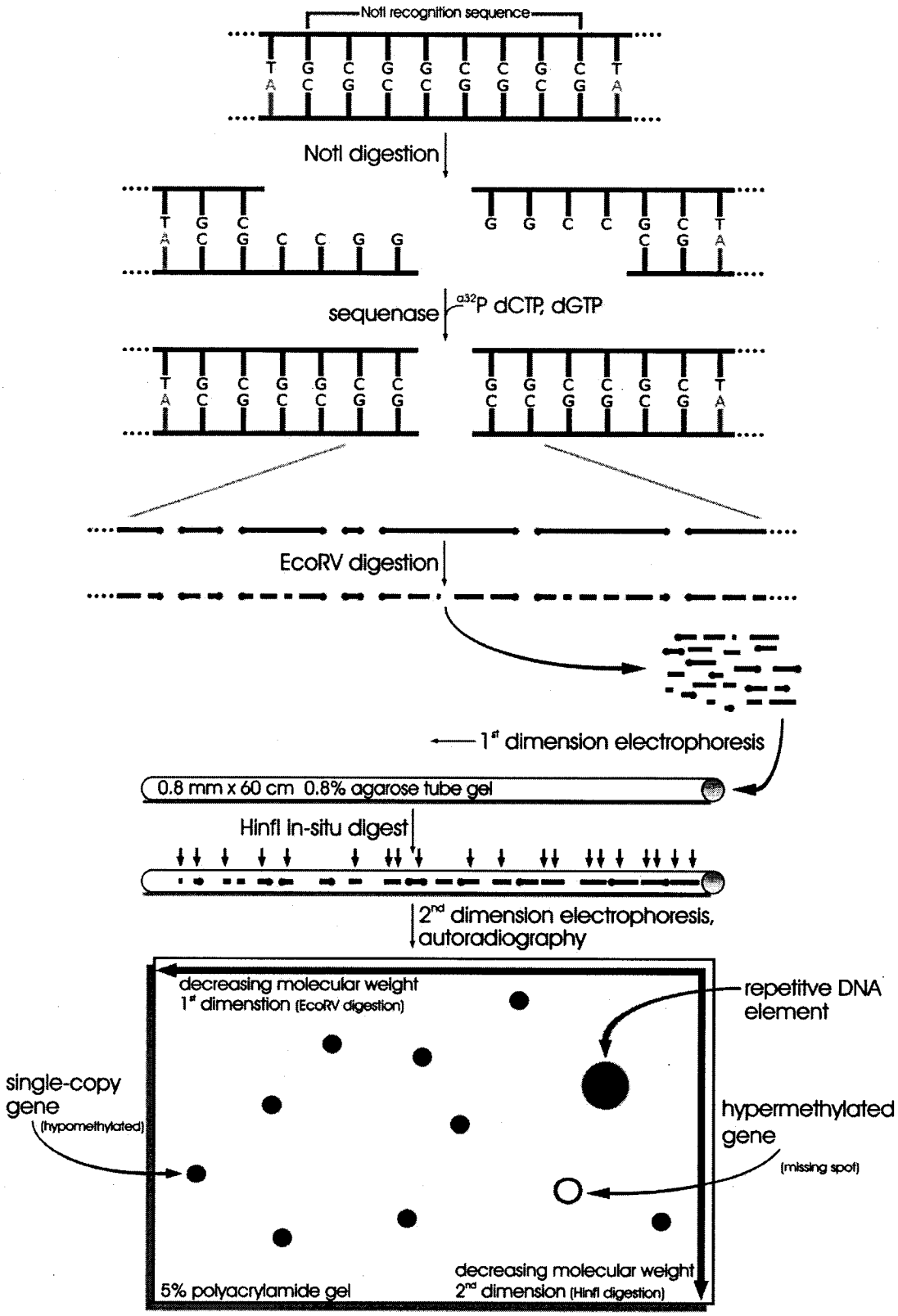
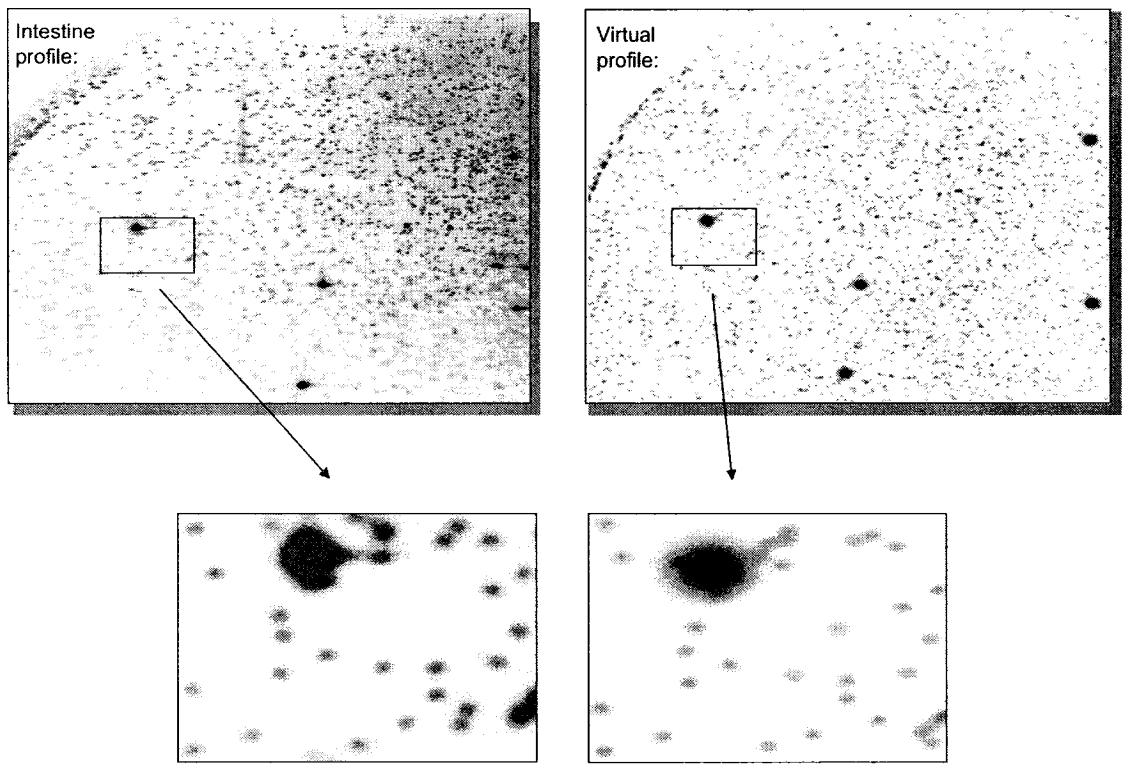

NotI recognition sequence
NotI digestion
sequenase
α32P dCTP, dGTP
EcoRV digestion
1st dimension electrophoresis
0.8 mm x 60 cm 0.8% agarose tube gel
HinfI in-situ digest
2nd dimension electrophoresis, autoradiography
decreasing molecular weight
1st dimenstion (EcoRV digestion)
repetitve DNA element
single-copy gene
(hypomethylated)
hypermethylated gene
(missing spot)
5% polyacrylamide gel
decreasing molecular weight
2nd dimension (HinfI digestion)

Figure 1.8: Virtual and actual RLGS profiles. Virtual RLGS profiles are generated by *in silico* digestion of an entire genomic sequence using the same recognition sequences cleaved by restriction enzymes used in the generation of actual profiles. An algorithm that predicts two-dimensional fragment migration based on fragment length, curvature and rigidity is used to produce the virtual spot pattern. Virtual RLGS can very accurately predict the positions of actual spots; differences arise from missing sequence in current genomic builds or hypermethylation of NotI sites in actual RLGS profiles.



CHAPTER II

Evaluation of a Quantitative DNA Methylation Analysis Technique using Methylation-Sensitive/Dependent Restriction Enzymes and Real-Time PCR

Christopher C. Oakes, Sophie La Salle, Bernard Robaire & Jacquetta M. Trasler

Epigenetics (2006) 1:146-152

ABSTRACT

DNA methylation in mammals has been shown to play many important roles in diverse biological phenomena. Several methods have been developed for the measurement of region-specific levels of DNA methylation. We sought a technique that could be used to quantitatively evaluate multiple independent loci in several tissues in a quick and cost-effective manner. Recently, a few quantitative techniques have been developed by employing the use of real-time PCR, though they require the additional step of sodium bisulfite conversion. Here we evaluate a technique that involves the digestion of non-sodium bisulfite-treated genomic DNA using methylation-sensitive and methylation-dependent restriction enzymes followed by real-time PCR. The utility of this method is tested by analyzing seventeen genomic regions of known tissue-specific levels of DNA methylation including three imprinted genes. We find that this approach generates rapid, reproducible and accurate results (range= ±5%) without the additional time required for bisulfite conversion. This approach is also adaptable for use with smaller amounts of starting material. We propose this method as a rapid, quantitative method for the analysis of DNA methylation at single sites or within small regions of DNA.

INTRODUCTION

In the genomes of mammals, the majority of cytosine molecules within the CpG dinucleotide sequence are methylated at the fifth carbon of the cytosine ring. DNA methylation is involved in normal biological phenomena such as X chromosome inactivation (Avner and Heard, 2001) and genomic imprinting (da Rocha and Ferguson-Smith, 2004). Modifications of DNA methylation patterns have been found to occur in cancer (Robertson, 2005) and aging (Issa, 2000)

Restriction landmark genomic scanning (RLGS) is a useful method that generates information on the DNA methylation status of multiple unique genomic loci from a single DNA sample. We sought a technique that could be used in concert with RLGS to confirm and further evaluate the DNA methylation status of the many loci that are explored by this method. To be efficient for this purpose, we required a rapid, flexible assay, yet one that could quantitatively measure a full range of DNA methylation.

A variety of techniques have been used to detect site-specific levels of DNA methylation (Tollefsbol, 2004). The bisulfite genomic sequencing method is unmatched in its ability to determine methylation at all CpG dinucleotides within a small region of DNA providing the highest level of resolution and information. However, this method is not appropriate for the investigation of large numbers of loci because of the costs and time requirements involved. Novel quantitative techniques that investigate a smaller number of CpG dinucleotides by using sodium bisulfite-conversion, such as Ms-SNuPE (Gonzalgo and Jones, 1997) and Bio-COBRA (Brena et al., 2006), are useful because they are faster, but the analysis requires gel electrophoresis followed by densitometry. QMSP (Lo et al., 1999) and QAMA (Zeschnigk et al., 2004) employ time-saving real-time PCR technology, but still require the additional step of bisulfite conversion. Other drawbacks of the use of sodium bisulfite include incomplete conversion and loss of the DNA sample (Grunau et al., 2001).

Here we describe an alternative technique that can be used to rapidly profile the DNA methylation status of numerous loci without the use of sodium bisulfite. This technique has three basic steps: 1) the digestion of a DNA sample

of interest with several methylation-sensitive enzymes (MSREs) and a methylation-dependent restriction enzyme (MDRE); 2) the designing of primers to specific genomic regions; 3) a real-time PCR amplification reaction to monitor the formation of the PCR product. A similar strategy combining a MSRE and MDRE has been successfully used previously (Yamada et al., 2004); however, the method was limited to a non-quantitative assessment. Although this strategy investigates fewer CpGs than the bisulfite-sequencing technique, by generating data in a single day this method is substantially faster and more economical. There are also several additional advantages over other bisulfite-based quantitative techniques that include a simplified oligonucleotide design strategy, no standard curve measurements and increased template stability.

In this study, we evaluate the ability of this technique to accurately assess site- and region-specific differences of DNA methylation between tissues with known levels of DNA methylation, including three imprinted genes. The results of these experiments demonstrate that this assay is a rapid and accurate way of determining levels of DNA methylation. We find that this method, termed here as quantitative analysis of DNA methylation using real-time PCR (qAMP), is a valuable analytical tool suitable for the investigation of site- and region-specific levels of DNA methylation.

MATERIALS & METHODS:

DNA Isolation

Adult male C57BL/6 mice were obtained from Charles River Laboratories (St-Constant, Quebec). All animal studies were conducted in accordance with the principles and procedures outlined in the Guide to the Care and Use of Experimental Animals prepared by the Canadian Council on Animal Care. Genomic DNA was isolated from liver, intestine, brain and testis using either proteinase K and phenol followed by dialysis (Okazaki et al., 1995) or the DNeasy Tissue Kit from Qiagen (Germantown, MD, USA).

Preparation of DNA templates

First, a set of PCR templates is created for each DNA sample; each set is composed of separate tubes of DNA of equal concentration that have been digested with a single restriction enzyme (Figure 2.1). Restriction enzyme digests include three groups: 1) a mock digestion with no enzyme (sham); 2) MSRE digests which cleave the DNA strand if the restriction site(s) are unmethylated; 3) a MDRE digest using the homing endonuclease, McrBC which cleaves the DNA strand only if it is methylated.

Each DNA sample is homogenized and sheared by repeatedly passing the sample through a 27½ gauge needle attached to a 1 ml syringe resulting in an average fragment size of ~4-5 kb. The DNA is then distributed equally to as many tubes as required for the desired number of individual digestions. Here we used amounts of DNA starting material ranging from 2 µg – 2 ng for each digestion reaction. For the template set using 2 µg per digestion, DNA was diluted to 50 ng/µl, distributed to separate tubes and digested in a volume of 50 µl with 25 units of NotI, HpaII, HhaI, McrBC, or no enzyme for 4-5 hours according to the manufacturer's suggested conditions. NotI and HpaII were purchased from Invitrogen (Carlsbad, CA, USA) and HhaI and McrBC were purchased from New England Biolabs (Ipswich, MA, USA). After digestion, each PCR template was diluted 8-fold in water and incubated at 65° for 30 minutes to inactivate the enzyme. For DNA amounts of 200 ng, 20 ng, and 2 ng digestions were done in a volume of 10 µl with 5 units of enzyme for 2 h, 40 min, and 20 min, respectively. Templates were kept at 4° for short-term usage and at -80° for long-term storage.

Primer Design

Primer design involves the placement of primer pairs that flank both the region of interest and a control region. The control primers are designed to a region that is devoid of any of the restriction sites of the enzymes used in the design of the experiment. This region can be located in unique (single-copy)

sequence anywhere in the genome. The other pair of primers is chosen to flank both MSRE and MDRE restriction sites.

PCR primers were chosen to flank specific regions based on the presence of informative restriction sites within the PCR amplified region with careful consideration of the special enzymatic properties of McrBC. McrBC recognizes two half-sites of 5'-G/A^{m}C-3'[13]. In mammalian DNA, this sequence must be followed by a G to make a CpG dinucleotide. The recognition sequence is a non-palindrome, thus, along with 5'-G^{m}CG-3' and 5'-A^{m}CG-3', McrBC will also recognize the sequences 5'-^{m}CGC-3' and 5'-^{m}CGT-3' (complementary to the recognition sequence). Optimal separation of the two half-sites is 55-103 bp and the enzyme cleaves the DNA in between the two sites approximately 30 bp from either site (Stewart et al., 2000). Best results are obtained when McrBC sites are separated by the optimal distance and located towards the center of the product so that both of the potential cut sites fall within the flanked region and complete cleavage can occur. MSRE sites should be between the primer annealing sequences. Primers were designed using the Primer3 software (http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi) according to the standard principles for successful quantitative PCR outlined in the QuantiTect™ SYBR® Green PCR Handbook from Qiagen. Product sizes range from 100-200 bp. Genomic sequence data was obtained from the University of California at Santa Cruz Genome Browser, version mm6 (http://www.genome.ucsc.edu). All primer pairs were tested to identify the annealing temperature for optimal efficiency and to check for the formation of non-specific products. Primer sequences and genomic loci are listed in Table 2.1.

Quantitative PCR

Cycle threshold (Ct) values are obtained by real-time PCR amplification of the various templates. PCR amplification is initially done on each PCR template within a given set using the control primers to ensure that the templates are of equal concentration and that non-specific cleavage of the DNA sample has not occurred. Each PCR reaction for both control and experimental reactions is done

in triplicate. Acceptable variability in mean template Ct values range from +/-0.3 cycles. Once it has been confirmed that all of the templates in a set fall within the acceptable range, Ct values are then obtained using the primers designed to the region of interest. Changes in the Ct values (ΔCt) of the digested templates are expressed relative to the sham digested template within a given set. Quantitative PCR was done using the QuantiTect™ SYBR® Green PCR Kit (Qiagen) according to the manufacturer's suggested conditions for use with the Mx3000P PCR machine from Stratagene (La Jolla, CA, USA). Reactions were mixed in a total volume of 20 µl with 2 µl of template DNA. Different templates from the same set were always run together in the same PCR run. Ct values were calculated using the MxPro v3.00 software (Stratagene) and expressed relative to ROX, a passive dye. Non-specific amplification was monitored by melting curve analysis of each reaction.

The percentage of methylation of a given site is determined from the change in Ct value by using the basic principle that each successive round of PCR amplification results in approximately a 2-fold increase in the amount of product. Thus, a ΔCt of 1.0 indicates that 50% of the template has been cleaved, 2.0 equals 75% cleavage, etc. For MSREs, the relationship of ΔCt to percent methylation can then be described using the formula percent methylation=$100(2^{-\Delta Ct})$; for MDREs, the relationship follows the inverse function, percent methylation=$100(1-2^{-\Delta Ct})$. Non-linear regression analysis was done using SigmaStat v3.0 software (SPSS).

RLGS and Bisulfite Sequencing

The identification of differentially methylated sites for use in the evaluation of the technique was done using RLGS (Okazaki et al., 1995). Densitometry of RLGS spots was done by exposing the RLGS gel to a phosphorimager screen from Kodak (Rochester, NY, USA). Images were analyzed using the ImageQuant v5.1 software from GE Healthcare (Piscataway, NJ, USA). Spot density values were obtained by comparing a spot of interest to approximately 10-15 surrounding spots of unchanged intensity. The identities of spots of

interest were determined by using a RLGS cloning library as described previously (Yu et al., 2004). Bisulfite sequencing was done as described previously (Warnecke et al., 1998). Primers used to amplify the region of the U2af1-rs1 gene from bisulfite-converted DNA were 5'-GTATAGGTTAGTTGTGTTAT/5'-ACCTACCTAAACAATCACCC as described previously (Zhang et al., 2006).

RESULTS

Identification of Differentially Methylated Sites

The evaluation of the qAMP method was done by comparing methylation levels in liver, intestine, brain and testis determined by RLGS. RLGS is a methylation-sensitive technique that reveals site-specific levels of DNA methylation of multiple genomic loci simultaneously, thus providing a system to test a variety of sequences. Initially, three RLGS spots demonstrating tissue-specific spot intensity were mapped to single NotI restriction sites within the genomic regions of the estrogen receptor-α *(Esr1)*, *Gata-4,* and the imprinted gene *U2af1-rs1* (Figure 2.2). Using a phosphorimager, the densities of these spots were measured relative to surrounding fully unmethylated spots that did not change. The relative spot density is inversely proportional to the percent of DNA methylation of the NotI site.

Analysis of Differentially Methylated Sites

Figure 2.3 depicts the DNA methylation analysis of the imprinted gene *U2af1-rs1* in liver and testis. Surrounding the differentially methylated NotI site visualized in Figure 2.2 are HhaI and McrBC restriction sites (Figure 2.3a). Amplification of this region results in different Ct values for templates digested with different enzymes, whereas control primers amplify at similar values (Figure 2.3b,c). A digestion curve ranging from ±5-fold the amount of enzyme does not result in a significant alteration in Ct values indicating that the DNA digestion is complete with the amounts of enzyme used (Figure 2.4).

To calculate the percentage of methylation from the ΔCt values, it was first necessary to test if the theoretical relationship (outlined in materials and methods) holds true in our experimental system. Approximately one-hundred measurements comparing RLGS spot density to ΔCt values using each of the two enzymes, NotI and McrBC, were done by designing primers that flank twelve additional differentially methylated sites found in the four different tissues in two animals (Figure 2.5 & Table 2.1). Spots were chosen to include a full range of methylation. A non-linear regression analysis of the total data set produces a curve that closely approximates the relationship for NotI (Figure 2.5a). The McrBC data set approximates the MDRE relationship, but deviates from the curve (Figure 2.5b). This occurs because although the NotI site also serves as a McrBC site, the resulting ΔCt is representative of an integrated value from multiple McrBC sites due to the high frequency of such sites in virtually all regions tested. The ΔCt value will be different if heterogeneity in the methylation state of neighboring sites exists within the region flanked by the primers. Furthermore, the value of the standard error estimate for both NotI and McrBC results, in part, from errors associated with RLGS densitometry.

To test the accuracy of the method, equal concentrations of brain and sperm DNA were mixed in a range of ratios (100:0, 90:10, 75:25, 50:50, 25:75, 10:90 & 0:100). Primers that flank a NotI site and several McrBC sites in the *AK142239* gene, which is methylated in brain and unmethylated in sperm DNA, were used to compare the percent methylation (calculated from ΔCt values using the relationship) to the expected values predicted from the ratios (Figure 2.5c). The %methylation values for the MSRE, NotI, closely coincide with the expected percentage in the range from 0-75%, but are less accurate in determining values between 75-100%. The inverse is true for the MDRE, McrBC. This loss of precision is observed in these defined ranges because of the nature of the curved relationship (Figures 2.5a,b). In these ranges, small variations in Ct values will result in a noticeable effect in % methylation. This demonstrates that by using both classes of enzymes, a full range of methylation can be accurately determined.

The imprinted nature of *U2af1-rs1* produces an overall methylation value of approximately fifty percent in somatic tissues within its differentially methylated region (DMR) (Shibata et al., 1996; Zhang et al., 2006). Using the established relationship, calculation of the percentage of methylation from the ΔCt values results in numbers very close to fifty percent (49-60%) for both MSREs and McrBC in all three somatic tissues tested (Figure 2.6a). Because *U2af1-rs1* is a maternally-methylated imprinted gene, the level of methylation within the male germ cells of the testis is very low. Somatic cells comprise only a small fraction of the adult testis cell population (<15%) (Bellve et al., 1977a), therefore, results in the range of 6-20% are expected. The findings also closely match the values determined by RLGS densitometry (Figure 2.2).

To further demonstrate the value of the qAMP method, the commonly used bisulfite sequencing technique was used to analyze the same region of the *U2af1-rs1* gene (Figure 2.6b). Bisulfite sequencing reveals the anticipated low levels of methylation in the testis and approximate 50% methylation in liver due to the imprinted nature of the region. Although more methylated than unmethylated strands were sequenced for liver, it is not statistically different from 50% ($p=0.3$) and previous evaluation of this region has determined the methylation to be specific to the maternal allele (Zhang et al., 2006). To appropriately compare the results, the low level of methylation in the testis is most accurately determined by MSREs. In the analyzed region there are three HhaI sites and one NotI site. All three HhaI sites must be methylated for amplification to occur. Bisulfite sequencing reveals that one strand in twenty is methylated at all three sites, which is equivalent to the qAMP result of 6%. Bisulfite sequencing also reveals a low amount of sporadically methylated CpGs, which substantiates the slightly increased level of methylation that is detected at the single NotI site. These results show that the percentage of methylation ascertained using qAMP is highly comparable to the proportion of methylated to unmethylated CpGs within the assayed restriction sites determined by the bisulfite sequencing method. Futhermore, at this locus, the level of methylation at the restriction enzyme sites is representative of the overall methylation level of the region.

Reproducibility and Sensitivity

The primers designed to the *U2af1-rs1* gene were also used to test the reproducibility and the sensitivity of the assay. Re-digestion of the same four DNA samples using the same amount of starting material (2 µg per digestion) results in very similar values (±5%) for all four tissues using HhaI and McrBC (Figure 2.6c). The sensitivity of the assay using lower amounts of starting material was tested by decreasing the amount of DNA to 200 ng, 20 ng and 2 ng per digestion. The results obtained demonstrate that the assay can accurately determine methylation levels down to 20 ng of DNA with an average range of values obtained for four experiments (all other than 2 ng) in all tissues was determined to be ±3.6% for HhaI and ±6.1% for McrBC. At 2 ng (roughly equivalent to the DNA content of 340 diploid cells), range values increase to ±6.3% for HhaI and ±12.4% for McrBC.

Analysis of CpG Island Methylation

Analysis of two other differentially methylated sites found within the CpG islands of *Gata4* and *Esr1* also reveal tissue-specific methylation levels by using this method (Figure 2.7b,c). The NotI sites found within each of the CpG islands (marked by 'N' in Figures 2.7b,c) are hypermethylated in liver to differing degrees (Figure 2.1). Primer design in CpG islands can be problematic due to the high GC content, however, a multitude of restriction sites within CpG islands provide many alternative design strategies. Although the flanked regions for *Gata4* and *Esr1* are approximately 150bp away from the NotI site, levels of methylation determined by this method for all MSREs are very similar to RLGS findings for both genes. McrBC values for liver also reflect the RLGS result; however, they are generally higher for other tissues and may reflect heterogeneous methylation in these regions.

Analysis of Other Imprinted Loci and Optimization of Primer Design

The method was also tested using primers designed to the previously established DMRs of the imprinted genes *Snrpn* (Shemer et al., 1997)and *H19* (Tremblay et al., 1995). *Snrpn*, like *U2af1-rs1*, is a maternally-methylated imprinted gene and displays the expected result using both MSREs in all tissues (Figure 2.8a). The paternally-methylated imprinted gene *H19* was chosen to demonstrate the effect of primer placement because of the lower McrBC site density found within the DMR (Figure 2.8b). *H19* is 50% methylated in somatic tissues and is expected to be approximately 90% methylated in whole testis. Using primer pair 1, HhaI and HpaII digests produce the expected result in the testis, however, the level is lower (74%) than expected for McrBC. Although there are two McrBC half-sites within the flanked region, they are closer than the optimal distance apart. Other half-sites of the optimal distance are outside the region flanked by primer pair 1, and if methylated, some cleavage would then be expected to occur outside the region. Primer pair 2 was designed to flank these adjacent sites and corrects the percentage of methylation to 90% as detected by the McrBC digest.

DISCUSSION

Our data show that the qAMP method can generate fast and accurate results without the use of sodium bisulfite. In addition to avoiding the drawbacks of using bisulfite, qAMP has several advantages over other methods that quantitatively assay region- and site-specific levels of DNA methylation. Compared to techniques that depend on densitometry for quantitative evaluation, such as Ms-SNuPE and Southern blotting, the use of real-time PCR greatly reduces the processing time and eliminates the need for the use of radionucleotides. In comparison to quantitative techniques that employ real-time PCR, these methods necessitate the design of two sets of oligonucleotides (whether they be primers or probes) because bisulfite-conversion creates both converted (unmethylated) and unconverted (methylated) template DNA. The use

of different oligonucleotides for the purpose of a quantitative measurement requires that differences in amplification/binding efficiencies be controlled for by producing standard curves for each region investigated. Instead, the qAMP assay uses the same primer pair to amplify all DNA templates, thus eliminating the need to relate measurements to constructed standard curve values which increases the accuracy and the throughput of the assay. Secondly, oligonucleotides are not as accurate in discerning methylation as are restriction enzymes. Oligonucleotide binding can occur despite a single nucleotide mismatch making detection of single CpGs problematic. Also, oligonucleotides designed to multiple methylated or unmethylated CpGs will not efficiently bind to heterogeneously methylated templates. The use of different MSREs that assay individual CpGs is a much better strategy to investigate single and multiple CpGs within a region. Thirdly, by using both MSREs and McrBC, the qAMP assay produces two independent and complementary measurements that are derived from both the methylated and unmethylated DNA strands. This allows for an accurate assessment of a full range of DNA methylation and greatly improves the reliability of the results. In contrast, values obtained for unmethylated and methylated templates are expressed as a single ratio in bisulfite-based assays. Finally, the single-stranded nature of bisulfite-treated DNA greatly reduces template stability, whereas, in the qAMP assay, digested DNA used remains double stranded, which is much more conducive for long term usage.

One of the useful advantages of this assay is the ability to examine many candidate loci in a rapid, cost-effective manner. Once loci of interest have been determined, the more thorough bisulfite sequencing method could be used. For screening purposes, inspecting the precise methylation state of every CpG in a given region is not necessary in the majority of cases. Thirty to forty percent of all CpGs in mammalian DNA can be surveyed by using 3-5 relatively inexpensive MSREs, respectively (Schumacher et al., 2006). This tactic would be greatly more efficient to manage large numbers of loci. Specifically, the investigation of a single locus using the bisulfite sequencing method, with sequencing 15 clones to obtain a semi-quantitative measurement, costs an estimated 5-fold more than

the qAMP method as described here using a four-enzyme digest with each template run in triplicate. In addition, processing times are reduced from 5 days to 1 by avoiding lengthy cloning and sequencing steps. Furthermore, once the digestions have been done, the digested DNA can conveniently serve as templates for many PCR reactions. For example, high quality results can be obtained with 1 µg of digested DNA diluted into 1000 µl, enough for 500 individual reactions (166 loci done in triplicate). Once the DNA has been digested, the processing time to re-use the same template to explore additional loci is only 2-3 hrs, whereas it is still ~4 days for re-use of bisulfite-converted DNA with the bisulfite sequencing method.

In order to fully interpret the data generated by this assay, a few caveats must be considered. Firstly, the overall accuracy of the assay is dependent upon the range of DNA methylation of the region. Because of the nature of the curved relationship between the percentage of methylation and changes in the Ct value, MSREs are more sensitive to changes in methylation in the lower percent range, whereas MDREs are more sensitive in the higher percent range (Figure 2.5). Although all PCR reactions are run in triplicate, small variations in mean Ct values of +/-0.3 cycles can occur, which can result in a significant change in the calculated level of methylation. For example, if a DNA region is 90% methylated, a 0.3 Ct variation will result in an error of 17% for MSREs, however the same variation will result in an error of only 2% for MDREs. Clearly, to evaluate a full range of methylation, a combination of data from both classes of enzymes must be considered. If the methylation level of the region is fifty percent, the added MSRE and MDRE ΔCt value is at its minimum possible value, making this range the most sensitive to Ct variation. For this reason, the analysis of imprinted genes within somatic tissue was chosen as the most appropriate test of the accuracy of the assay.

Another important aspect in the interpretation of the results is the consideration of the number of restriction sites for each enzyme within the amplified region. The percentage of methylation is a reflection of the integrated value of all the sites and is interpreted differently for the two classes of enzymes.

The McrBC enzyme only requires two half-sites that are methylated within the amplified region to cleave the DNA strand. As a result, the percentage of methylation is truly representative of the percent of DNA strands that are methylated at two or more sites, or inversely, the opposite of the total percentage of unmethylated strands. For MSREs, if there are multiple sites for the same enzyme, the percentage of methylation is representative of the percent of DNA strands that are methylated at all of the sites. Incongruent results from the two classes of enzymes reflect heterogeneity in the methylation state between neighboring restriction sites. In the case of *U2af1-rs1*, the values for both NotI and HhaI are similar although there are three HhaI sites compared to one NotI site. Because all the HhaI sites have to be fifty percent methylated for this result to occur, it would indicate that there is a low level of site to site variation within this region which is supported by bisulfite sequencing of the region. In this condition, the McrBC result exactly reflects the MSRE result. However, in the cases of *Gata4* and *Esr1*, where the MSREs produce similar results to the RLGS but McrBC is slightly higher especially in the testis, tissue-specific hypermethylation of a few of the McrBC sites could explain the discrepancy. If a higher level of accuracy is required in these particular cases where heterogeneity is suspected, use of the bisulfite genomic sequencing technique would be most appropriate.

The data presented here clearly demonstrate the utility of the qAMP method to determine quantitative levels of DNA methylation in a variety of sequences. Single sites or small regions of DNA can be analyzed without the use of bisulfite conversion while using relatively small amounts of source DNA. Primers can be designed to analyze virtually any restriction site(s) found in non-repetitive sequences. With the added convenience of being able to use the same templates to assay many different regions, this assay provides quality results requiring a modest investment of material and time.

ACKNOWLEDGEMENTS

We thank Nicole Darricarrere for performing the bisulfite sequencing assay on *U2af1-rs1* and Liyuan Deng for excellent technical support. C.C.O. and S.L. are recipients of Canadian Institutes of Health Research Doctoral Research Awards. J.M.T. is a William Dawson Scholar of McGill University and a Scholar of the Fonds de Recherches en Santé du Québec. B.R. is a James McGill Professor of McGill University.

Figure 2.1: Schematic diagram outlining the qAMP procedure.

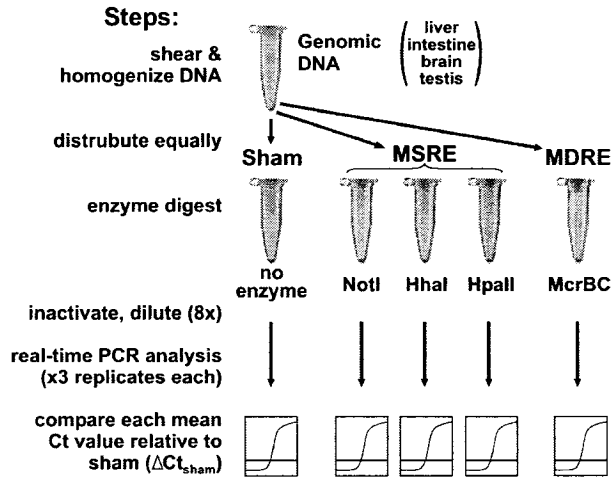


Figure 2.2: Enlargements of RLGS autoradiographs showing tissue-specific spot intensity. Arrows point to three identified spots within different regions of the two-dimensional RLGS gel. Individual spots represent a single NotI site found in a unique copy sequence. Spot densities were measured using phosphorimager screens and compared relative to the intensities of surrounding spots that do not display tissue-dependent changes in intensity.

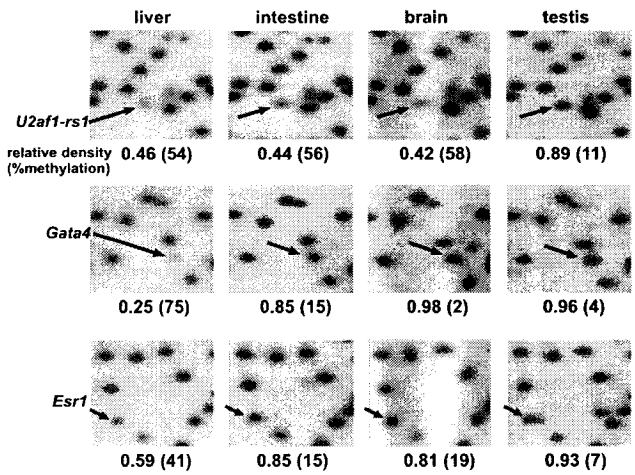
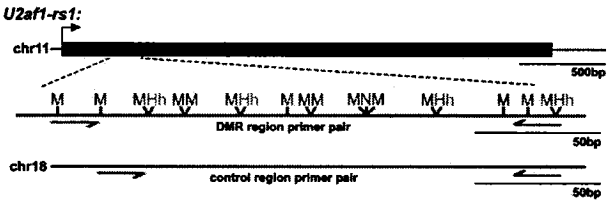
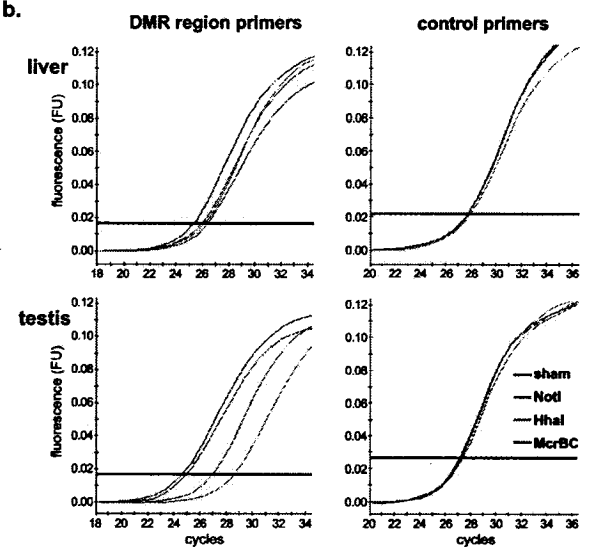


Figure 2.3: Analysis of the DMR region within the *U2af1-rs1* gene. **(a)** The NotI RLGS landmark (N) is found within the 5' region of the single exon and is surrounded by HhaI (Hh) and McrBC (M) restriction sites. Primers are designed to flank the DMR region as well as a second pair to a sequence devoid of NotI, HhaI, or McrBC restriction sites found elsewhere in the genome. **(b)** Liver and testis genomic DNA digested with either no enzyme (sham) or NotI, HhaI, or McrBC are used as templates for real-time PCR using either primer pair. **(c)** Restriction digests result in little or no change in the Ct values using the control primers, whereas differences are observed using the DMR region primers. A shift in the Ct value relative to the sham-digested template reflects a difference in the levels of DNA methylation between liver and testis. For clarity, in **(b)** only single values are shown, whereas values in **(c)** are means of three replicates.

a.



b.



c.

liver

	Ct		ΔCt(sham)	
Template Digest	DMR region	Control	DMR region	Control
sham	25.3	27.8	\	\
NotI	26.1	27.8	0.8	0.0
HhaI	26.0	28.0	0.7	0.2
McrBC	26.3	27.8	1.0	0.0

testis

	Ct		ΔCt(sham)	
Template Digest	DMR region	Control	DMR region	Control
sham	24.8	27.1	\	\
NotI	27.1	27.3	2.3	0.2
HhaI	28.8	27.1	4.0	0.0
McrBC	25.1	27.4	0.3	0.3

Figure 2.4: Effect of digestion conditions on the shift in Ct value. Decreased (5-fold less) and increased (2.5-fold more) amounts relative to the normal amount of enzyme used do not significantly change the results indicating that under- or over-digestion does not occur using 20 U of enzyme. Not more than 50 U are used in the reaction to ensure that the amount of glycerol does not exceed the maximum recommended level of 5% in the reaction. Increased digestion time also does not alter the shift in Ct value. Results presented as mean +/-SD for three replicates.

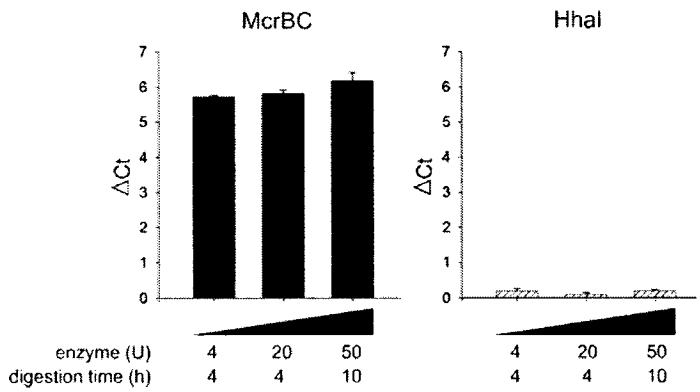


Figure 2.5: The correlation between the shift in Ct value and RLGS spot density. Thirteen genomic loci in liver, intestine, brain and testis digested with NotI **(a)** and McrBC **(b)**. Solid black lines represent a non-linear regression comparing differentially methylated sites in various tissues totaling ninety-nine individual measurements each run in triplicate. The dotted curve represents the theoretical relationship between differences in Ct value and %methylation for two classes of restriction enzymes. **(c)** Comparison of %methylation determined by qAMP versus expected values. NotI and McrBC sites in the *AK142239* gene are hypermethylated in brain and unmethylated in sperm DNA. A primer pair that flanks these sites was used to amplify DNA mixed in the ratios 100:0, 90:10, 75:25, 50:50, 25:75, 10:90 & 0:100. The expected value curve is adjusted for a small amount of residual methylation in sperm DNA at this locus. DNA methylation values closely coincide with the expected values for both enzymes except where an enzyme class-specific decrease in accuracy is predicted to occur. Results presented as mean +/-SD for four replicates.

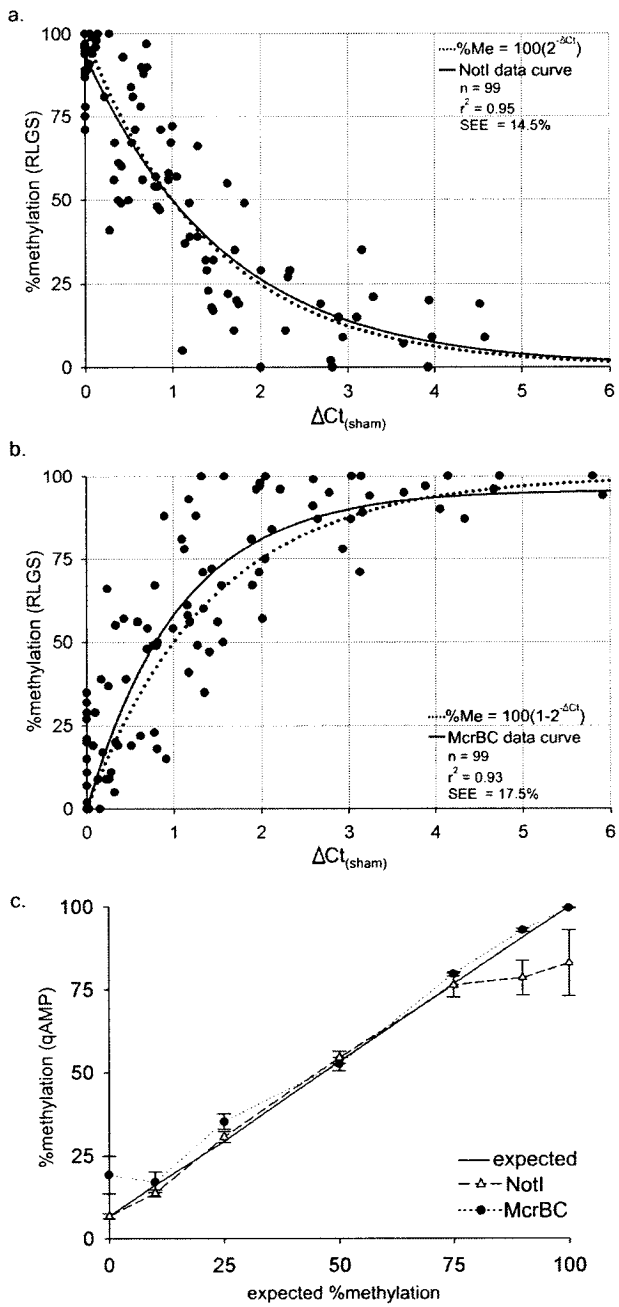

a.
%Me = 100(2^-ΔCt)
NotI data curve
n = 99
r² = 0.95
SEE = 14.5%
%methylation (RLGS)
ΔCt(sham)
b.
%Me = 100(1-2^-ΔCt)
McrBC data curve
n = 99
r² = 0.93
SEE = 17.5%
%methylation (RLGS)
ΔCt(sham)
c.
%methylation (qAMP)
expected
NotI
McrBC
expected %methylation

Figure 2.6: Calculation of the percent methylation of the *U2af1-rs1* gene. (a) Percent methylation was calculated from the ΔCt value for NotI, HhaI, and McrBC using the established relationship. **(b)** Analysis of the *U2af1-rs1* gene from the same DNA samples using the bisulfite sequencing technique. The methylation status of all CpGs in the amplified region is shown as either methylated (black circles) or unmethylated (open circles). The CpGs that are assayed by the various restriction enzymes used in the qAMP assay as well as the qAMP primer binding sites are shown. **(c)** Reproducibility and sensitivity of the assay. Digestion of the same DNA samples from all four tissues were repeated using 2 μg, 200 ng, 20 ng and 2 ng of DNA for each of the sham, HhaI and McrBC digests followed by PCR amplification using the *U2af1-rs1* primers.

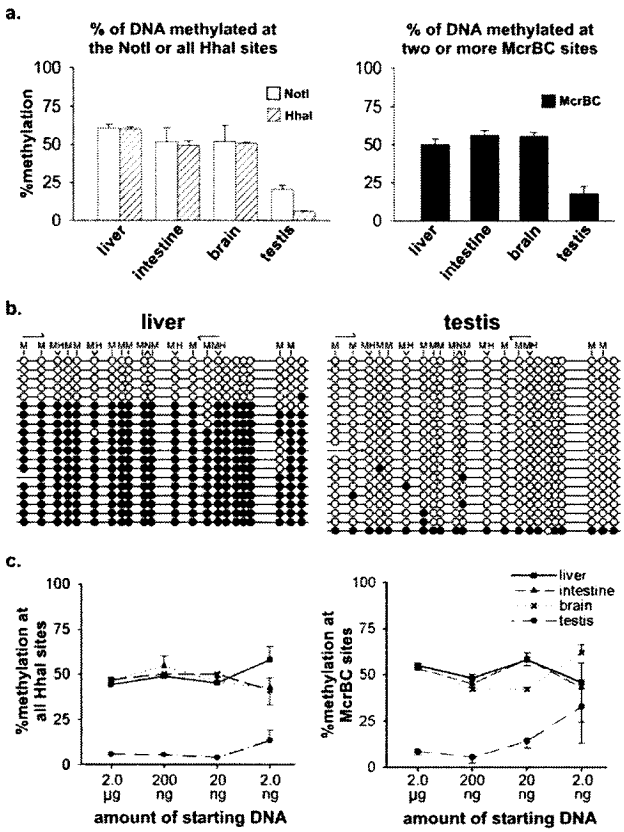

a.
% of DNA methylated at the NotI or all HhaI sites
% of DNA methylated at two or more McrBC sites
%methylation
NotI
HhaI
McrBC
liver
intestine
brain
testis
b.
liver
testis
c.
%methylation at all HhaI sites
%methylation at McrBC sites
amount of starting DNA
2.0 µg
200 ng
20 ng
2.0 ng

Figure 2.7: Analysis of differentially methylated sites in CpG islands. *Gata-4* **(a)** and *Esr1* **(b)** were analyzed using HpaII and HhaI as methylation-sensitive enzymes. The regions of amplification as well as the locations of the NotI RLGS landmarks (N) within the genes are shown. Other mapped restriction sites include HpaII (Hp), HhaI (Hh), and McrBC (M). The percent methylation for each tissue is calculated by using the established ΔCt vs %methylation relationship for each class of enzyme. Results presented as mean +/-SD for three replicates.

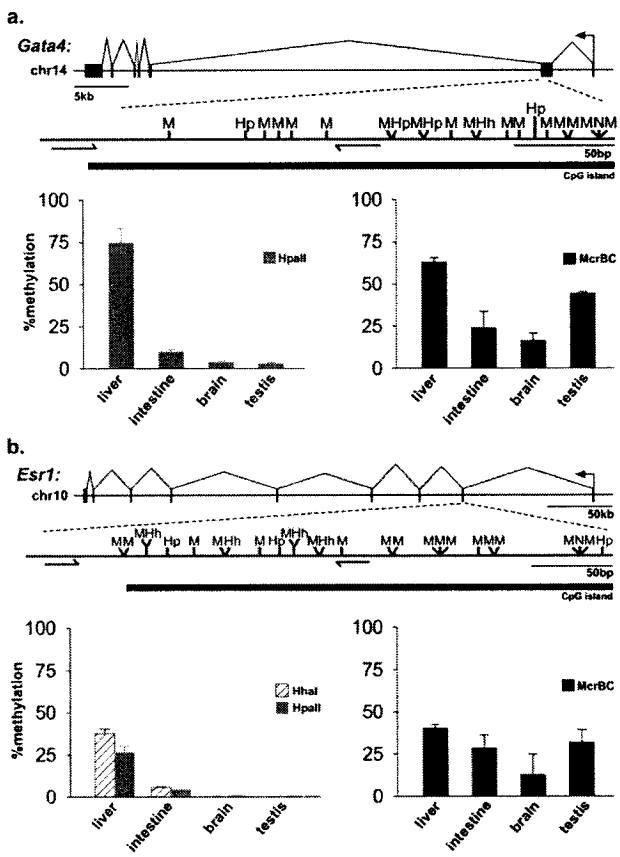


Figure 2.8: Analysis of established DMR regions of two known imprinted genes. The percent of DNA methylation was measured for a maternally-methylated gene, *Snrpn* **(a)**, and the paternally-methylated *H19* gene **(b)**. The results for two different primer design strategies are shown for *H19*. Results presented as mean +/-SD for three replicates. N, NotI; Hp, HpaII; Hh, HhaI; M, McrBC.

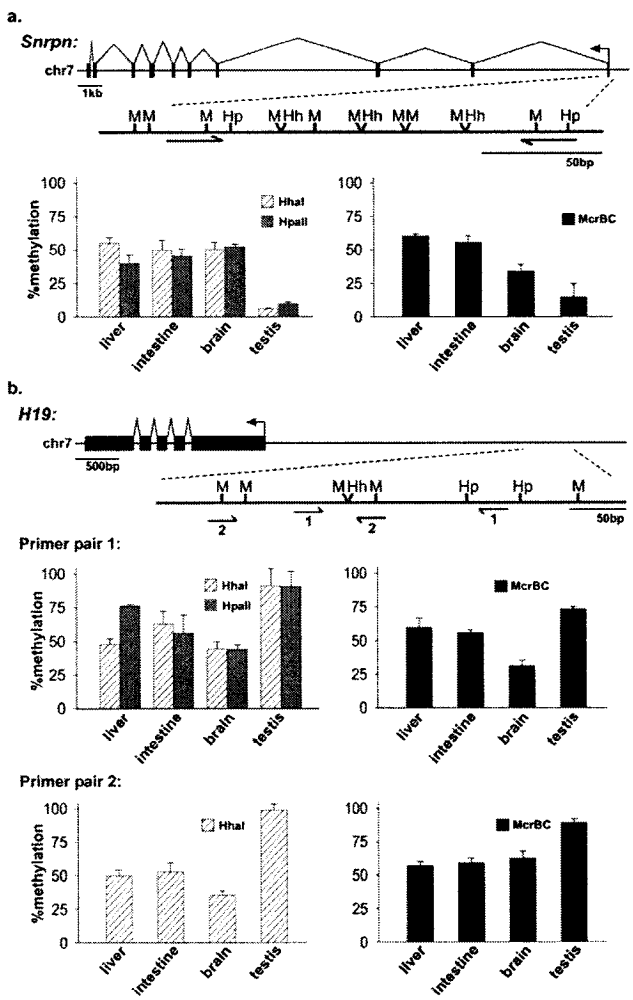


Table 2.1: Primers used for qAMP analysis

Sequence	Genomic Location (mm6)	Forward Primer (5'-3')	Reverse Primer (5'-3')
U2af1-rs1	chr11:22,866,725	GGCCGTACCACAGATAACCA	GCGCAGTTATCCGTCTATCCA
Gata4	chr14:57,770,881	GGCTCAAGGCCCTAAGGTAA	AGCCCCTACCCAGCCTACAT
Esr1	chr10:5,791,650	CTGCCAAAGAATCCTTTCTGAC	GCCCTACTACCTGGAGAACGA
Snrpn	chr7:53,781,658	CTCCTCAGAACCAAGCGTCT	ATTCCGGTCAGAGGGACAGA
H19	chr7:136,993,942	AAAAGCAGAAGGCAGGACAC	CTTTGGCTGTCTCTGGAACAT
AK016529	chr16_random:14,160,131	CTGGCAAGCTGTCCGATAGT	CCGAGATCTTACGCAGGAGATA
AK045080	chr18:47,714,811	TGATTAGCACTCGTTGGACA	CTTTGGAAACGTCTGCGATTA
4933401N24Rik	chr12:87,099,316	TGCCAACTGCTTTTCCAGAT	TAAAGCATCCCAGTGGCAGT
intergenic	chr14:95,005,924	CATCAACCCAGGTACGAAGA	GGAAGAGGAGGAGCCACAGT
intergenic	chr4:113,607,322	GCGCTCCCAAATGAAGGTTA	AGCGACAAGCTCCATTCTCT
AK142239	chr1:135,952,219	GAAGAGCATCACGATCAGCA	AATGCCACTAGCCCAAGGT
intergenic	chr4:52,412,674	GAGAGAGAGACTGTCACCCTGTTA	TTCAGGGTTGCAGCTAAAGA
Adcy5	chr16:34,013,972	AAGCAGGAAGTTGGCGTTTA	CTTCCTCCTCCTGCTTCCA
S76657	chr2:73,984,368	TCTGCAATAAGGGGCTGAAC	AAAGTGGCTTGTGAGGCATT
intergenic	chr15:85,727,534	TTAGGGCTGCCAGGAGTAAG	ACCCACGAGCATCTTTCG
AK136359	chr4:58,946,140	TAGTCCTGCGTGTGGACTGT	GCGTTGCTAGGAGAGAAGGA
6130401J04Rik	chr1:16,803,660	CCCTTTCTTCCCCCTAAGC	GAGACGAGGTGAGCGTCTG
control1	chr18:47,713,995	GCAATCAGGCTTGTAGCAGTT	CATATGCACCATGTGTCTTGG
control2	chr16_random:14,159,508	TCTGATCCCATGTTCCCTGT	TGGGAGGAGAAGCAAATGTC

CONNECTING TEXT

The genome-wide, locus-specific nature of the data generated by the RLGS technique necessitates that any technique used for follow-up studies to confirm and expand on RLGS results must be rapid and flexible enough to manage a large number of loci. The studies in chapter II also demonstrate that tissue-specific differences in DNA methylation often occur as partial differences in methylation; thus, the ability to perform measurements quantitatively is of great importance. The studies described in chapter II make the quantitative measurement of locus-specific DNA methylation possible on the scale necessary for further investigation in connection with the RLGS technique. In chapter III, we employ the RLGS technique to analyze tissue-specific patterns of DNA methylation on a genome wide scale. The qAMP technique is a tool that is extensively used in these studies.

CHAPTER III

A Unique Configuration of Genome-Wide DNA Methylation Patterns in the Testis

Christopher C. Oakes, Sophie La Salle, Dominic J. Smiraglia, Bernard Robaire, & Jacquetta M. Trasler

Proceedings of the National Academy of Sciences (2007) 104(1):228-33

ABSTRACT

In the mammalian lifecycle, the two periods of genome-wide epigenetic reprogramming are in the early embryo, when somatic patterns are set, and during germ cell development. Although some differences between the reprogrammed states of somatic and germ cells have been reported, overall patterns of genomic methylation are considered to be similar. Using restriction landmark genomic scanning (RLGS) to examine approximately 2600 loci distributed randomly throughout the genome, we find that the methylation status of testicular DNA is highly distinct, displaying eight-fold the number of hypomethylated loci relative to somatic tissues. Identification and analysis of more than 300 loci show that these regions are generally located within non-repetitive sequences that are away from CpG islands and 5' regions of genes. We show that a contributing factor for these differences is that the methylation state of non-CpG island DNA is correlated with the regional level of GC content within chromosomes, and that this relationship is inverted between the testis and somatic tissues. We also show that in *Dnmt3L*-deficient mice, which exhibit infertility associated with abnormal chromosomal structures in germ cells, this unique testicular DNA methylation pattern is not established. These special properties of testicular DNA point to a broad, distinct epigenetic state that may be involved in maintaining a unique chromosomal structure in male germ cells.

INTRODUCTION

Although germ cells faithfully carry parental genetic information to the next generation, they edit the epigenetic information that they inherit. Epigenetic marks in the form of DNA methylation are erased in primordial germ cells in favor of the establishment of a sex-specific, *de novo* epigenetic program (Reik et al., 2001). This initial programming is not only important for gametogenesis (Bourc'his and Bestor, 2004; Hayashi et al., 2005; Kaneda et al., 2004), but also provides the basis for genomic imprinting (Ferguson-Smith and Surani, 2001), as some sex-specific marks persist following a second wave of demethylation that occurs during early embryogenesis. Furthermore, it has been shown recently that some epigenetic marks can pass from one generation to the next by avoiding both periods of reprogramming in what is termed 'epigenetic inheritance' (Morgan et al., 1999; Rakyan et al., 2003). A clear appreciation of the genome-wide nature of the germ cell epigenetic program is paramount to the understanding of these important processes.

DNA methylation in mammals occurs at cytosine residues in CpG dinucleotides. CpG islands are generally thought to be devoid of methylation, while the rest of the genome is highly methylated in non-developing tissues(Reik et al., 2001). A recent study in humans confirmed that CpG islands represent the only major unmethylated genomic compartment in somatic (brain) DNA and implied that germ cells were similar, but this was not tested directly(Rollins et al., 2006). Although DNA methyltransferases are differentially regulated and essential during germ cell development (La Salle et al., 2004; La Salle and Trasler, 2006b), few studies to date have directly examined germ cell genome-wide DNA methylation patterns. Due to the larger number of germ cells produced in the male than in the female, their greater accessibility and their continual renewal, most germ cell studies focus on spermatozoa and testicular tissues. In the adult, the testis is composed of over 80% germ cells, mainly in the meiotic and post-meiotic phases of spermatogenesis (Bellve et al., 1977a), and thus will be referred to in this study as non-somatic. Although the data are limited, imprinted loci(Ferguson-Smith and Surani, 2001), some repetitive sequences

(Sanford et al., 1984) and tissue-specific genes (MacLean and Wilkinson, 2005), point to a unique state of DNA methylation in the testis. Previous studies using RLGS have found methylation differences between the testis and other tissues (Shiota et al., 2002; Song et al., 2005), but these studies either did not identify the sequences or the analysis was restricted to tissue-specific genes. In this study, we have used RLGS as well as other techniques to explore, in a non-biased fashion, the DNA methylation pattern in the testis.

RESULTS AND DISCUSSION

RLGS was used to survey approximately 2600 genomic *Not*I restriction sites that are mostly found in CpG islands, but also in non-coding unique and repetitive sequences not within CpG islands. Analysis of testis, liver, intestine and brain from adult mice revealed 241 spots that display tissue-specific intensity (Figure 3.1). Although each tissue demonstrates a unique spot profile, more than one-half of all differentially methylated spots are specific to the testis. The testis displays 8-fold the number of exclusively hypomethylated spots compared to the average of the somatic tissues. A comparison of adult testis and mature sperm isolated from the cauda of the epididymis display identical RLGS patterns, consistent with the fact that the adult testis is composed of a majority of mature germ cells (Figure 3.2).

Using a combination of the mouse RLGS spot cloning library (Yu et al., 2004), and the virtual RLGS method (Smiraglia et al., 2007), we identified the genomic location of approximately half (125) of the differentially methylated spots (Figure 3.3 & Table 3.2). We further identified the genomic locations of 48 'absent' spots that correspond to unchanged methylated *Not*I sites, and, combining 89 spots identified previously (Yu et al., 2004), 186 unchanged unmethylated spots. Over 90% of loci that are unmethylated and unchanged in all tissues are found in CpG islands and within the 5' region of genes; unchanged methylated loci are found in regions of low CpG content and away from 5' regions (Figure 3.4). Interestingly, differentially methylated (DM) loci are found in sequences of diverse CpG content and in various locations within and outside of

known genes. Focusing on the loci that are either DM between testis and somatic tissue (testis-specific DM) or DM between different somatic tissues (somatic-specific DM) reveals that testis-specific DM loci are less likely to be located in 5' regions of genes or to be found in strong CpG islands than are somatic-specific DM loci. Only 20% of testis-specific DM loci are within CpG islands and include genes that are previously known to be differentially methylated, such as imprinted genes (*Cdkn1c, U2af1-rs1, and Nap1l5*) and testis-specific genes (*Ddx4* and *Hspa1l*), whereas the remaining 80% (non-CpG island) represent novel DM loci.

Seventy-five percent of *Not*I sites are found within CpG islands (Fazzari and Greally, 2004). Due to this bias, by using RLGS we would expect to find a majority of DM loci in CpG islands if they are located randomly. The high proportion of DM loci being found in non-CpG islands is highly significant for both testis and somatic tissues ($P<1x10^{-17}$ and $P<1x10^{-7}$, respectively) clearly demonstrating that a unique state of DNA methylation exists outside of CpG island sequences, especially in the testis. This also implies that the numbers of testis-specific DM loci that are found away from CpG islands are vastly underrepresented on RLGS profiles.

In the mouse, *Not*I sites are found within some interspersed repetitive elements, predominantly in the internal sequence of intracisternal-A particle (IAP) and in the long-terminal repeats (LTRs) of early transposon (Etn) and other LTR-containing families of repeat sequences (Aota et al., 1987; Baust et al., 2003; Jurka et al., 2005). All unchanged unmethylated loci were found solely within non-repetitive sequences, whereas approximately one-half of the unchanged methylated loci were found to be within various interspersed repeats, particularly IAP (Table 3.1). An analysis of 240 and 60 *Not*I sites (present on vRLGS profiles) within IAP and Etn repeats, respectively, revealed consistent hypermethylation in all tissues tested, as revealed by absence on actual RLGS profiles (Figure 3.5). Despite the high level of methylation of repetitive DNA, a limited number of full-length LTR-containing repeats were found to be unmethylated in somatic tissues. Interestingly, all these repeats were methylated

in the testis. We did find several repetitive elements to be hypomethylated in the testis that are specific to solitary LTR fragments belonging to ERVK (class II) and MaLR (class III) of the LTR family of repeats. The differential methylation state of longer-length LTRs versus solitary LTRs demonstrates that some small repetitive sequences are not targets for methylation in the testis. Transcripts derived from these specific subfamilies of repetitive elements are present and are developmentally regulated in female germ cells and early embryos (Peaston et al., 2004). These sequences may serve a functional role in germ cells (Shapiro, 2005), or may not be of sufficient length and/or are too divergent to be targeted by the methylation machinery.

The germ cells of the testis have a highly unique global transcriptional profile owing to the vast number of different gene products that are produced during meiosis and spermiogenesis (Shima et al., 2004). Although testis-specific DM loci are generally not found in typical promoter regions, hypomethylation of loci found in or near to genes could be related to increased gene activity. Of the 72 DM loci found within expressed sequences, tissue-specific expression levels are known for 62 in the GNF SymAtlas database. Of these, only 2/21 loci found in 5' regions, *Ddx4* and *Cdkn1c*, and 1/44 non-5' loci, *Hspa1l*, demonstrated a correlation in all four tissues of increased expression with hypomethylation (Figure 3.4c). Only nine additional loci showed a correlation in three of four tissues. The poor correlation with expression, coupled with the fact that approximately a third of all DM loci are not found associated with known expressed sequences and many are within repetitive sequences, suggests that in addition to its role as a transcriptional regulator, DNA methylation has additional functions, especially outside of 5' regions.

The majority of loci that are found to be differentially methylated in this study occur within non-CpG island sequences in the testis. As a potential explanation, we observed that those loci that are hypermethylated in the testis are significantly more commonly located within R (reverse)-type bands on chromosomal ideograms, and that loci hypomethylated in the testis are commonly located in G (Giemsa) bands (Figure 3.6a and Figure 3.7).

Heterogeneous banding patterns reveal regional differences in chromosomal structure and are related to the disparity in the long-range G+C nucleotide composition that occurs along chromosomes (Bernardi et al., 1985). R-Bands generally contain above-average and G-bands below-average GC content. We tested if the sequence that surrounds each locus in these two groups differs in overall GC content. We found that loci that are hypomethylated specifically in the testis are found in regions of lower GC content (Figure 3.7). Consistent with the banding pattern results, loci that are hypomethylated in somatic tissue are located in regions of higher GC content, opposite to what is found in the testis. To further compare the differential nature of the relationship between methylation and GC content between testis and somatic tissues, the level of methylation of all identified non-CpG island hypomethylated loci in each tissue was compared to its regional GC content. In all somatic tissues, there is a significant negative relationship between the level of methylation and GC content demonstrating that lower GC content regions are more likely to be hypermethylated (Figure 3.6b). In the testis, this relationship is inverted, illustrating that lower GC content regions are specifically hypomethylated.

Due to the GC rich nature of the *Not*I recognition site (GCGGCCGC), RLGS sites are relatively rare in non-CpG island DNA and are biased towards sequences of higher GC content. To avoid this bias and the potential influence of nearby genes and repeat elements on DNA methylation, we focused on *Hha*I, *Hpa*II, and McrBC sites in non-5', non-CpG island, and non-repetitive sequences using the qAMP assay (Oakes et al., 2006). A survey of 104 small amplified regions evenly spaced along chromosomes 4, 10, 17 and X in liver and testis shows striking differences in the pattern of DNA methylation between the two tissues (Figure 3.8a and Figure 3.9). Consistent with the RLGS findings, lower levels of methylation are generally found in the testis (Figure 3.8b). In liver, non-CpG island sequences are generally highly methylated in GC-poor regions while lower levels are observed in some GC-rich regions. In the testis, the methylation status of these sites is again found to be inverted using this alternate approach. A comparison of the regional GC content with the difference in methylation

between the two tissues reveals a significant correlation between lower GC content and hypomethylation on these four chromosomes in the testis (Figure 3.8c).

These data show that by taking into account broad regional characteristics of chromosomes, we are able to explain a significant proportion of the differences of DNA methylation that occur at single sites or small regions in non-CpG island sequences at both a whole-chromosome and a genome-wide scale. Because these differences occur broadly, are more likely to occur away from regulatory regions of genes and correlate with broad attributes of chromosomes, it raises the possibility that the testicular DNA methylation pattern may have a role in maintaining a unique chromosomal state that is capable of undergoing germ cell-specific processes, such as meiosis. The *Dnmt3L* gene is highly expressed during the perinatal period of reprogramming in male germ cells; yet, the primary defect that results from the disruption of the *Dnmt3L* gene is observed after birth in meiotic germ cells in which highly abnormal chromosomal structures are seen (Bourc'his and Bestor, 2004). If the testis-specific methylation patterns we describe for non-CpG island sequences have a role in chromosome structure or meiosis in germ cells, we would predict they might be altered in the germ cells of *Dnmt3L*-/- mice. To address this, spermatogonia were isolated from *Dnmt3L*-/- mice and the methylation levels at various DM loci determined by RLGS were analyzed using qAMP (Figure 3.10). In the spermatogonia from *Dnmt3L*-/- mice, all loci examined that were hypermethylated in the testis failed to gain the normal levels of methylation found in wild-type spermatogonia and testis. Other DM loci are also hypomethylated. Testis-specific hypomethylated loci remained normal, indicating that the primary influence of *Dnmt3L* is to promote DNA methylation. Furthermore, levels of methylation in wild-type spermatogonia resemble the levels found in the adult testis demonstrating that the testis-specific pattern was acquired in early germ cell development at these non-CpG island sequences.

We have demonstrated the novel finding that male germ cells have highly unique patterns of DNA methylation with most of the hypomethylation found in non-repetitive, non-CpG island sequences. These hypomethylated sequences

do not correlate with expression patterns; this is consistent with the fact that they are generally not found in the regulatory regions of genes. Instead, they are correlated with regional chromosome features, such as banding patterns and GC content. Although a previous study inferred a paucity of non-CpG island hypomethylation in germ cells (Rollins et al., 2006), experimental evidence in the present study indicates a prominent role for regulation of DNA methylation in this compartment of the testis genome. Because greater than half of the total genomic sequence and CpG dinucleotide content belongs to non-repetitive, non-CpG island DNA (Fazzari and Greally, 2004), these modifications might be expected to have prevalent effects on the overall structure of chromosomes in the testis. We also find that in the spermatogonia of *Dnmt3L*-/- mice, a model that displays abnormal chromosomal structures, at the sites examined the unique testicular DNA methylation pattern is not established. We propose that the maintenance of a unique, germ cell-specific chromosomal structure may require a distinct configuration of DNA methylation patterns in non-CpG island DNA.

MATERIALS AND METHODS

RLGS and spot identification

Adult male C57BL/6 mice were from Charles River Laboratories (St-Constant, Quebec). Genomic DNA was isolated from tissues using proteinase K followed by phenol extraction(Okazaki et al., 1995). RLGS was done as described previously (Okazaki et al., 1995). Three replicate RLGS profiles were generated for each tissue. Each spot was assessed a score between 0-4 based on its intensity relative to surrounding, fully unmethylated spots. An intensity score of 0, 1, 2, 3 & 4 is representative of 100, 75, 50, 25, & 0 percent methylation, respectively. Visual assessment of spot intensity was confirmed by densitometry. All DM spots were observed to be changed in all three replicates. Spot identities were determined by using either the RLGS cloning library (Yu et al., 2004) or the virtual RLGS method (Smiraglia et al., 2007). Loci identified using virtual RLGS were confirmed by obtaining the corresponding BAC clone

(Roswell Park Microarray Core Facility, Buffalo, NY) and running RLGS mixing gels. The methylation status of 36 identified loci was confirmed by using the qAMP methylation assay (Oakes et al., 2006). Briefly, DNA is digested with various methylation-sensitive and methylation-dependent restriction enzymes and amplified using real-time PCR. Changes in cycle threshold values were used to calculate the percentage of methylation at each restriction enzyme site. For the analysis of chromosomes 4, 10, 17 and X, the percentage of methylation of each locus was determined by averaging the values obtained for all of the restriction enzyme sites analyzed within the amplified region. To obtain purified *Dnmt3L*-deficient spermatogonia, heterozygous *Dnmt3L*-deficient mice (Bourc'his et al., 2001) were crossed with GOF18deltaPE-Oct4/GFP mice (Yoshimizu et al., 1999), and spermatogonia were isolated from the testes of 6-dpp offspring using fluorescence-activated cell sorting, 3-5 mice were pooled in each group. All primers are listed in Tables 3.3 & 3.4, Supplementary Information.

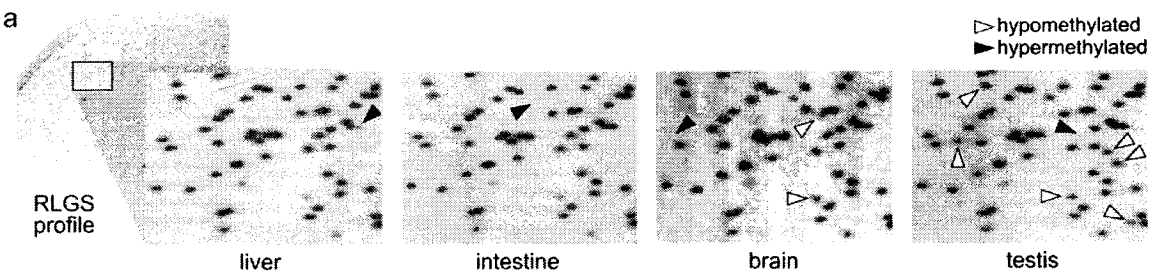
Data analysis

All positional and sequence information was obtained from the UCSC genome browser (Kent et al., 2002) Mouse August 2005 assembly, mm7 (http://genome.ucsc.edu/). Chromosomal bands were defined as a G-band if they appear black or grey, and as an R-band if white on ideograms. CpG Islands were defined by the criteria of Takai and Jones (Takai and Jones, 2002) (G+C≥55%, CpGobs/exp≥0.6, ≥500bp), 5'regions were defined as being within -3 to +1 kb of the transcriptional start site. DM loci were divided into the sub-categories of 'somatic-specific DM' and 'testis-specific DM' if the unique state of hyper- or hypomethylation occurred solely in a single somatic tissue or in the testis, respectively. SigmaStat v3.0 software (SPSS) was used for statistical analysis with $p<0.05$ considered to be significant. Gene expression data were obtained from the GNF SymAtlas v1.2.4 expression database (http://symatlas.gnf.org/SymAtlas/) using the GNF1M, gcRMA dataset.

ACKNOWLEDGEMENTS

We thank T. Bestor for the gift of the *Dnmt3L*-/- mice and H. Scholer for the gift of the GOF18deltaPE-Oct4/GFP mice. C.C.O. and S.L. are recipients of Canadian Institutes of Health Research (CIHR) Doctoral Research Awards and Montreal Children's Hospital Studentship Awards. J.M.T. is a William Dawson Scholar of McGill University and a Scholar of the Fonds de Recherches en Santé du Québec. B.R. is a James McGill Professor of McGill University. These studies were supported by funding from the CIHR.

Figure 3.1: **RLGS analysis of tissue-specific DNA methylation.** Two-dimensional spot profiles are produced by digestion of genomic DNA with the methylation-sensitive restriction enzyme, *Not*I, followed by radioactive end-labeling. A spot is produced if the site is unmethylated and absent if methylated. **(a)** Enlargements of liver, intestine, brain and testis RLGS profiles. Black and white arrowheads denote hyper- and hypomethylated spots, respectively. **(b)** Summary of RLGS spot data. A total of 1945 spots are visible over four tissues; 241 display differential intensity between tissues. Consistent methylation of approximately 1/3 of all spots is predicted by subtracting the total visible spots from the total number of spots calculated by virtual RLGS analysis of the mouse genome. **(c)** Tissue-specific breakdown of differentially methylated (DM) spots. Hyper- and hypomethylated spots can occur either in one tissue or can be shared. More than half of all observed DM spots are testis-specific.



b

Spot Behavior		No. of Spots	% of Total
Differentially Methylated		241	8.2
Unchanged	Unmethylated	1704	57.7
	Methylated*	1009	34.2
Total*		2954	100

*derived from the virtual RLGS profile

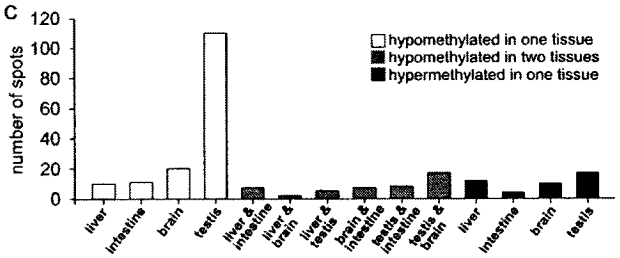


Figure 3.2: **Enlargements of RLGS profiles of testis and cauda epididymal sperm demonstrating identical spot patterns.**

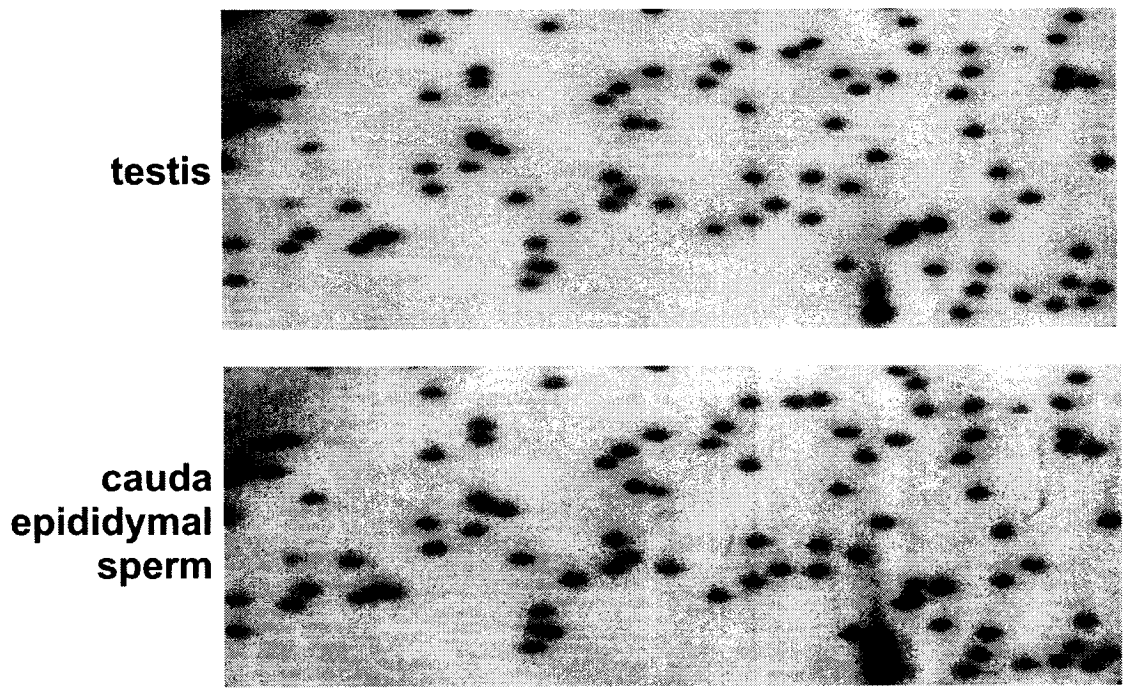
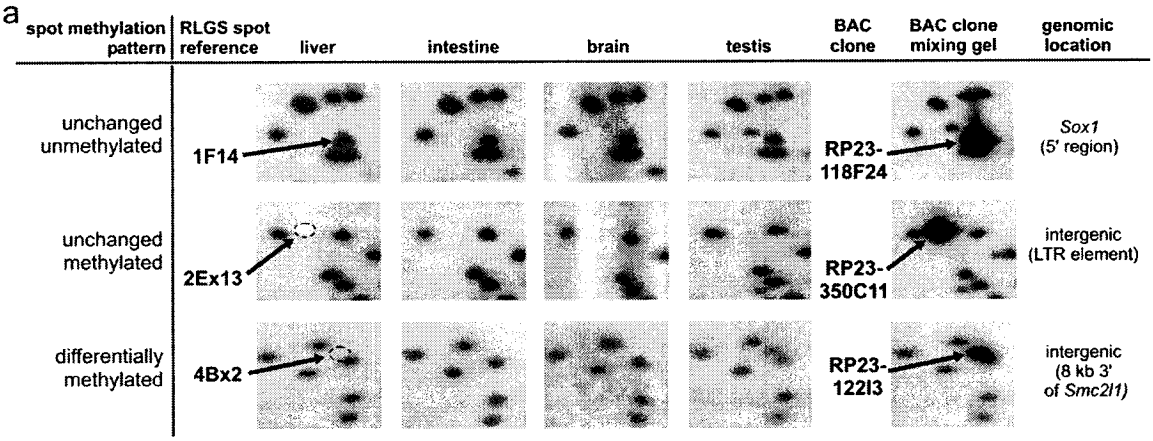


Figure 3.3: Identification and confirmation of the genomic locations of RLGS spots. **(a)** The identity of 260 unmethylated, methylated, or differentially methylated spots were established by observing enhanced spot intensity on RLGS gels produced by mixing plasmid or BAC clone DNA with genomic testis DNA as background. **(b)** Summary of identified spots and the corresponding loci. Forty spots were found to originate from the same *Not*I site as another spot and were removed from the dataset to ensure that each spot relates to a single locus. **(c)** Verification of the DNA methylation status of identified loci using the qAMP assay (Oakes et al., 2006). Real-time PCR amplification of the region after digestion with the methylation-sensitive enzymes *Hha*I(H) and *Not*I(N) and the methylation-dependent enzyme McrBC(M) confirms the testis-specific hypomethylation of the 4Bx2 locus using all three enzymes. The methylation status of a total of 36 identified loci was verified using this method.

a



b

Spot Methylation Pattern	No. of Identified Spots	No. of Identified Loci
unchanged unmethylated	186	161
unchanged methylated	48	47
differentially methylated	125	111
Total	359	319

c

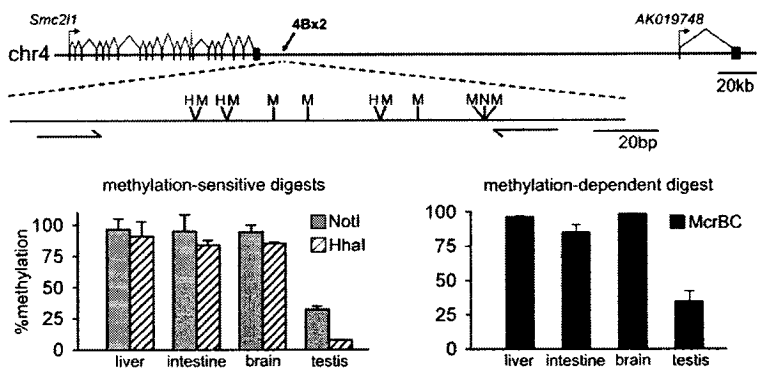


Figure 3.4: **The positional distribution of all identified loci with respect to CpG islands and 5' regions. (a,b)** All identified loci were grouped according to their DNA methylation pattern in different tissues as being either unmethylated, methylated or differentially methylated (DM). DM loci are sub-divided into somatic-specific and testis-specific DM loci. The percent of loci from each group that are found in CpG islands and within 5' regions are shown. Proportionally fewer testis-specific DM loci are within CpG islands and 5' regions ($p<0.05$). **(c)** The correlation of expression of known genes with tissue-specific methylation of DM loci. Approximately 2/3 of all DM loci are located within or near known genes. Relative levels of tissue-specific expression are known for 66 of these genes in the GNF gene expression database. Tissue-specific hypomethylation correlated with expression in only a minority of DM loci located in the vicinity of genes.

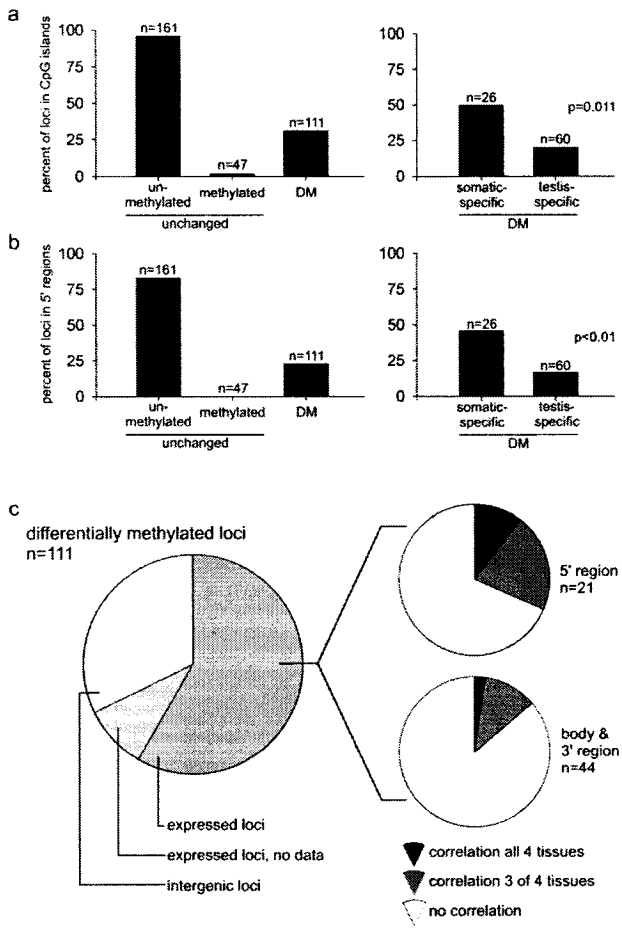

a
percent of loci in CpG islands
100
75
50
25
0
n=161
n=47
n=111
un-methylated
methylated
DM
unchanged
n=26
n=60
p=0.011
somatic-specific
testis-specific
DM
b
percent of loci in 5' regions
n=161
n=47
n=111
un-methylated
methylated
DM
unchanged
n=26
n=60
p<0.01
somatic-specific
testis-specific
DM
c
differentially methylated loci
n=111
5' region
n=21
body &
3' region
n=44
expressed loci
expressed loci, no data
intergenic loci
correlation all 4 tissues
correlation 3 of 4 tissues
no correlation

Figure 3.5: Absence of IAP and Etn repetitive element spots on RLGS profiles of liver, intestine, brain and testis. **(a)** Restriction maps showing *Not*I sites in the *gag* and LTR sequences of IAP and Etn, respectively. Vertical positions of RLGS fragments are dependent on the size of *Not*I-*Hin*fI fragments. The sizes of the vertical RLGS fragments are shown corresponding to both the 5' and 3' sides of the *Not*I cut site. Three of these fragments are of sufficient size to be visible on a RLGS profile and are labeled A, B and C. **(b)** There are ~240 and ~60 *Not*I sites visible by virtual RLGS in IAP and Etn repeats, respectively. The vertical positions of the repetitive fragments are apparent in the lower portion of the virtual RLGS gel. Spots are spread horizontally due to variation in the flanking sequence that generates the horizontal spot position **(c)** IAP and Etn spots are absent in RLGS profiles of all four tissues indicating that these repeats are constitutively hypermethylated.

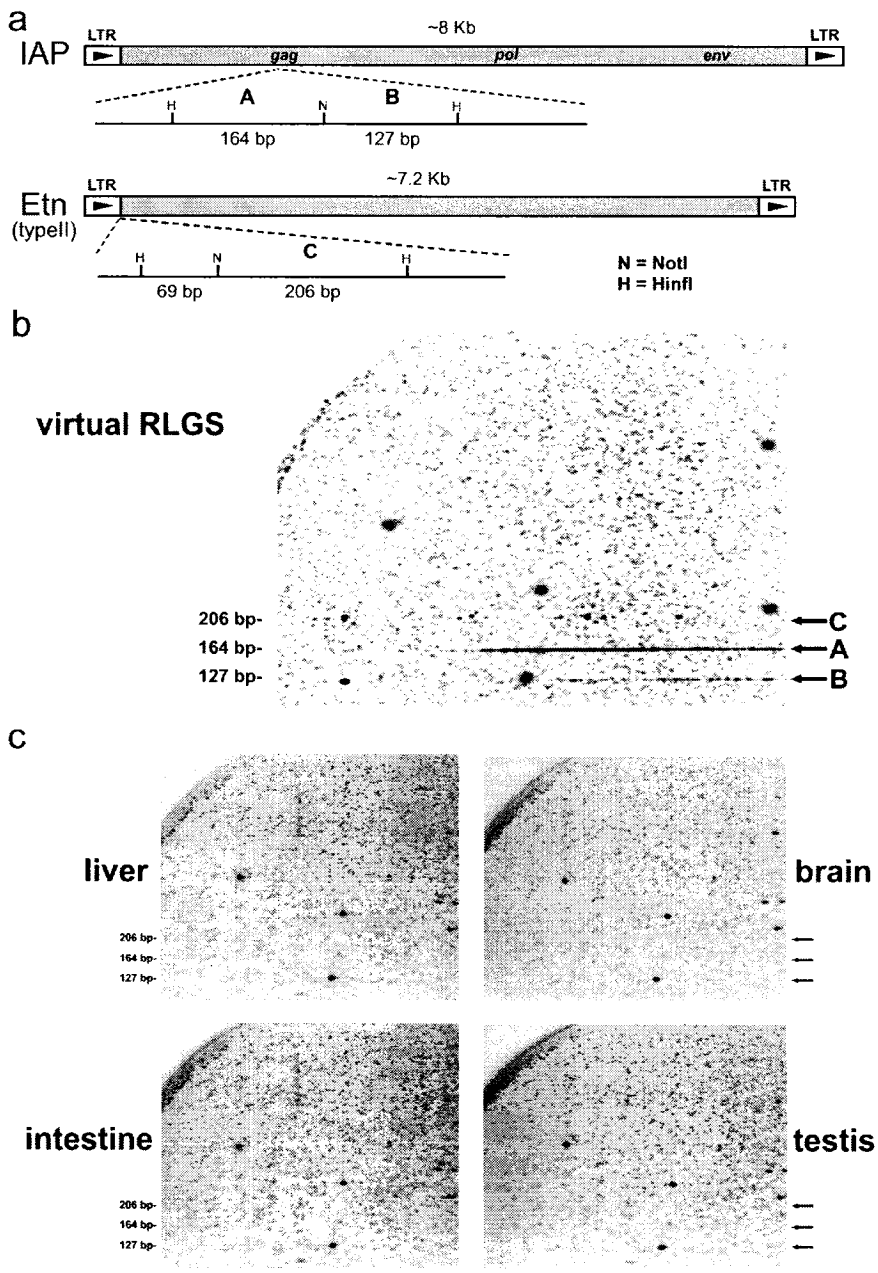

a
IAP
LTR
~8 Kb
LTR
gag
pol
env
A
B
H
N
H
164 bp
127 bp
Etn
(typeII)
LTR
~7.2 Kb
LTR
C
H
N
H
69 bp
206 bp
N = NotI
H = HinfI
b
virtual RLGS
206 bp-
164 bp-
127 bp-
C
A
B
c
liver
brain
intestine
testis
206 bp-
164 bp-
127 bp-

Figure 3.6: Analysis of non-CpG island hypomethylated loci. **(a)** The chromosomal locations of non-CpG island loci that are hypomethylated in testis, somatic tissues, or both (open, black & grey triangles, respectively) are shown. **(b)** Correlation analysis of all hypomethylated non-CpG island loci comparing the %GC of 50 Kb of flanking sequence with the level of methylation as assessed by visual inspection of each RLGS spot in different tissues. The relationship between methylation and GC content is inverted in testicular DNA. Ideograms adapted from the UCSC genome browser.

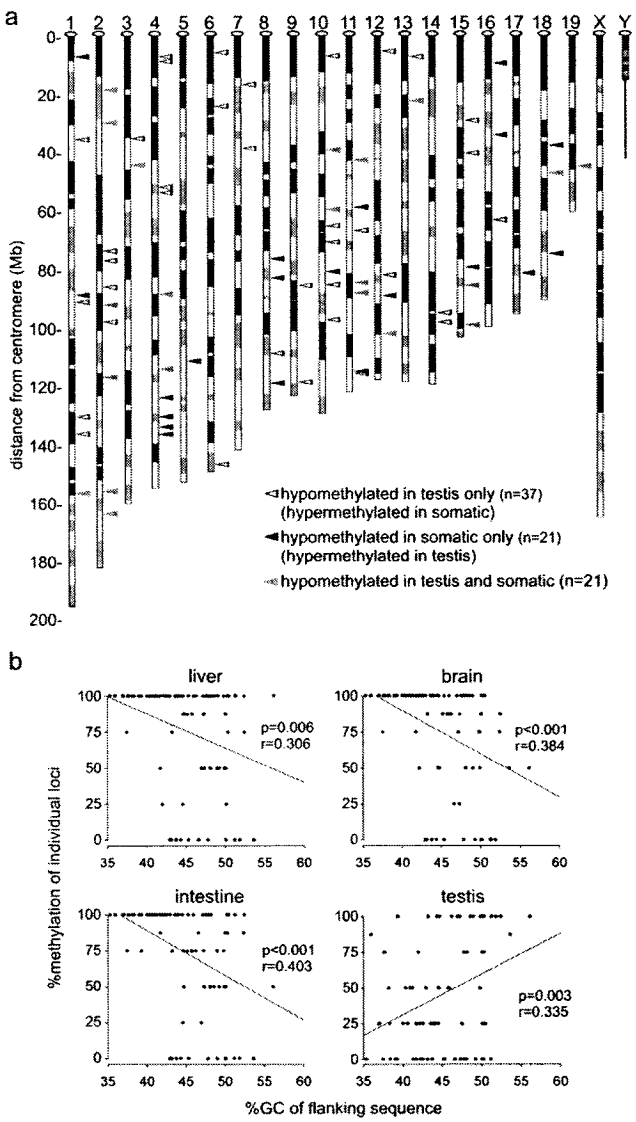

a
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 X Y
0-
20-
40-
60-
80-
100-
120-
140-
160-
180-
200-
distance from centromere (Mb)
hypomethylated in testis only (n=37)
(hypermethylated in somatic)
hypomethylated in somatic only (n=21)
(hypermethylated in testis)
hypomethylated in testis and somatic (n=21)
b
liver
brain
intestine
testis
p=0.006
r=0.306
p<0.001
r=0.384
p<0.001
r=0.403
p=0.003
r=0.335
%methylation of individual loci
%GC of flanking sequence

Figure 3.7: Chromosomal banding patterns and levels of GC content associated with non-CpG island loci. Chromosomal banding patterns and levels of GC content that are associated with differentially methylated loci that are hypomethylated in testis, somatic tissues or both. The percentage of loci in each group that are located in R-bands and the %GC of 50 Kb of sequence flanking each locus is shown. Loci that are hypomethylated in the testis are less likely to be found in R-bands and in regions of higher GC content. Statistical analysis done using χ^2 and ANOVA, error bars represent +/-SEM.

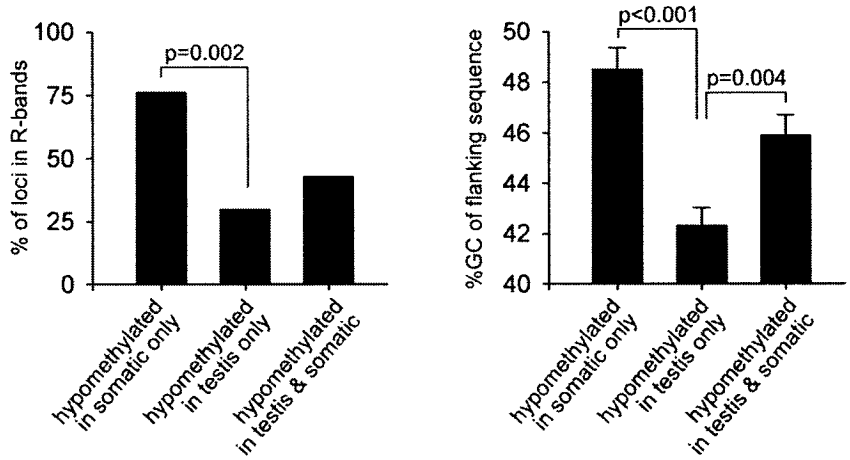


Figure 3.8: DNA methylation analysis of non-CpG island DNA on chromosomes 4, 10, 17 and X using the qAMP assay. Primers were designed to flank McrBC and *Hha*I or *Hpa*II restriction sites placed at roughly five Mb intervals along each chromosome. Amplified regions were chosen only on the basis of the sequence not being in the proximity of a CpG island, a known 5' region, or within a repetitive sequence. **(a)** Analysis of chromosome 4. The percent of DNA methylation of each amplified region in liver (black dash) and testis DNA (grey dash) is shown. Differences in methylation are shown by arrows. The ideogram and GC percent are shown as are the positions of hyper- and hypomethylated loci identified using RLGS. **(b)** The average difference in %methylation of all DM regions analyzed on each chromosome shows an overall decreased level of methylation in testis versus liver in 3 of 4 chromosomes. Unchanged regions (<10% difference) were excluded. **(c)** The correlation between the difference in methylation in each region on all three chromosomes and the GC percent of flanking sequence shows a lower relative level of methylation in the testis in sequences of lower GC content. Chromosome 4 ideogram and %GC were obtained from the UCSC genome browser.

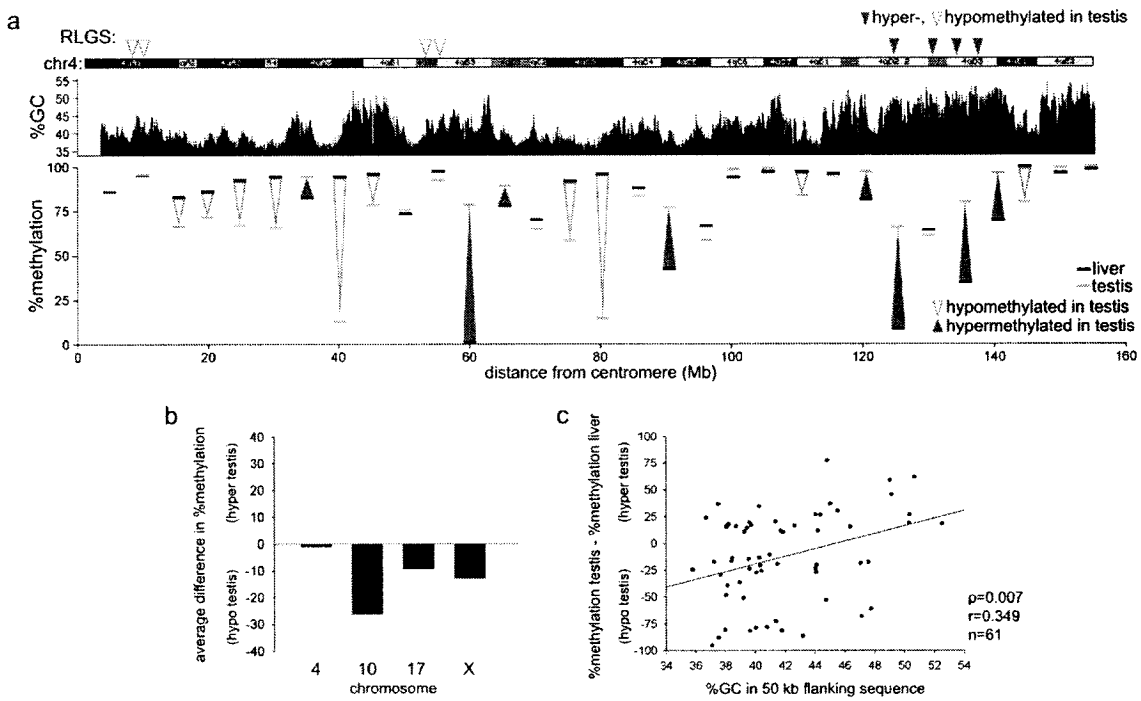

a
RLGS:
hyper-, hypomethylated in testis
chr4:
%GC
%methylation
liver
testis
hypomethylated in testis
hypermethylated in testis
distance from centromere (Mb)
b
average difference in %methylation
(hyper testis)
(hypo testis)
4
10
17
X
chromosome
c
%methylation testis - %methylation liver
(hyper testis)
(hypo testis)
ρ=0.007
r=0.349
n=61
%GC in 50 kb flanking sequence

Figure 3.9: **DNA methylation analysis of non-CpG island DNA on chromosomes 10, 17 and X using the qAMP assay.** DNA methylation analysis of non-CpG island DNA on chromosomes was done as described for chromosome 4 in Figure 3.8. The positions of known testis-specific genes that are hypomethylated in the testis are shown.

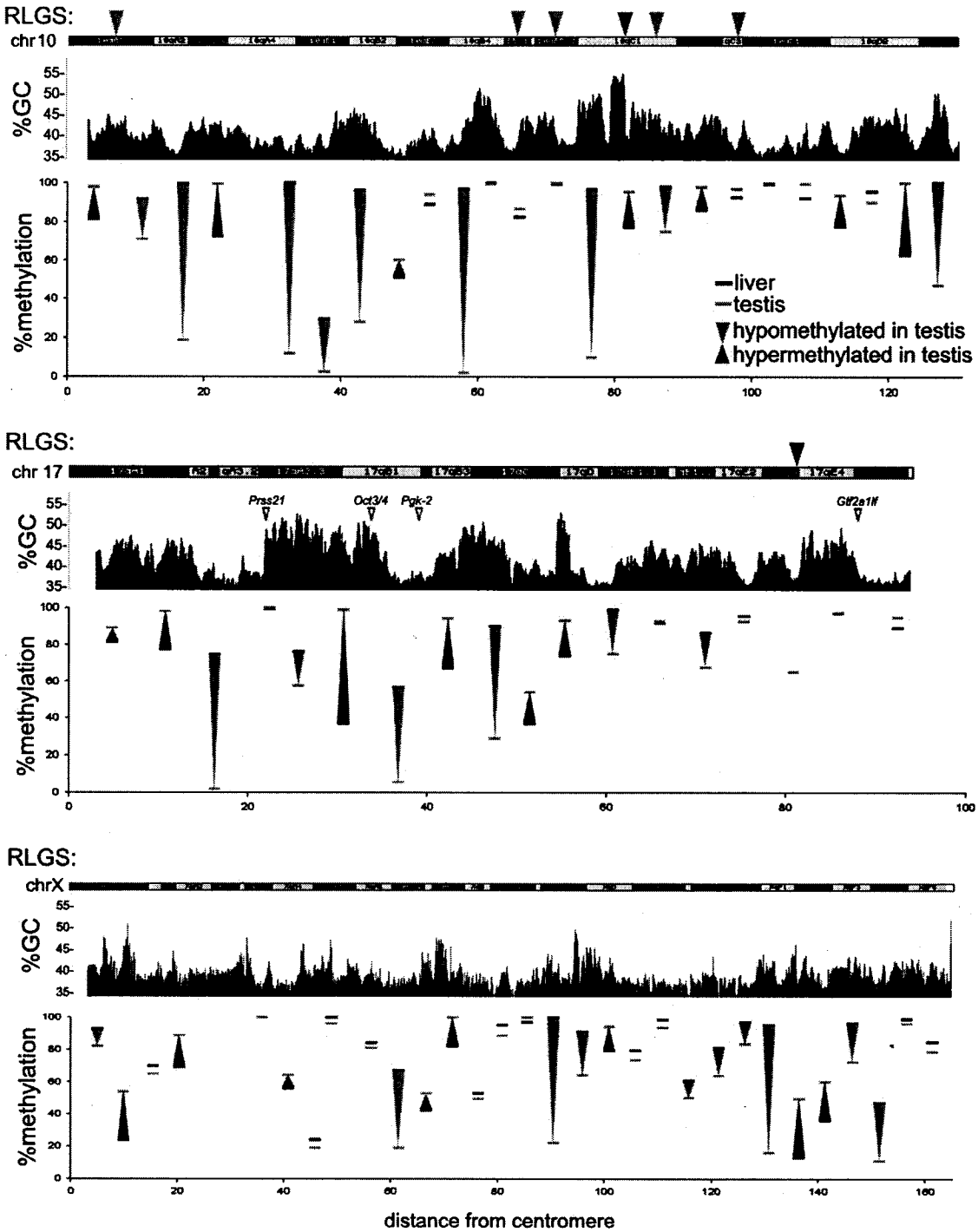

RLGS:
chr10
%GC
%methylation
liver
testis
hypomethylated in testis
hypermethylated in testis
RLGS:
chr 17
Prss21
Oct3/4
Pgk-2
Gtf2a1lf
%GC
%methylation
RLGS:
chrX
%GC
%methylation
distance from centromere

Figure 3.10: Levels of DNA methylation in *Dnmt3L*-deficient spermatogonia and adult tissues. The percent methylation of DM RLGS loci that are **(a)** hypermethylated in testis, **(b)** hypomethylated in testis and **(c)** somatic-specific DM was determined using qAMP. Loci that are hypermethylated in the testis fail to gain normal levels of methylation in *Dnmt3L*-deficient spermatogonia. Analyzed loci were all in non-5' regions. L, liver; I, intestine; B, brain; T,testis.

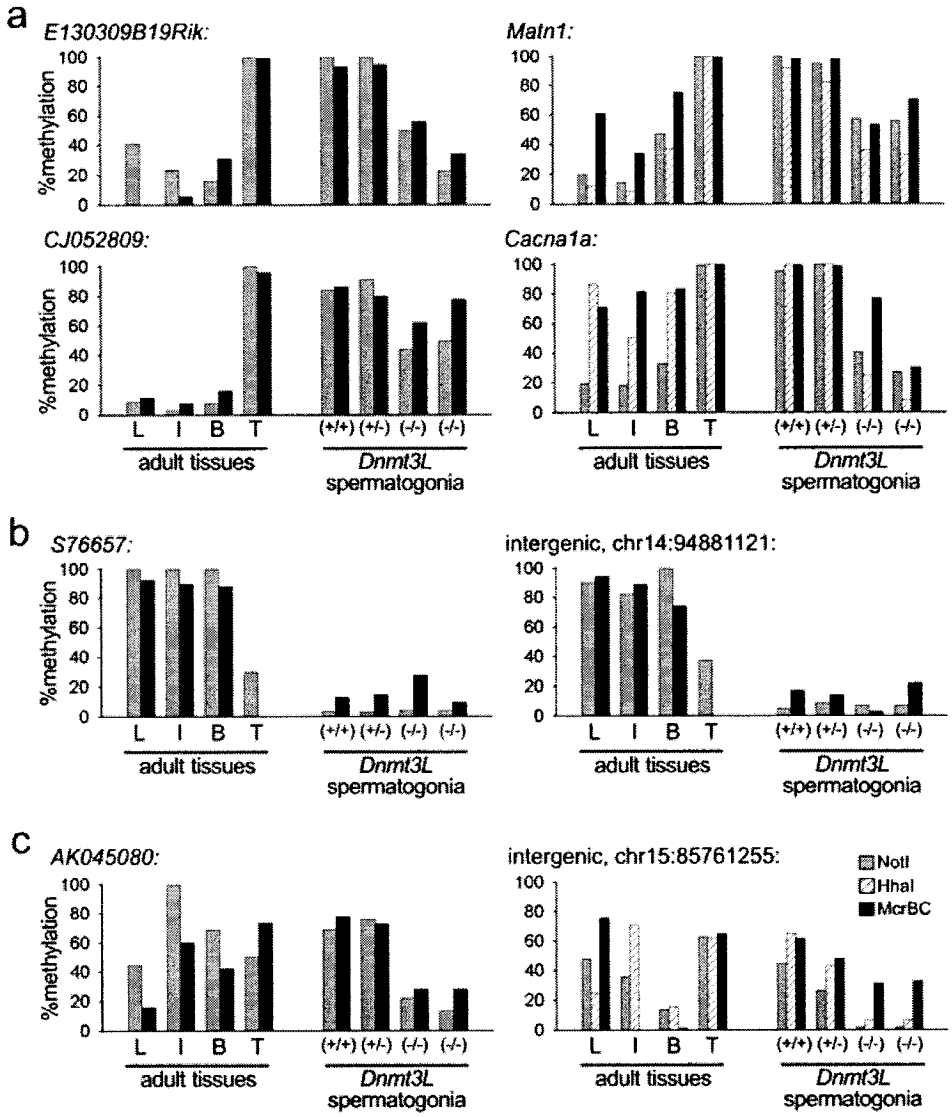

a
E130309B19Rik:
Matn1:
CJ052809:
Cacna1a:
%methylation
100
80
60
40
20
0
L I B T
(+/+) (+/-) (-/-) (-/-)
adult tissues
Dnmt3L
spermatogonia
b
S76657:
intergenic, chr14:94881121:
c
AK045080:
intergenic, chr15:85761255:
NotI
HhaI
McrBC

Table 3.1: The number and characterization of repetitive elements associated with identified RLGS loci

DNA Methylation Pattern	Total # of loci	repetitive (%)	hyper-methylated	hypo-methylated	classification of repetitive elements: non-LTR-containing: SINE	LINE	LTR-containing (solitary LTR): ERVK IAP	ETN	RLTR	MaLR MT
unchanged unmethylated	161	0 (0)	0	0	0	0	0	0	0	0
unchanged methylated	47	23 (49)	23	0	1	3	11	3(1)	4(2)	1(1)
DM testis vs. somatic	60	22 (37)	3	19	0	1	0	1	18(17)	2(2)
DM inter-somatic	26	3 (12)	0	3	0	0	0	3	0	0

Table 3.2: All identified RLGS loci

RLGS ID	NotI site position +/-500 bp (mm7)	gene	NotI site location			RLGS spot intensity (methylation)			
			position	CpG Island	Repeat	testis	liver	intestine	brain
1cx1	chr1:130281590-130282590	intergenic	\	LIKE	RLTR44C	2	0	0	0
3Gx3	chr1:135991532-135992532	LO6234	body	LIKE		4	0	0	0
3D38	chr1:156204026-156205026	intergenic	\	N		2	3	3	2
3ex3	chr1:36370867-36371867	AK137601	5'	N	MTEa, MaLR	1	0	0	0
3fx9	chr1:8058492-8059492	intergenic	\	N	RLTRETN	0	0	1	0
5dx4	chr1:89001109-89002109	Sh3bp4	body	LIKE	RLTRETN	0	0	2	0
4bx4	chr1:91312941-91313941	AK006182	5'	LIKE		1	0	0	0
3fx6	chr2:116796001-116797001	intergenic	\	LIKE		4	2	0.5	1
4E73	chr2:155482740-155483740	intergenic	\	N		2	1	1	1
4cx3	chr2:155613723-155614723	2410003P15Rik	body	N		0	0	1	1
1hx3	chr2:168061251-168062251	AK133086	5'	Y		4	0	0	0
2G75	chr2:19314558-19315558	Armc3	body	LIKE	RLTR44E	3	0	0	2
5F49	chr2:26098731-26099731	Btbd14a	5'	Y		4	0	4	4
4cx1	chr2:74130334-74131334	S76657	body	LIKE		4	0	0	0
5cx6	chr2:77255277-77256277	Zfp533	3'	LIKE		4	0	0	0
3gx1	chr2:86342178-86343178	intergenic	\	N	RLTR10D	0.5	0	0	0
4B18	chr2:92335401-92336401	AK036400	5'	LIKE		4	2	2	1
2fx2	chr2:98070124-98071124	intergenic	\	N	RLTR44D	3	0	0	0
3D31	chr3:10095927-10096927	BC060946	5'	Y		4	2	3	4
5cx4	chr3:36012850-36013850	Acad9	body	LIKE	RLTR10D	4	0	0	0
5gx1	chr3:45159893-45160893	intergenic	\	N	RLTR10D	4	1	1	1
3dx3	chr4:114024855-114025855	intergenic	\	LIKE		4	4	1	3
5cx3	chr4:123797266-123798266	intergenic	\	LIKE	RLTRETN	0	0	1	0
1D27	chr4:130166729-130167729	Matn1	body	LIKE		0	4	4	4
3D27	chr4:133538960-133539960	Ccdc21	body	LIKE	RLTR4	0	4	4	4
5gx2	chr4:136128890-136129890	Ephb2	body	N		0	0	1	1
5D36	chr4:48065246-48066246	Nr4a3	5'	Y		4	4	4	1
4bx2	chr4:52510613-52511613	intergenic	\	LIKE		4	0	0	0
6fx2	chr4:54418300-54419300	intergenic	\	LIKE	RLTR10A	3	0	0	0

RLGS ID	NotI site position +/-500 bp (mm7)	gene	NotI site location			RLGS spot intensity (methylation)			
			position	CpG Island	Repeat	testis	liver	intestine	brain
6bx1	chr4:8037009-8038009	AK035064	body	LIKE		2	0	0	0
3D22	chr4:88677758-88678758	Cdkn2a	3'	LIKE	LINE Lx3C	2	4	4	4
4dx1	chr4:9701490-9702490	AK035353	body	LIKE	RLTR44E	3	0	0	0
4E15	chr5:105630030-105631030	E130309B19Rik	3'	Y		0	3	3	4
4F43	chr5:105632198-105638870	E130309B19Rik	5'	Y		0	4	4	4
4D44	chr5:111294164-111295164	intergenic	\	N		0	1	1	1
2G63	chr5:59593352-59593751	CG869761	5'	Y		4	4	4	1
5D52	chr5:71955036-71956036	Zar1	5'	Y		3	2	3	2
6fx4	chr5:73043767-73044767	Spata18	body	Y		3	0	0	3
6G33	chr6:146248423-146249423	Itpr2	body	LIKE		3	0	0	0
6gx2	chr6:24967030-24968030	intergenic	\	LIKE	RLTR44A	1	0	0	0
3E09	chr6:58874742-58875742	Nap1l5	5'	Y		1	4	4	7
5dx1	chr6:6688292-6689292	intergenic	\	LIKE		3	0	0	0
2cx3	chr7:11557189-11558189	Zfp324	3'	Y		4	0	0	0
3G91	chr7:139888309-139888708	Cdkn1c	5'	Y		4	2	2	2
6ex1	chr7:17493013-17494013	Bcl3	body	LIKE		1	0	0	0
2dx7	chr7:39442553-39443553	intergenic	\	LIKE	RLTR44E	2	0	0	0
1gx8	chr7:75332581-75333581	Polg	5'	Y		1	0	0	0
6dx2	chr7:75332881-75333881	Polg	5'	Y		3	1	4	2
6dx4	chr7:75334206-75335206	Polg	5'	Y		2	1	4	2
4C13	chr7:75376878-75377277	AK032343	body	Y		1	2	2	0
2ex4	chr8:108647075-108648075	intergenic	\	N	RLTR44E	2	0	0	0
4gx4	chr8:118899502-118900502	CJ052809	body	LIKE		0.5	4	4	2
5cx1	chr8:68335566-68336566	Cilp2	3'	Y		4	2	2	2
2C28	chr8:68336816-68337816	Cilp2	3'	Y		4	2	2	2
3E61	chr8:76679487-76680487	intergenic	\	LIKE	RLTRETN	0	1	1	1
5F09	chr8:83198115-83199115	Cacna1a	body	N		0	2	2	2
2gx1	chr9:118558574-118559574	Itga9	body	LIKE		3	0	0	0
5bx5	chr9:15818207-15819207	CO043424	body	LIKE		4	0	0	2
3ex8	chr9:85800273-85801273	Ibtk	body	LIKE		2	0	0	0
2G35	chr10:33404605-33405605	A330019N05Rik	5'	LIKE		4	4	2	4

RLGS ID	NotI site position +/-500 bp (mm7)	gene	NotI site location			RLGS spot intensity (methylation)			
			position	CpG Island	Repeat	testis	liver	intestine	brain
2D30	chr10:39978903-39979903	intergenic	\	N		4	0.5	1	0.5
2F21	chr10:5633569-5634569	Esr1	body	Y		4	2	3	4
2dx1	chr10:62180628-62181628	AK133671	body	N	MTE2b, MaLR	4	0.5	0.5	0.5
2ex7	chr10:65580897-65581897	intergenic	\	N	RLTR44B	1	0	0	0
2dx4	chr10:71064227-71065227	intergenic	\	LIKE	RLTR44D	3	0	0	0
3cx2	chr10:7871549-7872549	intergenic	\	N	RLTR44D	4	0	0	0
5dx6	chr10:81050933-81051933	Matk	body	N		0	0	2	2
4gx5	chr10:85525697-85526697	Bpil2	body	LIKE		3	0	0	0
4cx2	chr10:97306373-97307373	AK134882	body	N	RLTR44D	3	0	0	0
2D21	chr11:102546032-102547032	Grn	3'	Y		4	1	2	1
3B29	chr11:108023826-108024826	AK034065	5'	Y		4	0.5	0.5	0.5
3F01	chr11:114789169-114790169	Ttyh2	body	N		0	0	0	4
5C43	chr11:115359508-115359907	Grin2c	3'	Y		4	4	4	2
2hx1	chr11:115577931-115578931	Slc16a5	body	LIKE		3	0	0	0
4C01	chr11:115967838-115968838	intergenic	\	Y		4	4	3	3
3E07	chr11:22923940-22924940	U2af1-rs1	5'	Y		4	2	2	2
2fx1	chr11:59211978-59212978	OBSCN	body	N	RLTR44E	0	0	1	2
3D30	chr11:64005264-64005663	Hs3st3b1	5'	Y		4	1	3	3
4fx2	chr11:67206493-67207493	intergenic	\	LIKE		4	0	0	0
4D02	chr11:84834099-84835099	intergenic	\	LIKE		4	2	2	4
5F66	chr11:88257584-88258584	Cuedc1	body	LIKE		2	0	1	1
2F08	chr12:102025532-102026532	intergenic	\	N		4	2	2	2
2G48	chr12:108961612-108962612	Tnfaip2	5'	Y		4	0	4	4
2cx2	chr12:29928868-29929868	intergenic	\	LIKE	RLTR44A	1	0	0	0.5
3dx6	chr12:6239843-6240843	intergenic	\	N	RLTR44B	3	0	0	0
5cx10	chr12:69261419-69262419	Dact1	3'	Y		3	0	0	0
3dx1	chr12:82147883-82148883	BC038335	3'	LIKE		4	0	0	0
1fx6	chr12:89165037-89166037	4933401N24Rik	body	N		0	1	2	2
4G103	chr12_random:413547-414547	Cables2	5'	Y		4	2	2	3
3bx5	chr13:109807817-109808817	Ddx4	5'	Y		4	0	0	0
3dx5	chr13:23144133-23145139	intergenic	\	LIKE	RLTR31	4	0	0	0

RLGS ID	NotI site position +/-500 bp (mm7)	gene	NotI site location			RLGS spot intensity (methylation)			
			position	CpG Island	Repeat	testis	liver	intestine	brain
2bx1	chr13:8081558-8082558	Adarb2	5'	LIKE		2	0	0	0
2F64	chr14:58104168-58105168	Gata4	body	Y		4	1	4	4
2cx1	chr14:94881121-94882121	intergenic	\	N		3	0	0	0
3fx2	chr14:98136003-98137003	BC079866	3'	LIKE		3	0	0	0
1fx5	chr15:29903539-29904539	intergenic	\	N	RLTR44B	4	0	0	0
3bx1	chr15:41100483-41101483	Zfpm2	3'	LIKE		4	0	0	0
1G05	chr15:79515671-79516670	Csnk1e	body	LIKE		0	1	0	1
4cx4	chr15:85761255-85762255	intergenic	\	N		1	2	4	4
4ax1	chr15:99084619-99085619	2810451A06Rik	5'	N		3	3	2	0
4C11	chr16:10316403-10317403	AK016529	body	LIKE		0	2	3	0
3E60	chr16:18170194-18171194	Gp1bb	5'	Y		4	4	2	4
4C31	chr16:34905665-34906665	Adcy5	body	N		0	2	4	0
3ax1	chr16:63543509-63544509	Epha3	body	N		4	0	0	0
4F84	chr17:33074114-33075114	Hspa1l	3'	Y		4	1	0	0
3G87	chr17:35098434-35099434	Rfn39	3'	Y		4	2	2	3
6dx1	chr17:81582462-81583462	intergenic	\	N		0	1	1	3
6E36	chr18:38473730-38474730	intergenic	\	N		0	1	1	0
3bx2	chr18:47870542-47871542	AK045080	3'	N		1	3	0	0
3G10	chr18:75074491-75075491	Myo5b	body	N		0	0	0	1
3C21	chr19:6323488-6324488	Nrxn2	5'	Y		4	4	4	0
UNCHANGED METHYLATED:									
4dx7	chr1:191694316-191695316	intergenic	\	N	SINE B3-B2 family	0	0	0	0
5fx3	chr1:8051385-8052385	intergenic	\	LIKE	RLTRETN	0	0	0	0
4fx4	chr1:8054188-8055188	intergenic	\	N	RLTRETN	0	0	0	0
2gx12	chr2:109031686-109032686	Mett5d1	body	N	RLTR10D	0	0	0	0
3gx7	chr2:155627541-155628541	2410003P15Rik	body	N		0	0	0	0
5fx4	chr2:25946146-25947146	Camsap1	body	N		0	0	0	0
3ex12	chr2:26469065-26470064	Notch1	body	LIKE		0	0	0	0
6cx3	chr2:26472365-26473364	Notch1	body	LIKE		0	0	0	0
1ex6	chr2:46,704,677-46,705,677	intergenic	\	N	IAPLTR2	0	0	0	0

RLGS ID	NotI site position +/-500 bp (mm7)	gene	NotI site location			RLGS spot intensity (methylation)			
			position	CpG Island	Repeat	testis	liver	intestine	brain
3bx6	chr2:5,519,825-5,520,825	Camk1d	body	N		0	0	0	0
5fx5	chr2:86,345,247-86,346,306	intergenic	\	N	RLTR10D	0	0	0	0
4hx2	chr3:103722382-103723382	Rsbn1	body	N		0	0	0	0
4fx6	chr3:43851799-43852799	BC099952	body	N	RLTR27	0	0	0	0
2ex12	chr3:87,841,451-87,842,451	Arhgef11	body	N		0	0	0	0
2dx11	chr4:124421742-124422742	intergenic	\	N		0	0	0	0
5ex8	chr4:155146333-155147333	intergenic	\	N	RLTR10D	0	0	0	0
6fx8	chr4:59072485-59073463	intergenic	\	N	IAPLTR2	0	0	0	0
1gx3	chr4:8,241,763-8,242,763	intergenic	\	LIKE	MMERVK10C	0	0	0	0
2cx6	chr5:144,319,244-144,320,244	AA137378	3'	N		0	0	0	0
3fx11	chr6:92038749-92039749	intergenic	\	LIKE	RLTRETN	0	0	0	0
2gx9	chr7:31091264-31092264	Gpi1	3'	N		0	0	0	0
6fx7	chr8:4857718-4858718	intergenic	\	N	RLTRIAP1a	0	0	0	0
2ex13	chr8:95,493,253-95,494,253	intergenic	\	N	RLTR27	0	0	0	0
1gx1	chr9:117,887,550-117,888,550	intergenic	\	N	MMERVK10C	0	0	0	0
2ex10	chr9:122,589,894-122,590,894	BC042569	5'	N	SINE-B2	0	0	0	0
4ex5	chr9:15829320-15830294	CO043424	body	N		0	0	0	0
2fx7	chr9:48672841-48673841	Plzf	body	N		0	0	0	0
4dx8	chr9:48741810-48742810	Plzf	body	N	MTE2b MaLR	0	0	0	0
1ex5	chr10:37,632,522-37,633,522	intergenic	\	N	IAPEz	0	0	0	0
2gx11	chr11:116717090-116718090	Rhbd16	body	N		0	0	0	0
7fx1	chr11:118782291-118783291	D11Bwg0517e	body	N		0	0	0	0
1fx9	chr11:119,175,699-119,176,698	intergenic	\	N		0	0	0	0
4fx5	chr11:119206448-119207447	intergenic	\	N		0	0	0	0
5dx9	chr11:90617696-90618696	Stxbp4	body	N	RLTR27	0	0	0	0
3fx12	chr12:108180863-108181863	Dnchc1	body	LIKE		0	0	0	0
4gx10	chr13:69052451-69053451	intergenic	\	LIKE		0	0	0	0
4fx3	chr13:99940745-99941745	AK090136	body	N		0	0	0	0
1gx2	chr14:108,593,648-108,594,643	intergenic	\	N	RLTR27	0	0	0	0
5ex9	chr14:58585165-58586164	AY534252	body	N		0	0	0	0
4dx6	chr14:68573674-68574674	intergenic	\	N		0	0	0	0

RLGS ID	NotI site position +/-500 bp (mm7)	gene	NotI site location			RLGS spot intensity (methylation)			
			position	CpG Island	Repeat	testis	liver	intestine	brain
2dx10	chr15:13,411,380-13,412,380	intergenic	\	N	RLTR44C	0	0	0	0
2gx10	chr16:22096541-22097541	intergenic	\	N		0	0	0	0
2fx6	chr16:49829798-49830798	B130050K08	body	N	RLTR10D	0	0	0	0
1gx1	chr16:52,497,054-52,498,054	intergenic	\	N	IAPLTR2	0	0	0	0
2ex11	chrX:96,244,173-96,245,173	Tex11	body	N	EtnERV2	0	0	0	0
UNCHANGED UNMETHYLATED:									
3C06	chr1:15469944-15470945	AW045422	5'	Y		4	4	4	4
4G106	chr1:16807772-16808772	6130401J04Rik	5'	Y		4	4	4	4
3A10	chr1:191,690,179-191,691,178	Rcor3	5'	Y		4	4	4	4
1E05	chr1:191690703-191691702	Rcor3	5'	Y		4	4	4	4
3F42	chr1:4487556-4488556	Sox17	3'	Y		4	4	4	4
5E65	chr1:58220251-58221251	6530404F10Rik	5'	Y		4	4	4	4
4F09	chr1:89,845,970-89,846,970	intergenic	\	Y		4	4	4	4
3G72	chr2:102251170-102252170	Trimm44	5'	Y		4	4	4	4
3C13	chr2:105853654-105854654	Zcs13	5'	Y		4	4	4	4
4G96	chr2:109768396-109769396	intergenic	\	Y		4	4	4	4
6D12	chr2:135399882-135400882	AK020944	5'	Y		4	4	4	4
3F14	chr2:140136204-140137204	AK013491	5'	Y		4	4	4	4
2B19	chr2:163222202-163223202	intergenic	\	LIKE		4	4	4	4
3G90	chr2:173854485-173855485	Npepl1	5'	Y		4	4	4	4
2C15	chr2:19722709-19723709	AK076525	5'	Y		4	4	4	4
2G81	chr2:30611927-30612927	Cstad	body	LIKE		4	4	4	4
3D03	chr2:64028398-64029398	Fign	5'	Y		4	4	4	4
3C27	chr2:74499025-74500025	Evx2	3'	Y		4	4	4	4
5G39	chr2:74502334-74503333	Evx2	5'	Y		4	4	4	4
1F02	chr2:74511886-74512886	Hoxd13	5'	Y		4	4	4	4
3D07	chr3:138550339-138551339	Tspan5	5'	Y		4	4	4	4
4E40	chr3:32516993-32517993	Actl6a	5'	Y		4	4	4	4
1D14	chr3:34457260-34458260	Sox2	5'	Y		4	4	4	4
3F09	chr3:45324156-45325156	Pcadh10	5'	Y		4	4	4	4

RLGS ID	NotI site position +/-500 bp (mm7)	gene	NotI site location			RLGS spot intensity (methylation)			
			position	CpG Island	Repeat	testis	liver	intestine	brain
2C31	chr3:57976407-57977407	Pfn2	5'	Y		4	4	4	4
3G60	chr3:62482883-62483883	4631416L12Rik	5'	Y		4	4	4	4
2D06	chr3:68594659-68595659	Schip1	5'	Y		4	4	4	4
5G16	chr3:84372431-84373431	AK079415	5'	Y		4	4	4	4
2C37	chr3_random:-163-837	X59474	5'	Y		4	4	4	4
5D50	chr4:11248741-11249741	AK144417\AK133116	5'\5'	Y		4	4	4	4
5H33	chr4:120089377-120090377	intergenic	\	Y		4	4	4	4
3D67	chr4:131327548-131328548	Oprd1	3'	Y		4	4	4	4
6D20	chr4:141617289-141618289	9030409G11Rik	5'	Y		4	4	4	4
4D09	chr4:146091219-146092219	AK089842	body	Y		4	4	4	4
2F89	chr4:152471366-152472366	BE985330	5'	Y		4	4	4	4
2F39	chr4:54958381-54959381	CJ147550	5'	Y		4	4	4	4
1G34	chr4:59044079-59045079	AK136359	5'	Y		4	4	4	4
5D51	chr4:59667448-59668448	E130308A19Rik	5'	Y		4	4	4	4
3G119	chr4:77615294-77616294	Ptprd	5'	Y		4	4	4	4
5B13	chr4:99043424-99044424	intergenic	\	LIKE		4	4	4	4
6C17	chr4:99050103-99051103	Foxd3	5'	Y		4	4	4	4
2C25	chr5:116586572-116587572	Wsb2	5'	Y		4	4	4	4
1F08	chr5:128257505-128258505	Ran	5'	Y		4	4	4	4
3D19	chr5:134364721-134365721	Pom121	5'	Y		4	4	4	4
2G05	chr5:148344808-148345808	6330406I15Rik	5'	Y		4	4	4	4
4D58	chr5:27103323-27104323	Insig1	5'	Y		4	4	4	4
3B24	chr5:32883436-32884436	AK038286	5'	Y		4	4	4	4
5C23	chr5:47308516-47309516	Slit2	5'	Y		4	4	4	4
3G35	chr5:73613244-73614244	BC083068	5'	Y		4	4	4	4
4F84	chr6:122,855,579-122,856,579	Foxj2	5'	Y		4	4	4	4
2E10	chr6:24898170-24899170	6332401O19Rik	5'	Y		4	4	4	4
3D49	chr6:38229607-38230607	B130055L09Rik	5'	Y		4	4	4	4
4D27	chr6:52127185-52128185	Hoxa2	5'	Y		4	4	4	4
4E29	chr6:52138663-52139663	5730596B20Rik	5'	Y		4	4	4	4
2G95	chr6:55166342-55167342	C330043M08Rik	5'	Y		4	4	4	4

RLGS ID	NotI site position +/-500 bp (mm7)	gene	NotI site location			RLGS spot intensity (methylation)			
			position	CpG Island	Repeat	testis	liver	intestine	brain
2C19	chr6:85541127-85542127	Egr4	5'	Y		4	4	4	4
4E47	chr6:91198541-91199541	Nup210	5'	Y		4	4	4	4
1E27	chr6:92172168-92173168	AK157305	5'	Y		4	4	4	4
1G14	chr7:105618761-105619761	Bca3	5'	Y		4	4	4	4
4B21	chr7:32104130-32105130	AK149337	5'	Y		4	4	4	4
6G67	chr7:60359386-60360386	Apba2	5'	Y		4	4	4	4
5G21	chr7:61228288-61229288	Tjp1	5'	Y		4	4	4	4
1G50	chr7:76559422-76560422	2610312F20Rik	5'	Y		4	4	4	4
5B30	chr7:96976931-96977931	CF723412	5'	Y		4	4	4	4
3D04	chr8:105656289-105657289	Cog8	5'	Y		4	4	4	4
1F14	chr8:12370578-12371577	Sox1	5'	Y		4	4	4	4
1D26	chr8:35013720-35014720	BF456945	5'	Y		4	4	4	4
4G73	chr8:90391925-90392925	Irx3	5'	Y		4	4	4	4
4E58	chr8:91448223-91449223	A330042H22	5'	Y		4	4	4	4
5G61	chr9:45387989-45388989	Dscaml1	5'	Y		4	4	4	4
5F59	chr9:47491453-47492453	Igsf4	5'	Y		4	4	4	4
1G51	chr9:49773317-49774317	Ncam1	5'	Y		4	4	4	4
3E11	chr9:52146492-52147492	BC042784	3'	Y		4	4	4	4
3D29	chr10:117355938-117356938	Rap1b	5'	Y		4	4	4	4
4D26	chr10:18041072-18042070	intergenic	\	Y		4	4	4	4
6C25	chr10:27855268-27856268	Ptprk	5'	Y		4	4	4	4
5B27	chr10:28926887-28927887	6330407J23Rik	5'	Y		4	4	4	4
5E24	chr10:28927649-28928649	6330407J23Rik	5'	Y		4	4	4	4
2C51	chr10:53568628-53569626	AK015334	5'	Y		4	4	4	4
3E44	chr10:60042860-60043860	Cdh23	body	N		4	4	4	4
5C16	chr10:62244082-62245082	AK028920	5'	Y		4	4	4	4
3E54	chr10:67138965-67139965	Egr2	5'	Y		4	4	4	4
6G41	chr10:80544550-80545550	Dot1l	5'	Y		4	4	4	4
3F27	chr10:81243984-81244984	intergenic	\	Y		4	4	4	4
2B31	chr10:81276209-81277209	AK132321	5'	Y		4	4	4	4
4F21	chr10:81353485-81354485	Aes	5'	Y		4	4	4	4

RLGS ID	NotI site position +/-500 bp (mm7)	gene	NotI site location			RLGS spot intensity (methylation)			
			position	CpG Island	Repeat	testis	liver	intestine	brain
2G09	chr10:9373956-9374956	E130306M17Rik	5'	Y		4	4	4	4
1F40	chr11:102425843-102426843	Ubtf	5'	Y		4	4	4	4
5G51	chr11:106721451-106722451	Tex2	5'	Y		4	4	4	4
4E72	chr11:11637804-11638804	AK077815	5'	Y		4	4	4	4
5F63	chr11:119146234-119147097	Cbx8	3'	Y		4	4	4	4
6B16	chr11:119149258-119150257	Cbx8	5'	Y		4	4	4	4
2B40	chr11:119150132-119151131	Cbx8	5'	Y		4	4	4	4
4F28	chr11:119161722-119162722	intergenic	\	Y		4	4	4	4
3F03	chr11:119164598-119165597	intergenic	\	Y		4	4	4	4
5E61	chr11:120051613-120052613	Baiap2	5'	Y		4	4	4	4
5B12	chr11:23207637-23208637	Xpo1	5'	Y		4	4	4	4
3B27	chr11:23208412-23209412	Xpo1	5'	Y		4	4	4	4
2B23	chr11:43393725-43394725	d11ertd730e	body	LIKE		4	4	4	4
2F73	chr11:59257728-59258728	Gja12	3'	Y		4	4	4	4
4B19	chr11:61489409-61490409	intergenic	\	Y		4	4	4	4
2H54	chr11:66606861-66607861	AK147504	5'	Y		4	4	4	4
4E24	chr11:70295636-70296636	Slc16a11	5'	Y		4	4	4	4
1F30	chr11:8576172-8577172	AK089717	body	Y		4	4	4	4
4F52	chr11:90747059-90748058	Cox11	5'	Y		4	4	4	4
5G22	chr11:96573227-96574227	AK037877	5'	Y		4	4	4	4
5H105	chr11:99046635-99047635	Rara	5'	Y		4	4	4	4
3F61	chr12:20917137-20918137	LOC245297	5'	Y		4	4	4	4
2F61	chr12:33881603-33882603	Tspan13	5'	Y		4	4	4	4
3F56	chr12:47396880-47397880	Foxg1	5'	Y		4	4	4	4
4G54	chr12:5317394-5318394	BC065168	5'	Y		4	4	4	4
2G50	chr12:65162828-65163828	Mamdc1	5'	Y		4	4	4	4
5G90	chr12:78095338-78096338	AK053779	5'	Y		4	4	4	4
3D63	chr12:99963951-99964951	Golga5	5'	Y		4	4	4	4
1E25	chr13:110143397-110144508	AK053890	5'	Y		4	4	4	4
5G11	chr13:23076015-23077015	Hfe	5'	LIKE		4	4	4	4
3H19	chr13:23104309-23105315	Hist1h1c	5'	Y		4	4	4	4

RLGS ID	NotI site position +/-500 bp (mm7)	gene	NotI site location			RLGS spot intensity (methylation)			
			position	CpG Island	Repeat	testis	liver	intestine	brain
1F15	chr13:24635336-24636336	Vmp	5'	Y		4	4	4	4
1D01	chr13:33353087-33354087	Serpinb6a	5'	Y		4	4	4	4
2G41	chr13:47743959-47744959	Id4	5'	Y		4	4	4	4
2C22	chr13:53822640-53823640	Gprin1	?	GAP		4	4	4	4
5C30	chr13:59283846-59284846	Gas1	5'	Y		4	4	4	4
2C33	chr13:59284746-59285746	Gas1	5'	Y		4	4	4	4
3F70	chr14:111327498-111328498	Gpc6	5'	Y		4	4	4	4
2G96	chr14:12515714-12516714	Slc4a7	5'	Y		4	4	4	4
3E75	chr14:27151399-27152399	Cacna2d3	5'	Y		4	4	4	4
4F06	chr14:27782976-27783976	AK147415	5'	Y		4	4	4	4
1H21	chr14:58470731-58471735	AY534252	5'	Y		4	4	4	4
5F75	chr14:65536009-65537009	Gfra2	5'	Y		4	4	4	4
2F57	chr14:71095432-71096432	Tgfb1i4	5'	Y		4	4	4	4
2F72	chr15:64073316-64074316	BC011343	5'	Y		4	4	4	4
2E14	chr15:68445727-68446727	BB645359	5'	Y		4	4	4	4
3B38	chr15:85753215-85754215	intergenic	\	Y		4	4	4	4
5E02	chr15:91654853-91655853	A630029G22Rik	5'	Y		4	4	4	4
6G69	chr15:99073349-99074348	C1q14	5'	Y		4	4	4	4
5C29	chr15:99078538-99079538	intergenic	\	Y		4	4	4	4
2F31	chr15:99081949-99082948	2810451A06Rik	5'	Y		4	4	4	4
1C13	chr16:10066697-10067697	AK014352	5'	Y		4	4	4	4
4F59	chr16:20853694-20854694	EphB3	5'	Y		4	4	4	4
1E24	chr16:21916518-21917518	Sfrs10	5'	Y		4	4	4	4
3D36	chr16:32736583-32737583	Lmln	5'	Y		4	4	4	4
2F53	chr16:34830959-34831959	Adcy5	5'	Y		4	4	4	4
6D06	chr16:8411311-8412311	AK135814	body	Y		4	4	4	4
5C25	chr16:84833713-84834713	Gabpa/Atp5j	5'/5'	Y		4	4	4	4
2D41	chr17:24321508-24322508	Axin1	5'	Y		4	4	4	4
5D11	chr17:30380436-30381436	AK010021	5'	Y		4	4	4	4
5E57	chr17:81508970-81509969	AW548124	5'	Y		4	4	4	4
2D47	chr18:23694400-23695400	X95227	5'	Y		4	4	4	4

RLGS ID	NotI site position +/-500 bp (mm7)	gene	NotI site location			RLGS spot intensity (methylation)			
			position	CpG Island	Repeat	testis	liver	intestine	brain
4D36	chr18:36663249-36664249	Pura	3'	Y		4	4	4	4
5F70	chr18:47745768-47746768	Sema6a	5'	Y		4	4	4	4
2F12	chr18:51494854-51495854	AK046456	5'	Y		4	4	4	4
1E23	chr18:53854762-53855762	AK133910	body	Y		4	4	4	4
2E30	chr19:45556841-45557841	Fbxw4	5'	Y		4	4	4	4
6D10	chr19:45572506-45573506	intergenic	\	LIKE		4	4	4	4
6D08	chr19:4650792-4651792	4921506I22Rik	5'	Y		4	4	4	4
1G31	chrX:103142779-103143779	2610002M06Rik	5'	Y		4	4	4	4
3G99	chrX:104161360-104162360	AK050539/AK086455	5'/5'	Y		4	4	4	4
2G34	chrX:21797681-21798681	Klhl13	5'	Y		4	4	4	4
6D16	chrX:51855137-51856137	6330419J24Rik	5'	Y		4	4	4	4
1F21	chrX:53511527-53512527	Zic3	5'	Y		4	4	4	4
2F37	chrX:56286481-56287481	Sox3	5'	Y		4	4	4	4
4C18	chrX:69060075-69061075	Slc6a8	5'	Y		4	4	4	4

NOTES: CpG Islands: Y = meets criteria of Takai & Jones
LIKE = meets criteria of Gardiner-Garden & Frommer, not Takai & Jones
N = meets neither

Site Position: 5' = 3kb upstream - 1kb downstream of transcriptional start site
3' = 3kb downstream - 1kb upstream of transcriptional termination site
body = anything inside
Ig = anything outside

Methylation:	RLGS	approximate %methylation
	4 =	0%
	3 =	25%
	2 =	50%
	1 =	75%
	0.5 =	90%
	0 =	100%

Table 3.3: qAMP primer sequences used to confirm the DNA methylation status of identified RLGS loci and to analyze *Dnmt3L* spermatogonia.

RLGS spot	Forward Primer (5'-3')	Reverse Primer (5'-3')
1D27	CTTTCGCAGGTTGTCCATCT	CCTCAGCTCTTCCTCCACAG
1F55	TGCCAACTGCTTTTCCAGAT	TAAAGCATCCCAGTGGCAGT
1G34	TAGTCCTGCGTGTGGACTGT	GCGTTGCTAGGAGAGAAGGA
1H64	CTAAGCTGCGCCTCAGACA	TACTTGATGCCACCCACGTT
2B19	CACACCTGGCCTGATTGAT	CCCAGGGATGCTCCATAGT
2C56	CATCAACCCAGGTACGAAGA	GGAAGAGGAGGAGCCACAGT
2C58	GCTCCAACCTCAGTCAGCA	CACAGAGGGAGCAGGCATA
2F21	CTGCCAAAGAATCCTTTCTGAC	GCCCTACTACCTGGAGAACGA
2F64	GGCTCAAGGCCCTAAGGTAA	AGCCCCTACCCAGCCTACAT
3B50	TGATTAGCACTCGTTGGACA	CTTTGGAAACGTCTGCGATTA
3D31	GGTAGCAGAGTTGAGCCATT	CTCGGACTTGTCGGCATC
3D38	CAAGCTGCATGCTAAATTCG	TCAAGCAAGTGGAGCCTTTT
3D70	GCGCTCCCAAATGAAGGTTA	AGCGACAAGCTCCATTCTCT
3E07	GGCCGTACCACAGATAACCA	GCGCAGTTATCCGTCTATCCA
3E09	GTCCAGAGTGCGTCTGAAGAT	TCGGAGATAGCAGCATAGCA
3E80	CTCCCCCATTCTCCCTCAC	CCTAACTTCTTGCCGTGCTC
3E85	ACTCTCCTGCCTTGCACCT	GCTGTCCACTCGGTGTCAT
3F27	GTGGCTCTCTTTTTCCAAAGC	AAACCAGGCAGGAGGGATT
3F91	CCTTTTCCCTTTCTGCCACT	AAGTTGATCCTGTGGTGATGC
3G122	GAAGAGCATCACGATCAGCA	AATGCCACTAGCCCAAGGT
3G91	TACCTGGCTGATTGGTGATG	ACTGAGAGCAAGCGAACAGG
3G91	TACCTGGCTGATTGGTGATG	ACTGAGAGCAAGCGAACAGG
4B46	GAGAGAGAGACTGTCACCCTGTTA	TTCAGGGTTGCAGCTAAAGA
4C11	CTGGCAAGCTGTCCGATAGT	CCGAGATCTTACGCAGGAGATA
4C13	ACAGTCTTGCCCGGATTGAT	GTGTGAGAAGCCGAGGAGAA
4C31	AAGCAGGAAGTTGGCGTTTA	CTTCCTCCTCCTGCTTCCA
4C40	TCTGCAATAAGGGGCTGAAC	AAAGTGGCTTGTGAGGCATT
4C43	TTAGGGCTGCCAGGAGTAAG	ACCCACGAGCATCTTTCG
4D72	GTGCGATTTGTGAACTTCGT	GGGCAAGGCAGTAAACATCT
4E15	CAGTGGCTCTCCTCACCTTC	CGTGAGAGCCAAGAAACCTC
4F97	ATCTGTGTCGGTGGTGAGGT	GAAGGGATGGTAGCAAGTGG
4G106	CCCTTTCTTCCCCCTAAGC	GAGACGAGGTGAGCGTCTG
4G110	TAAACAAACCCCACCACTGC	CATGTTGCCAGGAGCTTTG
5C54	GGCATCTGGTGAAAGCTCA	CTGGGAAACGGCACTTCTT
5D58	ATTCTCCTGGGCATGACATC	CAGCTGCTGCCTTCTAACTTG
5F09	AAGGCCACTGCCTAGAACCT	CGAGAGGGGGCGATATTACT

Table 3.4: qAMP Primer sequences used to determine methylation levels of non-CpG island sequences on chromosomes 4, 10, 17 and X.

Position (mm7)	Forward Primer (5'-3')	Reverse Primer (5'-3')
chr4:4872334-4872813	ACCCTTCAAAACCCGTGAAT	GCCTGAATCTTGCTCTTTGCAT
chr4:9927753-9928179	GAAAGGGGAACAGGGGAGTA	GGCACCTAGCATCTTGGAGA
chr4:15477713-15478091	GTGTTGGCTAATGAGGAGGA	GAAGGAGAAAGGATGCTGGA
chr4:19850460-19850826	CTATGACTCCCCACGTCACA	GTCATGCGGAAGACATCTGA
chr4:24806042-24806420	GGAAACAGAGCTCTCTGGAA	ACAGCTAACCCAATGGCTCT
chr4:30339739-30340307	TTGAACAGCATGCCTCTCT	TTTAACTGCGCTGTGGAGAA
chr4:35145276-35146275	TCATCAAGGGCAGAGGAAAT	TTTCGAGAAGGACGGAGGT
chr4:40155437-40155738	ACCACACAGACCTCCTCTCA	CTCAAAGCAGCCACGACTGT
chr4:45282104-45282572	TCATCAGTGACCCCTCTTCC	CTGGACCAGCTCTTCCTCAT
chr4:50216637-50217105	TGGCTAGGGAAGAGGTGAGA	CTTCTTCCCTTGTGGCTTGA
chr4:55202844-55203145	TTTGAGAGAAGGCAGCATGA	AAGGCCTTCGTCGTTAGACA
chr4:59969750-59970164	GTGCCACATGGTGTGGTAAA	ATATGCCGTATTGCACAACC
chr4:65403458-65403793	CCCAGGGTAAAAAGGATCA	AATCGTCTCGAACTCGCTCA
chr4:70217560-70217910	GGGGCTTTAAATGGGAAACA	TCAAGCAGGAAGAGCTGGATA
chr4:75238764-75239513	CCCAGATACCAAGGTGTGTCT	GGCTGACAGGTGAACTGAGA
chr4:80266828-80267395	CATGTGTCCCCGTTTCTTGT	CAGCTTGGTCACAACCATCA
chr4:85800348-85800847	CACCCCATCTCCCATTTCTA	AGGATCACCACGAAACAGGT
chr4:90473664-90474084	GGACAAGGGGGCTTTCTTTCT	GGGAATGGAGCTGTATGGT
chr4:96111617-96112110	AGCTTCCCACTTTCCAACAA	GCCTTTCAGCTACAGTTCCAA
chr4:100277854-100278187	AAAACCAACAGGCCTGAGAA	TCGTCGTCAAAAAGGTCAGA
chr4:105596524-105596856	AGCAACAGCAGCAACTGAAC	TCCCTGGTTGATCCTGTGTA
chr4:110703297-110703796	GAGAGGAACCTGAGGCTTGA	CAGCAGAGACGGGAGACAAT
chr4:115472605-115473015	TTCATGGCATCCCTACCAGT	TCTTGCTGTGACTGCATCCT
chr4:120466404-120466860	CTGGAAGAACATGGCAAGTGA	TCCTCCCTCTGTTCTCTGGT
chr4:125407419-125407757	ACCATGGAAAAGGAGCAACA	GCCAGGTCTGGATACAAGGA
chr4:129975322-129975744	TCCACTGTGTTCAGAAGCAA	GGAATAACCGGTCATCCAAA
chr4:135498472-135498890	CCCGAGGTCATGAGAAAGAA	CTGTTTCCTGGGGTTGTGAT
chr4:140499164-140499582	TGCACTGGAACAGGACTGAG	GAGGGGATGTAAACGGGAGT
chr4:144580754-144581253	TCTGCTAGCTCCTCCTGCTT	GGAGGTGTTGTGGCTAGCTT
chr4:150486904-150487236	CGAACGGAAGAAAAGGACTG	TGTAGTGTCTGGGAGGGACA
chr4:155169477-155169776	AATGTGAGGGTCGTTGCACT	TGTAGGCCAGCTTCCTTCAT
chr10:3022525-3023024	GGGTAGGTGGAAGATGGAAA	TCTGTGTTCCCCTCCTGTTT
chr10:10171776-10172275	AGGCACTGCGTGATAAAGGT	ACAGCAGACGCTGGAAGTCT
chr10:16148635-16150301	AAACCGGTTCTGGTGGCTAT	CAGCCAGTCCAAATTGTGTG
chr10:21168346-21168845	AGTGGCCGAGTCCACTCATA	GCCTCTCTAGGAAAGCCACA
chr10:31696268-31696767	CCAGCCTTCATGCTTTATCC	GGTAGCAGCCATCAGGTGAG
chr10:36795939-36797049	CAAAACCAAGCTGGGTGAGA	GTGGCTTCCTCTCTGGAAGA
chr10:41919793-41920292	GGGCCCTTTCCACTGCTAT	ACATCGATGGTGAGGTGGA
chr10:47643544-47644043	GAGAAGGGAAAGAGCTTCACTC	GGCTTTTACAGGGGAGAAATG
chr10:52120732-52121231	CCTTCTGTGCAGAACTGCTAAT	TGGCTGGGAATGATGGTAAT

chr10:57109907-57110406	AGGAGGCACTTCACGAGCA	ACCTGGTAGCTGGGAAGGAG
chr10:61051807-61053306	ATCCTGGGTAATGCAGCAGA	GGCTAGAGAGCCAAAGAGCA
chr10:65305221-65305970	TCAGAGAGCACAGGGAAGGT	CCATTTCAGGCTAGATGTCCT
chr10:70667883-70668382	GTCTCCCCTGAATGTTGTCT	CATGAAGACAAAGAGGCACA
chr10:75781133-75781632	GCTATGGAATGGGATGTGGT	GCCCTGCCCTATTGGAGTT
chr10:81242447-81242946	AACATCCCCGCCTCTGACT	TGCTGTCTGTGTGCTCAGGA
chr10:86594635-86595134	CCAACATACACCACTGCCTCT	GGAAGAGCCTCTGCATCAAG
chr10:91796647-91797146	TCAAACCCAGACTCAAGCAA	GATGGGCAGCACTTATTTCA
chr10:96942710-96943209	TGGCTAAGGCTTGAATACACA	GAAGCTCTGCCCATGTCAAT
chr10:101731736-101732235	CATTTTCAAGGAAGTGTTTTACAGT	TGCTATCAGCTAAAGCAGCATAAT
chr10:106985289-106986955	TGCTGTGTTAAGTGGTCCAGA	AGGATGTGTCTGTATCAGACGTG
chr10:112041194-112041693	AAAAGGACATGCTAACCTGTGT	GCTCTTTAATACAAGGGCAGTT
chr10:116708169-116708668	CTCACAGTGGCAAGCACCT	GACCCTGAGATGGACTTTGC
chr10:121511044-121511543	TCTGCCACCCTACCTGATCT	GGAAGTGGTGCATCTTGTGA
chr10:126312457-126312956	TTCCAGACGATGCTGAAAGA	AGTTAATACCCACGGCCACA
chr17:4233200-4233699	TCTACATGCAACAGTCCAGAGA	AGTTTCTGCACAGCACAGTCTT
chr17:10131737-10132236	GCATGTGGCCTGAATTATCA	GGACCTGACATCTCCTTCCA
chr17:15662405-15663529	GGAAGGGAAATGGCTTCCTT	TCTCCAAGTCCTCTGAGTCCA
chr17:21803071-21803570	GGTCAGAAAGCTGCACCATA	TTGTAGGCATGTGCACTGGT
chr17:25013796-25014295	CAGACCATCACACACCCAGA	CCTCAGCCCACAGGATAAAG
chr17:30062796-30063545	CGGCAGCCTCATTACACAAT	CCTTCATCAGGGGCAGCTAT
chr17:36167472-36167971	TTGCCCTGCTATCCTGTGAT	GGACCACAAGGGACCTTACA
chr17:41678272-41678771	GGCTGACCTCAGCTCTCAGT	TGCTGGTAGGAGTTGGCTTT
chr17:46937248-46937747	CCACACTTGGGTCCTCTGTT	CCTGCACTTCCTGACACAGT
chr17:50849373-50849872	GAGCCATGGATCCTTCAAGT	TGTCAGGCCTGGACTCCTT
chr17:54763523-54764022	TCAGGCAGAGGCCTAGAGTT	GAGAGTCTCAGGCTGGCAAT
chr17:60097848-60098347	CTGAGGAGTGAGCAAAACAAT	CACAGATAATTTCCTCTGTTGG
chr17:65306561-65307060	TGGAGACTGGTGAGCTAGAGA	AAGACATGATGCTTGCCATCT
chr17:70393712-70394211	TTAAGCAAAATCCCCCATCC	AAGGCTGTGGAGGGTTTTAAG
chr17:74730937-74731436	CCCACCGTGATTTTTACAGAA	TGTGAGGAATGTTGTCAGCAA
chr17:80208923-80209422	TTGACATTTCTTTCAGGGGACT	GGCTTGTAAACTTAAACTGGAGGA
chr17:85207048-85207547	CTCACAGCGGCTCCTGATA	TTGAGGGAAGGGTGAGAGAA
chr17:91821561-91822060	ATAACCCAAGGGGTCTGTGA	AAAGACAATGAGAGGGCAAA
chrX:4890124-4890623	GGCAGTAAGGGGACAGGTTA	CCTTGACAGACCAGCTCTGAA
chrX:9827238-9827737	AGCGCTGACTGGAAGTCATT	CTCCTCCCTCACCTTTTGAA
chrX:15319447-15319946	ACCGGTAAGTGGCAGTAAGT	TTCTCCAGATGGTGCTCCT
chrX:20292767-20293099	AGTAATAGACGGGGCCAGGA	CTGCTTTGGGAGAAGTGGTCT
chrX:35736464-35736818	TTTCCAGCGCCAATAGTTGT	TGTGGTGATCACATCACTTCAG
chrX:40656582-40656936	CTGGCATCACTGCATCTATGA	CAGAGTGGTTGTGCTCCTGAT
chrX:45749431-45750930	AAGAATCTAGGGCCCTCTTTG	GTGGACCGAGACAAAGGAAA
chrX:48789755-48790109	TCCCCAGGAACCTTCTACCT	GTGTAGGAGGATGGCAGGAA
chrX:56195598-56196098	CAGACTCCAGCAACAAAGGA	CCCTTGAGAGCAAATCGAAG
chrX:61330784-61331547	AGCACGGGCTGCTACAAA	CACTGCCTCTTCTGTGTCCA
chrX:66502472-66502918	AATTGCCCAACTGCATACAA	AATGCTGGGCTTCACCAATA

chrX:71408206-71408702	TAGCTAGCGGGGTATGACTT	CCTTGAAGGAGAACAGCATA
chrX:76236203-76236748	TGCTCACTGTGCTACAACCA	GTCCATGAGCCTCTCCCTTA
chrX:80709428-80709859	AGTCCAGGAGCAACTGGCTA	CGAGGCAGCTCAAAGGTAGA
chrX:85404289-85404720	ATTGGGTTCATCTGGCTACA	CCTGATTATCACAAGCATTCATC
chrX:90325997-90326375	CTTTTGCCAACGGTCCTCTA	CCCGAGGAAAATGAACCTGT
chrX:95723738-95724169	TGTGTCCCGAGCATTGATTA	GGGCTGAGCAAACCTTTCTT
chrX:100788840-100789271	GACCCACATGATCAAAGTGCT	AAGAAAATCCTGGCTGCTGA
chrX:105645103-105645481	TCAGAGCCATGGAATGTGTT	TTTGGAGCTCATTCAGACCT
chrX:110765173-110765604	GACAAATGCTATCTTTAATGGCTGT	TTGAGAGAAACGCACCAAAA
chrX:115633810-115634309	AGGATTTCGTCCAATCTCA	TGTATGAAATGCACTGTCTTGT
chrX:121249040-121249469	TGAAGGTGGAAAAATCTGCTGT	TGCATGTGATACTTACTCCCTATGA
chrX:126210018-126210447	CTTCCTTCTCTTTTGCCTTTGA	GCAGAATTACTGGGAGAGAAAT
chrX:130693473-130695722	GGCTCCTTTGCTCATCTTCT	GGGTGCCTGTGTCTGATTCT
chrX:136254525-136254954	GCCTTAGAAGCCATTATAGTTTTGTC	TCTCTGAACCCTTTGAAAGCA
chrX:141149803-141150371	CCAGCTGCTTGATTGCTTCT	GCTTTTAAAGGAGGGGCTTC
chrX:146231833-146232219	TCCTAAATACTCTTGGGGTTTGG	TCAGCAATAAAAGGGGAATAG
chrX:151368505-151369254	CCTTCAGAGACTTTCTCCCTTT	GGCAGTCTGCTAGAATGTGCT
chrX:156408103-156408651	CAGCCTTGCTGATCTGACAT	TTGTGTCACTGCCTCCATAA
chrX:161246279-161246665	AGCATGAGCCCACTCTTGA	TGGACAGGTGTCACTTGCAT

APPENDIX TO CHAPTER III

Appendix 3.1: The relationship between testis-specific DNA methylation and the proximity scaffold/matrix attachment regions (S/MARs).

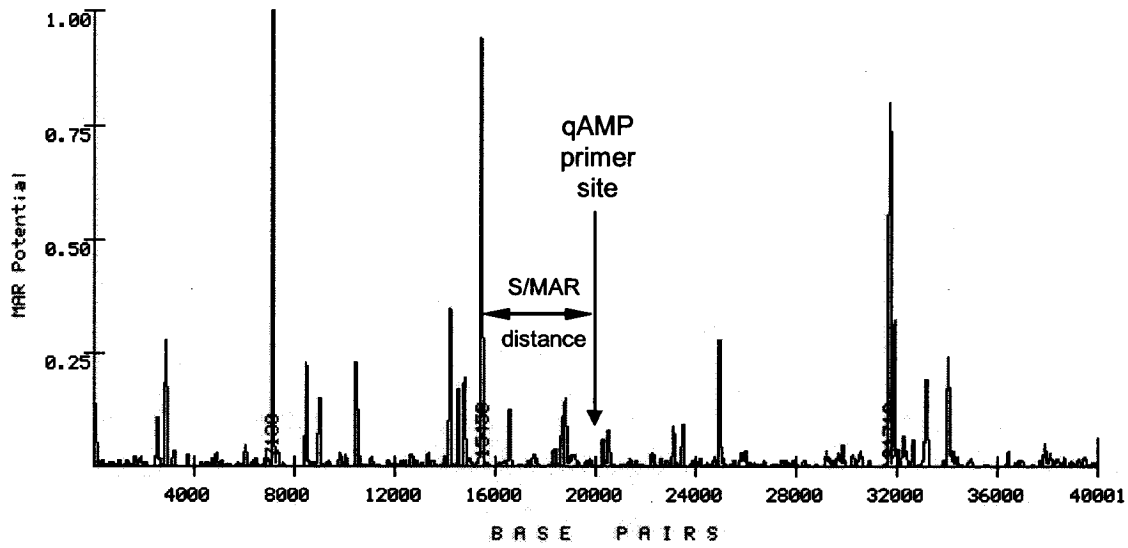
The organization of chromosomes within interphase nuclei is dependent on interactions between chromatin and a proteinatious component of the nucleus called the nuclear matrix. The nuclear matrix is made up of structural filamentous proteins that make up a nuclear scaffold and proteins that link the scaffold to chromatin (Gruenbaum et al., 2005). In this manner, the matrix is associated with the organization of chromatin into functional and non-functional compartments within the nucleus. Interactions between chromatin and nuclear matrix proteins occur in a DNA sequence-specific manner, in regions called S/MARs (Boulikas, 1993). S/MARs are distributed throughout the genome and are usually found in evolutionarily conserved non-coding DNA sequences (Glazko et al., 2003).

Although the exact locations of interactions between the nuclear matrix and chromatin are not clearly defined, regions of high S/MAR-forming potential can be identified using a bioinformatic approach (Figure 3.11a). Using this approach, regions of high S/MAR potential occur on average approximately every 20 kb. Comparing the difference between methylation of liver and testis DNA at non-CGI, non-repetitive sites on several chromosomes to the S/MAR proximity of each site reveals that there is an increased likelihood of observing a difference in methylation between the tissues if the sites are closer to a region of high S/MAR potential (Figure 3.11b).

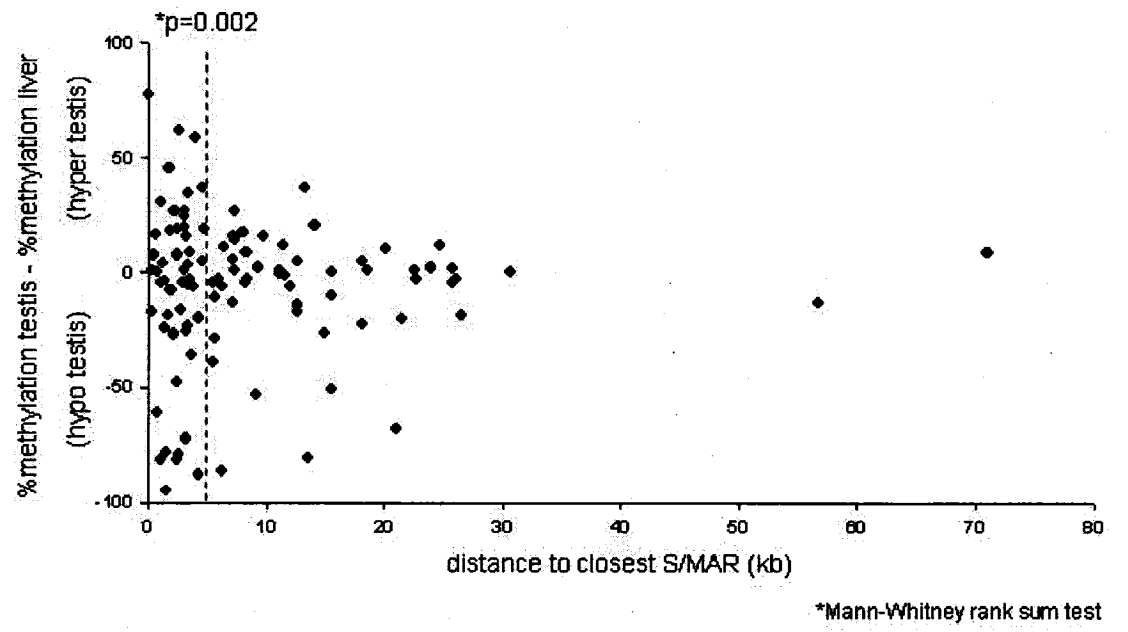
This may indicate that methylation of non-5' sites may play a role in the re-organization of chromatin architecture. Germ cells may use DNA methylation to modulate which S/MARs are bound to the matrix to produce an architecture that is different than somatic cells and is permissive for meiotic processes.

Figure 3.11: Testis-specific DNA methylation and S/MAR proximity. (a) The MAR-Wiz v1.5 output file of a 40 kb region of DNA showing the SMAR-forming potential of the sequence in 100 bp windows (http://futuresoft.org/mar-wiz). Any region showing a potential over 0.75 was considered a SMAR. The distance from the analyzed region to the closest SMAR was recorded. **(b)** The comparison of the difference in DNA methylation between liver and testis compared to the proximity of the analyzed site to the closest S/MAR. All data generated using the qAMP assay on chromosomes 4, 10, 17 and X was included in the comparison. The dashed vertical line represents the median S/MAR distance of all analyzed sites. The average absolute difference in DNA methylation between sites closer or farther away than the median distance was compared. There is a statistically greater likelihood of an analyzed site showing a difference in methylation if it is closer to S/MAR (P=0.002). Statistical analysis performed by Mann-Whitney rank sum test (SigmaStat v2.03).

a


1.00
0.75
0.50
0.25
MAR Potential
qAMP
primer
site
S/MAR
distance
4000
8000
12000
16000
20000
24000
28000
32000
36000
40001
BASE PAIRS

b

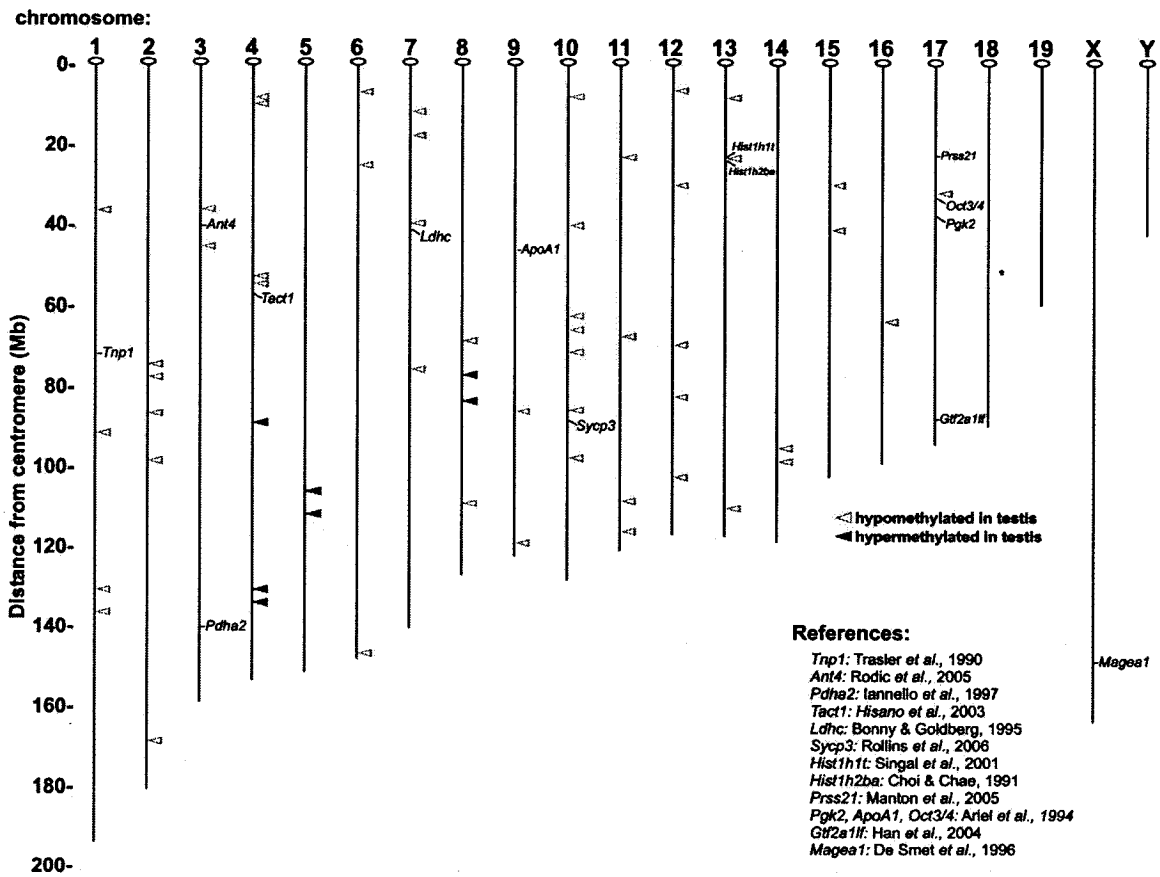

*p=0.002
100
50
0
-50
-100
%methylation testis - %methylation liver
(hyper testis)
(hypo testis)
0
10
20
30
40
50
60
70
80
distance to closest S/MAR (kb)
*Mann-Whitney rank sum test

Appendix 3.2: The relationship between the genomic locations of hypomethylated testis-specific genes and testis-specific differentially methylated loci.

There is much evidence that supports a functional relationship between promoter-specific hypomethylation and gene expression. Although in this chapter we have concluded that much of the testis-specific hypomethylation is not likely associated directly with gene expression, it is possible that the expression of some testis-specific genes is associated with hypomethylated loci. Only three genes of the 66 analyzed demonstrated the highest level of expression in the tissue that was hypomethylated, and all three occurred in the testis. This is an agreement with the finding that the vast majority of genes that are known to display this association are found in the testis and not other tissues (MacLean and Wilkinson, 2005).

We find that several of the testis-specific genes that are known to be specifically hypomethylated in the testis are located near genomic loci that are found to be differentially methylated between the testis and somatic tissues (Figure 3.12). Half of these genes are located between 100 kb and 4 Mb from a locus identified by RLGS to be hypomethylated in the testis; none are associated with hypermethylated sites. The distances separating several of these genes from an observed hypomethylated site is larger than what would be expected to be influenced by the activity of the gene itself. It is possible that regional chromatin changes that may occur on a larger scale would allow gene activity to be increased if the gene were located in a hypomethylated region.

Figure 3.12: The relationship between the locations of testis-specific genes and hypomethylated genomic loci determined by RLGS. All testis-specific genes that are known to be specifically hypomethylated in the testis in mouse are shown (MacLean and Wilkinson, 2005). Half of these genes are located between 100 kb and 4 Mb from a locus identified by RLGS to be hypomethylated in the testis. Hyper- and hypomethylated loci are shown as closed red or open blue triangles, respectively.



CONNECTING TEXT

The primary objective of the work described in chapter III was to determine the status of DNA methylation in the testis. The adult testis is made up many distinct types of germ cells in various stages of spermatogenic development; the germ cells in several of these stages possess striking differences in their chromatin structure. Once the unique nature had been described, we wanted to determine if the very complicated developmental process of spermatogenesis is associated with changes to the unique pattern of DNA methylation. In chapter IV, by isolating purified populations of male germ cells, we determine the changes in DNA methylation that occur during spermatogenesis and elucidate the origins of the unique pattern of DNA methylation in the testis.

CHAPTER IV

Developmental Acquisition of Genome-Wide DNA Methylation Occurs Prior to Meiosis in Male Germ Cells

Christopher C. Oakes, Sophie La Salle, Dominic J. Smiraglia, Bernard Robaire, & Jacquetta M. Trasler

Manuscript submitted to Developmental Biology

ABSTRACT

The development of germ cells is a highly ordered process that begins during fetal growth and is completed in the adult. Epigenetic modifications that occur in germ cells are important for both germ cell function and post-fertilization embryonic development. We have previously shown that male germ cells in the adult mouse have a highly distinct epigenetic state, as revealed by a unique genome-wide pattern of DNA methylation. Although it is known that these patterns begin to be established during fetal life, it is not known to what extent DNA methylation is modified during spermatogenesis. We have used restriction landmark genomic scanning (RLGS) and other techniques to examine DNA methylation at multiple sites across the genome during postnatal germ cell development in the mouse. Although a significant proportion of the distinct germ cell pattern is acquired prior to the type A spermatogonial stage, we find that both *de novo* methylation and demethylation occurs during spermatogenesis, mainly in spermatogonia and spermatocytes in early meiotic prophase I. Alterations include predominantly non-CpG island sequences from both unique loci and repetitive elements. These modifications are progressive and are almost exclusively completed by the end of the pachytene spermatocyte stage. These studies define the developmental timing of genome-wide DNA methylation pattern acquisition during spermatogenesis.

INTRODUCTION

Epigenetic marks in the form of DNA methylation are involved in the development of germ cells and important in the maintenance of fertility. Catalyzed by a family of DNA methyltransferase (DNMT) enzymes, mammalian DNA is commonly modified by the addition of a methyl group to the 5th position of the cytosine ring in CpG dinucleotides. DNA methylation is thought to act by promoting heterochromatin formation that can lead to gene repression when present in regulatory regions of genes (Klose and Bird, 2006). The importance of DNA methylation in male germ cells has been highlighted by studies investigating the properties of DNMTs. DNMTs are expressed in male germ cells in a developmentally regulated fashion, and some are expressed as germ cell-specific alternative transcripts (La Salle et al., 2004; La Salle and Trasler, 2006a; Mertineit et al., 1998; Shovlin et al., 2006). *Dnmt3L*, a DNMT lacking catalytic activity, is expressed at especially high levels in the gonocytes in fetal testes beginning at 15.5-18.5 days *post coitum* (dpc) (Bourc'his et al., 2001; La Salle et al., 2004). Males lacking *Dnmt3L* are infertile due to a complete lack of mature germ cells (Bourc'his et al., 2001; Hata et al., 2002) and display structural abnormalities of meiotic chromosomes (Bourc'his and Bestor, 2004; Webster et al., 2005). DNA methylation is not fully acquired in germ cells of these mice at several repetitive and non-repetitive sequences, including imprinted and non-imprinted loci (Bourc'his and Bestor, 2004; Hata et al., 2006; Kaneda et al., 2004; La Salle et al., 2007; Oakes et al., 2007b; Webster et al., 2005). Germ cell-specific deletion of *Dnmt3a* results in infertility and a loss of methylation at imprinted genes but not at repeat sequences (Kaneda et al., 2004).

The importance of DNA methylation in male germ cells is also inferred by its distinct patterns of DNA methylation. We have shown that genome-wide DNA methylation patterns involving an array of sequence types are highly unique in spermatozoa compared to somatic tissues in mouse (Oakes et al., 2007b). Recently, a study in humans concluded that DNA methylation in sperm is highly distinct from all cells and tissues of somatic origin (Eckhardt et al., 2006). Other studies have found distinct DNA methylation states in repetitive sequences

(Sanford et al., 1984) and some testis-specific genes in male germ cells (MacLean and Wilkinson, 2005). This unique state of DNA methylation arises from a genome-wide reprogramming event that occurs specifically in the primordial germ cells (PGCs) of the developing embryo (Reik et al., 2001). Between 10.5 and 12.5 dpc, patterns of DNA methylation are erased in PGCs in imprinted and testis-specific genes (Hajkova et al., 2002; Maatouk et al., 2006). Repetitive elements, such as the intracisternal A particle (IAP), LINE-1 (L1) and minor satellites undergo a similar demethylation although not to the same extent as is seen for single-copy genes (Hajkova et al., 2002; Lees-Murdock et al., 2003; Walsh et al., 1998). In the male, DNA methylation begins to be reestablished around 15.5 dpc for imprinted genes (Davis et al., 2000; Li et al., 2004; Ueda et al., 2000). The repeat sequences IAP and L1 are remethylated by 17.5 dpc (Lees-Murdock et al., 2003). This also coincides with the developmental time point where germ cells stain strongly using an antibody directed against 5-methylcytosine (Coffigny et al., 1999).

Although several studies have addressed the acquisition of DNA methylation during fetal germ cell development, few have investigated the behavior of DNA methylation patterns during spermatogenesis. Spermatogenesis is a well defined, complex developmental process whereby morphologically distinct, haploid spermatozoa that are capable of fertilization are produced from diploid germ cell precursors (Russell et al., 1990). In mammals, this process continually produces a supply of spermatozoa for the duration of the life of the adult animal. Spermatogonial stem cells occupy the seminiferous tubules of the testis and continually provide a pool of undifferentiated diploid cells called type A spermatogonia. These cells undergo several rounds of mitosis before entering meiotic prophase I, where tetraploid spermatocytes pair and recombine homologous chromosomes. After recombination, spermatocytes are reduced to haploid spermatids that undergo morphological changes from round spermatids to elongating spermatids and finally to spermatozoa. These spermatozoa exit the testis and complete their maturation process during epididymal transit.

Studies of the paternally methylated imprinted gene, *H19*, shows that although initial acquisition occurs before birth, complete levels of DNA methylation are not achieved until the pachytene spermatocyte phase of spermatogenesis (Davis et al., 1999). The two other known paternally methylated imprinted genes, *Rasgrf1* and *Gtl2*, similarly acquire most of their DNA methylation in the prenatal window, but have yet to acquire the levels found in spermatozoa (Li et al., 2004). Other data show that some sequences have fully acquired their DNA methylation status before the beginning of spermatogenesis, including some repetitive elements, such as IAP, L1 and satellite sequences (Bourc'his and Bestor, 2004; Lees-Murdock et al., 2003; Walsh et al., 1998). The hypomethylated state of *Pgk-2*, a testis-specific gene expressed in spermatocytes, is also established prior to spermatogenesis (Geyer et al., 2004). Although limited data point to the acquisition of DNA methylation patterns beyond the fetal development window, a comprehensive study of the timing and the range of sequences involved has not been done.

Restriction landmark genomic scanning (RLGS) is a highly reproducible technique that is used to investigate genome-wide patterns of DNA methylation in a variety of sequences. In combination with a recently developed second generation virtual RLGS resource (Smiraglia et al., 2007) that uses genomic sequence to produce simulated RLGS profiles, individual genomic loci that display alterations of DNA methylation can be identified. In this study, we produce a non-biased, detailed view of the patterns of DNA methylation in a variety of sequences as male germ cells progress though spermatogenesis. We find that both *de novo* methylation and demethylation occurs during spermatogenesis in a sequence-specific manner. Most importantly, we establish that, in addition to prenatal acquisition, patterns of DNA methylation at multiple sites across the genome are acquired postnatally and are complete prior to meiosis in male germ cells.

MATERIALS AND METHODS

Isolation of purified spermatogenic cells

Adult male C57BL/6 mice were obtained from Charles River Laboratories (St-Constant, Quebec). All animal studies were conducted in accordance with the principles and procedures outlined in the Guide to the Care and Use of Experimental Animals prepared by the Canadian Council on Animal Care. Purified populations of type A spermatogonia, early pachytene and pachytene spermatocytes, and round and elongated spermatids were obtained from the testes of mice using the sedimentation velocity method (Bellve et al., 1977a). Type A spermatogonia were obtained from 8-day *post partum* (dpp) mice with an average purity of 86% (n=2, 100 mice pooled per cell separation). Early pachytene spermatocytes were obtained from 17 dpp mice with a purity of 75% (n=1 cell separation, 100 mice pooled). Pachytene spermatocytes (avg. purity=85%), round spermatids (avg. purity=95%) and elongated spermatids (avg. purity of nucleated cells=97%) were obtained from 70 dpp mice (n=3, 12 mice pooled per cell separation). Spermatozoa were isolated from the cauda epididymidis of 70 dpp mice (n=3, 12 mice pooled per purification, avg. purity=99%) as described previously (Alciviar, 1989). Primitive type A spermatogonia were isolated from the testes of 6 dpp GOF18deltaPE-Oct4/GFP (Yoshimizu et al., 1999) mice that have been bred into C57BL/6 background for 3+ generations. Germ cells were isolated using fluorescence-assisted cell sorting (FACS) as described in (La Salle et al., 2007) (n=3, 3-4 mice pooled per purification). The qAMP method was used (see below) to determine the level of inter-strain variability in DNA methylation levels. Primitive type A spermatogonia were isolated from the testes of 6 dpp GOF18deltaPE-Oct4/GFP bred to CD1 mice and methylation levels for approximately 20 randomly chosen genomic loci were found to be similar (<15% variation) to GOF18deltaPE-Oct4/GFP mice bred into C57BL/6. Liver, intestine and brain tissues were isolated from adult C57BL/6 mice. Genomic DNA was isolated using proteinase K and phenol followed by

dialysis for the RLGS and Southern blotting experiments or the DNeasy Tissue Kit from Qiagen (Germantown, MD, USA) for qAMP analysis.

RLGS and spot identification

RLGS was done as described previously (Okazaki et al., 1995). Densitometry of RLGS spots was done by exposing the RLGS gel to a phosphorimager screen from Kodak (Rochester, NY, USA). Images were analyzed using the ImageQuant v5.1 software from GE Healthcare (Piscataway, NJ, USA). Spot density values were obtained by comparing a spot of interest to approximately 10-15 surrounding spots of unchanged intensity. In order for a spot to be identified as having altered DNA methylation, the alteration had to be consistent in all RLGS profiles of the same cell type and spot densitometry had to reveal a difference of greater than 25%. Genomic location of spots was identified using a virtual RLGS resource (Smiraglia et al., 2007). All identified spots were confirmed by the BAC mixing gel method (Oakes et al., 2007b). CpG islands were defined as done previously (Gardiner-Garden and Frommer, 1987) and 5' regions were defined as being within 1 kb of the transcriptional start site or within 200 bp of a CpG island that was found within the first exon or up to 5kb from the transcriptional start site. Spermatogenic cell type-specific gene expression data were obtained from the Mammalian Reproductive Genetics database (http://mrg.genetics.washington.edu/).

DNA methylation analysis using qAMP

The qAMP method was done as described previously (Oakes et al., 2006). Briefly, genomic DNA is digested in separate reactions with either no enzyme (sham digest), methylation sensitive restriction enzymes (MSREs) and the methylation-dependent restriction enzyme, McrBC. Primers are designed to flank restriction sites of interest and individually-digested DNA samples are amplified using real-time PCR. Shifts in Ct value (ΔCt) between the sham- and enzyme-digested samples are used to calculate the percentage of methylation at the various CpG sites within the amplified region (MSREs: %methylation=100(2^{-}

$^{\Delta Ct}$); McrBC: %methylation=100(1-$2^{-\Delta Ct}$). All ΔCt values are the means of triplicate reactions. Due to the curved relationship between ΔCt and percent methylation, MSREs are more accurate in the low % methylation range (<50%) and McrBC is more accurate in high % methylation ranges (>50%). Primers used to analyze genomic regions identified by RLGS are as follows: *AK137601*, 5'-CTCCCCCATTCTCCCTCAC 5'-CCTAACTTCTTGCCGTGCTC; *Polg*, 5'-CAGACCTCCACGTCGAACA 5'-CAGAGCCTGCCTTACTTGGA; *Abt1*, 5'-CCATGGGCGTGTTATGTAGA 5'-TGCTTGATGGGATGTTCATT; *Ibtk*, 5'-ACTCTCCTGCCTTGCACCT 5'-GCTGTCCACTCGGTGTCAT; *Tcf3*, 5'-GCAAGGGCCTGGATAGGA 5'-GCTACCCACTCCGAGCAA. Primers used to analyze differentially methylated regions (DMRs) of imprinted genes are as follows: *H19*, 5'-AAAAGCAGAAGGCAGGACAC 5'-ATGTTCCAGAGACAGCCAAAG; *H19* (McrBC), 5'-AGCCGTTGTGAGTGGAAAGA 5'-CATAGCGGCTTCGGACATT; *Rasgrf1*, 5'-CTGCACTTCGCTACCGTTTC 5'-CAGCAGCAGCAGTAGCAGTC; *Gtl2*, 5'-CCGTGAACTAGCGAGGAGGT 5'-ATAATGCAGCCCTTCCCTCA. In the chromosome-wide survey, a region was considered to be different if at least one of the enzyme digests detected a reproducible difference of 15% or greater in each replicate (n=2 for both primitive type A spermatogonia and spermatozoa), and that the enzyme used to detect the difference was within its accurate percent range. The difference in both enzyme digests had to be in the same direction or unchanged. Primers used to analyze regions on chromosome 7 are listed in Table 4.4 primers for chromosomes 4, 10, 17 and X are listed in Table 3.4. Primers were designed using the Primer3 software (http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi). Genomic sequence data was obtained from the University of California at Santa Cruz Genome Browser, version mm7 (http://www.genome.ucsc.edu).

Southern blotting

Southern blots were performed as described (Trasler et al., 1990) and visualized by autoradiography. Minor satellite probes were constructed by PCR

amplification of mouse genomic DNA using primers 5'-CATGGAAAATGATAAAAACC and 5'-CATCTAATATGTTCTACAGTGTGG (Lehnertz et al., 2003). The ribosomal DNA (rDNA) repeat probe was constructed using primers 5'-CGTTATGGGGTCATTTTTGG and 5'-CAGACCCAAGCCAGTAAAAAG to analyze HpaII sites located in the proximal promoter of the rDNA repeat. The IAP probe has been used previously (Michaud et al., 1994; Walsh et al., 1998). DNA was digested completely with either MspI or its methylation-sensitive isoschizomer HpaII. The membrane was stripped and reprobed according to the manufacture's recommended conditions (Hybond, GE Healthcare).

RESULTS

Detection of alterations of DNA methylation during spermatogenesis using RLGS

RLGS investigates genome-wide patterns of DNA methylation by separating genomic DNA that has been digested with the methylation-sensitive restriction enzyme, NotI, by two-dimensional gel electrophoresis. In the mouse, NotI sites occur in a variety of sequence types. To determine if the pattern of genome-wide DNA methylation in spermatozoa is acquired during spermatogenesis, RLGS profiles of purified populations of type A spermatogonia, pachytene spermatocytes from two developmental time points (early and mid-late pachytene) as well as post-meiotic round and elongating spermatids were generated (Figure 4.1a). The intensity of a total of 19 RLGS spots is observed to be different between these cell types; 11 demonstrate increased methylation (*de novo* methylation) and 8 are demonstrate decreased methylation (demethylation) during spermatogenesis, as indicated by a loss or a gain of spot intensity, respectively (Figure 4.1b). The majority of the changes in individual spot intensities, in *de novo* methylation and demethylated directions, occur between type A spermatogonia and early pachytene spermatocytes. The intensity of some spots continues to change between early and mid-late pachytene, and

always occurs in the same direction. With the exception of one spot, all spots do not gain or lose measurable amounts of methylation after the pachytene stage. Other than the progressive changes that occur between type A spermatogonia and spermatocytes, no *de novo* or demethylation events are observed in any of the cell types tested. Virtual RLGS analysis (Smiraglia et al., 2007) reveals that, in the analyzable window of the RLGS gel, there are 2954 potential RLGS spots that originate from approximately 2600 NotI sites (Table 4.1). This indicates that only a small fraction (<0.7%) of the assayable NotI sites display modified DNA methylation during spermatogenesis, leaving greater than 99% unchanged.

Germ cell-specificity of spots that show altered methylation during spermatogenesis

To determine if the spots that display altered methylation during spermatogenesis have a methylation status that is unique to male germ cells, the intensity of changed spots was examined in three somatic tissues: liver, intestine and brain (Figure 4.2a). Thirteen of the 19 spots were hypermethylated (absent) in all three somatic tissues studied, and none of them were hypomethylated in all three tissues. A minority (6/19) of spots that are hypomethylated in somatic tissues demonstrate tissue-specific patterns of methylation (Figure 4.2b). Spots that were *de novo* or demethylated during spermatogenesis demonstrated equal levels of germ cell-specificity, indicating that the unique hypomethylated state of these loci is not related to the methylation states in a particular phase of spermatogenesis. The dissimilarity between spermatozoa and somatic profiles (Oakes et al., 2007b), versus the relative similarity between type A spermatogonia and spermatozoa, indicates that the bulk of germ cell-specific methylation pattern is acquired prior to the type A spermatogonia stage.

Identification of spots that show alterations during spermatogenesis

Identification of the genomic location of spots of interest was accomplished by using a second-generation virtual RLGS resource to identify candidate loci (Smiraglia et al., 2007), and confirming the identity of each spot

using the BAC mixing gel method (data not shown). Using these methods, we identified 5/11 spots that were *de novo* methylated and 3/8 spots that were demethylated during spermatogenesis (Table 4.2). Spots are found on several chromosomes, are located within a variety of positions relative to known genes and within various sequence types (i.e. CpG islands (CGIs), repeats, etc.). All sites are in the vicinity of expressed sequences; either in the 5' region, body, or 3' end of genes. Only 2/8 identified spots are within CGIs. Interestingly, these two sites are found only 42 kb apart on chromosome 7, one within the 5' CpG island of the *Polg* (mitochondrial DNA-directed polymerase) gene and a mRNA, *AK032343*, located upstream. Due to the GC-rich nature of the NotI recognition site, 75% of the approximate 8000 NotI sites found throughout the mouse genome are found within CGI. Although the group size is small, 2/8 (25%) is less than the expected proportion and may suggest that altered DNA methylation occurs more commonly in non-CGI sequences. Most interestingly, all three identified sites that are demethylated during spermatogenesis are found in small solitary long terminal repeats (LTRs) that belong to the mammalian retroposon-like (MaLR) and endogenous retroviral-K (ERVK) families of the LTR class of repetitive sequences, whereas all identified *de novo* methylated sites are in unique sequences.

Male germ cells possess a global gene expression profile that is highly unique from somatic tissues (Shima et al., 2004). These transcripts are found to be highly regulated during spermatogenesis. DNA methylation has been proposed to function as a transcriptional regulator by causing gene repression when present in 5' regulatory sequences (Klose and Bird, 2006). To investigate if the status of DNA methylation in specific spermatogenic cell types correlates with transcriptional activity, the DNA methylation status of all identified RLGS loci located in 5' regions of known genes was compared to known levels of gene expression in these same cell types. Gene expression data were obtained from the Mouse Reproductive Genetics Database. Approximately 400 spots have been identified on mouse RLGS profiles that are located within the 5' regions of transcribed sequences (Smiraglia et al., 2007). Of these, expression levels have

been determined for 166 known genes in type A spermatogonia, pachytene spermatocytes and round spermatids (Shima et al., 2004). Despite greater than 90% (140) of these genes demonstrating a greater than 1.5-fold difference in expression between spermatogenic cell types (66% show greater than 2-fold expression differences), greater than 99% (165/166) show no detectable change in methylation status (Table 4.3). In addition, for the only spot belonging to a 5' region that does demonstrate a change in methylation, *Polg*, increased methylation is correlated with an increase in expression, the opposite of what would be expected.

Quantitative DNA methylation analysis of selected identified loci

To determine whether changes that are observed by RLGS at NotI sites are representative of the DNA methylation status of neighboring CpGs, sites within small regions (~200 bp) flanking the NotI sites were chosen for analysis by the qAMP method. To confirm and expand upon the results found in type A spermatogonia, an additional cell type, primitive type A spermatogonia, was analyzed. These cells are derived from a time point two days earlier in spermatogonial development (6 dpp). NotI sites that are found in various positions relative to genes were chosen for analysis: 3' end (*Tcf3*), body region (*Ibtk*), 5' upstream region (*AK137601*), and 5' CGI (*Polg*). Percent methylation values at NotI sites determined using RLGS correspond to the value determined by qAMP in each of the cell types investigated (Figure 4.3). Neighboring CpG sites generally showed similar levels of methylation to the NotI site and gained or lost methylation in a similarly progressive manner. Exceptions to this are HpaII sites found in *Tcf3* and *Polg*, where no increase in methylation is detected, demonstrating that heterogeneous methylation occurs between neighboring CpGs in various cell types examined. Some small differences are observed between primitive type A and type A spermatogonia, especially at the NotI site in the *Polg* gene. *De novo* methylation and demethylation changes are found to be virtually complete by the pachytene spermatocyte stage, supporting the RLGS findings.

Acquisition of DNA methylation at paternally methylated imprinted DMRs

Imprinted genes acquire a parent-specific pattern of methylation during germ cell development. Previously, the *H19* gene was shown to possess an incomplete level of methylation in spermatogonia, specifically on the maternal allele, which was later completed by the pachytene stage (Davis et al., 1999). There are two other well described regions that possess paternally methylated DMRs. Unfortunately, none of the paternally methylated DMRs are present on our RLGS profiles. To determine if DNA methylation is being acquired during spermatogenesis, the qAMP method was used to investigate DMRs in the *Dlk1-Gtl2* region (Takada et al., 2002) and the *Rasgrf1* gene (Yoon et al., 2002). The previously defined *H19-Igf2* region (Tremblay et al., 1995) was used as control. This analysis reveals that the majority of CpGs investigated have acquired their full methylation status by the primitive type A spermatogonial stage; however, a small amount of methylation is acquired up to the pachytene stage (Figure 4.4). The HhaI enzyme digest of the *Rasgrf1* DMR displays the largest percentage increase of DNA methylation during spermatogenesis. There are three HhaI sites in the amplified region of *Rasgrf1*; if only one of the three is unmethylated, the strand will not amplify and contribute to the percentage of unmethylated strands. This particular digest reveals that a small proportion of CpGs are unmethylated in the *Rasgrf1* DMR in type A spermatogonia and that by the pachytene stage, DNA strands gain their fully methylated status. Changes in DNA methylation at other sites were minor.

Chromosome-wide survey of non-CpG island unique sequences

RLGS analysis has a strong bias towards CGIs; however, a higher proportion of non-CGI sites were shown to display altered patterns of DNA methylation during spermatogenesis. To determine the prevalence of DNA methylation changes occurring at non-CGI sites, small groups of CpGs (regions) were chosen for quantitative analysis by the qAMP method at approximately 5 Mb intervals across chromosomes 4, 7, 10, 17 and X. These regions were

chosen at random other than not being proximal (>10 kb) to a CGI, or the transcriptional start site of a known gene. Regions were also chosen to be solely within non-repetitive sequences. Analysis of 125 total regions in primitive type A spermatogonia and spermatozoa reveals differences in DNA methylation in a region-specific manner (Figure 4.5). Regions displaying high, intermediate (partial), and low levels of DNA methylation are detected, revealing that a full range of methylation levels can be found in germ cells in a site-specific manner. Differences were observed in 12 regions on 4 of 5 chromosomes. The number and/or extent of methylation differences were observed to be similar between the autosomes and the X chromosome, chromosomal position (telomeric versus centromeric), G- and R-banding patterns and flanking GC content (data not shown). Interestingly, all 12 regions that showed a difference were gaining methylation during spermatogenesis. Differences are in the same range (<60%) as previously detected at RLGS sites.

A closer examination of the changes to six of the 12 sites reveals that, like other changes observed, methylation acquisition during spermatogenesis is complete by the pachytene stage (Figure 4.6). Several of the changes are specific to the CpGs investigated by a particular restriction enzyme, indicating that heterogeneous methylation exists at these sites. Further examination of the methylation of these regions in primitive type A spermatogonia shows that most CpGs investigated have methylation states that are similar to type A spermatogonia; however, significant increases are observed at sites in two of the six regions.

Examination of repetitive sequences

The observation that all demethylated sequences identified by RLGS were of repetitive origin suggests that changes might be occurring in repetitive sequences during spermatogenesis. We chose to analyze three different types of repeat sequences that have been previously determined to have different levels of methylation: the minor satellite repeat, the ribosomal DNA repeat, and IAP, an interspersed LTR-containing endogenous retroviral sequence (Figure

4.7). Equal amounts of DNA isolated from type A spermatogonia to spermatozoa were digested with HpaII, along with somatic tissues as control. Differences are observed to occur between somatic and germ cells for both the minor satellite repeat and IAP; however, no changes are observed to occur during spermatogenesis for these classes of repeat sequences. No change in the methylation status of full-length IAP repeats is also found by a comparison of real and virtual RLGS profiles. Several hundred IAP and early transposon (ETn) repeats of the LTR class of repeat sequences are distinctly visible on virtual RLGS profiles (Oakes et al., 2007b). The spots corresponding to these repeat sequences are not observed in any spermatogenic cell type, indicating invariable hypermethylation during spermatogenesis (data not shown).

DISCUSSION

We have examined a wide variety of sequence-types to determine the development of DNA methylation patterns during spermatogenesis. Our findings demonstrate that *de novo* and demethylation events occur in a sequence-specific manner. Through the use of several methods, we have shown that sequences which undergo changes in methylation during spermatogenesis include CGI and non-CGI sequences that are found within various positions within known genes or in intergenic sequences. These modifications occur in a specific developmental window during spermatogenesis. Both *de novo* and demethylation events occur in the early phases of spermatogenesis, and, regardless of the direction of the change or the sequence type, are complete by the end of the pachytene stage. During spermatogenesis, the reported *de novo* DNA methyltransferase enzymes, DNMT3a and DNMT3b, display their highest levels of expression in spermatogonia (La Salle and Trasler, 2006a; Shima et al., 2004), and are probable candidates to facilitate *de novo* methylation events in early germ cell types. Germ cells in the early phases of spermatogenesis undergo frequent DNA replication, thus, demethylation may occur passively. Demethylation does not occur in spermatogenic cell types that are not replicating DNA. Sequences that acquire *de novo* methylation during spermatogenesis are

generally non-repetitive. Demethylated sequences are observed in solitary LTR fragments, a category of small, divergent interspersed repeat sequences that are the remnants of transposition events involving full-length LTR repeats. The specific fragments identified are from MaLR and ERVK families of repeats, sequences that have previously been found to be expressed in oocytes and early embryos (Peaston et al., 2004). Although the numbers of identified demethylated sequences are low, it is interesting to observe this dichotomy between the behavior of small repetitive and non-repetitive sequences. This difference is probably confined to a subset of repeat sequence types, as no other types of repeats tested demonstrated this behavior.

We also find that DNA methylation changes are only present in a minor proportion of the sequences investigated. Using RLGS, where 75% of all sites examined are within CGIs, only 0.7% (19/2954) of NotI sites were observed to change during spermatogenesis. The survey of non-CGI, non-repetitive sequences across five chromosomes revealed a higher 9.6% (12/125) proportion of sequences that were changing during spermatogenesis. A few possibilities exist to explain this discrepancy. Firstly, non-CGI sequences demonstrate a higher level of variability between tissues (Oakes et al., 2007b) and a similar phenomenon may be present between developing spermatogenic cells. Because only a fraction of NotI sites in the mouse genome are in non-CGI, non-repetitive sequences (Fazzari and Greally, 2004), fewer changes are expected to be observed using RLGS. Secondly, due to the random nature of spot positions on two-dimensional RLGS profiles, spots displaying altered intensity can be overlapped or obscured by others and would be missed. Thirdly, the qAMP method is more sensitive to small-scale changes (error range ±5%) (Oakes et al., 2006) than is RLGS. For these reasons, the fraction of loci found to be changed using RLGS would be considered to be an underestimate. However, even with these caveats considered, it remains clear that changes occur only in a limited proportion of the sequences examined.

Some small differences are observed between primitive type A and type A spermatogonia. One explanation is that these differences are developmental,

although we cannot exclude the possibility that some of these small differences are representative of the differences in purity levels of these cells. Evidence of developmental differences is supported by progressively decreased levels of methylation in *AK137601* (Figure 4.3) and by methylation levels increased beyond the 20% level that would be predicted by somatic contamination (*Polg*, Figure 4.3 & Figure 4.6). Type A spermatogonia have an average purity of 85%, whereas primitive type A spermatogonia isolated by flow cytometry are more highly purified. Contaminating cells in type A spermatogonia, isolated by sedimentation velocity, are likely to be immature Sertoli cells.

Previously, acquisition of DNA methylation during spermatogonial development has been demonstrated at DMRs of paternally methylated imprinted genes (Davis et al., 1999; Lees-Murdock et al., 2003). To compare our approach with previous findings, we investigated the developmental acquisition of methylation at the *H19* DMR. Our results indicate that a low amount of DNA methylation is acquired during the phases up to the pachytene stage in paternally methylated DMRs. Because the HhaI restriction enzyme in the *Rasgrf1* amplified region has the most restriction sites of all of the MSREs used, this region is the most sensitive for detecting DNA strands that are incompletely methylated. This particular measurement reveals that complete methylation is achieved by the pachytene stage. Based on the allele-specific differential acquisition of DNA methylation at *H19* (Davis et al., 1999), we predict that unmethylated CpGs would most likely be found on alleles of maternal origin.

The prevailing view of the primary biological role of DNA methylation involves the promotion of heterochromatin formation in gene promoter regions leading to a transcriptional repression. The global transcriptional profile of individual spermatogenic cell types is highly distinct, especially in the pachytene spermatocyte and round spermatid cell types where a burst of unique transcripts are produced presumably to facilitate meiotic and spermiogenic processes (Shima et al., 2004). Although the expression of a limited number of testis-specific genes has been shown to be correlated with testis-specific hypomethylation of 5' regions (Sanford et al., 1984), a primary role for DNA

methylation in the global direct control of spermatogenic cell type-specific levels of gene expression is not supported by our data. Of 166 genes examined, no 5' region was shown to be hyper- or hypomethylated in a cell type where expression was repressed or increased, respectively. Furthermore, changes in DNA methylation during spermatogenesis were more commonly found away from the regulatory (5') regions of genes. These results do not challenge the prevailing view that DNA methylation and gene expression are mechanistically linked, rather they point to potential alternative functions for DNA methylation in germ cells. In addition, there are several explanations that could contribute to the lack of correlation, including RNA stabilization. There is evidence to suggest that changes in DNA methylation play a role in establishing an epigenetic state in the early stages of germ cell development that is permissive for transcription to occur at a later stage (Geyer et al., 2004). It is likely that other regulatory mechanisms, such as transcription factor regulation, are responsible for the variability observed in transcript levels between spermatogenic cell types.

An alternate role for these modifications of DNA methylation is their involvement in the organization of a germ cell-specific chromatin configuration. Alternate roles for DNA methylation have been described, and include silencing of repetitive elements and chromatin stability/organization (Bestor and Tycko, 1996). The results of the present study indicate that the bulk of the unique germ cell-specific pattern that is achieved by meiosis has already been established in primitive type A spermatogonia. Thus, one possible explanation is that these changes represent the final modifications of an epigenetic program that is important to the organization of a specialized, genome-wide chromatin configuration necessary for passage through meiosis. There are a few observations that support this hypothesis: firstly, the majority of modifications are non-5'. The involvement of non-5' methylation in meiotic chromosomal organization is suggested by the abnormal chromosomal structures in *Dnmt3L*-null spermatocytes (Bourc'his and Bestor, 2004). These germ cells fail to gain normal methylation patterns at interspersed repetitive and intergenic/intronic loci (Bourc'his and Bestor, 2004; La Salle et al., 2007; Oakes et al., 2007b).

Although the numbers of pre-meiotic *Dnmt3L*-null germ cells are decreased, they do not completely fail until meiotic entry, indicating that male germ cell DNA methylation patterns are more essential for meiosis than mitosis (La Salle et al., 2007). Secondly, very few changes were observed to occur after meiosis, despite highly dynamic chromatin modulations in spermatid stages. Thirdly, most changes that occur are partial (20-60%) changes, indicating that some methylation has been acquired in prior stages at these sites. Finally, changes are generally restricted to sites with germ cell-specific (non-somatic) methylation states, supporting a connection to the distinct, post-meiotic patterns. An example of this connection is that the three identified loci that are demethylated during spermatogenesis are of the same family of repetitive sequences as 21 others identified on mouse RLGS profiles, 19 of which are already hypomethylated in germ cells despite being hypermethylated in somatic tissues (Oakes et al., 2007b). It is reasonable to believe that the selective demethylation of these repeat sequences during spermatogenesis reflects a requirement for male germ cells to have sequences of this type hypomethylated.

In summary, we find that in addition to the acquisition of DNA methylation that occurs in prenatal male gonocytes, patterns continue to be acquired during spermatogenesis in a sequence-specific manner. These studies raise the possibility that male germ cells may be especially sensitive to potential 'epimutations'; further studies will be required to test if these processes render male germ cells particularly sensitive to environmental influences.

ACKNOWLEDGEMENTS

We thank Liyuan Deng for excellent technical support. The IAP probe was a gift of T. Bestor and the GOF18deltaPE-Oct4/GFP mice were a gift of H. Scholer. C.C.O. and S.L. are recipients of Canadian Institutes of Health Research Doctoral Research Awards. J.M.T. is a William Dawson Scholar of McGill University and a Scholar of the Fonds de Recherches en Santé du Québec. B.R. is a James McGill Professor of McGill University.

Figure 4.1: Examination of RLGS profiles from purified spermatogenic cell types. **(a)** RLGS profiles are produced by digestion of genomic DNA with the methylation-sensitive enzyme, NotI; these cleavage sites are radiolabeled and DNA fragments are separated by two-dimensional gel electrophoresis. Visible spots reveal hypomethylated sites, absent spots are hypermethylated. Enlargements of RLGS profiles produced from type A spermatogonia, early and mid-late pachytene spermatocytes, round and elongated spermatids and spermatozoa are shown. Selected enlargements are representative of areas throughout the two-dimensional RLGS profile. Spots that are *de novo* methylated and demethylated relative to type A spermatogonia are indicated by open and black arrows, respectively. The identified genes that contain the differentially methylated spots are shown. **(b)** RLGS densitometry of spots that are *de novo* methylated and demethylated during spermatogenesis. Cell-type specific spot intensity of spots was determined by comparing the intensity of spots of interest with unchanged, surrounding spots. Percent methylation values are determined by the inverse of the relative density.

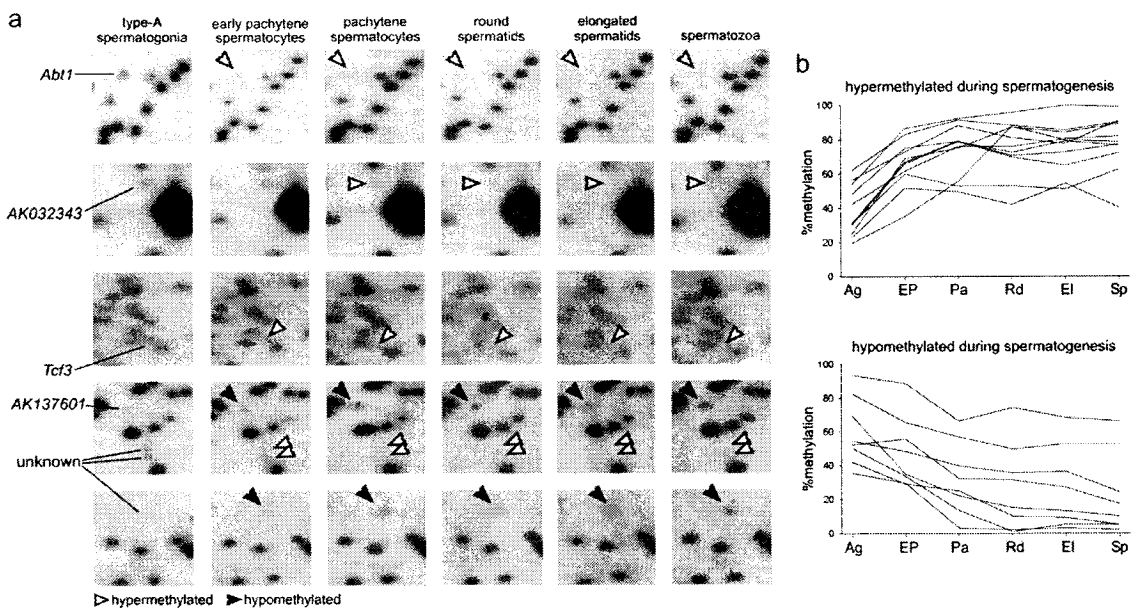

a
type-A spermatogonia
early pachytene spermatocytes
pachytene spermatocytes
round spermatids
elongated spermatids
spermatozoa
Abt1
AK032343
Tcf3
AK137601
unknown
hypermethylated
hypomethylated
b
hypermethylated during spermatogenesis
%methylation
Ag
EP
Pa
Rd
El
Sp
hypomethylated during spermatogenesis
%methylation
Ag
EP
Pa
Rd
El
Sp

Figure 4.2: Determination of the methylation state of spots that are differentially methylated during spermatogenesis in somatic tissues. (a) Enlargements of selected portions of RLGS profiles produced from type A spermatogonia, spermatozoa, liver, intestine and brain are shown. Hypomethylated spots that are differentially methylated during spermatogenesis are indicated by black arrows. The known genes associated with the differentially methylated spots are shown. **(b)** A comparison of the methylation status of individual spots that are differentially methylated during spermatogenesis with their methylation status in somatic tissues. The number of spots that are differentially methylated during spermatogenesis that are hypermethylated in 3/3, 2/3, 1/3, or 0/3 somatic tissues, demonstrating that spots that are differentially methylated during spermatogenesis are largely germ cell-specific.

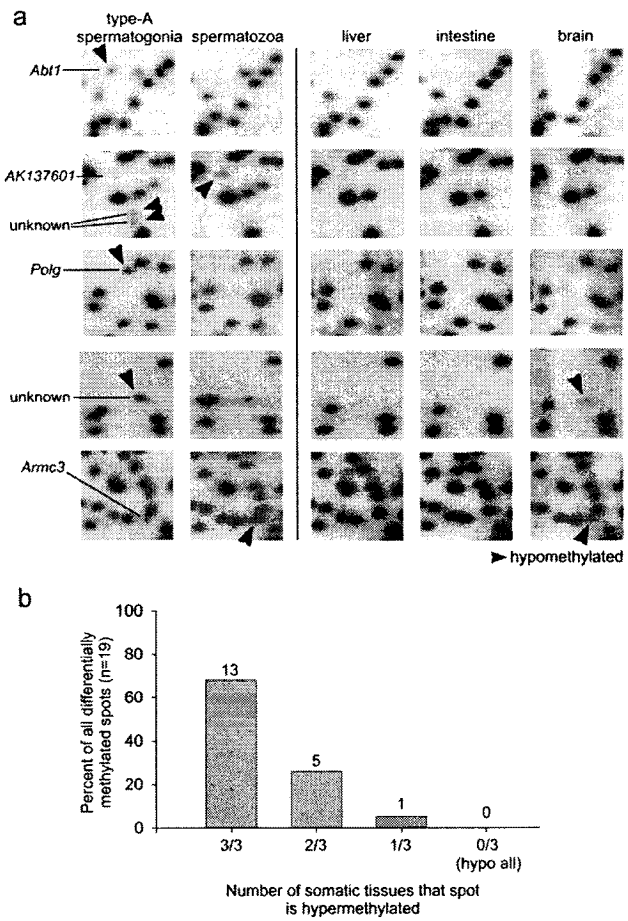

a
type-A spermatogonia
spermatozoa
liver
intestine
brain
Abl1
AK137601
unknown
Polg
unknown
Armc3
hypomethylated
b
Percent of all differentially methylated spots (n=19)
100
80
60
40
20
0
13
5
1
0
3/3
2/3
1/3
0/3 (hypo all)
Number of somatic tissues that spot is hypermethylated

Figure 4.3: Detailed examination of differentially methylated loci using the qAMP method. DNA is digested using methylation-sensitive restriction enzymes and the methylation-dependent enzyme, McrBC. Primers are designed to flank the NotI site along with neighboring restriction sites (assayable CpGs) and are amplified using real-time PCR. The positions of the assayed regions relative to known genes are shown. All genomic sequences are orientated from centromere to telomere in gene diagrams. The percent methylation at different CpG sites (or groups of sites) determined by independent enzyme digests is shown in primitive type A and type A spermatogonia, pachytene spermatocytes and spermatozoa. N, NotI; Hh, HhaI; Hp, HpaII; M, McrBC.

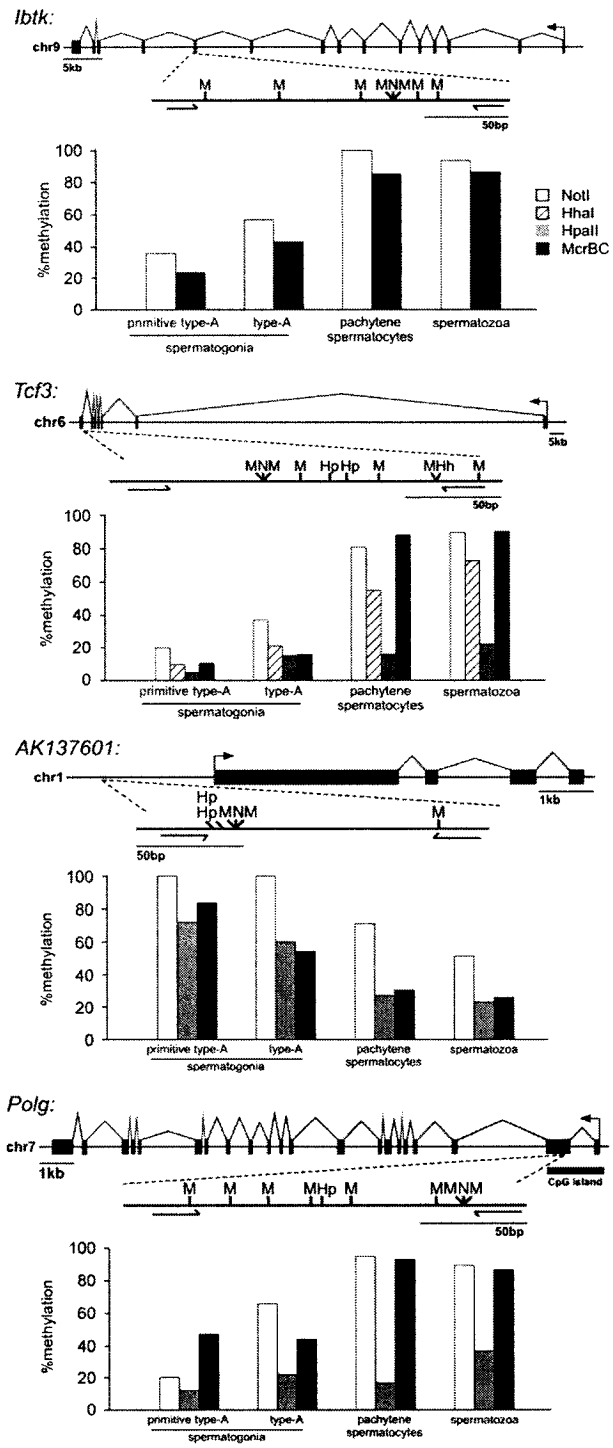

Ibtk:
chr9
5kb
M
M
M MNMM
50bp
100
80
60
40
20
0
%methylation
NotI
HhaI
HpaII
McrBC
primitive type-A
type-A
spermatogonia
pachytene spermatocytes
spermatozoa
Tcf3:
chr6
5kb
MNM M HpHp M MHh M
50bp
AK137601:
chr1
1kb
Hp
HpMNM
M
50bp
Polg:
chr7
1kb
M M M MHp M MMNM
CpG island
50bp

Figure 4.4: Examination of paternally methylated imprinted DMRs. Primers were designed to flank restriction enzyme sites within the DMRs of *H19-Igf2*, *Rasgrf1* and *Dlk1-Gtl2*. The location of the primers used to assay the regions and the restriction enzyme sites examined are shown for each DMR. All genomic sequences are orientated from centromere to telomere in gene diagrams. The percent methylation at different CpG sites (or groups of sites) determined by independent enzyme digests is shown in primitive type A and type A spermatogonia, pachytene spermatocytes and spermatozoa. N, NotI; Hh, HhaI; Hp, HpaII; M, McrBC.

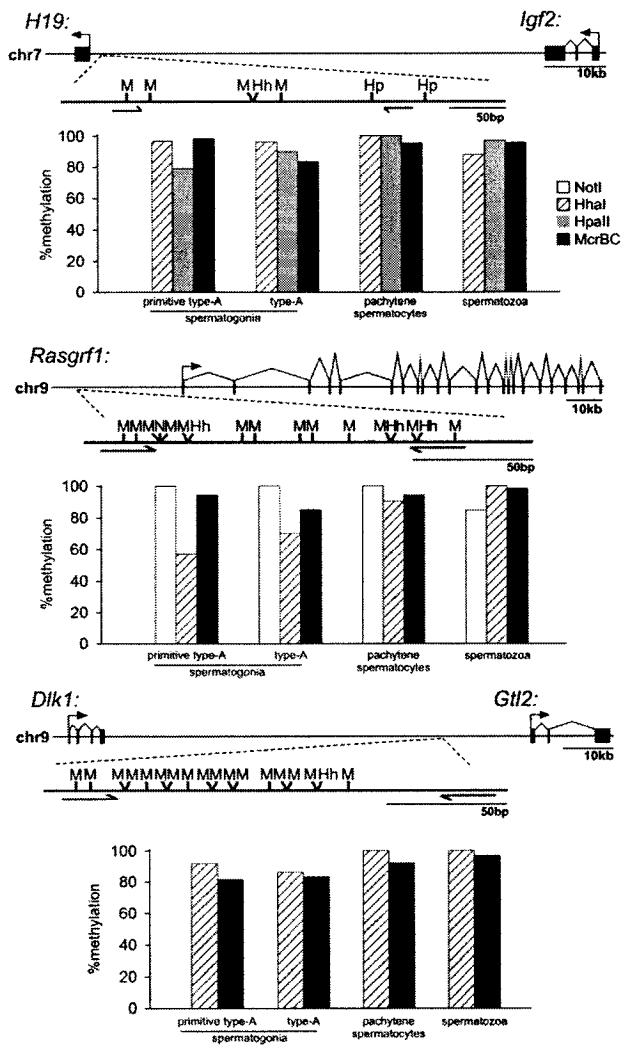

H19:
Igf2:
chr7
10kb
M M
MHhM
Hp
Hp
50bp
100
80
60
40
20
0
%methylation
NotI
HhaI
HpaII
McrBC
primitive type-A
type-A
pachytene
spermatocytes
spermatozoa
spermatogonia
Rasgrf1:
chr9
10kb
MMMNMMHh
MM
MM
M
MHhMHh
M
50bp
100
80
60
40
20
0
%methylation
primitive type-A
type-A
pachytene
spermatocytes
spermatozoa
spermatogonia
Dlk1:
Gtl2:
chr9
10kb
MM
MMMMMMMMMM
MMM
MHh
M
50bp
100
80
60
40
20
0
%methylation
primitive type-A
type-A
pachytene
spermatocytes
spermatozoa
spermatogonia

Figure 4.5: Chromosome-wide analysis of non-CGI, non-repetitive sequences. Using the qAMP method, HhaI and McrBC sites were randomly chosen for examination at approximately 5 Mb intervals across chromosomes 4,7,10,17 and X. Chosen sites were within non-repetitive sequences and >10 kb from a CpG island or the transcriptional start site of a known gene. Percent methylation values for the two digests were averaged to give a single value for primitive type A spermatogonia (light blue dash) or spermatozoa (dark blue dash). Regions that demonstrated a change in methylation during spermatogenesis are indicated (red arrows).

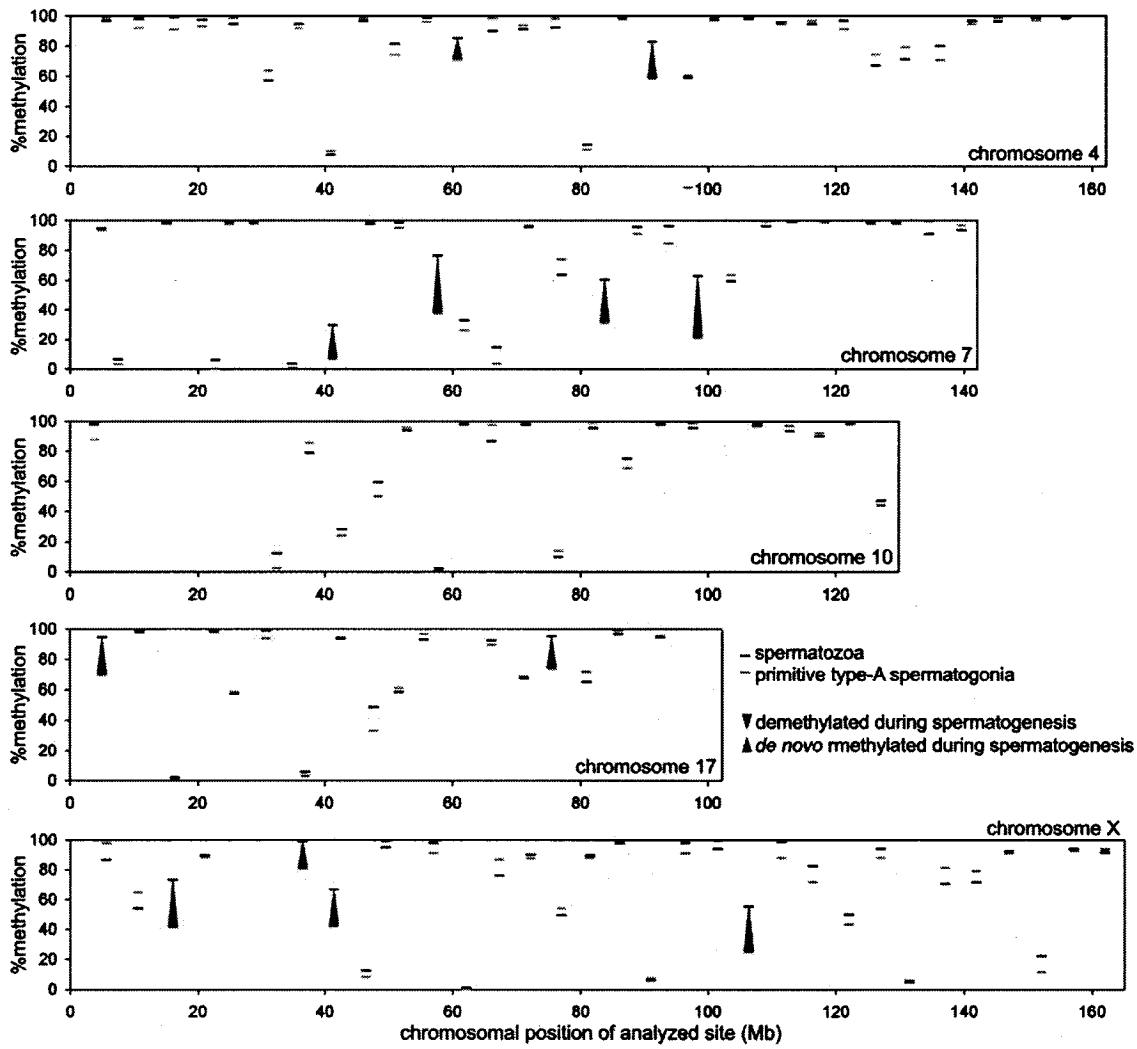

%methylation
chromosome 4
chromosome 7
chromosome 10
chromosome 17
chromosome X
– spermatozoa
– primitive type-A spermatogonia
▼ demethylated during spermatogenesis
▲ de novo methylated during spermatogenesis
chromosomal position of analyzed site (Mb)

Figure 4.6: Detailed analysis of differentially-methylated CpGs identified in the chromosome-wide analysis. Percent methylation determined by HhaI and McrBC individual restriction enzyme digests are shown for primitive type A (P-Ag) and type A (Ag) spermatogonia, pachytene spermatocytes (Pa) and spermatozoa (Sp).

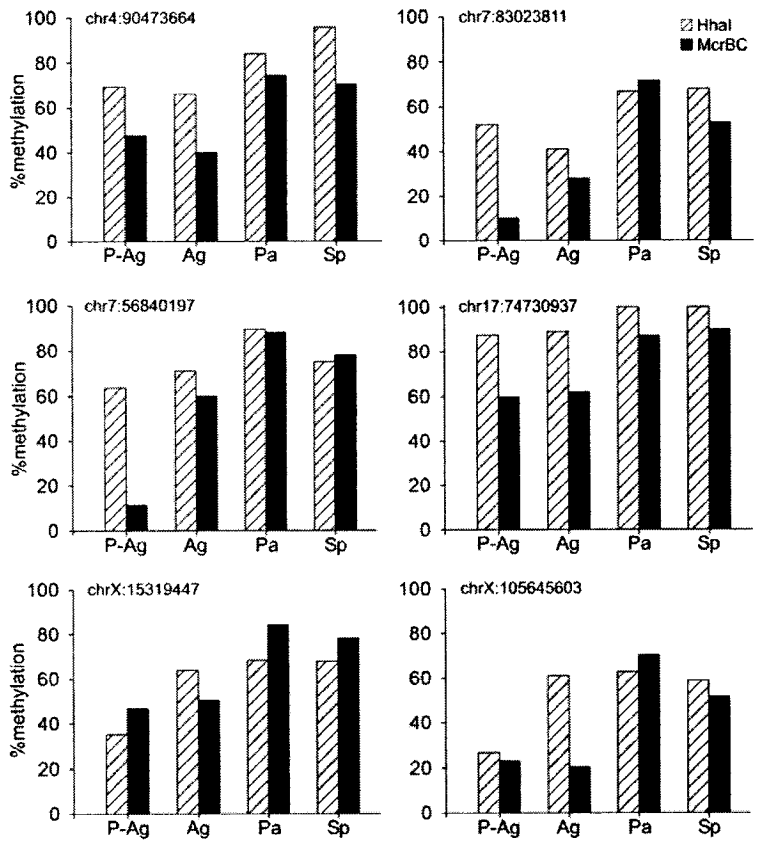


Figure 4.7: DNA methylation of repetitive elements during spermatogenesis and in somatic tissue. Genomic DNA was digested with MspI (lane 1) or HpaII (all other lanes) and hybridized to probes specific for the minor satellite, ribosomal DNA and IAP repeats. Each blot was produced by stripping and hybridizing the same membrane to each respective probe. Ag, type A spermatogonia; EP, early pachytene spermatocytes; Pa, pachytene spermatocytes; Rd, round spermatids; El, elongated spermatids; Sp, spermatozoa; T, testis; L, liver; I, intestine; B, brain.

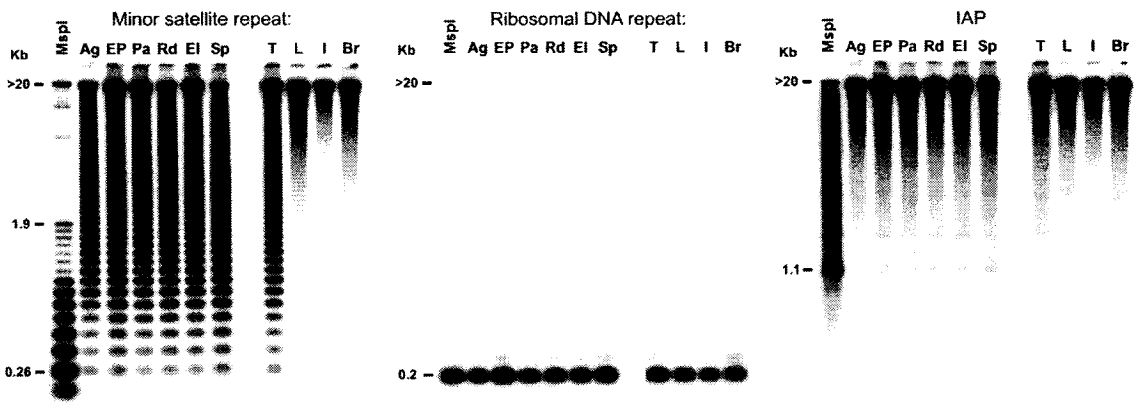


Table 4.1: RLGS spot summary

Methylation during spermatogenesis	# of spots	% of total
hypermethylated during spermatogenesis	11	0.37
hypomethylated during spermatogenesis	8	0.27
unchanged	2935	99.4
total*	2954	100

*derived from virtual RLGS profile

Table 4.2: Characteristics of identified loci

Methylation during Spermatogenesis	RLGS spot	Gene	NotI site position	CpG island	Repeat (family)	Genome position
Hypermethylated	5ax2	*Tcf3*	3'	N	non-repetitive	chr6:72645454
Hypermethylated	1gx8	*Polg*	5'	Y	non-repetitive	chr7:75333081
Hypermethylated	4C13	*AK032343*	body	Y	non-repetitive	chr7:75377378
Hypermethylated	3ex8	*Ibtk*	body	N	non-repetitive	chr9:85800773
Hypermethylated	2F36	*Abt1*	body	N	non-repetitive	chr13:22791131
Hypomethylated	3ex3	*AK137601*	5'	N	LTR (MaLR)	chr1:36371367
Hypomethylated	2G75	*Armc3*	body	N	LTR (ERVK)	chr2:19315058
Hypomethylated	4dx1	*AK035353*	body	N	LTR (ERVK)	chr4:9701990

Table 4.3: DNA methylation of all identified 5' RLGS loci compared to developmental expression in male germ cells (Shima *et al.* 2004)

RLGS Spot	NotI site position (mm7)	CpG island	Gene symbol	Methylation during spermatogenesis	Gene Expression: Type-A spermatogonia	Pachytene spermatocytes	Round spermatids	Type-A/ Pachytene	Type-A/ Round	Pachytene/ Round
>2.0 -fold increase or decrease in gene expression between type-A spermatogonia, pachytene spermatocytes and round spermatids:										
4B09	chr3:121195825	Y	Cnn3	**unchanged**	868.5	52.5	26.3	16.54	33.02	2.00
4F21	chr10:81353988	Y	Aes	**unchanged**	675.2	60.9	30.1	11.09	22.43	2.02
3B25	chr3:30907176	Y	Skil	**unchanged**	108.2	5	8.8	21.64	12.30	0.57
2C06	chr1:182004461	Y	Enah	**unchanged**	189.9	45.4	11.4	4.18	16.66	3.98
5B33	chr7:59688248	Y	Klf13	**unchanged**	188.2	29.8	12.5	6.32	15.06	2.38
6E41	chr10:74622936	Y	Gnaz	**unchanged**	145.5	10.6	21.3	13.73	6.83	0.50
2B16	chr7:34910692	Y	C80913	**unchanged**	77.6	11.3	5.8	6.87	13.38	1.95
1F40	chr11:102426346	Y	Ubtf	**unchanged**	489.3	73.3	40.1	6.68	12.20	1.83
3D10	chr5:42554991	Y	Cpeb2	**unchanged**	66.9	404.2	35.9	0.17	1.86	11.26
4D11	chr14:100327082	Y	Spry2	**unchanged**	835.9	158.7	79.9	5.27	10.46	1.99
1F06	chr9:65490698	Y	Spg21	**unchanged**	315	70.9	30.2	4.44	10.43	2.35
6G05	chr10:79496639	Y	Hcn2	**unchanged**	24.8	2.4	4	10.33	6.20	0.60
4E34	chr5:103650201	Y	Pkd2	**unchanged**	273	48.2	28.9	5.66	9.45	1.67
6B09	chr3:76062107	Y	Golph4	**unchanged**	42.6	11.3	4.9	3.77	8.69	2.31
3D55	chr17:54484543	Y	Jmjd2b	**unchanged**	156.6	18.6	81.4	8.42	1.92	0.23
4C12	chr10:70950199	Y	Ipmk	**unchanged**	135.9	17.1	79.2	7.95	1.72	0.22
3B05	chr9:88582905	Y	Syncrip	**unchanged**	248	60.7	34.7	4.09	7.15	1.75
1D06	chr13:48354380	Y	Phf2	**unchanged**	248.6	1737.9	266.5	0.14	0.93	6.52
4C18	chrX:69060578	Y	Slc6a8	**unchanged**	38	6.1	7	6.23	5.43	0.87
4E47	chr6:91199044	Y	Nup210	**unchanged**	383.4	62.7	123.7	6.11	3.10	0.51
2G31	chr7:139888509	Y	Cdkn1c	**unchanged**	136.7	23	157.6	5.94	0.87	0.15
2C39	chr4:47367145	Y	Tgfbr1	**unchanged**	10.1	1.8	5.2	5.61	1.94	0.35
2G14	chr6:23775774	Y	Cadps2	**unchanged**	253.6	47.6	93.4	5.33	2.72	0.51
3F33	chr16:96463994	Y	Hmgn1	**unchanged**	3143.8	609.5	592.6	5.16	5.31	1.03
1C12	chr4:136735561	Y	Cdc42	**unchanged**	689.6	176.4	131.1	3.91	5.26	1.35
4D57	chr4:28933944	Y	Epha7	**unchanged**	4.7	24.4	4.7	0.19	1.00	5.19
3F63	chr12:71861462	Y	Hif1a	**unchanged**	1068.3	206.9	207.5	5.16	5.15	1.00
2E44	chr16:46275446	Y	Pvrl3	**unchanged**	27.9	70.6	14.2	0.40	1.96	4.97
2C33	chr13:59285249	Y	Gas1	**unchanged**	223.4	45.9	46	4.87	4.86	1.00
3B15	chr13:31173603	Y	Foxc1	**unchanged**	2.8	3.4	0.7	0.82	4.00	4.86

RLGS Spot	NotI site position (mm7)	CpG island	Gene symbol	Methylation during spermatogenesis	Gene Expression					
					Type-A spermatogonia	Pachytene spermatocytes	Round spermatids	Type-A/ Pachytene	Type-A/ Round	Pachytene/ Round
4D69	chr10:82305166	Y	Nfyb	**unchanged**	246.9	70.9	50.9	3.48	4.85	1.39
4E04	chr4:43537404	N	Tpm2	**unchanged**	75	18.8	15.7	3.99	4.78	1.20
5G11	chr13:23076518	N	Hfe	**unchanged**	46.6	36.8	10.1	1.27	4.61	3.64
3C02	chr18:5759245	Y	Zfhx1a	**unchanged**	214	47.6	72.9	4.50	2.94	0.65
3A04	chr19:28951323	Y	Jak2	**unchanged**	51.1	11.4	23.3	4.48	2.19	0.49
1D16	chr7:33589283	Y	Zfp537	**unchanged**	227.8	51.5	92.9	4.42	2.45	0.55
2G96	chr14:12516217	Y	Slc4a7	**unchanged**	440.3	216.4	101.3	2.03	4.35	2.14
6D07	chr10:90641565	Y	Apaf1	**unchanged**	83.4	19.3	33.6	4.32	2.48	0.57
5C23	chr5:47309019	Y	Slit2	**unchanged**	154.9	38.4	96.5	4.03	1.61	0.40
4F59	chr16:20854197	Y	Ephb3	**unchanged**	24.8	13	6.2	1.91	4.00	2.10
4E70	chr11:75254073	Y	Hic1	**unchanged**	119.5	121	30.3	0.99	3.94	3.99
5G21	chr7:61228791	Y	Tjp1	**unchanged**	570.2	245.2	144	2.33	3.96	1.70
3E54	chr10:67139468	Y	Egr2	**unchanged**	296.1	96.3	77.4	3.07	3.83	1.24
1D14	chr3:34457763	Y	Sox2	**unchanged**	15.6	4.1	10.1	3.80	1.54	0.41
1E20	chr8:68540083	Y	Hapln4	**unchanged**	25.7	90.7	23.9	0.28	1.08	3.79
3A03	chr4:8617986	Y	Chd7	**unchanged**	54.2	14.7	17.8	3.69	3.04	0.83
4D28	chr10:126812970	Y	Mars	**unchanged**	602.9	292	163.7	2.06	3.68	1.78
2C27	chr9:61396680	Y	Tle3	**unchanged**	240.9	73	66.6	3.30	3.62	1.10
1G11	chr9:31111394	Y	Aplp2	**unchanged**	708.6	856.3	242.6	0.83	2.92	3.53
1E24	chr16:21917021	Y	Sfrs10	**unchanged**	557.4	1274.9	367.3	0.44	1.52	3.47
2C55	chr18:15032382	Y	Ss18	**unchanged**	325.4	205.5	94.2	1.58	3.45	2.18
4E02	chr8:118546524	Y	Usp10	**unchanged**	226.8	147.7	66.1	1.54	3.43	2.23
1F27	chr2:180543938	Y	Psma7/Ss18l1	**unchanged**	310.3	123	90.5	2.52	3.43	1.36
2G41	chr13:47744462	Y	Id4	**unchanged**	53.9	173	51.1	0.31	1.05	3.39
5H105	chr11:99047138	Y	Rara	**unchanged**	67.3	27.8	20.2	2.42	3.33	1.38
4C22	chr16:91456904	Y	Il10rb	**unchanged**	203.8	74.9	61.5	2.72	3.31	1.22
4B03	chr12:106313454	Y	Yy1	**unchanged**	151.4	161.5	49	0.94	3.09	3.30
5F57	chr15:76793102	Y	Recql4/Lrrc14	**unchanged**	175.6	60.2	63	2.92	2.79	0.96
4D29	chr8:69847664	Y	Nr2f6	**unchanged**	254.4	90	104.8	2.83	2.43	0.86
1C10	chr19:38537602	Y	Tbc1d12	**unchanged**	73.4	188.7	69.5	0.39	1.06	2.72
3F19	chr5:105053964	Y	Zfp326	**unchanged**	133.1	93.3	49.1	1.43	2.71	1.90
2E39	chr17:32034148	Y	Rps18	**unchanged**	4823.6	1963.6	2177.8	2.46	2.21	0.90
1D22	chr7:105875598	Y	Rab6ip1	**unchanged**	382.7	222.2	155.9	1.72	2.45	1.43
4B27	chr13:95424603	Y	Foxd1	**unchanged**	24.5	13.4	10.1	1.83	2.43	1.33
6E43	chr6:87925903	Y	Cnbp1	**unchanged**	1001.3	772.8	416.5	1.30	2.40	1.86

RLGS Spot	NotI site position (mm7)	CpG island	Gene symbol	Methylation during spermatogenesis	Gene Expression					
					Type-A spermatogonia	Pachytene spermatocytes	Round spermatids	Type-A/ Pachytene	Type-A/ Round	Pachytene/ Round
4F52	chr11:90747562	Y	Stxbp4/Cox11	**unchanged**	681.6	460.5	287.2	1.48	2.37	1.60
1F21	chrX:53512030	Y	Zic3	**unchanged**	40.8	17.2	60.5	2.37	0.67	0.28
5F59	chr9:47491956	Y	Igsf4a	**unchanged**	407.6	174.6	406.4	2.33	1.00	0.43
5G109	chr11:98136023	Y	Rpl19	**unchanged**	2148.7	924.6	1101.5	2.32	1.95	0.84
4B07	chr3:144463229	Y	Hs2st1/Sep15	**unchanged**	56.2	25.5	36.3	2.20	1.55	0.70
6D04	chr5:31167823	Y	Fosl2	**unchanged**	53	24.5	71.6	2.16	0.74	0.34
5D42	chr5:51939647	Y	Lgi2	**unchanged**	317.6	174	153.6	1.83	2.07	1.13
2F64	chr14:58104671	Y	Gata4	**unchanged**	192.1	94.7	96.7	2.03	1.99	0.98
7G10	chr12:84079990	N	Tgfb3	**unchanged**	177.1	89.1	92.8	1.99	1.91	0.96
3E74	chr7:130181507	Y	Dhx32/Fank1	**unchanged**	8.1	3570.7	4619.2	0.00	0.00	0.77
1E22	chr10:22488086	Y	Tbpl1	**unchanged**	95.6	1237.9	1841.9	0.08	0.05	0.67
3C09	chr9:87831050	Y	Tbx18	**unchanged**	3.3	9	42.2	0.37	0.08	0.21
1E05	chr1:191691206	Y	Rcor3	**unchanged**	62.1	386	239.9	0.16	0.26	1.61
3B16	chr6:8736466	Y	Ica1	**unchanged**	107.4	93.4	565.3	1.15	0.19	0.17
3C10	chr2:84558866	Y	Zdhhc5	**unchanged**	9.7	38.4	56.5	0.25	0.17	0.68
1F14	chr8:12371081	Y	Sox1	**unchanged**	11.6	29.1	61.5	0.40	0.19	0.47
3C20	chr15:69012599	Y	Khdrbs3	**unchanged**	26.7	128.5	105.1	0.21	0.25	1.22
2F69	chr12:70016617	Y	Gpr135	**unchanged**	24.5	19.8	94.3	1.24	0.26	0.21
5F77	chr9:96005241	Y	Atp1b3	**unchanged**	1242	1058	4685	1.17	0.27	0.23
6C29	chr6:124991072	Y	Cops7a	**unchanged**	114.6	414.9	294	0.28	0.39	1.41
4D58	chr5:27103826	Y	Insig1	**unchanged**	2.3	2.1	7.3	1.10	0.32	0.29
5E05	chr19:11498970	Y	Stx3	**unchanged**	4	13.9	10.6	0.29	0.38	1.31
2F79	chr15:82438623	Y	Ndufa6	**unchanged**	64.9	151.1	224.2	0.43	0.29	0.67
3E60	chr16:18170697	Y	Gp1bb	**unchanged**	4.5	15.5	8.5	0.29	0.53	1.82
2C31	chr3:57976910	Y	Pfn2	**unchanged**	222.7	740.3	658.8	0.30	0.34	1.12
3E75	chr14:27151902	Y	Cacna2d3	**unchanged**	4.1	3.7	11.8	1.11	0.35	0.31
5E61	chr11:120052116	Y	Baiap2	**unchanged**	180.2	405.5	574.5	0.44	0.31	0.71
3C27	chr2:74499528	Y	Evx2	**unchanged**	7.8	17.2	24.2	0.45	0.32	0.71
5G45	chr11:106721954	Y	Tex2	**unchanged**	103.2	155.6	305.6	0.66	0.34	0.51
6D30	chr7:118894584	Y	Rbbp6	**unchanged**	405.7	269	780.4	1.51	0.52	0.34
3F06	chr1:87054948	Y	Ecel1	**unchanged**	145.3	310.8	420.9	0.47	0.35	0.74
5F75	chr14:65536512	Y	Gfra2	**unchanged**	63.2	58	165.5	1.09	0.38	0.35
5F70	chr18:47746271	Y	Sema6a	**unchanged**	11.1	16	31.6	0.69	0.35	0.51
5H01	chr11:120782221	Y	Aspscr1	**unchanged**	1.2	3	1.8	0.40	0.67	1.67
4D65	chr5:120894922	Y	Brap	**unchanged**	274.1	584.8	675.4	0.47	0.41	0.87

RLGS Spot	NotI site position (mm7)	CpG island	Gene symbol	Methylation during spermatogenesis	Gene Expression					
					Type-A spermatogonia	Pachytene spermatocytes	Round spermatids	Type-A/ Pachytene	Type-A/ Round	Pachytene/ Round
5C25	chr16:84834216	Y	Atp5j/Gabpa	**unchanged**	397.1	663.6	976.3	0.60	0.41	0.68
3E34	chr6:99781479	Y	Gpr27	**unchanged**	197.1	335.6	475.8	0.59	0.41	0.71
4B29	chr10:79764334	Y	ORF61	**unchanged**	426.5	1017.1	712.6	0.42	0.60	1.43
4E27	chr9:68685762	Y	Rora	**unchanged**	7.4	16.6	16.2	0.45	0.46	1.02
2G93	chr15:85817232	N	Ppara	**unchanged**	30.4	52.9	65.2	0.57	0.47	0.81
2D48	chr4:6380904	Y	Nsmaf	**unchanged**	95.4	50.1	102.5	1.90	0.93	0.49
2C41	chr8:68982926	Y	Fkbp8	**unchanged**	328.4	655.7	669.3	0.50	0.49	0.98
5D38	chr7:141322497	Y	Fgf15	**unchanged**	4.7	3.8	7.7	1.24	0.61	0.49
2F90	chr4:11118407	Y	Ccne2	**unchanged**	21.6	42.9	28.6	0.50	0.76	1.50
1gx8	chr7:75333081	Y	Polg	**hypermethylated**	53.2	75.5	36.8	0.70	1.45	2.05
3bx5	chr13:109808317	Y	Ddx4	**unchanged**	609.5	1810.9	1656.3	0.34	0.37	1.09
2.0-1.5 -fold increase or decrease in gene expression between type-A spermatogonia, pachytene spermatocytes and round spermatids:										
3F56	chr12:47397383	Y	Foxg1	**unchanged**	2.3	3.7	1.9	0.62	1.21	1.95
3B29	chr11:108024329	N	Cacng5	**unchanged**	182.6	95.3	112.3	1.92	1.63	0.85
5C10	chr17:25808890	Y	Nudt3	**unchanged**	413.6	624	326.8	0.66	1.27	1.91
3D28	chr12:71498554	Y	Tmem30b	**unchanged**	9.5	8.4	5.1	1.13	1.86	1.65
5C17	chr12:52831045	Y	Cfl2	**unchanged**	69.9	38.4	51.3	1.82	1.36	0.75
2B45	chr16:17638759	Y	Prodh	**unchanged**	44.5	27.2	24.7	1.64	1.80	1.10
1G37	chr13:101060069	Y	Erbb2ip	**unchanged**	115.3	64	112.3	1.80	1.03	0.57
2B40	chr11:119150635	Y	Cbx8	**unchanged**	5.9	3.3	5.8	1.79	1.02	0.57
2C19	chr6:85541630	Y	Egr4	**unchanged**	373.2	394.5	224.7	0.95	1.66	1.76
4E24	chr11:70296139	Y	Slc16a11	**unchanged**	92.1	52.7	60.5	1.75	1.52	0.87
2B23	chr11:43394228	N	D11Ertd730e	**unchanged**	120.1	70.4	126.4	1.71	0.95	0.56
6G25	chr16:20282499	Y	Chrd	**unchanged**	9.7	12.5	7.4	0.78	1.31	1.69
6C22	chr13_ran:397372	Y	Zfp131	**unchanged**	151.9	162.8	96.6	0.93	1.57	1.69
6C25	chr10:27855771	Y	Ptprk	**unchanged**	98.4	73.5	59.8	1.34	1.65	1.23
4D45	chr2:71372755	N	Dlx1	**unchanged**	14.4	20.4	12.4	0.71	1.16	1.65
5A03	chr7:32007761	Y	Cebpa	**unchanged**	182.2	111.9	120	1.63	1.52	0.93
4D36	chr18:36663752	Y	Pura	**unchanged**	211.3	141.3	130.7	1.50	1.62	1.08
4G27	chr11:98131652	Y	Cacnb1	**unchanged**	520.4	324	463.2	1.61	1.12	0.70
2C09	chr4:99049166	Y	Foxd3	**unchanged**	5.4	3.4	6.1	1.59	0.89	0.56
2E30	chr19:45557344	Y	Fbxw4	**unchanged**	64.9	77.5	48.8	0.84	1.33	1.59
7D10	chr1:189454671	Y	Smyd2	**unchanged**	19.4	20.5	13.2	0.95	1.47	1.55
1D01	chr13:33353590	N	Serpinb6a	**unchanged**	426.9	815.3	705.4	0.52	0.61	1.16

RLGS Spot	NotI site position (mm7)	CpG island	Gene symbol	Methylation during spermatogenesis	Gene Expression					
					Type-A spermatogonia	Pachytene spermatocytes	Round spermatids	Type-A/ Pachytene	Type-A/ Round	Pachytene/ Round
3D63	chr12:99964454	Y	Golga5	**unchanged**	158.9	118.8	221.1	1.34	0.72	0.54
4E72	chr11:11638307	Y	Zfpn1a1	**unchanged**	13.7	20.8	24.9	0.66	0.55	0.84
2C29	chr13:40659482	Y	Elovl2	**unchanged**	33.9	59.2	51.6	0.57	0.66	1.15
3E55	chr1:37782382	Y	Mgat4a	**unchanged**	71.9	67.1	115.7	1.07	0.62	0.58
5A19	chr13:57902535	Y	Ntrk2	**unchanged**	3.7	6.3	5.3	0.59	0.70	1.19
7F24	chr19:15834161	Y	Gnaq	**unchanged**	30.2	50.1	51.4	0.60	0.59	0.97
4D27	chr6:52127688	Y	Hoxa2	**unchanged**	34.8	59	43.4	0.59	0.80	1.36
4D54	chr10:107795510	Y	Pawr	**unchanged**	92.7	81.5	137.9	1.14	0.67	0.59
4G04	chr2:76250268	Y	Osbpl6	**unchanged**	5.8	9.8	8.9	0.59	0.65	1.10
5E17	chr19:10069558	Y	Syt7	**unchanged**	14.7	22.8	24.8	0.64	0.59	0.92
2B07	chr3:17780620	Y	Bhlhb5	**unchanged**	26.3	24.9	41.7	1.06	0.63	0.60
3F18	chr19:3895560	Y	Nudt8	**unchanged**	12.1	15	20.2	0.81	0.60	0.74
5D52	chr5:71955539	Y	Zar1	**unchanged**	177.5	203	291.7	0.87	0.61	0.70
4G102	chr10:43391874	Y	Cd24a	**unchanged**	110.9	80.9	132.4	1.37	0.84	0.61
2B17	chr8:90951656	Y	Irx5	**unchanged**	49	37.3	60.9	1.31	0.80	0.61
6G41	chr10:80545053	Y	Dot1l	**unchanged**	474	599.7	758.1	0.79	0.63	0.79
5F52	chr1:132036873	Y	Pctk3	**unchanged**	59.6	83.4	89	0.71	0.67	0.94
<1.5 -fold increase or decrease in gene expression between type-A spermatogonia, pachytene spermatocytes and round spermatids:										
3C21	chr19:6323991	Y	Nrxn2	**unchanged**	8.4	9.1	11.5	0.92	0.73	0.79
2D25	chr19:43334330	Y	Cnnm1	**unchanged**	57.9	78.4	62.5	0.74	0.93	1.25
2F28	chr4:109053461	Y	Faf1	**unchanged**	64.6	81	80.5	0.80	0.80	1.01
1G51	chr9:49773820	Y	Ncam1	**unchanged**	55.7	69.2	64	0.80	0.87	1.08
1F15	chr13:24635839	Y	Vmp	**unchanged**	18	16.4	20.3	1.10	0.89	0.81
3A01	chr10:128147538	N	Zfpn1a4	**unchanged**	62	73.2	76.7	0.85	0.81	0.95
3F09	chr3:45324659	Y	Pcdh10	**unchanged**	15	13.6	16.7	1.10	0.90	0.81
1E30	chr15:102996927	Y	Hoxc5	**unchanged**	22.7	26	20.6	0.87	1.10	1.26
2F37	chrX:56286984	Y	Sox3	**unchanged**	16.6	13.1	14.5	1.27	1.14	0.90
2F53	chr16:34831462	Y	Adcy5	**unchanged**	16.1	17.6	13.3	0.91	1.21	1.32
3H19	chr13:23104812	Y	Hist1h1c	**unchanged**	13.5	14.3	13.6	0.94	0.99	1.05
5G43	chr3:104581892	Y	Rhoc	**unchanged**	145.1	101.3	103.8	1.43	1.40	0.98
1F02	chr2:74512389	Y	Hoxd13	**unchanged**	4.5	4.5	4.1	1.00	1.10	1.10
2D41	chr17:24322011	Y	Axin1	**unchanged**	256.4	253.5	189.4	1.01	1.35	1.34
6G67	chr7:60359889	Y	Apba2	**unchanged**	39.8	32.4	30.9	1.23	1.29	1.05
3E07	chr11:22924443	Y	U2af1-rs1	**unchanged**	105.9	97	89.8	1.09	1.18	1.08

Table 4.4: qAMP primer sequences for the analysis of chromosome 7

Position (mm7)	Forward Primer (5'-3')	Reverse Primer (5'-3')
chr7:4108913	CCAGACAGGATTTTGGTGTG	CTGCAGACATTCACCCTCCT
chr7:6746813	CCCTGGGATCTGGAATCTAA	GTCGGTCTCAGTGGCAATTT
chr7:14306726	GCCAGGGAGAACATCTGACA	GGGTCAGCCTCTGTGCTAAT
chr7:21977178	ATCACAGTCAACGCTCTGGA	GAAGCCAGGTAGGGTTGAAA
chr7:24130951	CTGCACTGAAGCCCACATAC	GGGTGAGGTCTTGCTGTGTT
chr7:28023804	TCTTCAAACCCACTGGCACT	CAGCGACAGACAGATCTCCA
chr7:34071401	CCGCTCTTTAGTGTGCCTCT	CTTCATCAAGTGGCTGCTGA
chr7:40447675	CCATTTTCAAAGCCTGCAT	AGCAAGTTGTCCTCCAGCA
chr7:46339059	CCTCCTCCTTGGACACTTCT	AAGACACTGGGTAATCCGTCA
chr7:50861196	GCGTCCAACAATATACCATTGA	GGATAAGCATCCAGGGCTAA
chr7:56864697	GTGCACACACACAACCTTGG	AGTACAGGGGCAGATGATGG
chr7:61063773	GGGTCCAAGTGGAACAGCTA	ACTGATGTCATTGCCCTGCT
chr7:66063823	GAGTGGTGGGGATCCAAATA	AGTTCCATTTCCCACACCAG
chr7:71097048	GAAATCTCCCAAGGCCTGAT	CCCATTTTGGGTTTCTTGAG
chr7:76285711	TCAGGTCCAGGTCTGAGGAT	GTGGGGCTGTCACTTGAGTT
chr7:83048811	TCAGAAATGGCAATCACCAA	ACATTGACAGCGAGAACAGATT
chr7:88111373	AATCTTGAGTGCCAGGCTGA	CCCCTCCCCACATATTCCTA
chr7:93098898	CTGTCACAGATCAGCAAGCA	AACGAGCAGGCAGAGCTATT
chr7:97669003	CTCCGAGCCCTAGTCTCTGT	AGCGGAACACCCATTAACAT
chr7:102854935	CTTTATGCCTACTCCATTTTGC	TTGCTTTCCTCTTGGATGTCT
chr7:108398636	CCGGGCAATAATTACAGCAT	TCAGGTCTTCTCTCCCCTTG
chr7:112379886	CACAGAGCTAGAGCTGTCTCAAA	AACATTCCAGCATTCCAAGC
chr7:117545186	GGCAATATTGATGGCTGTGA	AGTGGGAAATGCAGGCTTCT
chr7:124654460	TGAAGGTCTCAGCTGTGGAG	CAGTGGTGCCAACCTGAGT
chr7:128659310	GGTGATCAAGCTACAGACACGA	AGTGATAAGCGACCACAGCA
chr7:133765185	AGCAAACCTTGATGGGACAT	CATTGCCAGTGTGCTTGAAA
chr7:138855398	CCTCTGTGAGAAAGAGGGTCA	AAGGCTGAACCAGCCAGACT

CONNECTING TEXT

The studies described in chapter IV determine that patterns of DNA methylation are being acquired during spermatogenesis. Previous studies performed in our laboratory have demonstrated that the administration of the hypomethylating drug, 5-aza-2'deoxycytidine (5-azaCdR), to adult mice and rats causes a disruption of spermatogenesis. 5-azaCdR inhibits DNA methyltransferase activity in a dose-dependent manner. As animal models of DNA hypomethylation caused by genetic ablation of DNA methyltransferases result in a cessation of sperm production, 5-azaCdR becomes a very useful tool to partially perturb DNA methyltransferase activity and determine the resulting sperm-specific effects. Previously determined effects of the drug include decreased sperm production, abnormalities in the development of embryos sired by treated animals and, at higher doses, a reduction in global levels of sperm DNA methylation. In chapter V, we describe the effects of perturbing DNA methyltransferase activity to the health of germ cells in greater detail and determine the genome-wide status of DNA methylation in sperm from 5-azaCdR treated animals.

CHAPTER V

Adverse effects of 5-aza-2'-deoxycytidine on spermatogenesis include reduced sperm function and selective inhibition of *de novo* DNA methylation

Christopher C. Oakes, Tamara L.J. Kelly, Bernard Robaire & Jacquetta M. Trasler

Manuscript submitted to The Journal of Pharmacology and Experimental Therapeutics

ABSTRACT

The anti-cancer agent, 5-aza-2'-deoxycytidine (5-azaCdR, Decitabine), causes DNA hypomethylation and a robust, dose-dependent disruption of spermatogenesis. Previously we have shown that altered testicular histology and reduced sperm production in 5-azaCdR-treated animals is associated with decreased global sperm DNA methylation and an increase in infertility and/or a decreased ability to support preimplantation embryonic development. The goal of this study was to determine potential contributors to 5-azaCdR-mediated infertility including alterations in sperm motility, fertilization ability, early embryo development and sequence-specific DNA methylation. We find that although sperm from 5-azaCdR-treated animals displayed decreased motility, altered morphology and produced embryos that were less likely to survive to the blastocyst stage, the major contributor to infertility was a marked (56-70%) decrease in fertilization ability. Sperm DNA methylation was investigated using Southern blot, restriction landmark genomic scanning (RLGS), and quantitative analysis of DNA methylation by real-time PCR (qAMP). Interestingly, hypomethylation was restricted to genomic loci that have been previously determined to acquire methylation during spermatogenesis, demonstrating that 5-azaCdR selectively inhibits *de novo* methylation activity. Similar to previous studies, we show that mice that are heterozygous for a non-functional *Dnmt1* gene are partially protected against the deleterious effects of 5-azaCdR; however, methylation levels are not restored in these mice, suggesting that adverse effects are due to other mechanism(s) in addition to DNA hypomethylation. These results demonstrate that clinically relevant doses of 5-azaCdR specifically impair *de novo* methylation activity in male germ cells; however, genotype-specific differences in drug responses suggest that adverse reproductive outcomes are mainly mediated by the cytotoxic properties of the drug.

INTRODUCTION

DNA methylation is an essential modification of mammalian DNA and occurs at 60 to 80 percent of CpG dinucleotides in the genome. DNA methylation is known to play important roles in several cellular processes and is often associated with transcriptional repression and increased genomic stability (Bird, 2002). In cancer, deregulation of DNA methylation is often observed (Das and Singal, 2004; Egger et al., 2004). DNA methylation is catalyzed by a family of DNA methyltransferase enzymes (DNMTs). The principal enzyme, DNMT1, primarily maintains established methylation patters during DNA replication; whereas, patterns are established by the *de novo* methyltransferases DNMT3a and DNMT3b (Li et al., 1992; Okano et al., 1999).

5-AzaCdR is currently used clinically as an anti-cancer agent for the treatment of myelodysplastic syndromes and other types of cancer due to its ability to demethylate tumour-suppressor genes and cause replication-dependent cytotoxicity. Following incorporation into replicating DNA, DNMTs become irreversibly bound to 5-azaCdR as a covalent adduct (Gabbara and Bhagwat, 1995; Santi et al., 1984; Taylor and Jones, 1982). Hypomethylation occurs during subsequent rounds of DNA replication because the depleted cellular pool of DNMTs is insufficient to maintain established genomic methylation patterns. Adducts are cytotoxic and induce apoptosis (Juttermann et al., 1994) in a p53-dependent manner (Schneider-Stock et al., 2005). Notably, decreases in methylation can occur at non-cytotoxic concentrations of 5-azaCdR that do not significantly impair DNA synthesis (Glazer and Knode, 1984; Haaf, 1995; Mondal and Heidelberger, 1980). Use of 5-azaCdR inhibits all known methyltransferases (Weisenberger et al., 2004).

Mounting evidence points to an important role for DNA methylation in the process of male germ cell development. Within the male germ line, germ cell-specific methylation patterns are initiated before birth and are completed during spermatogenesis by the pachytene phase of meiosis I (Davis et al., 1999; La Salle and Trasler, 2006b; Oakes et al., 2007a). Methylation patterns in male germ cells are highly distinct from those found in somatic tissues (Eckhardt et al.,

2006; Oakes et al., 2007b). Expression of various DNMTs is highly regulated throughout spermatogenesis (Benoit and Trasler, 1994; Jue et al., 1995; La Salle and Trasler, 2006a; Trasler et al., 1992), and inactivation of the DNMTs through gene-targeting results in male infertility (Bourc'his et al., 2001; Kaneda et al., 2004). Due to the embryonic lethality and/or sterility in DNMT-deficient mice, the use of 5-azaCdR becomes a useful alternative to further investigate the role of DNA methylation in germ cells.

Previous studies on the effects of 5-azacytidine in rats (Doerksen et al., 2000; Doerksen and Trasler, 1996) and 5-azaCdR in mice (Kelly et al., 2003) have demonstrated that treatment results in a robust disruption of spermatogenesis. Spermatogenesis is distinctively sensitive to the effects of the drug as these adverse reproductive effects occur at doses where body weights and haematological parameters remain unaffected (Kelly et al., 2003). Treatments used in these studies were of sufficient duration to expose developing germ cells from the spermatogonial stem cell stage through spermatogenesis and epididymal transit. 5-AzaCdR treatment resulted in several reproducible effects. Treatment caused a dose-dependent decrease in testis weights, lowered sperm counts and an increased level of abnormalities in testicular histology. Interestingly, mating of treated mice with control females resulted in increased preimplantation loss (reduction of the number of implantation sites minus the number of oocytes ovulated). An increase in preimplantation loss could be the result of either a failure of sperm from 5-azaCdR-treated animals to fertilize oocytes or a reduction in the survival of embryos during preimplantation development. Furthermore, global DNA methylation analysis reveals a dose-dependent decrease in DNA methylation in sperm from treated animals.

In this study, we determine the cause of the 5-azaCdR-dependent preimplantation loss by performing a detailed analysis of sperm function and preimplantation development. The relationship of these effects to drug-dependent alterations in sperm DNA methylation is determined by using a variety of techniques that assess sequence-specific levels of DNA methylation.

Previous studies have indicated that the treatment of mice that have reduced levels of DNMT1 results in a greater level of DNA hypomethylation and less cellular toxicity compared to wild-type mice in somatic tissues (Juttermann et al., 1994; Laird et al., 1995). We have also previously found that in *Dnmt1*$^{c/+}$ mice, animals that are heterozygous for a targeted mutation in the catalytic domain of DNMT1 (Lei et al., 1996), some of the adverse spermatogenic effects of 5-azaCdR are attenuated in comparison to *Dnmt1*$^{+/+}$ males (Kelly et al., 2003). Genotype-dependent responses in the extent of the hypomethylation in sperm DNA and the level of adverse spermatogenic effects between *Dnmt1*$^{+/+}$ and *Dnmt1*$^{c/+}$ animals allude to the roles played by cytotoxic adducts versus abnormal DNA methylation. Thus a further aim of the current study was to examine the effects of 5-azaCdR in *Dnmt1*$^{+/+}$ and *Dnmt1*$^{c/+}$ mice.

MATERIALS AND METHODS

Animals

Male *Dnmt1*$^{+/+}$ and *Dnmt1*$^{c/+}$ mice, heterozygous for a deletion within the catalytic domain of the primary mammalian DNA methyltransferase, DNMT1 (Lei et al., 1996), were bred and raised in our own facilities (McGill University – Montreal Children's Hospital Research Institute) on a C57BL/6 background. Male mice of both genotypes were obtained through crosses of *Dnmt1*$^{c/+}$ males and C57BL/6 females; PCR genotyping of mice was performed as described (Kelly et al., 2003). Adult virgin C57BL/6 and CD1 females were obtained from Charles River, Canada (St. Constant, Canada). All mice were maintained on a 12:12 hour light/dark cycle and were provided with food and water *ad libutum*. Animal experiments were carried out according to the principles and procedures detailed in the Guide to the Care and Use of Experimental Animals, by the Canadian Council on Animal Care.

Treatment

Dnmt1$^{+/+}$ and *Dnmt1*$^{c/+}$ males (age 7 to 10 weeks) were randomly assigned to one of two treatment groups (Saline: *Dnmt1*$^{+/+}$, n = 13; *Dnmt1*$^{c/+}$, n = 15. 5-AzaCdR: *Dnmt1*$^{+/+}$, n = 14; *Dnmt1*$^{c/+}$, n = 15). Males were treated 3 times a week for 7 weeks, by intraperitoneal injection (IP), with either saline or 0.1 mg/kg 5-azaCdR to expose male germ cells throughout their development. Throughout the treatment, males were weighed twice per week. After 7 weeks of treatment, males were mated with four virgin superovulated CD1 females (age 8 weeks) and then sacrificed. The testes, epididymides, seminal vesicles and spleen were removed, weighed, snap-frozen and stored at -80°C. A section of liver was also removed and frozen. Spermatozoa from the caudal epididymides were isolated and purified as described (Alcivar et al., 1989) with modifications (Kelly et al., 2003), and were stored at -80°C.

Mating and Embryo Culture

Adult female CD1 mice aged 8 weeks were superovulated by administration of 5 IU of pregnant mare serum gonadotropin (PMSG; Sigma) followed by 5 IU of human chorionic gonadotropin (hCG; Sigma) 48 hours later. To obtain fertilized embryos, each male (n = 7-9/treatment group), after 7 weeks of treatment, was mated, overnight, with 1 superovulated virgin CD1 female per night for 4 nights, for a total of 4 females per male. The next morning females were examined for the presence of a vaginal plug. One-cell embryos and unfertilized oocytes were isolated at 27 hours post-hCG and cumulus cells removed by hyaluronidase treatment (1mg/ml) (Sigma) in HEPES-buffered M2 medium (Sigma). Oocytes were examined for the presence of two pronuclei indicating that fertilization had taken place. Oocytes were classified as fertilized, unfertilized or fragmented; the majority of fragmented oocytes could not be evaluated as to fertilization status, and thus, were not subcategorized. Fertilized embryos were washed three times, using a mouth-controlled drawn-out glass pipette, and placed into pre-equilibrated bicarbonate-buffered kSOM medium (Erbach et al., 1994) with gentomycin, under oil, and cultured under an

atmosphere of 5%O_2, 5%CO_2, in nitrogen at 37°C in a humidified modular incubator (Billups-Rothenberg, Del Mar, CA). Embryos were examined daily, on a heated stage, and scored for development through to the blastocyst stage. Data are presented on a per male basis; to avoid skewing of data males were removed from all data sets if less than 10 eggs, in total, were collected from females mated to that male; only two males were removed, one saline *Dnmt1*$^{+/+}$ male and one 5-azaCdR-treated *Dnmt1*$^{c/+}$ male.

Sperm Motility Analysis

Sperm motility of treated *Dnmt1*$^{+/+}$ and *Dnmt1*$^{c/+}$ male mice (n = 6/group) was analysed using an IVOS semen analyser (Hamilton-Thorne Research, Beverly, MA) with parameters determined by the Jackson Laboratory (courtesy of Hamilton-Thorne). All dishes and slides were kept at 37°C during all steps. Briefly, the cauda epididymidis was tied off, both proximally and distally, removed from the epididymis and rinsed in 3 ml of warmed M199 medium with Hank's salts (Sigma, St. Louis, MO) supplemented with 0.5% w/v BSA, pH 7.4 (GIBCO, Mississauga, ON) in a 35 mm Petri dish at 37°C. The cauda epididymidis was then moved to a new Petri dish containing 3 ml of warm supplemented M199 medium, minced and the epididymal tissue removed. Sperm were allowed to disperse for 5 minutes. The sperm suspension was diluted 1:10 in warm supplemented medium prior to motility analyses such that concentration did not impair motility. The diluted suspension (20µl) was loaded into a pre-warmed 2X-CEL Sperm Analysis Chamber (80µm deep) (Hamilton-Thorne Research). Movement characteristics analysed were: percent motility – motile sperm divided by the sum of the motile and immotile sperm within the analysis field; percent progressive motility – progressively motile sperm divided by the sum of motile and immotile sperm within the field; average path velocity (VAP) – the average velocity of the smoothed cell path; progressive/straight line velocity (VSL) – the average velocity measured in a straight line from the beginning to end of the track; curvilinear velocity (VCL) – the sum of the distances moved in each frame along the sampled path divided by the time taken to cover the track; amplitude of

lateral head displacement (ALH); beat cross frequency (BCF) – the frequency with which the sperm track crosses the sperm path; straightness (STR) – the departure of the cell path from a straight line and; linearity (LIN) – the departure of the cell track from a straight line.

Tracks were digitally recorded at 60Hz under 4X dark-field illumination. Analysis was completed using the following IVOS settings: stage temperature, 37°C; frames acquired, 30; frame rate, 60 Hz; minimum contrast, 30; minimum cell size, 4 pixels; magnification, 0.81; cell intensity, 75; static size, 0.13-2.43; static intensity, 0.10-1.52. Five slides were analysed for each mouse and each slide was sampled five times such that a minimum of 300 sperm were analysed per slide. The mean of the five slides was calculated for each mouse.

DNA Methylation Analysis

DNA was extracted from purified spermatozoa using proteinase K followed by phenol extraction for Southern blot and RLGS analysis as described previously (Okazaki et al., 1995). Southern blots were done as described previously (Trasler et al., 1990) and visualized by autoradiography. Major and minor satellite probes were constructed by PCR amplification of mouse genomic DNA using primers described previously (Lehnertz et al., 2003). The probe for the intracisternal A-particle (IAP) has been used previously (Michaud et al., 1994; Walsh et al., 1998). Three to four replicate RLGS profiles were generated for each treatment group. Each profile was visualized by autoradiography for identification of changed spots and phophorimager screen for spot densitometry analysis. Visual assessment of changes in spot intensity was confirmed by densitometric analysis by comparing the intensity of the spot of interest in comparison to the intensity of 8-10 surrounding spots of unchanged intensity. All spots showing differential intensity were observed to be changed in all replicates except one spot in a *Dnmt1*$^{c/+}$ saline-treated profile. The genomic location of spots of interest was determined by using either the mouse RLGS cloning library method (Yu et al., 2004) or a second-generation virtual RLGS resource (Smiraglia et al., 2007). Loci identified using virtual RLGS were confirmed by

obtaining the corresponding BAC clone (Roswell Park Microarray Core Facility, Buffalo, NY) and running RLGS mixing gels. The methylation status of paternally methylated imprinted DMRs and RLGS spots was determined by using the qAMP methylation assay (Oakes et al., 2006). Briefly, DNA is digested with various methylation-sensitive restriction enzymes (MSREs) and a methylation-dependent restriction enzyme, McrBC, followed by amplification using real-time PCR. Shifts in Ct value (ΔCt) between the sham- and enzyme-digested samples are used to calculate the percentage of methylation at the various CpG sites within the amplified region (MSREs: %methylation=$100(2^{-\Delta Ct})$; McrBC: %methylation=$100(1-2^{-\Delta Ct})$. All ΔCt values are the means of triplicate reactions. Primers are designed to flank CpG/restriction sites of interest. Primers used were described previously (Oakes et al., 2007a).

Statistical Analysis

Statistical analysis was performed using SigmaStat 2.03 software (SPSS, Chicago, IL). Significant differences ($p<0.05$) between treatment groups with respect to the various motility parameters, morphological characteristics, fertilization ability, and percent methylation were detected using two-way ANOVA, with a post-hoc Tukey test. Embryo data are expressed on a per male basis and were evaluated for significance using three-way ANOVA, with a post-hoc Tukey test.

RESULTS

General Results

Mice were treated with 5-azaCdR for 7 weeks, a sufficient period for the exposure of developing germ cells throughout the entire window of spermatogenesis (from spermatogonial stem cell to spermatozoa) and epididymal transit. As observed in a previous study (Kelly et al., 2003), treatment with 5-azaCdR elicited no obvious changes in behaviour and weight, although

initial and final body weights of the *Dnmt1*$^{+/+}$ and *Dnmt1*$^{c/+}$ genotypes were significantly different ($p<0.05$) (Table 5.1). Again, testis weights were significantly decreased in treated males, regardless of genotype ($p<0.001$), and similar to our previous studies, the extent of reduction was considerably less in *Dnmt1*$^{c/+}$ males than in *Dnmt1*$^{+/+}$ males ($p<0.05$). However, testis weight decline was greater for both genotypes than in our previous study (Kelly et al., 2003).

Sperm Motility, Morphology and Fertilization Ability

To assess the effects of 5-azaCdR-treatment on sperm movement, multiple motility characteristics were assayed by computer assisted sperm analysis (CASA) using a Hamilton-Thorne IVOS semen analyser. Both the proportion of sperm found to be motile (sperm motility) and the fraction that were progressively motile were significantly lowered in sperm from 5-azaCdR-treated *Dnmt1*$^{+/+}$ males, but the extent of the decrease was less in sperm from *Dnmt1*$^{c/+}$ males (Figure 5.1a; Table 5.2). The curvilinear velocity (VCL), a measure of total movement of sperm, was reduced in both treated groups. While several other parameters concerning various aspects of sperm motility were significantly reduced in 5-azaCdR-treated *Dnmt1*$^{+/+}$ males, significant reductions were observed in only two parameters (VCL and ALH) in *Dnmt1*$^{c/+}$ mice. CASA analysis also allows for a measurement of the morphology the analyzed sperm. The average sperm head size (area) from 5-azaCdR-treated *Dnmt1*$^{+/+}$ males was smaller and be abnormally shaped (i.e. elongated), indicated by a change in the ratio of the minor axis to major axis (Figure 5.1b). These abnormalities in morphology were found only in sperm from *Dnmt1*$^{+/+}$ males.

To determine if the ability of sperm from 5-azaCdR-treated males to fertilize oocytes is reduced, each male was mated to four superovulated females (total matings = 132), and embryos were collected at the one-cell stage. The presence of a vaginal plug on the morning after mating indicated successful copulation, and the copulation rate was similar for all treatment groups; only females with vaginal plugs were used for embryo collection. Approximately 500 oocytes on average were scored per treatment group. An oocyte was

considered fertilized if two pronuclei were present. Whereas the incidence of fragmented oocytes/embryos was low (<5%) and was similar in all groups, the proportion of fertilized oocytes was dramatically reduced by 70% and 56% ($p<0.001$) after matings with treated *Dnmt1*$^{+/+}$ and *Dnmt1*$^{c/+}$ males, respectively (Figure 5.1c). These results show that although a significant proportion of sperm remain motile after treatment, the majority of these sperm are unable to successfully fertilize oocytes.

In Vitro Embryonic Development

To assess the ability of embryos sired by treated males to progress normally through pre-implantation development, fertilized oocytes were placed into culture and scored daily for survival to advanced preimplantation stages. Embryo viability was calculated as the percent of embryos that survived from the previous stage. As Figure 5.2 illustrates, there was no change in the progression of embryos throughout preimplantation development, with the exception of survival to the blastocyst stage. At this point, approximately 50% of embryos from saline-treated groups survived; however, the proportion of surviving embryos sired by treated *Dnmt1*$^{+/+}$ males was significantly reduced by an additional 25% ($p<0.05$). No such decrease was observed for embryos sired by treated *Dnmt1*$^{c/+}$ males. These results demonstrate that the ability of the paternal genome from 5-azaCdR-treated *Dnmt1*$^{+/+}$ animals to support normal embryonic growth is reduced; in contrast, blastocyst development was similar to saline-treated controls for the *Dnmt1*$^{c/+}$ mice.

DNA Methylation Analysis of Repetitive Elements

Previous analysis of sperm from 5-azaCdR-treated animals has demonstrated a dose-dependent reduction in global levels of DNA methylation in rats (Doerksen et al., 2000) and mice (Kelly et al., 2003). In mice, only a dose higher (0.15 mg/kg, 3x/wk IP) than the lower dose (0.1 mg/kg, 3x/wk IP) used in this study resulted in a significant reduction in global sperm DNA methylation, despite adverse spermatogenic effects observed at the lower dose in both

previous and current work. The global assessment of DNA methylation may not be sensitive enough to detect sequence-specific changes in DNA methylation at the lower dose. To address the possibility that aberrant DNA methylation of sperm from treated males is associated with the 5-azaCdR-dependent effects, we investigated methylation status of three types of repetitive elements in sperm from 5-azaCdR-treated males using Southern blotting. Although the methylation status of the major and minor satellite repeats (structural elements mainly found in centromeric regions), and the interspersed LTR-containing retroviral element, IAP, have different average levels of methylation in sperm, no difference was detected in sperm from 5-azaCdR-treated males (Figure 5.3).

DNA Methylation Analysis of Paternally-Methylated Imprinted Regions

Imprinted genes display an allele-specific pattern of DNA methylation that is acquired in a sex-specific manner in germ cells. Loci that display this property are termed differentially methylated regions (DMRs). There are three imprinted genes with well characterized DMRs that are methylated in sperm. We investigated if 5-azaCdR treatment could affect the ability to maintain these patterns. Using primers that target restriction sites within the DMRs of *H19-Igf2* (Tremblay et al., 1995), *Dlk1-Gtl2* (Takada et al., 2002), and *Rasgrf1* (Yoon et al., 2002), we used the qAMP assay to determine the percent methylation in these regions. No changes were observed for *H19-Igf2* or *Dlk1-Gtl2*; however, digestion of DNA with the HhaI restriction enzyme reveals a significant reduction in the percentage of methylation in sperm from 5-azaCdR-treated animals in both $Dnmt1^{+/+}$ and $Dnmt1^{c/+}$ groups (Figure 5.4). The methylation at these sites is also reduced in the $Dnmt1^{c/+}$ saline-treated group. Because the HhaI digest of the *Rasgrf1* DMR contains the most methylation-sensitive restriction enzyme sites of all the digests tested and only one of the three HhaI sites is required to be unmethylated for the enzyme to cleave the DNA strand, it is the most sensitive in detecting a reduction in methylation (Figure 5.4a). Interestingly, previous studies using the same approach have determined that the HhaI digest of *Rasgrf1* was the sole digest to reveal a substantial gain in methylation during

spermatogenesis compared to the other DMRs investigated (Oakes et al., 2007a).

Genome-Wide Analysis of Single-Copy Sequences using RLGS

RLGS investigates genome-wide patterns of DNA methylation by separating genomic DNA that has been digested with the methylation-sensitive restriction enzyme, NotI, by two-dimensional gel electrophoresis. Hypomethylated sites generate spots that are visible on RLGS profiles; a change in spot intensity is inversely proportional to the methylation status of individual loci in the genome. In the mouse, NotI sites occur in a variety of sequence types, including unique and interspersed repetitive sequences. Previously, RLGS analysis found that the methylation status of 19 spots is modified during spermatogenesis (Oakes et al., 2007a). To further examine the relationship between 5-azaCdR-dependent alterations in sperm DNA methylation and the acquisition of patterns of DNA methylation during spermatogenesis, RLGS profiles were generated from sperm from each treatment group (n=2-4/group). This analysis reveals that a subset of spots are consistently hypomethylated in all profiles from both *Dnmt1*$^{+/+}$ and *Dnmt1*$^{c/+}$ 5-azaCdR-treated groups (Figure 5.5; Table 5.3). A total of 9 spots were observed to change with treatment, and all were hypomethylated. All changed spots in these groups are hypomethylated in all profiles investigated. One spot, belonging to a NotI site upstream of the *AK032343* gene, was observed to be hypomethylated in 1/3 profiles generated from a *Dnmt1*$^{c/+}$ saline-treated animal. The vast majority (>99%) of the total number of spots (both hyper and hypomethylated) remain unaffected. Using a second-generation virtual RLGS resource (Smiraglia et al., 2007) to assess the amount of hypermethylated NotI sites visible on real RLGS profiles, reveals that greater than 1000 hypermethylated 'spots' are not hypomethylated with 5-azaCdR treatment. Included in these hypermethylated sites are ~240 and ~60 spots originating from the IAP and Etn (early transposon) interspersed repeats, respectively (Oakes et al., 2007b). Hypomethylation of these repeats is highly visible on RLGS profiles and was not observed in any profile from any group

(data not shown). This result is consistent with the Southern blot analysis of IAP (Figure 5.3). We also investigated if changes were specific to sperm or were also hypomethylated in somatic tissues from 5-azaCdR-treated animals. We found that all of the sites observed to be hypomethylated in sperm did not change in brain or liver (Table 5.3; Figure 5.6).

Most interestingly, of the 11 spots that in our previous study were shown to gain appreciable levels of methylation between the type A spermatogonial and sperm stages, 8 of them were consistently hypomethylated in sperm from 5-azaCdR-treated animals (Table 5.3). Only 1/9 5-azaCdR-dependent hypomethylated spots was not observed to acquire methylation during spermatogenesis. The eight spots that are normally hypomethylated during spermatogenesis were unaffected by treatment. Combined with the knowledge that the methylation status of >99% of spots are unchanged both during spermatogenesis and in sperm from 5-azaCdR-treated animals, our results indicate that maintenance methylation and demethylation processes proceed normally in the presence of the drug, despite a selective inhibition of *de novo* methylation activity in germ cells.

Analysis of 5-azaCdR-Responsive Single-Copy Sequences using qAMP

The genomic location of 5/9 RLGS spots that are hypomethylated in sperm from 5-azaCdR-treated animals were determined previously (Oakes et al., 2007a). Identified NotI sites that show altered methylation are found near or within genes and within non-repetitive sequences, but show varied locations within genes and variable CpG island status (Table 5.4). To further define the specificity and extent of the hypomethylation effect observed in sperm from 5-azaCdR-treated animals and the relationship to the inhibition of *de novo* methylation, we used the qAMP assay to quantitatively measure the percentage of methylation in sperm from 5-azaCdR-treated mice. The qAMP assay allows for the examination of DNA methylation at multiple neighbouring CpGs in addition to the NotI site. We chose to examine the methylation of three loci that were detected by RLGS to be hypomethylated at their respective NotI sites, *Tcf3*, *Abt1*

and *Ibtk* (Figure 5.7a). These three loci were chosen due to the varied location of the NotI sites within their respective gene (3', 5' and body regions, respectively). We found that all restriction enzyme digests detect a significant reduction in the percentage of methylation in response to treatment with 5-azaCdR (n=5-6 mice/group) (Figure 5.7b-d). This indicates that the change in methylation status observed by RLGS at NotI sites is representative of changes present in neighbouring CpG sites. The magnitude of the reduction in methylation is less in the methylation-dependent McrBC digests, due to this particular digest being less sensitive to demethylation (Oakes et al., 2006). Slightly less methylation is observed in *Dnmt1*$^{c/+}$ animals; however, the reduction is small relative to the effect of 5-azaCdR. In many cases, 5-azaCdR reduces the level of methylation in sperm from treated animals to levels that are similar or close to the levels previously found in type-A spermatogonia. Analysis of a fourth locus, a region shown to be hypomethylated during spermatogenesis located upstream of the *AK137601* gene, is largely unaffected by treatment, demonstrating that hypomethylation is observed to proceed normally in the presence of 5-azaCdR (Figure 5.7e). These data, taken together with the RLGS data, indicate again that *de novo* methylation activity is selectively inhibited, while maintenance activities and demethylation events that occur during spermatogenesis are unaffected.

DISCUSSION

In this study, we have determined several novel effects of 5-azaCdR treatment on male reproductive physiology and the epigenetic integrity of male germ cells. We find that although sperm from 5-azaCdR treated animals display somewhat decreased motility, altered morphology and decreased preimplantation embryonic development, embryo loss most likely results from a sharply decreased ability of sperm to complete fertilization. Interestingly, the investigation of sperm DNA methylation using a variety of quantitative techniques reveals that hypomethylation is restricted to genomic loci that have been

previously determined to acquire methylation during spermatogenesis, indicating that *de novo* methylation activity is selectively inhibited. We also show that *Dnmt1*$^{c/+}$ mice are partially protected from the adverse physiological effects of 5-azaCdR; yet, levels of DNA methylation in sperm are as low or lower than levels found in sperm from *Dnmt1*$^{+/+}$ mice.

Previous studies have shown that mating 5-azaCdR treated males with untreated females results in an increase in preimplantation embryo loss (Kelly et al., 2003). This could be the result of a failure of the sperm to fertilize or reduced survival of preimplantation embryos sired by 5-azaCdR-treated fathers. In this study, we have observed a greater than 50% reduction in the ability of the sperm of 5-azaCdR-treated males to successfully fertilize oocytes versus saline-treated males. The magnitude of the reduction in fertilization ability in the present study was similar to the level of preimplantation loss observed in previous studies; whereas, in contrast, in the current study, the decrease in the survival of embryos developing from the 2-cell to the blastocyst stage following treatment was minimal. Furthermore, the level of preimplantation loss and the reduction in the fertilization ability is similar in both *Dnmt1*$^{c/+}$ and *Dnmt1*$^{+/+}$ animals; whereas decreased survival of preimplantation embryos is specific to only *Dnmt1*$^{+/+}$ animals. These results suggest that the primary cause of the 5-azaCdR-dependent preimplantation loss noted in our previous study is mainly due to the inability of sperm to successfully fertilize the oocyte. The slight reduction observed in sperm number and motility is unlikely to be the primary reason for the failure of the sperm to fertilize, as the reduced parameters are within the range sufficient to maintain normal fertility. Changes in sperm head morphology are also unlikely to be the cause as measurable changes occur only in *Dnmt1*$^{+/+}$ animals. The primary cause for the failure to fertilize is most likely an additional parameter of sperm function not addressed in these studies, which may include a failure of the acrosome reaction, capacitation or sperm-egg recognition.

Changes to sperm head shape and area in wild-type animals signify a perturbation in chromatin packaging, indicating that 5-azaCdR may affect sperm chromatin organization or the mechanisms that direct morphological changes to

the nucleus during spermatogenesis. This may or may not be related to 5-azaCdR-dependent changes in DNA methylation, as very few loci throughout the genome were observed to be hypomethylated. Treatment with 5-azaCdR is also associated with a decreased capacity of the paternal genome to support embryonic growth several days post-fertilization in wild-type animals. These results support the idea that the sperm chromatin quality is affected as a result of 5-azaCdR treatment. Furthermore, previous work in combination with these studies demonstrates that, despite the problems associated with fertilization and embryonic development, embryos that are sired by 5-azaCdR-treated fathers can progress into later stages of development (Doerksen and Trasler, 1996). Other studies have shown that DNA methylation in the male pronucleus is erased shortly after fertilization (Oswald et al., 2000); however, the full extent of this reprogramming event is not known. Future studies will be required to determine if 5-azaCdR-dependent alterations in DNA methylation are present in developing embryos and if adverse effects are present in offspring. The fact that acquisition of DNA methylation in male germ cells occurs during adulthood, added with the knowledge that germ cell methylation can be influenced, raises the possibility that exogenous epigenetic insults could potentially be passed on to the progeny. Recent studies have shown that the treatment of animals with the antiandrogenic fungicide, vinclozolin, results in a trangenerational disease phenotype that is associated with the trangenerational transmission of epigenetic abnormalities (Anway and Skinner, 2006; Chang et al., 2006).

Investigation of the genome-wide methylation status of individual sequences in sperm DNA from treated animals reveals that male germ cells are able to maintain methylation patterns at repetitive elements and single-copy sequences in the presence of 5-azaCdR. This demonstrates that at the doses used here, not enough DNMT1 is covalently trapped to inhibit the maintenance of previously established patterns of methylation. This is supported by the finding that levels of DNMT1 are highest in the gonads of adult mice (Numata et al., 1994; Singer-Sam et al., 1990; Trasler et al., 1992). Furthermore, *Dnmt1*$^{c/+}$ germ cells are still able to maintain methylation patterns at these sequences in the

presence of 5-azaCdR despite having only half the normal amount of functional enzyme. Previously, we have established that in addition to maintaining established methylation patterns, developing germ cells possess *de novo* methylation activities (Oakes et al., 2007a). We find that the ability of germ cells to acquire methylation at CpG sites which normally gain methylation during spermatogenesis is selectively inhibited by 5-azaCdR treatment. Almost all CpG sites we have previously demonstrated to gain methylation during spermatogenesis are hypomethylated in sperm from 5-azaCdR treated animals.

Two mechanisms have been reported to mediate the effects of 5-azaCdR. Firstly, adducts can induce apoptosis via a p53-dependent mechanism. Secondly, covalent trapping of a sufficient amount of DNMT proteins results in replication-dependent hypomethylation in surviving cells. Figure 5.8 details how treatment with 5-azaCdR could potentially result in germ cell cytotoxicity and the selective inhibition of *de novo* methylation, while in the *Dnmt1*$^{c/+}$ background, some adverse effects are attenuated but *de novo* methylation activity is not restored. In spermatogonia, both maintenance and *de novo* methylation occur during mitotic divisions (Figure 5.8a). In the presence of 5-azaCdR, some cytosine residues are replaced creating potential sites for the formation of DNA-protein adducts (Figure 5.8b). At the dose used in this study, enough adducts are formed to result in the reduction of approximately half of the germ cell population (Kelly et al., 2003), while the other half survive but contain adducts in their DNA causing adverse effects to reproductive physiology. Due to the fact that the level of Dnmt1 expression is 5 to10-fold higher compared to Dnmt3a or Dnmt3b in spermatogonia (Shima et al., 2004) and 5-azaCdR has a more prominent effect on Dnmt3a or Dnmt3b compared to Dnmt1 (Oka et al., 2005), maintenance activity is maintained in the presence of 5-azaCdR while insufficient *de novo* methylation activity causes a loss of methylation at CpG sites that normally receive methylation during spermatogenesis. In *Dnmt1*$^{c/+}$ animals, 50% of DNMT proteins have a mutation in their catalytic domain which prevents the association with incorporated 5-azaCdR (Figure 5.8c). Due to the higher levels of Dnmt1 expression in germ cells, DNMT1 proteins are the major contributors to

adduct formation. Fewer adducts are formed in the *Dnmt1*$^{c/+}$ background, leading to less germ cell death and less adverse effects in surviving cells. Sufficient functional DNMT1 remains to maintain methylation patterns. Because *de novo* methylases are not mutated, their potential for adduct formation is not changed, causing the same level of sequestration and loss of *de novo* methylation activity.

To our knowledge, this is the first evidence of selective inhibition of *de novo* methylation activity by 5-azaCdR. The ability of 5-azaCdR at high doses to induce replication-dependent demethylation of hypermethylated loci *in vitro* is well documented. The examination of a developmental process where known hypermethylation events occur at distinct CpG sites in the genome has allowed for this novel property of the drug to be identified (Oakes et al., 2007a). Extrapolation of the number of CpG sites shown to be gaining methylation during spermatogenesis to the total number of CpG sites located in similar sequences ($12x10^6$ non-repetitive, non-CpG island CpGs; (Fazzari and Greally, 2004)) suggests the possibility of greater than one million CpGs potentially affected by 5-azaCdR treatment.

Is DNA hypomethylation the cause of the 5-azaCdR-dependent adverse effects on sperm function and embryonic development? Interestingly, some of the adverse effects of 5-azaCdR examined previously, such as testis weight, sperm counts and abnormal seminiferous tubule morphology, as well as effects examined in this study, such as sperm motility, morphology and the ability to support embryonic development are improved in *Dnmt1*$^{c/+}$ animals. Despite the protective effects of the *Dnmt1*$^{c/+}$ genotype, losses of DNA methylation are not attenuated in the germ cells of these animals; rather, DNA methylation is equally diminished or slightly lower at some CpG sites examined. This suggests that mechanisms, such as DNA-protein adduct formation, are more prominent in the mediation of these effects than DNA hypomethylation. However, some parameters, like fertilization ability, are equally effected by 5-azaCdR in *Dnmt1*$^{c/+}$ and *Dnmt1*$^{+/+}$ animals, indicating that hypomethylation may potentially have a role in mediating some of the effects.

ACKNOWLEDGEMENTS:

We would like to thank Tara Barton and Mary Gregory for their assistance with the CASA system and Keith Latham and Patricia Françon for their help with the details of embryo culture. Thank you to Stephanie Grénon for superb technical assistance. C.C.O. and T.L.J.K. are recipients of Canadian Institutes of Health Research (CIHR) Doctoral Awards and Montreal Children's Hospital Research Institute Studentships. B.R. and J.M.T. are James McGill Professors of McGill University. This work was supported by grants from the Canadian Institutes of Health Research (CIHR).

FIGURES

Figure 5.1: Effects of chronic 5-azaCdR treatment on sperm motility and velocity, sperm head morphology and fertilization ability. **(a,b)** Sperm motility and head morphology parameters were measured using computer assisted sperm analysis (CASA). **(c)** Fertilization ability was determined by isolating one-cell embryos and unfertilized oocytes after mating. Open bars represent *Dnmt1*$^{+/+}$ males, grey bars represent *Dnmt1*$^{c/+}$ males, and diagonal striped bars represent males treated with 5-azaCdR. Data are shown on a per male basis. Error bars represent ±SEM, asterisks indicate a significant difference between 5-azaCdR and saline treatment in genotype-matched groups, *p*-values for each parameter are indicated.

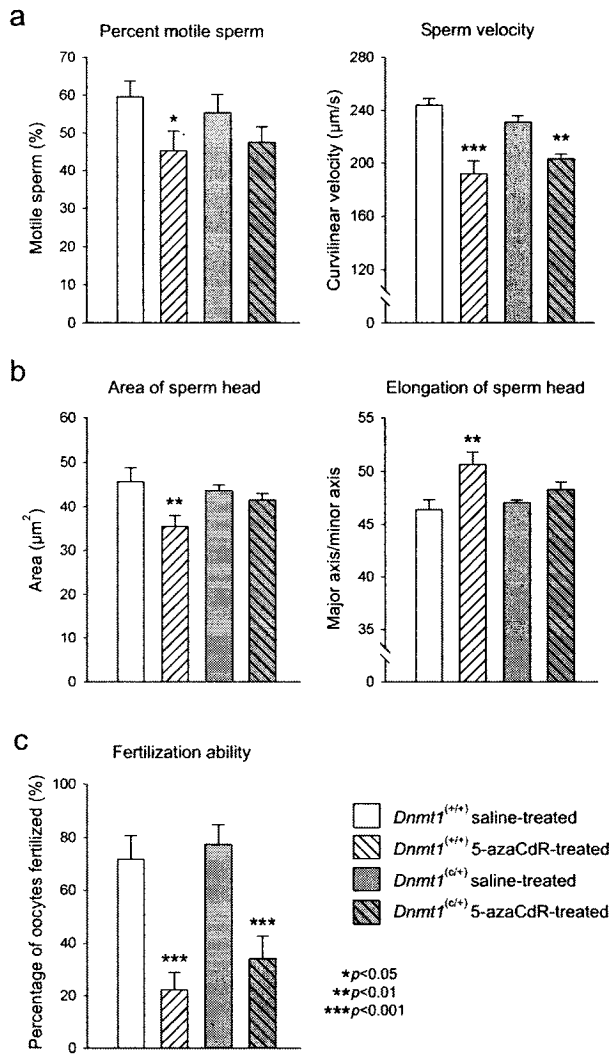

a
Percent motile sperm
Motile sperm (%)
70
60
50
40
30
20
10
0
*
Sperm velocity
Curvilinear velocity (µm/s)
280
240
200
160
120
0
***
**
b
Area of sperm head
Area (µm²)
60
50
40
30
20
10
0
**
Elongation of sperm head
Major axis/minor axis
55
50
45
40
35
0
**
c
Fertilization ability
Percentage of oocytes fertilized (%)
100
80
60
40
20
0
***
***
Dnmt1(+/+) saline-treated
Dnmt1(+/+) 5-azaCdR-treated
Dnmt1(c/+) saline-treated
Dnmt1(c/+) 5-azaCdR-treated
*p<0.05
**p<0.01
***p<0.001

Figure 5.2: Viability of embryos sired by 5-azaCdR-treated males through the stages of preimplantation development. Percentages represent the survival of embryos from the previous stage. Open bars represent *Dnmt1*$^{+/+}$ males, grey bars represent *Dnmt1*$^{c/+}$ males, and diagonal striped bars represent males treated with 5-azaCdR. Data are shown on a per male basis. Error bars represent ±SEM, asterisks indicate a significant difference between 5-azaCdR and saline treatment in genotype-matched control groups.

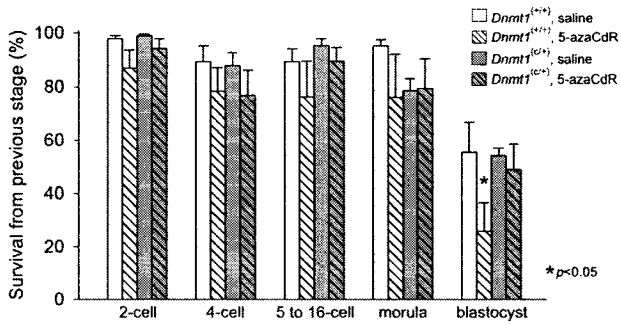


Figure 5.3: DNA methylation analysis of repetitive elements in sperm using Southern blot. Sperm DNA isolated from *Dnmt1*$^{+/+}$ and *Dnmt1*$^{c/+}$ males treated with either saline or 5-azaCdR was digested with either MspI to reveal the fully unmethylated pattern or HpaII to determine the amount of methylation in each sample. Equal amounts of DNA were loaded into each lane and hybridized to probes for the IAP interspersed repeat and the major and minor satellite centromeric repeat sequences. Although varying amounts of methylation can be found in these sequences in sperm, a measurable decrease in methylation is not induced by 5-azaCdR treatment.

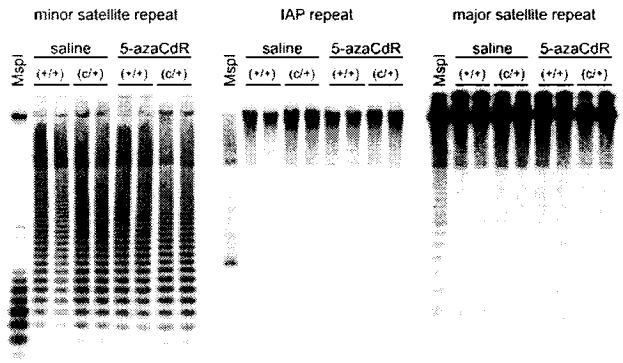


Figure 5.4: Quantitative DNA methylation analysis of paternally methylated imprinted DMRs in sperm using qAMP. Sperm DNA isolated from *Dnmt1*$^{+/+}$ and *Dnmt1*$^{c/+}$ males treated with either saline (sal) or 5-azaCdR (aza) was digested with either NotI (N), HhaI (Hh), HpaII (Hp), or McrBC (M) and amplified using real-time PCR. **(a)** Primer binding sites and the locations of flanked restriction sites are displayed for each DMR. **(b-d)** Percent methylation values at restriction sites in the DMRs of *H19*, *Dlk1-Gtl2*, and *Rasgrf1*. Only the HhaI digest of the *Rasgrf1* DMR reveals a decrease in the percentage of methylation, decreasing in response to both 5-azaCdR treatment (asterisks) and the *Dnmt1*$^{c/+}$ genotype (daggers). Error bars represent ±SEM, n=6-8 animals/group.

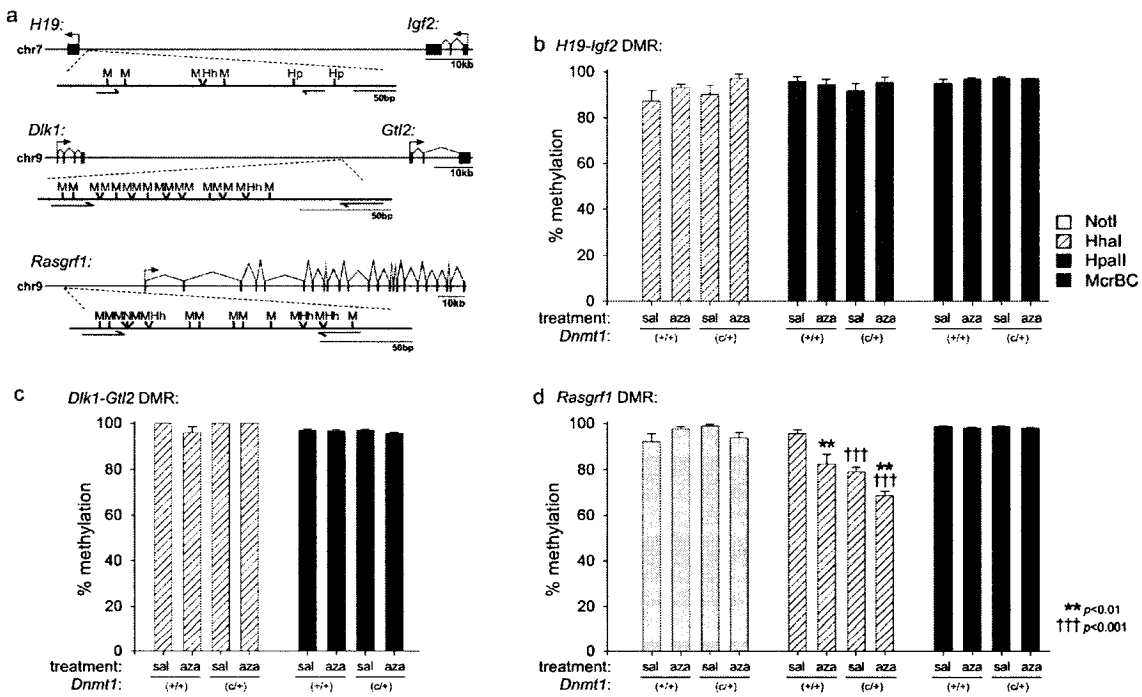


Figure 5.5: Genome-wide DNA methylation analysis of multiple loci using RLGS. RLGS determines the methylation status of approximately 3000 unique NotI sites located throughout the mouse genome by producing two-dimensional spots profiles by gel electrophoresis. A visible spot indicates a hypomethylated site. **(a)** Enlargements of small areas of RLGS profiles that display three of nine spots that are hypomethylated in response to 5-azaCdR treatment. **(b)** Measurement of the relative density of the three spots displayed in (a) reveals a decrease in DNA methylation in sperm from 5-azaCdR-treated (5-aza) versus saline-treated (sal) animals.

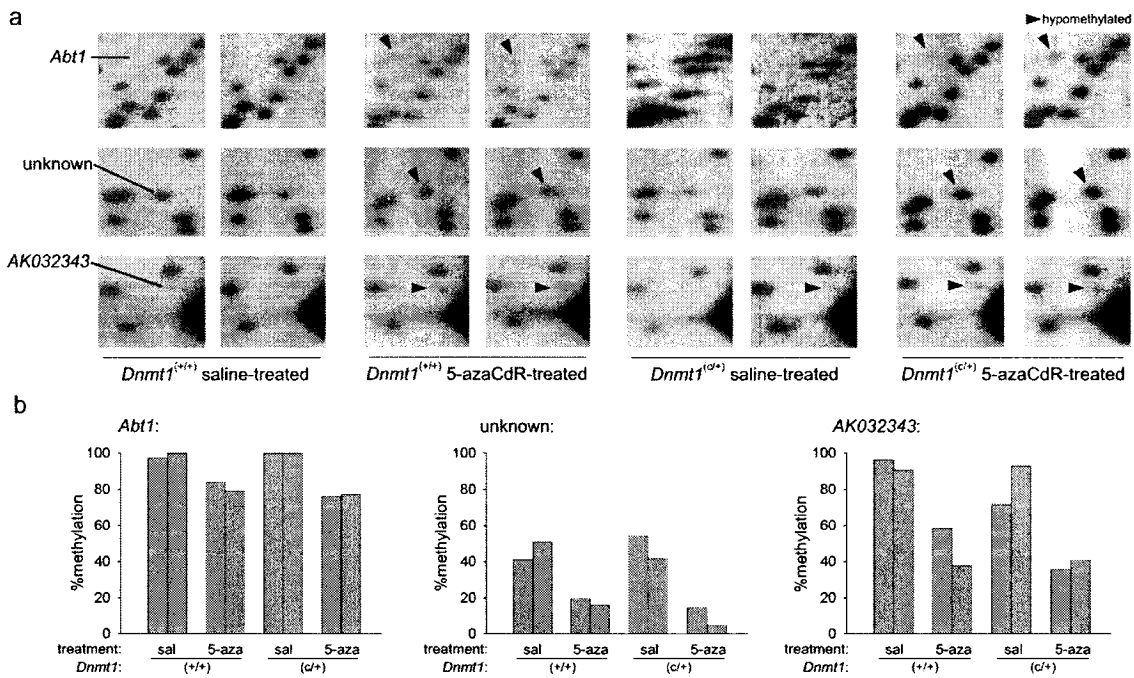


Figure 5.6: Enlargements of RLGS profiles showing that hypomethylation is observed in sperm and not in liver or brain. Hypomethylation of the NotI site found in the *Abt1* gene is visible on sperm RLGS profiles from 5-azaCdR-treated animals from *Dnmt1*$^{+/+}$ and *Dnmt1*$^{c/+}$ animals but not in liver or brain profiles from the same animals.

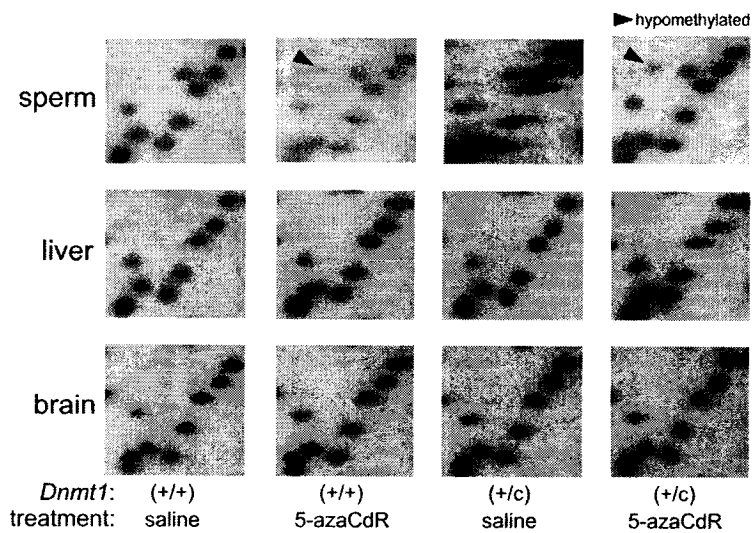


Figure 5.7: qAMP analysis of the percentage of DNA methylation at restriction sites within hypomethylated loci determined by RLGS in sperm and a comparison to the level of methylation previously found in spermatogonia. **(a)** The genes that harbour a NotI (N) site that is hypomethylated by 5-azaCdR treatment, primers binding sites and the location of surrounding HhaI (Hh), HpaII (Hp), or McrBC (M) sites are shown. **(b-d)** qAMP analysis of hypomethylated loci within *Tcf3*, *Abt1* and *Ibtk* reveals hypomethylation at all restriction sites examined. Horizontal grey lines on bar graphs show the percentage of methylation previously determined in type A spermatogonia (Oakes et al., 2007a). **(e)** qAMP analysis of a loci that is demethylated during spermatogenesis, *AK137601*, illustrates that 5-azaCdR does not impede demethylation during spermatogenesis. Error bars represent ±SEM, n=5-6 animals/group. Statistically significant differences between 5-azaCdR treatment vs. saline treatment in genotype-matched animals (asterisks) and between $Dnmt1^{c/+}$ vs. $Dnmt1^{+/+}$ genotypes in treatment-matched animals (daggers) are shown.

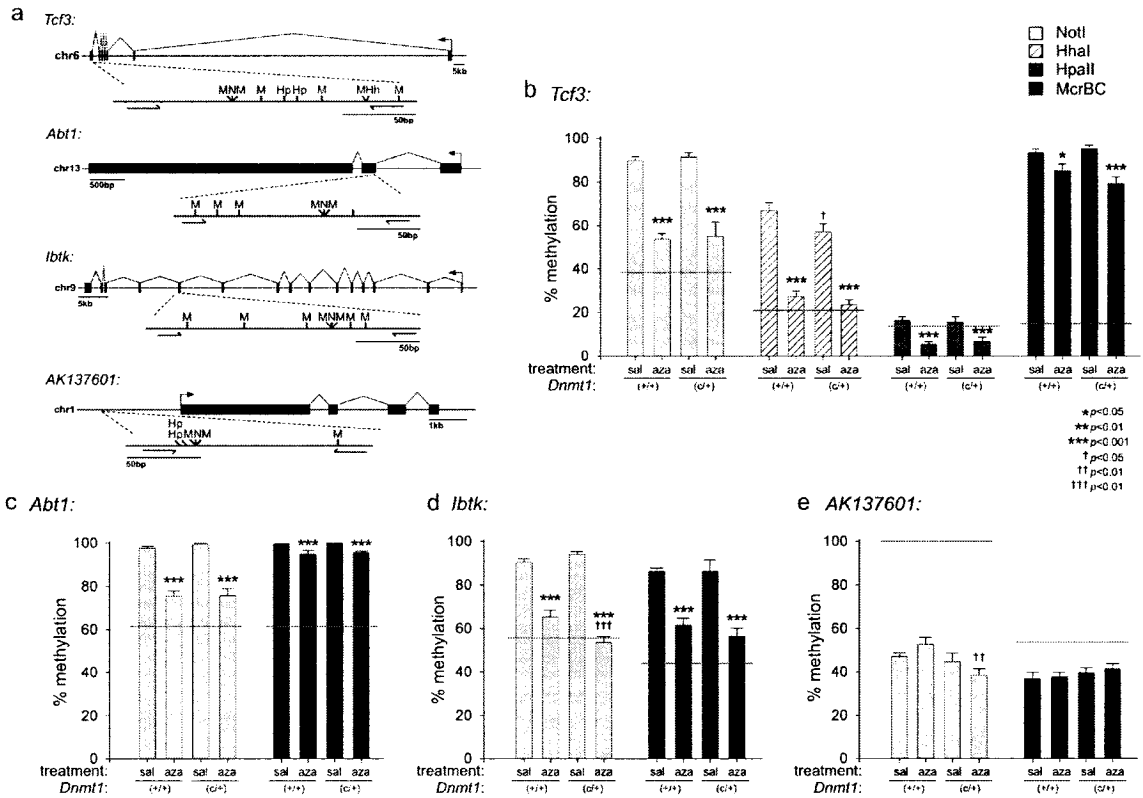
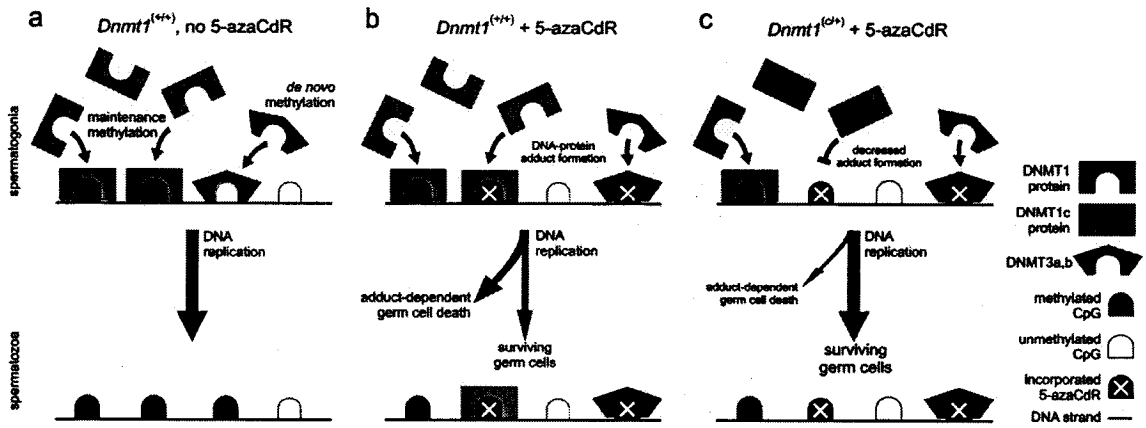

a
Tcf3:
chr6
5kb
MNM M HpHp M MHh M
50bp
Abt1:
chr13
500bp
M M M MNM
50bp
Ibtk:
chr9
5kb
M M M MNMM M
50bp
AK137601:
chr1
1kb
Hp
HpMNM
M
50bp
b Tcf3:
NotI
HhaI
HpaII
McrBC
% methylation
treatment:
Dnmt1:
sal aza sal aza
(+/+) (c/+)
*p<0.05
**p<0.01
***p<0.001
†p<0.05
††p<0.01
†††p<0.01
c Abt1:
d Ibtk:
e AK137601:

Figure 5.8: Diagram illustrating potential mechanisms that underlie changes in DNA methylation and adverse reproductive effects in 5-azaCdR-treated germ cells. **(a)** DNMT1 maintains established patterns of DNA methylation in spermatogonia; *de novo* DNA methyltransferases act on CpG sites programmed for methylation. **(b)** 5-AzaCdR incorporation causes adducts DNMT-DNA adducts, sequestering enough *de novo* DNA methyltransferase activity to prevent their action; sufficient DNMT1 is present to maintain established patterns of methylation. A high level of adducts in sperm DNA cause cell death and abnormalities in surviving cells. **(c)** A heterozygous mutation in the catalytic domain of DNMT1 prevents half of the pool of DNMT1 from forming adducts. Fewer adducts lead to a reduction in cell death and an attenuation of adverse effects in sperm.



TABLES

Table 5.1: Effect of chronic 5-azaCdR treatment on *Dnmt1*$^{+/+}$ and *Dnmt1*$^{c/+}$ mean body[a] and reproductive organ weights[b].

	Dnmt1$^{+/+}$		*Dnmt1*$^{c/+}$	
	Saline	**5-azaCdR**	**Saline**	**5-azaCdR**
Final body weight	26.07 ± 0.53	26.45 ± 0.61	24.27 ± 0.77*	23.85 ± 0.58
Testes	6.78 ± 0.21	3.87 ± 0.11*	6.57 ± 0.19	4.40 ± 0.11†
Seminal Vesicles	2.71 ± 0.10	2.80 ± 0.15	3.36 ± 0.20	3.11 ± 0.14

[a] weight in g (± SEM)
[b] Relative weight of paired organs (mg/g body weight) (± SEM)
* *p<0.05,* vs. saline-treated *Dnmt1*$^{+/+}$
† *p<0.05,* vs. saline-treated *Dnmt1*$^{c/+}$

Table 5.2: Effects of chronic 5-azaCdR treatment on *Dnmt1*$^{+/+}$ and *Dnmt1*$^{c/+}$ sperm motility parameters (as measured by CASA).

Sperm Parameter	*Dnmt1*$^{+/+}$		*Dnmt1*$^{c/+}$	
	Saline	5-azaCdR	Saline	5-azaCdR
VCL (μm/s) (track speed)	243.93 ± 4.91	191.54 ± 9.88*	231.04 ± 4.88	203.04 ± 3.66*
VSL (μm/s) (progressive velocity)	88.61 ± 3.13	68.17 ± 5.88*	84.74 ± 1.76	77.48 ± 2.39
VAP (μm/s) (path velocity)	125.22 ± 3.42	99.68 ± 5.40*	118.16 ± 2.51	105.61 ± 2.09$^{\phi}$
STR (%) (straightness)	69.65 ± 0.67	65.80 ± 2.78	70.82 ± 0.95	71.76 ± 1.22
LIN (%) (linearity)	37.49 ± 0.50	36.29 ± 1.62	38.04 ± 0.68	39.48 ± 0.65
Rapid Sperm (%)	52.76 ± 4.39	37.23 ± 5.09*	47.48 ± 4.73	39.04 ± 2.26
Medium Sperm (%)	6.65 ± 0.52	7.75 ± 0.92	7.76 ± 1.21	8.56 ± 2.16
Slow Sperm (%)	0.34 ± 0.12	0.30 ± 0.06	0.56 ± 0.14	0.44 ± 0.17
Static Sperm (%)	40.07 ± 4.15	54.55 ± 5.07*	44.06 ± 4.87	51.12 ± 4.13
BCF (beat cross frequency)	27.36 ± 0.58	27.30 ± 0.53	28.84 ± 0.89	29.17 ± 0.94
ALH (amplitude of lateral head movement)	13.24 ± 0.11	10.98 ± 1.29*	12.67 ± 0.17	11.28 ± 0.22*
Area	45.64 ± 3.07	35.41 ± 2.62*	43.57 ± 1.37	41.57 ± 1.40
Elongation	46.37 ± 0.97	50.65 ± 1.11*	47.04 ± 0.26	48.28 ± 0.71

*significantly different vs. saline, $p \leq 0.05$
$^{\phi}p=0.059$ vs. saline *Dnmt1*$^{c/+}$

Table 5.3: Tally of RLGS spots showing altered methylation in sperm.

Spot	*Dnmt1(c/+)*			5-azaCdR			5-azaCdR & *Dnmt1(c/+)*			Type-A	
Methylation	Liver	Brain	Sperm	Liver	Brain	Sperm	Liver	Brain	Sperm	spermatogonia*	Overlap†
Hypermethylated	0	0	0	0	0	0	0	0	0	8	0
Hypomethylated	0	0	1	0	0	9	0	0	9	11	8
Unchanged‡	2954	2954	2953	2954	2954	2945	2954	2954	2945	2935	NA

*number of spots showing altered methylation in type-A spermatogonia versus spermatozoa (Oakes, *et al.* 2007a)

†number of spots that show identical alterations in methylation between 5-aza & spermatogenesis RLGS studies

‡derived from virtual RLGS profiles

Table 5.4: Characteristics of identified RLGS spots that are hypomethylated by 5-aza-2'-deoxycytidine.

Genome position	RLGS spot	Gene	NotI site position	CpG island	Repeat
chr6:72645454	5ax2	*Tcf3*	3'	N	non-repetitive
chr7:75333081	1gx8	*Polg*	5'	Y	non-repetitive
chr7:75377378	4C13	*AK032343*	body	Y	non-repetitive
chr9:85800773	3ex8	*Ibtk*	body	N	non-repetitive
chr13:22791131	2F36	*Abt1*	body	N	non-repetitive

CHAPTER VI

DISCUSSION

DISCUSSION

6.1 DNA Methylation of the Male Germ Cell Genome

It was proposed more than thirty years ago that epigenetic modifications in DNA methylation functioned in the control of developmental processes (Holliday and Pugh, 1975; Riggs, 1975). Due to the significant role played by germ cells in development, several early studies attempted to address the status of DNA methylation in these cells (Ariel et al., 1991; Ariel et al., 1994; Kafri et al., 1992; Monk et al., 1987; Sanford et al., 1984; Sanford et al., 1987; Trasler et al., 1990). The use of a variety of different methods both to purify germ cells and to measure methylation, combined with the enormous complexity and size of the mammalian genome, has led to discrepancies between studies and to confusion in the literature. Without a comprehensive study providing a clear overall view of DNA methylation in germ cells, the prevailing view in the literature is that the status of DNA methylation between somatic and germ cells is similar (Ariel et al., 1991; Reik et al., 2001; Rollins et al., 2006). The studies described in this thesis have assessed DNA methylation in germ cells at a level of detail not previously achieved. Many of the findings challenge current dogma about DNA methylation and germ cells.

6.1.1 Global versus Sequence-Specific Detection of DNA Methylation

Previous attempts to determine the overall methylation status of male germ cells involved the study of either global methylation or a limited number of sequences. Due to the sequence-specific nature of DNA methylation patterning, extrapolation of data generated on a few sequences is of low value in the interpretation of overall methylation. Global assays suffer from low sensitivity and do not determine the methylation status of specific sequences. In previous studies of global DNA methylation levels, we have found the levels of methylation in somatic and male germ cell genomes are relatively similar on a global scale. Using a thin layer chromatography (TLC) assay, that assesses the percentage of DNA methylation of CCGG sites (~1.5 million sites in the mouse genome that are

approximately evenly distributed to both unique and repetitive sequences (Fazzari and Greally, 2004)), we find that spermatozoa are only 6-7% undermethylated compared to liver in both mice (C. Oakes, unpublished data) and rats (Oakes et al., 2003). High-performance liquid chromatography, (HPLC) the other technique commonly used to assess global levels of DNA methylation, measures a ratio of total genomic cytosine versus 5mC; thus, the sensitivity is too low to detect small differences in overall DNA methylation.

The primary finding of this thesis is that there are striking differences in DNA methylation between male germ cells and somatic cells. The discrepancy between this result and the prevailing view can be explained in part by the methods used previously and by the difference in DNA methylation between somatic and germ cells being a reorganization of methylation, rather than an overall change affecting all sites equally. Differences may affect a large proportion of the genome, but the extent is masked on a global scale by simultaneous hyper- and hypomethylation in different regions of the genome. In addition, we found high levels of methylation in interspersed repetitive elements of high copy number in both somatic and germ cells by RLGS and Southern blot experiments. Significant differences in methylation in the unique sequence portion of the genome are masked because of the substantial proportion of CpGs in the genome that are found in these repeats.

Despite the overall levels of DNA methylation being similar between somatic and male germ cells, the slightly lower level of DNA methylation found in spermatozoa by previous TLC assays is consistent with the findings presented in this thesis using other techniques. Using RLGS, qAMP analysis and Southern blot experiments, germ cells have overall lower amounts of methylation in comparison to somatic tissues. In locus-specific experiments, the total number of loci that have a lower amount of methylation in male germ cells is greater than the number of loci that have a higher amount. One explanation for there being an overall lower amount of methylation in male germ cells may relate to the overall level of GC content in the genome. Analysis of the frequency distribution of isochores in mammals (Costantini et al., 2006), reveals that the genome is

composed of more low GC regions than high GC regions. If low GC regions are hypomethylated in germ cells as equally as high GC regions are hypermethylated, it would be expected to find overall lower levels of methylation in germ cells, consistent with the findings of these four independent methods.

6.1.2 Germ Cell DNA Methylation and the Evolution of Genomic Cytosine Content

These findings also challenge a predominant, widely-held theory that describes how germ cell-specific DNA methylation is responsible for sculpting the mammalian genome into regions of high and low CpG content throughout evolutionary time. The basis of this theory is the reduced stability of the 5-methylcytosine molecule versus cytosine. 5mC will undergo spontaneous deamination to form thymine at 2-3 times the rate compared to unmethylated cytosine, which will result in a 5mC→T transition mutation if not properly repaired (Shen et al., 1994). An accumulation of these mis-repaired point mutations in the germ line of a species that contains DNA methylation would inevitably lead to a loss of cytosine content over evolutionary time. Regions containing methylated CpGs would lose cytosines and regions containing unmethylated cytosines would maintain an expected level of cytosine content. As the regional distribution of CpG and GC content is correlated in the genome, the basis for the chromosomal isochore structure may also be due to this mechanism (Fryxell and Zuckerkandl, 2000). Theories of the evolution of GC content and isochores are controversial (Duret et al., 2006), but it is generally thought that overall GC content has been decreasing in mammalian genomes (Belle et al., 2004). The underlying reason for the correlation of regional CpG and GC depletion is unclear; however, the presence of non-CpG methylation in germ cells (Haines et al., 2001; Imamura et al., 2005) would permit the theory to apply to GC content generally. Previously, authors of one study have used the spontaneous deamination theory to conclude that the DNA methylation status of the germ cell genome is revealed by sequences enriched or depleted in CpG dinucleotides without generating any experimental evidence to support this assumption (Rollins et al., 2006).

There are several points that support this theory. Firstly, CGIs are a dramatic example of sequences with high CpG and GC content, and their constitutively hypomethylated state in both somatic and germ cells has been thoroughly demonstrated. Secondly, C→T transitions at CpG sites represent approximately one-third of all described mutations in humans and occur 18-fold more frequently than the mean of other point mutations (Cooper and Youssoufian, 1988). Thirdly, it is widely-accepted that in somatic cells, heterochromatic regions are both depleted in CpG/GC content and highly methylated. Without any data to show otherwise, this pattern of methylation would be assumed to be the same for germ cells.

Results from the experiments presented in this thesis clearly disagree with the underlying assumption that low GC regions are hypermethylated in germ cells. Not only are low GC regions less likely to contain methylation, regions that are supposed to be unmethylated because of their above-average GC content are actually more likely to be hypermethylated. This result makes spontaneous deamination of hypermethylated regions in the male germ line unlikely to be the cause of region-specific CpG depletion. Although genome-wide patterns of DNA methylation are not known in female germ cells due to current technical limitations, female PGCs are demethylated early in germ cell development and methylation is not re-established until the oocyte growth phase. For the vast majority of the time the genome spends in the female germ line, it is likely that relatively little methylation is present. It seems reasonable to speculate that female germ cells are also unlikely to contribute to CpG depletion via spontaneous deamination caused by DNA methylation.

An alternate theory to the spontaneous deamination theory that could explain the striking relationship between GC content, heterochromatin and DNA methylation involves reprogramming in PGCs. Recently, a new mechanism of active demethylation has been proposed that involves site-directed deamination of methylcytosines coupled to DNA repair (Walsh and Xu, 2006). In this mechanism, an endogenously encoded family of DNA mutases, the Apobec/AID proteins, originally described to function in defense of viral DNA through directed

mutation of cytosines (Sheehy et al., 2002), is granted access to genomic DNA. Methylcytosines are mutated to thymine residues and are immediately repaired by thymidine DNA glycosylase (TDG) (Neddermann et al., 1996) or MBD4 (Hendrich et al., 1999). The result is a net conversion of 5mC to cytosine providing the desired process of demethylation; however, in this mechanism, the slightest inefficiency in repair would introduce the possibility that 5mC residues would mutate to thymine. This process would occur every generation in both male and female germ lines presumably across the entire genome making the possibility of C→T transitional mutations at demethylated cytosines prominent.

This alternate theory explains the striking relationship between GC content and DNA methylation in somatic cells. It has been shown that when PGCs enter the gonadal ridges, they retain a somatic pattern of DNA methylation at single copy genes and repeats (Hajkova et al., 2002; Maatouk et al., 2006). The pattern in these cells would also presumably retain the characteristic somatic pattern of hypermethylation of heterochromatic regions of below average GC content and hypomethylation of regions of high GC content. Thus, genome-wide demethylation of methylated heterochromatic regions would inevitably lead to a loss of CG content in these regions and not in above-average GC regions. Apobec/AID proteins also catalyze the deamination of cytosine to uracil, which may explain the regional association between CpG depletion and low GC content. The ability of PGCs to direct widespread active demethylation has been described (Tada et al., 1997), although the mechanism of active demethylation is unknown. Demethylation in these cells is likely to occur via a directed deamination/T-G repair pathway as mice deficient in the only reported demethylase that directly removes the methyl group from cytosine, MBD2 (Bhattacharya et al., 1999), develop normally and are fertile (Sansom et al., 2003).

Although these two theories are not mutually exclusive, there are a few lines of evidence that support the directed deamination/T-G repair theory over the spontaneous deamination theory. Firstly, in a study on mouse cells, it was noted that C→T transitions occur non-randomly (Steinberg and Gorman, 1992).

They found that a surprisingly high number of alleles that had a C→T mutation carried an additional C→T transition originating from a nearby CpG nucleotide. When both mutations were present they were always on the same allele, affecting CpG dinucleotides that were always on the same DNA strand. Furthermore, the second mutation was never observed in the absence of the first mutation. These observations are very difficult to reconcile with current views of spontaneous deamination, as mutations in this theory should occur randomly. Whereas it is possible that in a mechanism that involves a local repair process, if repair at one CpG is not performed, a neighbouring CpG may be similarly affected. Secondly, male germ cells re-establish methylation in their genomes very soon after it is removed. This is in stark contrast to the female where DNA methylation is not re-established until just prior to fertilization. Thus, male germ cell genomes spend much more time in a methylated state than the female germ cells per generation. Because the Y chromosome exclusively dwells in the male germ line compared to the X chromosome that occupies the male germ line only one-third of the time, the spontaneous deamination theory would predict that the Y chromosome would be more GC-depleted. However, these two chromosomes have equally depleted levels of both GC content and CpG ratios (Fazzari and Greally, 2004).

6.1.3 Potential Roles for the Distinct Pattern of DNA Methylation in Male Germ Cells: Gene Expression versus Chromatin Organization

Although the role of DNA methylation in the direct control of gene expression is controversial (Baylin and Bestor, 2002), there are a few genes that are clearly governed by DNA methylation in their 5' regions. Interestingly, most of these genes are exclusively expressed or are highest-expressing in the testis (MacLean and Wilkinson, 2005). This is consistent with the findings presented in this thesis, where the few genes that demonstrated a strong correlation between hypomethylation of 5' regions and increased expression were hypomethylated only in the testis. Surprisingly, a much higher proportion of 5' regions and CGIs were differentially methylated between somatic tissues; however, none of these

demonstrated the highest level of expression in the tissue that was hypomethylated. Confusingly, some of these genes demonstrated the opposite correlation: several somatic-specific differentially methylated genes that contained CGIs were more highly expressed in the tissues in which they were most highly methylated. This high methylation/high expression relationship has recently been noted for other somatic tissues (Eckhardt et al., 2006) and in cancer (Ordway et al., 2006). Some testis-specific genes also demonstrate this inverse correlation (Choi et al., 1997; Muller-Tidow et al., 2001; Trasler et al., 1990). The purpose or mechanism behind this phenomenon is unknown.

One of the primary functions of DNA methylation is the suppression of transposable element activity (Bestor and Tycko, 1996; Walsh et al., 1998). We find that transposable elements are generally methylated in all tissues examined. Interestingly, we find that a significant proportion of the sites that are hypomethylated in the testis are of repetitive origin. These repeat elements are almost exclusively from small, solitary LTRs that belong to class II and class III LTR repeats. These sequences are missing key elements required for retrotransposition, thus would not be expected to be expressed. Many other types of repeats were investigated in the course of these studies, and these were the only repeats to demonstrate this property.

Solitary LTRs occur in high frequency in the genome. They are remnants of a transposition event where homologous 5' and 3' LTRs are paired and recombined, removing the viral sequence as a circular piece of DNA and leaving behind an intact LTR. This obviously renders the retrovirus inactive; however, the remaining LTR retains promoter activity. Unmethylated solitary LTRs can thus recruit RNA polymerase and initiate transcription producing novel and modified transcripts. Genome-wide demethylation of solitary LTRs in germ cells would have a broad effect on transcription. Although the level of expression of these sequences in male germ cells is not currently known, these particular repeats are known to compose significant proportions of the total transcript population in female germ cells and early embryos (Peaston et al., 2004). This study demonstrates that the activity of LTR elements found upstream of genes or

within introns produces alternative transcripts in these cells. Some are translated into novel proteins. When considering that these cells may have a non-somatic pattern of DNA methylation, it is unclear whether the transcription of these sequences is causal or secondary to a change in chromosome structure. It has been argued that the unmasking of these LTR-driven promoters may provide a global regulatory network that shifts the overall transcription profile in these cells (Shapiro, 2005). An alternative argument is, due to the fact that the majority of these transcripts are degraded and/or blocked from translation, this transcription is a consequence of a primary reorganization of chromatin structure. These transcripts are tolerated because of the benefit an alternative chromosome structure provides for germ cell-specific processes.

Due to the fact that the RLGS technique investigates randomly positioned NotI sites, the methylation status of a variety of sites inside and outside of genes is investigated. This novel approach to investigating methylation in male germ cells has revealed that despite the RLGS technique being biased towards 5' regions and CGIs, we have found that there are more sites that are differentially methylated away from these regions than what would be expected at random (P=1x10^{-17}). The methylation status of non-5' sites that are found in the vicinity of genes (body & 3') is not representative of the expression status of the gene. This simple observation about the distribution of differentially methylated sites makes the argument that there are additional roles for DNA methylation other than gene expression.

Non-CGI loci are most commonly differentially methylated in the testis compared to other tissues. Germ cells are known to retain highly specialized chromosomal structures that occur during meiosis and during spermiogenesis in the male. Compared to somatic chromosomes, meiotic chromosomes have a different ultra-structure (Fang and Jagiello, 1981). Somatic chromosomes have a distinct hetero/euchromatic structure that serves to promote transcriptional activity in specific regions and repress activity in others. In addition to this function, meiotic chromosomes must maintain a structure that permits homologous pairing, recombination and reduction to haploid cells. During

meiosis, complex and unidentified processes direct some sequences to associate with the synaptonemal complex and others to the intervening chromatin loops that are not complexed with the core (Moens et al., 1998). A re-configuration of the methylation pattern in germ cells may be an attempt to position some genomic regions for either loop-specific or core-specific interactions. Another possibility is that hypomethylation of low GC regions that are normally heterochromatic and hypermethylation of above-average GC regions functions to provide a uniform linear configurational status of the chromosome. This may be necessary so that sequences are available for pairing along the entire chromosome and not just in less-condensed, euchromatic regions.

Another possible role for the unique pattern of DNA methylation in germ cells involves controlling the association of chromatin with the nuclear matrix. The nuclear matrix is intimately associated with the organization of chromatin into functional and non-functional compartments of the nucleus. Interactions between chromatin and nuclear matrix proteins occur in regions of DNA called scaffold/matrix attachment regions (S/MARs) (Boulikas, 1993). S/MARs are distributed throughout the genome and are usually found in conserved non-coding regions (Glazko et al., 2003). The discovery that MeCP2, a protein that specifically binds to methylated DNA, associates with known nuclear matrix proteins has provided a link between the methylation state of DNA and interactions with the nuclear matrix (Weitzel et al., 1997). Data presented in this thesis suggest the possibility that the methylation status of potential S/MARs may play a role in the re-organization of chromatin architecture.

Data presented in this thesis demonstrates a role for DNA methylation in male germ cells outside of 5' regions of genes. Several possible roles in connection with the modulation of chromatin structure have been discussed, such as the modulation of loop versus core interactions, nuclear matrix interactions and control linear configuration of the chromatin for homologous pairing. Future studies could specifically target sequences known to be associated with these phenomena and study both the DNA methylation status

and the chromatin marks in the form of various histone modifications. For example, DNA that associates with the chromatin loops versus DNA associated with the synaptonemal complex can be isolated individually (Moens et al., 1998). The DNA methylation status and the histone modifications present could be determined for these two groups of sequences. Furthermore, the behavior of these sequences in a model of abnormal synapsis, such as the *Dnmt3L*-deficient model, could be assessed.

It is very difficult to assess whether germ cell-specific transcription or a reorganization of germ cell chromatin structure is in the dominant purpose for the highly distinct genomic pattern of DNA methylation in male germ cells. These two purposes are not mutually exclusive; changes in transcription and changes in chromatin structure are clearly co-dependent. If promoters of testis-specific genes are found in regions of the genome that are hypomethylated in the testis, they are likely to have an increased potential for transcription. In support of this connection, testis-specific genes in the mouse and human genomes are found to be clustered at a higher frequency than by chance alone (Li et al., 2005). Regional changes in chromatin structure may facilitate the clustering effect. In a recent study of the transcriptional profile of developing oocytes, a time where DNA methylation is being acquired and reorganization of chromatin is presumably taking place, it was found that the most significant predictor of a change in the transcription of a particular gene was the chromosomal region in which the gene was located (Pan et al., 2005). On average, genes located within a particular chromosomal region demonstrated either increased or decreased activity concurrently. As it is unlikely that that the majority of all of the genes in a particular region are required to be activated or inhibited for oocyte growth specifically, this suggests that regional changes in chromatin structure occurring in the acquisition phase of oocyte maturation is having an influence in overall transcription in these regions. Finally, we have found that approximately half of all testis-specific genes that are hypomethylated in the testis are positioned within 3-4 Mb of hypomethylated sequences found by RLGS. As transcriptional activation in the testis of these specific genes are unlikely to influence the

chromatin structure of DNA several Mb distant, large-scale regional changes may be influencing the transcriptional capacity of these genes in the testis.

6.2 Establishment of DNA Methylation Patterns in Germ Cells

The findings in this thesis further define the timing of the acquisition of the unique pattern of DNA methylation in germ cells. Although it is regarded that male germ cell DNA methylation patterns continue to be acquired after birth up until the pachytene phase of spermatogenesis, progressive acquisition of DNA methylation has reliably been shown for a single locus, the imprinted gene *H19* (Davis et al., 1999; Ueda et al., 2000). Furthermore, in these studies it is only a proportion of methylation on the maternal allele that has not fully been established prior to spermatogenesis. The two other paternally-methylated imprinted genes, *Rasgrf1* and *Gtl2*, have been studied and may not have acquired their fully-methylated status before birth (Li et al., 2004); however, it is not clear whether the acquisition of DNA methylation occurs during spermatogenesis. All other sequences studied have acquired the methylation status by birth that is achieved by the pachytene spermatocyte stage. This begs three questions: 1) to what extent are patterns of DNA methylation being acquired during spermatogenesis? 2) Are only imprinted genes acquiring DNA methylation during spermatogenesis, or are other sequences involved? 3) What is the relationship between the amounts of methylation being acquired during spermatogenesis and the amount of DNA methylation being acquired before birth? Until now, all these questions remained largely unexplored.

Studies completed in this thesis provide data that addresses all three questions. Clearly, acquisition of DNA methylation occurs at several sequences, supporting and extending the previous findings on *H19*. Also in a similar fashion to the previous finding, these changes are almost always complete by the pachytene spermatocyte stage. Data presented in this thesis also shows that pattern acquisition involves demethylation of specific sequences, a result not previously shown. We have shown that there are regions of the genome beyond imprinted genes that progressively acquire DNA methylation in male germ cells.

Although it has not been exhaustively tested to ensure that these regions displaying novel DNA methylation patterns are not imprinted in all tissues, patterns of methylation of these regions in liver, intestine, brain, testis and spermatozoa do not show the characteristic methylation behavior of imprinted regions. We have also shown that only a minority of sequences in the genome participate in pattern acquisition during spermatogenesis, determined by the similarities of DNA methylation between spermatogonia and spermatozoa. Depending on the method used, partial acquisition of DNA methylation is detected in less than 10% of sequences. Although the data presented in this thesis are vastly more extensive than the studies done previously, only a minority of the genome is interpreted. Future studies employing tiling array technology will be able to interpret a larger proportion of the genome.

6.2.1 The Purpose for the Timing of Pattern Acquisition

Although the numbers of identified sites are low, there seems to be a connection between the types of modified sequences in spermatogenesis and the unique state of DNA methylation in the testis as compared to somatic tissues. The changes are both more likely to occur in non-CGI, non-5' regions. Also, a few of the solitary LTR sequences, a group of sequences that are that are hypomethylated in the testis, are in the process of demethylation during spermatogenesis. This supports the theory that these changes may represent the final modifications of changes to the chromatin structure before entry into meiosis (discussed in chapter IV).

The majority of modifications of DNA methylation are made during the fetal developmental window compared to those made during spermatogenesis. Establishment of DNA methylation patterns during fetal development precedes the creation of the spermatogonial stem cell population. Due to the heritability of established DNA methylation patterns, this strategy would necessitate the establishment of patterns at only a single time during development instead of during each wave of spermatogenesis. This strategy is more efficient and less prone to error, as exogenous influences such as exposure to toxins and

fluctuations in nutrient availability could affect the ability to continually establish patterns during adult life. This may be partly the reason why other than the limited number of sequences that gain their methylation early in spermatogenesis, other spermatogenic processes, such as meiotic divisions and spermiogenesis, are not associated with modification to DNA methylation patterns.

A few studies on the testis-specific gene, *Pgk-2*, have observed a post-meiotic hypermethylation event (Ariel et al., 1991; Kafri et al., 1992) that was later recognized to specifically occur during epididymal transit (Ariel et al., 1994; Geyer et al., 2004). Although this gene was not included in our study, we do not observe any modifications to the pattern of DNA methylation in this late developmental window. Based on the sequences investigated here, the occurrence of a global epigenetic reprogramming event occurring in the post-meiotic stages of spermatogenesis (Geyer et al., 2004) is not supported by our data.

6.3 Perturbation of DNMT Function and Male Germ Cell DNA Methylation

Recent studies have established that the perturbation of DNMT function specifically in male germ cells results in catastrophic effects. Knock-out of *Dnmt3L* and the germ cell-specific conditional knock-out of *Dnmt3a* results in a complete loss of male germ cells in adult mice (Bourc'his et al., 2001; Hata et al., 2002; Kaneda et al., 2004). Although this underscores the importance of DNA methylation in male germ cells, the lack spermatogenesis in these models makes studying the germ cells from these models problematic. In this thesis, we have taken two approaches to ascertain the effect of perturbations DNMT function to germ cell methylation. Firstly, as described in chapter III, we have taken advantage of the difference in the timing of the defect and the resulting effect in the *Dnmt3L* model. The *Dnmt3L* gene is highly expressed during late fetal development; however, *Dnmt3L*-deficent germ cells survive through this developmental window until they undergo apoptosis after birth during the spermatocyte stages of spermatogenesis (Bourc'his and Bestor, 2004). Isolation

of spermtogonia from *Dnmt3L*-deficient post-natal mice allows for the attainment of germ cells that have been perturbed by the absence of the DNMT3L protein, but have not yet begun to undergo the secondary effects associated with cell death. A second strategy, described in chapter V, is the treatment of male mice with 5-azaCdR, a general inhibitory agent of DNMT function. Although sufficient doses of this drug can cause a full disruption of spermatogenesis (Doerksen et al., 2000; Doerksen and Trasler, 1996; Kelly et al., 2003), the use of a lesser dose permits for the production of spermatozoa; yet, the spermatozoa have developed in the presence of the drug. These studies allow for the assessment of the effects of DNMT perturbation on sperm function, sperm quality and ability of sperm to support embryonic growth. In addition, alterations in germ cell DNA methylation that are associated with perturbations of DNMT function can be measured.

6.3.1 *Dnmt3L*-Deficient Male Germ Cells

In *Dnmt3L*-deficient germ cells, paternally-methylated imprinted genes are hypomethylated to varying extents; however, it is unclear how methylation changes at these few loci could result in such catastrophic, genome-wide effects. Two studies have produced additional data showing that male germ cells deficient in *Dnmt3L* demonstrate demethylation of interspersed transposable elements (Bourc'his and Bestor, 2004; Hata et al., 2006). This has led the authors to conclude that the genome-wide meiotic structural failure is the result of the expression of these elements that leads to a destabilization of the genome. The non-homologous synapsis that characterizes meiotic prophase in *Dnmt3L*-deficient spermatocytes may result from: 1) a perturbation of meiosis-specific gene expression; 2) single-strand or double-strand breaks produced during replicative retrotransposition; and/or, 3) illegitimate interactions between dispersed repetitive loci that were unmasked by demethylation (Maloisel and Rossignol, 1998). The studies presented here, in combination with other recent findings (La Salle et al., 2007), clearly demonstrate that in addition to the hypomethylation of imprinted and transposable element loci, undermethylated

loci include non-imprinted, non-repetitive sequences. This suggests a fourth possibility: the testicular DNA methylation pattern is not being acquired on a genome-wide scale, which may lead to an inability to maintain a chromosomal structure that permits proper homologous pairing interactions. This fourth possibility does not exclude the other possible reasons of the meiotic failure, as the failure may be a cumulative effect involving other causes, such as an alteration of expression of a key meiotic gene.

6.3.2 5-AzaCdR Treatment

Previous studies have demonstrated that 5-azaCdR has a robust effect on the development of male germ cells (Doerksen et al., 2000; Doerksen and Trasler, 1996; Kelly et al., 2003). These studies compose the only series of experiments that have shown that perturbing DNMT function using a method other than gene-targeting directly affects the development and function of male germ cells. Data presented in this thesis furthers the knowledge of these studies, demonstrating that sperm produced in the presence of 5-azaCdR are of reduced quality. This reduced quality includes decreased motility and fertilization ability, abnormal morphology and a reduced ability to support embryonic growth. These studies firmly establish that in addition to causing a reduction in the number of germ cells produced, the male germ cells that survive 5-azaCdR treatment are functionally non-equivalent. These experiments also further demonstrate the interesting protective aspects of reducing the amount of functional DNMT1 levels in the prevention of some decreases in sperm function.

Previous assessment of spermatozoa in 5-azaCdR-treated animals found a dose-dependent decrease in global levels of DNA methylation that was correlated with a dose-dependent increase in testicular abnormalities. Effects on germ cells were noticed at lower doses before significant decreases in global methylation could be detected. These results beg three questions: 1) are 5-azaCdR-dependent modifications occurring at lower doses, but are undetectable using a global assay? 2) If 5-azaCdR is causing a modification of DNA methylation, what specific loci or sequence-types are affected? 3) Are

abnormalities in germ cell function caused by alterations in germ cell DNA methylation or by other effects of the drug such as covalent adduct formation?

The dose of 5-azaCdR used in this study corresponds to the lowest dose used in previous studies (Kelly et al., 2003). Significant differences in DNA methylation between treated and non-treated animals were not observed at the global level using the TLC assay despite adverse testicular effects detected at this dose in these studies and those previously done. In this study, using the more sensitive and comprehensive RLGS technique, some differences were observed. However, differences represented only a very minor proportion of the sequences investigated by RLGS, in part supporting the lack of a significant decrease in global methylation found by the TLC assay. Sequences known to remain unchanged include all the interspersed transposable elements visible on the RLGS and the repetitive elements investigated using Southern blots.

The low dose of 5-azaCdR selectively causes an inhibition of *de novo* methylation activity, resulting in the specific hypomethylation of loci in spermatozoa that normally gain methylation during spermatogenesis. The plethora of sites of the genome that maintain their methylation indicates that maintenance methylation activity, generally thought to be performed by DNMT1, remains sufficiently uninhibited and continues to maintain methylation. Some sites are also normally demethylated during spermatogenesis, and this process was found to proceed unabated. This is consistent with demethylation occurring via a passive mechanism, although it is not known if 5-azaCdR would inhibit an active demethylation processes. The selective action of the drug most likely occurs because *de novo* and maintenance activities are provided by separate enzymes (DNMT3a and DNMT3b versus DNMT1) and, for reasons that are unclear, *de novo* enzyme activity is more sensitive to the drug.

5-azaCdR causes both DNA hypomethylation and DNMT-DNA adducts (Juttermann et al., 1994). These studies suggest that adduct formation contributes more to adverse effects to male reproduction than does DNA hypomethylation. The future development of DNMT inhibitory agents that do not cause the formation of adducts will be of great assistance in elucidating the exact

role played by alterations of DNA methylation in the generation adverse reproductive effects.

To our knowledge, the selective nature of 5-azaCdR has not been previously described. This discovery was made possible by observing the effect of the drug in a system where developmental modifications of DNA methylation of specific loci are known to occur. It would be interesting to know if the selective action of the drug is restricted to germ cells or if it causes a similar effect in other tissues, namely in the hematopoetic system. The current primary indication for the clinical use of 5-azaCdR is for the treatment of myelodysplastic syndromes. Selective inhibition of *de novo* methylation might contribute to the drug's known anti-cancer properties. It is possible that the beneficial anti-cancer effects of 5-azaCdR involve the prevention of *de novo* methylation of tumor suppressor genes in pre-cancerous cells in addition to its cytotoxic effects.

These studies also raise the possibility that due to evolving patterns of DNA methylation during spermatogenesis, male germ cells may be especially sensitive to potential 'epimutations'. It has been demonstrated that disruptions in epigenetic programming of germ cells during the fetal developmental window can lead to effects in progeny of the next generation (Anway et al., 2005; Anway and Skinner, 2006; Chang et al., 2006). This is possibly the result of an inhibition of the acquisition of germ cell epigenetic patterns that normally occurs in fetal development. Work presented in this thesis demonstrates that epigenetic patterns in male germ cells continue to be acquired beyond the pre-natal window; similar disruptions may occur in the reproductive life of the individual. This raises the possibility that offspring could be affected by environmental insults via alterations in germ cell epigenetic states. Several genomic loci are demethylated in the male pronucleus shortly after fertilization (Oswald et al., 2000) which would prevent the transmission of paternally-derived epimutations; however, the full extent of the demethylation is not known. Paternally-methylated imprinted genes retain their methylation status throughout preimplantation embryonic development (Reik et al., 2001), demonstrating that some regions of the genome are not demethylated. Further studies will be required to test if environmental

influences experienced during the adult life of the male can affect progeny via an alteration in the epigenetic program of the offspring.

All past and current methods used for the purpose of isolating large numbers of different types of spermatogonial cells require their isolation from pre-pubertal testis during the first wave of spermatogenesis. This is due to technical limitations combined with the small proportion of early spermatogenic cells in the adult testis. There is a concern that the spermatogonial cells isolated from the first round of spermatogenesis, due to the immature state of the testis and the lack of later stages of germ cells, are distinct from spermatogonial cells in adults (Jahnukainen et al., 2004). As almost all studies to date that have analyzed molecular aspects of these early spermatogenic cell types employ the use of pre-pubertal mice, there is a concern that data generated on these spermatogonia are not representative of adult spermatogonia. Our studies suggest that pre-pubertal and adult spermatogonia are of roughly equivalent maturity in their stage of acquisition of DNA methylation. The pattern of DNA methylation observed in spermatogonia isolated from pre-pubertal mice and the modifications that result from 5-azaCdR treatment of adult mice are highly correspondent. This result suggests that spermatogonia in adult mice retain a very similar pattern of DNA methylation to their pre-pubertal counterparts. This result also supports the concept of an early establishment of a spermatogonial stem cell population and that most germ cells in the pre-pubertal testis are progressing through spermatogenesis (Yoshida et al., 2006).

6.5 Implications of the Findings and Future Considerations

There are many questions that are raised by the experiments done in this thesis. Although it would be impossible to discuss all the potential implications of this work, I will highlight and elaborate on three concepts that I think are the most important.

6.4.1 Sequence-Specific Targeting of DNA Methylation to Genomic Loci

At several times during mammalian development, distinct cell type-specific epigenetic patterns are established. Data presented in this thesis in combination with the work of others firmly establishes that distinct patterns of DNA methylation occur on a genome-wide scale in various somatic, germ and embryonic cells. These results highlight the most fundamental unanswered question in the field of epigenetics: how is DNA methylation specifically targeted to distinct genomic loci. The enormous size and number of sequence types in the genome makes the task of accurate genome-wide reprogramming exceedingly complex.

There are several different possible mechanisms that may contribute to various extents to allow for site-specific DNA methylation to be either maintained or changed during a reprogramming event. Firstly, it is possible that some tissue-specific DNA methylation is simply the result of gene activity. In this mechanism, the gene products required for the function of a particular cell type drive the modifications to the epigenetic state of the DNA. The epigenetic modifications that are then established in any cell type, due to their heritable nature, in turn provide a stable environment for the long existence of a differentiated cell lineage. The information for the establishment of gene-specific epigenetic marks is provided by higher-order gene regulatory networks, perhaps stimulated by cellular environmental cues or developmental timing. Site-specific DNA methylation is the result of RNA polymerase, transcription factors, enhancers and repressor proteins complexed with DNA in a sequence-specific manner. These proteins may enhance or block the access of the DNA methylation machinery to DNA. This is unlikely to be a primary mechanism due to the fact that genes and regulatory sequences compose a very small proportion of the total genomic sequence, and here we have demonstrated extensive reprogramming in non-genic regions.

A second possible mechanism that may be involved in the direction of site-specific DNA methylation involves relationship to the other epigenetic factor, histone modifications. The multitude of histone modifications provide a multi-layered structure to the information encoded in chromatin. Thus, when

methylation is observed to be established in a sequence-specific manner, this sequence specificity may be a reflection of an epigenetic state maintained by one or more of these other modifications. Evidence for this comes from the allele-specificity of the timing of establishment of DNA methylation marks on imprinted genes in both male and female germ cell development. Imprinted alleles that originate from the parent that is the opposite sex to the embryo acquire their methylation later than alleles that come from the parent of the same sex as the embryo (Davis et al., 1999; Hiura et al., 2006; Lucifero et al., 2004). In these studies, alleles exhibit 'epigenetic memory' of a previous programmed state. Although experiments have not been done to explore the mechanism of allelic epigenetic memory in these cases, it is possible that chromatin conformations, most likely dictated by histone modifications, persist through the reprogramming phase. It is abundantly clear that future studies will have to evaluate multiple epigenetic factors in order to fully understand the process of site-directed DNA methylation that occurs during reprogramming.

A third interesting candidate for sequence-specific targeting of DNA methylation involves small RNA molecules. The emerging field of RNA interference describes how small single and double stranded RNA molecules perform previously uncharacterized key roles in cellular biology which mainly include gene regulation at both transcriptional and post-transcriptional levels (Bayne and Allshire, 2005). Small RNA molecules can target other RNA molecules using sequence homology-dependent interactions which can lead to RNA degradation, sequestration, stabilization or translational blocking. It has also been shown that these RNA-dependent targeting complexes can induce repressive chromatin conformational changes when homologies are targeted to DNA sequences (Lippman and Martienssen, 2004). This has been well characterized in several plant species and lower organisms. Although the existence of this process is currently controversial in mammals, RNA-based chromatin targeting based on sequence homology is an appealing mechanism that may prove to be a major force in site-specific direction of DNA methylation. The defined timing and sequence specificity of the acquisition of DNA

methylation in male germ cells, as illustrated in the studies presented here, makes male germ cell development a useful system to explore these mechanisms.

6.4.2 Developmental Dimorphism between Female and Male Germ Cells: Relationship to Epigenetic Patterns

The primary question that is raised by the work done in this thesis is the purpose of the unique pattern of DNA methylation in male germ cells. Why does the somatic pattern have to be changed? Slightly different versions of the same somatic pattern are used by several different types of somatic tissues. Furthermore, germ cells arrive at the gonadal ridges with a somatic pattern before they are reprogrammed. It would appear to be much simpler for male germ cells to use this pattern and to not have evolved mechanisms for the erasure and re-establishment of a distinct new pattern. Therefore, one must presume that the re-established pattern must serve an important purpose. We have suggested that meiosis is a plausible purpose because it is unique to germ cells, involves genome-wide changes in chromosomal configurations, modifications to the pattern occur prior to meiosis and the perturbation of the genome-wide pattern leads to meiotic failure.

One way to further explore the relationship between patterns of DNA methylation is to study the patterns in oocytes. If the patterns were similar, it would further support the hypothesis that the patterns are important for meiosis. Although current technical limitations do not permit genome-wide analyses to be done on oocytes, there is a strong indication that patterns of DNA methylation in meiotic male and female germ cells are different. As mentioned previously, while male germ cells re-establish their methylation patterns preceding meiosis, female germ cells enter meiosis shortly after being demethylated, and do not reacquire methylation until later on during oocyte growth. Thus, male and female germ cells undoubtedly have different patterns of methylation at meiotic entry and during meiotic prophase I where the important processes of chromosomal condensation, pairing and recombination occur. In *Dnmt3L*-deficient mice, male

germ cells undergo meiotic failure that results from a lack of methylation establishment; however, at the equivalent phase of meiosis, female germ cells are yet to acquire methylation and proceed unabated through meiosis. Therefore, the question is: why can female germ cells proceed through meiosis in a demethylated state, whereas, in the case of *Dnmt3L*-deficient male germ cells, demethylated male germ cells fail during meiotic passage? The answer may relate to the dimorphism in the events that occur between male and female germ lines. As a result, the purpose of the configuration of the male germ cell pattern may not be solely meiosis.

The purpose for the evolution of the striking dimorphism between the timing of male and female epigenetic reprogramming is unclear. Understanding the basis of this dimorphism is key to the understanding of the roles of DNA methylation in the germ cells of both sexes. One possible explanation of the differences in the timing of reprogramming and meiosis is to consider the functional commonalities and differences between the germ lines. The germ cells of both sexes need to accomplish three fundamental processes:

1) Genomic reprogramming
2) Recombination
3) Functional differentiation into fertile gametes

In addition to these processes, males need to perform a critically important extra function:

4) The production of a vast, perpetual supply of gametes

While females produce relatively few mature gametes throughout their reproductive lifespan, males will produce millions to billions fold more. This requirement has placed strong evolutionary pressure on the male germ line to evolve a more complex developmental process to facilitate this necessity. Higher output of male gametes is achieved through an expansion of the germ cell

population not seen in the female germ line. Meiosis is reductive (no new DNA is synthesized after the pre-leptotene stage); therefore, the expansion must occur mitotically preceding meiotic entry. Fetal testicular resources would obviously not be able to support such an expansion, thus, germ cell amplification must occur in the adult. In addition to the inclusion of an expanded mitotic window into male germ cell development, the perpetual supply of male germ cell precursors are provided by the creation of a pool of spermatogonial stem cells that fuels the mitotic expansion. Thus, meiosis in the male is delayed until the testis has developed and a suitable pool of spermatogonial stem cells and mitosis have expanded the germ cell population.

It is possible that gamete output is the primary reason for the differences in the timing and reordering of the events in the male versus female germ lines (Figure 6.1). Although gamete differentiation is strikingly different between male and female, it is unclear why this would necessitate a rearrangement of events. Genomic reprogramming is reordered to occur before the creation of the stem cell pool for purposes of efficient and reliable establishment of the epigenetic program. A post-reprogramming mitotic expansion of the germ cell pool requires that the epigenetic program needs to be dually compatible for mitosis and meiosis; relatively little modification in DNA methylation occurs between the mitotic and meiotic phases. Compared to female germ cells, male germ cells must provide the means for maintenance of methylation patterns as well as their initiation. By comparing the differences in the complexity between female and male germ lines, it appears that meiosis may not be the sole purpose of the epigenetic pattern in male germ cells. A detailed comparison of the patterns of DNA methylation in female and male germ cells will be of the utmost importance to the understanding of the roles of DNA methylation in germ cells.

6.4.3 The Role of the Paternal Epigenetic Pattern in Embryonic Development

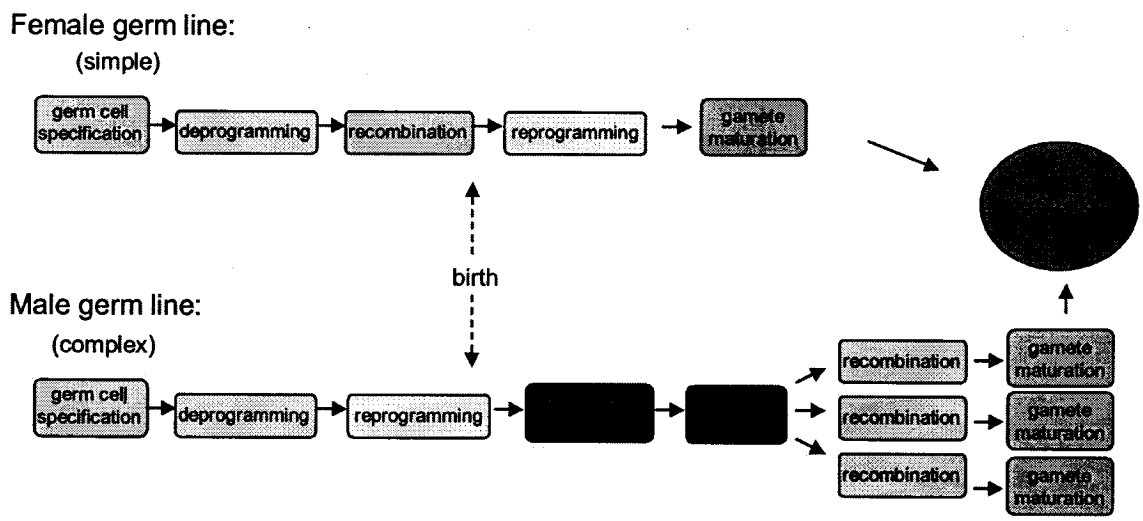
There is one possible role for the distinct pattern of DNA methylation in male germ cells that has yet to be discussed: the role of paternal epigenetic information in mammalian embryonic development. It is generally thought that all gametic epigenetic information is erased during preimplantation development and *de novo* patterns are re-established for all of the embryonic lineages, erasing any gametic patterns. The only current exceptions to this rule are the imprinted genes. As there are very few imprinted genes that are paternally-methylated, there is not much current evidence for paternal epigenetic contributions to the embryo. The active demethylation of the male pronucleus shortly after fertilization reduces the likelihood even further, preventing paternal epigenetic information from contributing to the development of preimplantation embryos.

Despite the lesser possibility of a contribution to the embryo, there is some evidence to support the concept of gametic epigenetic information being important for the development of the extraembryonic lineages. Several studies have described vast differences in epigenetic patterns between the extraembryonic lineages derived from the trophectoderm (TE) and the embryonic lineages derived from the inner cell mass (ICM) of the blastocyst (Chapman et al., 1984; Rossant et al., 1986). This is likely to be the result of genomic reprogramming occurring selectively in the ICM and not in the TE during the development of the blastocyst. The methylation status of some DNA sequences in extraembryonic tissues resembles germ cell methylation patterns, such as the demethylation of major and minor satellite centromeric repeats (Sanford et al., 1984), sequences that are hypermethylated in all other tissues tested.

In addition to the similarities between germ cell and extraembryonic methylation patterns, paternal-specific epigenetic information selectively survives into the extraembryonic versus the embryonic lineages. The best current evidence of this involves the epigenetic status of the paternal X chromosome in XX embryos. In all extraembryonic lineages of the XX embryo, only the paternal X chromosome is inactivated. The pre-programmed inactivation arises from the

inactivation of the X chromosome during XY-body formation during meiosis in the male germ line (Huynh and Lee, 2003; Namekawa et al., 2006). The transmission of the paternally pre-programmed inactivated state is of great importance because XX embryos that inherit a paternal X chromosome that carries a disrupted *Xist* gene, the gene responsible for the initiation of X-inactivation in *cis*, are unable to inactivate a maternally inherited wild-type copy of the X chromosome (Marahrens et al., 1997). This leads to a failure of X chromosome dosage compensation in the extraembryonic tissues. Another key piece of evidence supporting a paternal epigenetic contribution to the extraembryonic tissues relates to founding work in the field of mammalian epigenetics. Diploid mouse embryos with two male pronuclei (biparental androgenotes) develop differently than embryos with two female pronuclei (biparental gynogenotes) with respect to the extent of the development of embryonic versus extraembryonic tissues (Barton et al., 1984; McGrath and Solter, 1984). Extraembryonic tissue development is very limited in gynogenotes; whereas, extraembryonic lineages develop relatively well in androgenotes. This clearly demonstrates that not only does paternal-specific epigenetic information survive the period of demethylation that occurs shortly after fertilization, but that the paternal epigenetic pattern is important for normal development. Future studies addressing the behavior of parent-specific epigenetic information in early embryonic development will provide key insights into the understanding of genomic imprinting, cell lineage selection, and the mechanisms of embryonic and extraembryonic tissue development.

Figure 6.1: Diagram of the order of key events that occur in male and female germ lines.



ORIGINAL CONTRIBUTIONS

1. Site- and region-specific levels of DNA methylation can accurately and quantitatively be determined by using a strategy of combining methylation-sensitive and methylation-dependent restriction enzymes combined with real time PCR.

2. The methylation status of testicular DNA is highly distinct, displaying eight-fold the number of hypomethylated loci relative to somatic tissues. Differentially methylated loci are generally located within sequences that are away from CpG islands and 5' regions of genes. Several repetitive elements are specifically hypomethylated in the testis that originate from solitary LTR fragments belonging to ERVK (class II) and MaLR (class III) of the LTR family of repeats.

3. Tissue-specific hypomethylation of the vast majority of the differentially methylated loci identified using the RLGS technique does not correlate with increased levels of tissue-specific gene expression.

4. The methylation state of non-CpG island loci is correlated with the chromosomal banding pattern and the regional level of GC content. The relationship between DNA methylation and regional levels of GC content is inverted in the testis compared to somatic tissues.

5. In the spermatogonia from *Dnmt3L*$^{(-/-)}$ mice, all loci examined that were methylated in the testis failed to gain the normal levels of methylation found in wild-type spermatogonia and testis. Testis-specific hypomethylated loci remained normal in the spermatogonia from *Dnmt3L*$^{(-/-)}$ mice, indicating that the primary influence of *Dnmt3L* is to promote DNA methylation.

6. During spermatogenesis, a specific subset of genomic loci undergoes *de novo* methylation and demethylation; these changes occur mainly in the spermatogonia and spermatocyte stages.

7. Alterations of DNA methylation during spermatogenesis predominantly include non-CpG island sequences; unique loci are usually hypermethylated during spermatogenesis and the few loci that are demethylated during spermatogenesis are found within solitary LTR repetitive elements.

8. 5-azaCdR administration did not cause the demethylation of IAP and the major and minor satellite repeats in sperm DNA; however, 5-azaCdR did cause the demethylation of the imprinted gene *Rasgrf1* in sperm DNA.

9. A subset of single-copy loci are demethylated in sperm from 5-azaCdR-treated animals; these loci correspond to those that are hypermethylated during spermatogenesis. The lack of a change in maintenance methylation and demethylation suggests that 5-azaCdR selectively inhibits *de novo* methylation activity in germ cells.

10. The partial rescue of adverse effects in $Dnmt^{(c/+)}$ animals is associated with no restoration of methylation levels in sperm suggesting that 5-azaCdR defects are primarily mediated by the level of covalent adducts over hypomethylation.

REFERENCES

Aapola,U., Kawasaki,K., Scott,H.S., Ollila,J., Vihinen,M., Heino,M., Shintani,A., Kawasaki,K., Minoshima,S., Krohn,K. et al. (2000). Isolation and initial characterization of a novel zinc finger gene, DNMT3L, on 21q22.3, related to the cytosine-5-methyltransferase 3 gene family. *Genomics* **65,** 293-298.

Aapola,U., Lyle,R., Krohn,K., Antonarakis,S.E., and Peterson,P. (2001). Isolation and initial characterization of the mouse Dnmt3l gene. *Cytogenet. Cell Genet.* **92,** 122-126.

Alcivar,A.A., Hake,L.E., Millette,C.F., Trasler,J.M., and Hecht,N.B. (1989). Mitochondrial gene expression in male germ cells of the mouse. *Dev. Biol.* **135,** 263-271.

Amann,R.P. and Howards,S.S. (1980). Daily spermatozoal production and epididymal spermatozoal reserves of the human male. *J. Urol.* **124,** 211-215.

Anway,M.D., Cupp,A.S., Uzumcu,M., and Skinner,M.K. (2005). Epigenetic transgenerational actions of endocrine disruptors and male fertility. *Science* **308,** 1466-1469.

Anway,M.D. and Skinner,M.K. (2006). Epigenetic transgenerational actions of endocrine disruptors. *Endocrinology* **147,** S43-S49.

Aoki,A., Suetake,I., Miyagawa,J., Fujio,T., Chijiwa,T., Sasaki,H., and Tajima,S. (2001). Enzymatic properties of de novo-type mouse DNA (cytosine-5) methyltransferases. *Nucleic Acids Res.* **29,** 3506-3512.

Aota,S., Gojobori,T., Shigesada,K., Ozeki,H., and Ikemura,T. (1987). Nucleotide sequence and molecular evolution of mouse retrovirus-like IAP elements. *Gene* **56,** 1-12.

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

Ariel,M., McCarrey,J., and Cedar,H. (1991). Methylation patterns of testis-specific genes. *Proc. Natl. Acad. Sci. U. S. A* **88,** 2317-2321.

Avner,P. and Heard,E. (2001). X-chromosome inactivation: counting, choice and initiation. *Nat. Rev. Genet.* **2,** 59-67.

Bakken,A.H. and McClanahan,M. (1978). Patterns of RNA synthesis in early meiotic prophase oocytes from fetal mouse ovaries. *Chromosoma* **67,** 21-40.

Ballestar,E. and Wolffe,A.P. (2001). Methyl-CpG-binding proteins. Targeting specific gene repression. *Eur. J. Biochem.* **268,** 1-6.

Barton,S.C., Surani,M.A., and Norris,M.L. (1984). Role of paternal and maternal genomes in mouse development. *Nature* **311,** 374-376.

Baust,C., Gagnier,L., Baillie,G.J., Harris,M.J., Juriloff,D.M., and Mager,D.L. (2003). Structure and expression of mobile ETnII retroelements and their coding-competent MusD relatives in the mouse. *J. Virol.* **77,** 11448-11458.

Baylin,S. and Bestor,T.H. (2002). Altered methylation patterns in cancer cell genomes: cause or consequence? *Cancer Cell* **1,** 299-305.

Bayne,E.H. and Allshire,R.C. (2005). RNA-directed transcriptional gene silencing in mammals. *Trends Genet.* **21,** 370-373.

Belle,E.M., Duret,L., Galtier,N., and Eyre-Walker,A. (2004). The decline of isochores in mammals: an assessment of the GC content variation along the mammalian phylogeny. *J. Mol. Evol.* **58,** 653-660.

Bellve,A.R., Cavicchia,J.C., Millette,C.F., O'Brien,D.A., Bhatnagar,Y.M., and Dym,M. (1977a). Spermatogenic cells of the prepuberal mouse. Isolation and morphological characterization. *J. Cell Biol.* **74,** 68-85.

Bellve,A.R., Millette,C.F., Bhatnagar,Y.M., and O'Brien,D.A. (1977b). Dissociation of the mouse testis and characterization of isolated spermatogenic cells. *J. Histochem. Cytochem.* **25,** 480-494.

Benoit,G. and Trasler,J.M. (1994). Developmental expression of DNA methyltransferase messenger ribonucleic acid, protein, and enzyme activity in the mouse testis. *Biol. Reprod.* **50,** 1312-1319.

Bernardi,G., Olofsson,B., Filipski,J., Zerial,M., Salinas,J., Cuny,G., Meunier-Rotival,M., and Rodier,F. (1985). The mosaic genome of warm-blooded vertebrates. *Science* **228,** 953-958.

Bestor,T., Laudano,A., Mattaliano,R., and Ingram,V. (1988). Cloning and sequencing of a cDNA encoding DNA methyltransferase of mouse cells. The carboxyl-terminal domain of the mammalian enzymes is related to bacterial restriction methyltransferases. *J. Mol. Biol.* **203,** 971-983.

Bestor,T.H., Hellewell,S.B., and Ingram,V.M. (1984). Differentiation of two mouse cell lines is associated with hypomethylation of their genomes. *Mol. Cell Biol.* **4,** 1800-1806.

Bestor,T.H. and Tycko,B. (1996). Creation of genomic methylation patterns. *Nat. Genet.* **12,** 363-367.

Bhattacharya,S.K., Ramchandani,S., Cervoni,N., and Szyf,M. (1999). A mammalian protein with specific demethylase activity for mCpG DNA. *Nature* **397,** 579-583.

Bird,A. (2002). DNA methylation patterns and epigenetic memory. *Genes Dev.* **16,** 6-21.

Bird,A.P. (1980). DNA methylation and the frequency of CpG in animal DNA. *Nucleic Acids Res.* **8,** 1499-1504.

Boulikas,T. (1993). Nature of DNA sequences at the attachment regions of genes to the nuclear matrix. *J. Cell Biochem.* **52,** 14-22.

Bourc'his,D. and Bestor,T.H. (2006). Origins of extreme sexual dimorphism in genomic imprinting. *Cytogenet. Genome Res.* **113,** 36-40.

Bourc'his,D. and Bestor,T.H. (2004). Meiotic catastrophe and retrotransposon reactivation in male germ cells lacking Dnmt3L. *Nature* **431,** 96-99.

Bourc'his,D., Xu,G.L., Lin,C.S., Bollman,B., and Bestor,T.H. (2001). Dnmt3L and the establishment of maternal genomic imprints. *Science* **294,** 2536-2539.

Brena,R.M., Auer,H., Kornacker,K., Hackanson,B., Raval,A., Byrd,J.C., and Plass,C. (2006). Accurate quantification of DNA methylation using combined bisulfite restriction analysis coupled with the Agilent 2100 Bioanalyzer platform. *Nucleic Acids Res.* **34,** e17.

Burgers,W.A., Fuks,F., and Kouzarides,T. (2002). DNA methyltransferases get connected to chromatin. *Trends Genet.* **18,** 275-277.

Caspersson,T., Farber,S., Foley,G.E., Kudynowski,J., Modest,E.J., Simonsson,E., Wagh,U., and Zech,L. (1968). Chemical differentiation along metaphase chromosomes. *Exp. Cell Res.* **49,** 219-222.

Chang,H.S., Anway,M.D., Rekow,S.S., and Skinner,M.K. (2006). Transgenerational epigenetic imprinting of the male germline by endocrine disruptor exposure during gonadal sex determination. *Endocrinology* **147,** 5524-5541.

Chapman,V., Forrester,L., Sanford,J., Hastie,N., and Rossant,J. (1984). Cell lineage-specific undermethylation of mouse repetitive DNA. *Nature* **307,** 284-286.

Chedin,F., Lieber,M.R., and Hsieh,C.L. (2002). The DNA methyltransferase-like protein DNMT3L stimulates de novo methylation by Dnmt3a. *Proc. Natl. Acad. Sci. U. S. A* **99,** 16916-16921.

Chen,T., Ueda,Y., Xie,S., and Li,E. (2002). A novel Dnmt3a isoform produced from an alternative promoter localizes to euchromatin and its expression correlates with active de novo methylation. *J. Biol. Chem.* **277,** 38746-38754.

Choi,Y.C., Aizawa,A., and Hecht,N.B. (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.

Cisneros,F.J. (2004). DNA methylation and male infertility. *Front Biosci.* **9,** 1189-1200.

Clark,S.J., Harrison,J., Paul,C.L., and Frommer,M. (1994). High sensitivity mapping of methylated cytosines. *Nucleic Acids Res.* **22,** 2990-2997.

Clermont,Y. (1972). Kinetics of spermatogenesis in mammals: seminiferous epithelium cycle and spermatogonial renewal. *Physiol Rev.* **52,** 198-236.

Clermont,Y. and Bustos-Obregon,E. (1968). Re-examination of spermatogonial renewal in the rat by means of seminiferous tubules mounted "in toto". *Am. J. Anat.* **122,** 237-247.

Coffigny,H., Bourgeois,C., Ricoul,M., Bernardino,J., Vilain,A., Niveleau,A., Malfoy,B., and Dutrillaux,B. (1999). Alterations of DNA methylation patterns in germ cells and Sertoli cells from developing mouse testis. *Cytogenet. Cell Genet.* **87,** 175-181.

Cooke,H.J. and Saunders,P.T. (2002). Mouse models of male infertility. *Nat. Rev. Genet.* **3,** 790-801.

Cooper,D.N. and Youssoufian,H. (1988). The CpG dinucleotide and human genetic disease. *Hum. Genet.* **78,** 151-155.

Costantini,M., Clay,O., Federico,C., Saccone,S., Auletta,F., and Bernardi,G. (2006). Human chromosomal bands: nested structure, high-definition map and molecular basis. *Chromosoma.*

da Rocha,S.T. and Ferguson-Smith,A.C. (2004). Genomic imprinting. *Curr. Biol.* **14,** R646-R649.

Dadoune,J.P., Siffroi,J.P., and Alfonsi,M.F. (2004). Transcription in haploid male germ cells. *Int. Rev. Cytol.* **237,** 1-56.

Das,P.M. and Singal,R. (2004). DNA methylation and cancer. *J. Clin. Oncol.* **22,** 4632-4642.

Davis,T.L., Trasler,J.M., Moss,S.B., Yang,G.J., and Bartolomei,M.S. (1999). Acquisition of the H19 methylation imprint occurs differentially on the parental alleles during spermatogenesis. *Genomics* **58,** 18-28.

Davis,T.L., Yang,G.J., McCarrey,J.R., and Bartolomei,M.S. (2000). The H19 methylation imprint is erased and re-established differentially on the parental alleles during male germ cell development. *Hum. Mol. Genet.* **9,** 2885-2894.

de Rooij,D.G. and Russell,L.D. (2000). All you wanted to know about spermatogonia but were afraid to ask. *J. Androl* **21,** 776-798.

De Smet,C., De,B.O., Faraoni,I., Lurquin,C., Brasseur,F., and Boon,T. (1996). The activation of human gene MAGE-1 in tumor cells is correlated with genome-wide demethylation. *Proc. Natl. Acad. Sci. U. S. A* **93,** 7149-7153.

Dev,V.G., Warburton,D., and Miller,O.J. (1972). Giemsa banding of chromosomes. *Lancet* **1,** 1285.

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

Doerksen,T. and Trasler,J.M. (1996). Developmental exposure of male germ cells to 5-azacytidine results in abnormal preimplantation development in rats. *Biol. Reprod.* **55,** 1155-1162.

Donovan,P.J., Stott,D., Godin,I., Heasman,J., and Wylie,C.C. (1987). Studies on the migration of mouse germ cells. *J. Cell Sci. Suppl* **8,** 359-367.

Duret,L., Eyre-Walker,A., and Galtier,N. (2006). A new perspective on isochore evolution. *Gene* **385,** 71-74.

Eckhardt,F., Lewin,J., Cortese,R., Rakyan,V.K., Attwood,J., Burger,M., Burton,J., Cox,T.V., Davies,R., Down,T.A. et al. (2006). DNA methylation profiling of human chromosomes 6, 20 and 22. *Nat. Genet.* **38,** 1378-1385.

Egger,G., Liang,G., Aparicio,A., and Jones,P.A. (2004). Epigenetics in human disease and prospects for epigenetic therapy. *Nature* **429,** 457-463.

Ehrlich,M. (2006). Cancer-linked DNA hypomethylation and its relationship to hypermethylation. *Curr. Top. Microbiol. Immunol.* **310,** 251-274.

Erbach,G.T., Lawitts,J.A., Papaioannou,V.E., and Biggers,J.D. (1994). Differential growth of the mouse preimplantation embryo in chemically defined media. *Biol. Reprod.* **50,** 1027-1033.

Ewing,L.L., Davis,J.C., and Zirkin,B.R. (1980). Regulation of testicular function: a spatial and temporal view. *Int. Rev. Physiol* **22,** 41-115.

Fang,J.S. and Jagiello,G. (1981). A pachytene map of the mouse spermatocyte. *Chromosoma* **82,** 437-445.

Fazzari,M.J. and Greally,J.M. (2004). Epigenomics: beyond CpG islands. *Nat. Rev. Genet.* **5,** 446-455.

Feinberg,A.P. (2004). The epigenetics of cancer etiology. *Semin. Cancer Biol.* **14,** 427-432.

Ferguson-Smith,A.C. and Surani,M.A. (2001). Imprinting and the epigenetic asymmetry between parental genomes. *Science* **293,** 1086-1089.

Fryxell,K.J. and Zuckerkandl,E. (2000). Cytosine deamination plays a primary role in the evolution of mammalian isochores. *Mol. Biol. Evol.* **17,** 1371-1383.

Gabbara,S. and Bhagwat,A.S. (1995). The mechanism of inhibition of DNA (cytosine-5-)-methyltransferases by 5-azacytosine is likely to involve methyl transfer to the inhibitor. *Biochem. J.* **307 (Pt 1),** 87-92.

Gaillard,C., Doly,J., Cortadas,J., and Bernardi,G. (1981). The primary structure of bovine satellite 1.715. *Nucleic Acids Res.* **9,** 6069-6082.

Gama-Sosa,M.A., Midgett,R.M., Slagel,V.A., Githens,S., Kuo,K.C., Gehrke,C.W., and Ehrlich,M. (1983). Tissue-specific differences in DNA methylation in various mammals. *Biochim. Biophys. Acta* **740,** 212-219.

Gardiner-Garden,M. and Frommer,M. (1987). CpG islands in vertebrate genomes. *J. Mol. Biol.* **196,** 261-282.

Geyer,C.B., Kiefer,C.M., Yang,T.P., and McCarrey,J.R. (2004). Ontogeny of a demethylation domain and its relationship to activation of tissue-specific transcription. *Biol. Reprod.* **71,** 837-844.

Gilbert,S.F. (2000). *Developmental Biology*. Sunderland, MA: Sinauer Associates, Inc.

Glazer,R.I. and Knode,M.C. (1984). 1-beta-D-arabinosyl-5-azacytosine. Cytocidal activity and effects on the synthesis and methylation of DNA in human colon carcinoma cells. *Mol. Pharmacol.* **26,** 381-387.

Glazko,G.V., Koonin,E.V., Rogozin,I.B., and Shabalina,S.A. (2003). A significant fraction of conserved noncoding DNA in human and mouse consists of predicted matrix attachment regions. *Trends Genet.* **19,** 119-124.

Goll,M.G. and Bestor,T.H. (2005). Eukaryotic cytosine methyltransferases. *Annu. Rev. Biochem.* **74,** 481-514.

Goll,M.G., Kirpekar,F., Maggert,K.A., Yoder,J.A., Hsieh,C.L., Zhang,X., Golic,K.G., Jacobsen,S.E., and Bestor,T.H. (2006). Methylation of tRNAAsp by the DNA methyltransferase homolog Dnmt2. *Science* **311,** 395-398.

Gonzalgo,M.L. and Jones,P.A. (1997). Rapid quantitation of methylation differences at specific sites using methylation-sensitive single nucleotide primer extension (Ms-SNuPE). *Nucleic Acids Res.* **25,** 2529-2531.

Griswold,M.D. (1995). Interactions between germ cells and Sertoli cells in the testis. *Biol. Reprod.* **52,** 211-216.

Griswold,M.D. and Kim,J.S. (2001). Site-specific methylation of the promoter alters deoxyribonucleic acid-protein interactions and prevents follicle-stimulating hormone receptor gene transcription. *Biol. Reprod.* **64,** 602-610.

Grootegoed,J.A., Siep,M., and Baarends,W.M. (2000). Molecular and cellular mechanisms in spermatogenesis. *Baillieres Best. Pract. Res. Clin. Endocrinol. Metab* **14,** 331-343.

Gruenbaum,Y., Margalit,A., Goldman,R.D., Shumaker,D.K., and Wilson,K.L. (2005). The nuclear lamina comes of age. *Nat. Rev. Mol. Cell Biol.* **6,** 21-31.

Grunau,C., Clark,S.J., and Rosenthal,A. (2001). Bisulfite genomic sequencing: systematic investigation of critical experimental parameters. *Nucleic Acids Res.* **29,** E65.

Haaf,T. (1995). The effects of 5-azacytidine and 5-azadeoxycytidine on chromosome structure and function: implications for methylation-associated cellular processes. *Pharmacol. Ther.* **65,** 19-46.

Haines,T.R., Rodenhiser,D.I., and Ainsworth,P.J. (2001). Allele-specific non-CpG methylation of the Nf1 gene during early mouse development. *Dev. Biol.* **240,** 585-598.

Hajkova,P., Erhardt,S., Lane,N., Haaf,T., El Maarri,O., Reik,W., Walter,J., and Surani,M.A. (2002). Epigenetic reprogramming in mouse primordial germ cells. *Mech. Dev.* **117,** 15-23.

Hata,K., Kusumi,M., Yokomine,T., Li,E., and Sasaki,H. (2006). Meiotic and epigenetic aberrations in Dnmt3L-deficient male germ cells. *Mol. Reprod. Dev.* **73,** 116-122.

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

Hattori,N., Abe,T., Hattori,N., Suzuki,M., Matsuyama,T., Yoshida,S., Li,E., and Shiota,K. (2004). Preference of DNA methyltransferases for CpG islands in mouse embryonic stem cells. *Genome Res.* **14,** 1733-1740.

Hayashi,K., Yoshida,K., and Matsui,Y. (2005). A histone H3 methyltransferase controls epigenetic events required for meiotic prophase. *Nature* **438,** 374-378.

Hayashizaki,Y., Hirotsune,S., Okazaki,Y., Hatada,I., Shibata,H., Kawai,J., Hirose,K., Watanabe,S., Fushiki,S., Wada,S. et al. (1993). Restriction landmark genomic scanning method and its various applications. *Electrophoresis* **14,** 251-258.

Hecht,N.B. (1998). Molecular mechanisms of male germ cell differentiation. *Bioessays* **20,** 555-561.

Hellmann-Blumberg,U., Hintz,M.F., Gatewood,J.M., and Schmid,C.W. (1993). Developmental differences in methylation of human Alu repeats. *Mol. Cell Biol.* **13,** 4523-4530.

Hendrich,B., Hardeland,U., Ng,H.H., Jiricny,J., and Bird,A. (1999). The thymine glycosylase MBD4 can bind to the product of deamination at methylated CpG sites. *Nature* **401,** 301-304.

Hiura,H., Obata,Y., Komiyama,J., Shirai,M., and Kono,T. (2006). Oocyte growth-dependent progression of maternal imprinting in mice. *Genes Cells* **11,** 353-361.

Hogan,B., Constantini,F., and Lacy,E. (1986). *Manipulating the Mouse Embryo: A Laboratory Manual.* Cold Spring Harbor, NY: Cold Spring Harbor Press.

Holliday,R. and Pugh,J.E. (1975). DNA modification mechanisms and gene activity during development. *Science* **187,** 226-232.

Hsieh,C.L. (2005). The de novo methylation activity of Dnmt3a is distinctly different than that of Dnmt1. *BMC. Biochem.* **6,** 6.

Huckins,C. (1971). The spermatogonial stem cell population in adult rats. I. Their morphology, proliferation and maturation. *Anat. Rec.* **169,** 533-557.

Huynh,K.D. and Lee,J.T. (2003). Inheritance of a pre-inactivated paternal X chromosome in early mouse embryos. *Nature* **426,** 857-862.

Iannello,R.C., Young,J., Sumarsono,S., Tymms,M.J., Dahl,H.H., Gould,J., Hedger,M., and Kola,I. (1997). Regulation of Pdha-2 expression is mediated by proximal promoter sequences and CpG methylation. *Mol. Cell Biol.* **17,** 612-619.

Imamura,T., Kerjean,A., Heams,T., Kupiec,J.J., Thenevin,C., and Paldi,A. (2005). Dynamic CpG and non-CpG methylation of the Peg1/Mest gene in the mouse oocyte and preimplantation embryo. *J. Biol. Chem.* **280,** 20171-20175.

Ioshikhes,I.P. and Zhang,M.Q. (2000). Large-scale human promoter mapping using CpG islands. *Nat. Genet.* **26,** 61-63.

Ishida,C., Ura,K., Hirao,A., Sasaki,H., Toyoda,A., Sakaki,Y., Niwa,H., Li,E., and Kaneda,Y. (2003). Genomic organization and promoter analysis of the Dnmt3b gene. *Gene* **310,** 151-159.

Issa,J.P. (2000). CpG-island methylation in aging and cancer. *Curr. Top. Microbiol. Immunol.* **249,** 101-118.

Jahnukainen,K., Chrysis,D., Hou,M., Parvinen,M., Eksborg,S., and Soder,O. (2004). Increased apoptosis occurring during the first wave of spermatogenesis is stage-specific and primarily affects midpachytene spermatocytes in the rat testis. *Biol. Reprod.* **70,** 290-296.

Jenuwein,T. and Allis,C.D. (2001). Translating the histone code. *Science* **293,** 1074-1080.

Jost,J.P. and Jost,Y.C. (1995). Mechanism of active DNA demethylation during embryonic development and cellular differentiation in vertebrates. *Gene* **157,** 265-266.

Jue,K., Bestor,T.H., and Trasler,J.M. (1995). Regulated synthesis and localization of DNA methyltransferase during spermatogenesis. *Biol. Reprod.* **53,** 561-569.

Jurka,J., Kapitonov,V.V., Pavlicek,A., Klonowski,P., Kohany,O., and Walichiewicz,J. (2005). Repbase Update, a database of eukaryotic repetitive elements. *Cytogenet. Genome Res.* **110,** 462-467.

Juttermann,R., Li,E., and Jaenisch,R. (1994). Toxicity of 5-aza-2'-deoxycytidine to mammalian cells is mediated primarily by covalent trapping of DNA methyltransferase rather than DNA demethylation. *Proc. Natl. Acad. Sci. U. S. A* **91,** 11797-11801.

Kafri,T., Ariel,M., Brandeis,M., Shemer,R., Urven,L., McCarrey,J., Cedar,H., and Razin,A. (1992). Developmental pattern of gene-specific DNA methylation in the mouse embryo and germ line. *Genes Dev.* **6,** 705-714.

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

Kelly,T.L., Li,E., and Trasler,J.M. (2003). 5-aza-2'-deoxycytidine induces alterations in murine spermatogenesis and pregnancy outcome. *J. Androl* **24,** 822-830.

Kent,W.J., Sugnet,C.W., Furey,T.S., Roskin,K.M., Pringle,T.H., Zahler,A.M., and Haussler,D. (2002). The human genome browser at UCSC. *Genome Res.* **12,** 996-1006.

Kidwell,M.G. (2002). Transposable elements and the evolution of genome size in eukaryotes. *Genetica* **115,** 49-63.

Klose,R.J. and Bird,A.P. (2006). Genomic DNA methylation: the mark and its mediators. *Trends Biochem. Sci.* **31,** 89-97.

Kochanek,S., Renz,D., and Doerfler,W. (1993). DNA methylation in the Alu sequences of diploid and haploid primary human cells. *EMBO J.* **12,** 1141-1151.

Kondo,T., Bobek,M.P., Kuick,R., Lamb,B., Zhu,X., Narayan,A., Bourc'his,D., Viegas-Pequignot,E., Ehrlich,M., and Hanash,S.M. (2000). Whole-genome methylation scan in ICF syndrome: hypomethylation of non-satellite DNA repeats D4Z4 and NBL2. *Hum. Mol. Genet.* **9,** 597-604.

La Salle,S., Mertineit,C., Taketo,T., Moens,P.B., Bestor,T.H., and Trasler,J.M. (2004). Windows for sex-specific methylation marked by DNA methyltransferase expression profiles in mouse germ cells. *Dev. Biol.* **268,** 403-415.

La Salle,S., Oakes,C.C., Neaga,O.R., Bourc'his,D., Bestor,T.H., and Trasler,J.M. (2007). Wide-spread DNA methylation defects and loss of spermatogonia in DNMT3L-deficient neonatal male mice. *Submitted.*

La Salle,S. and Trasler,J.M. (2006b). Epigenetic patterning in male germ cells: importance to progeny outcome. In *The Sperm Cell: Production, Maturation, Fertilization, Regeneration.* (ed. De Jonge,C. and Barratt,C.), New York: Cambridge University Press.

La Salle,S. and Trasler,J.M. (2006a). Dynamic expression of DNMT3a and DNMT3b isoforms during male germ cell development in the mouse. *Dev. Biol.* **296,** 71-82.

Laird,P.W., Jackson-Grusby,L., Fazeli,A., Dickinson,S.L., Jung,W.E., Li,E., Weinberg,R.A., and Jaenisch,R. (1995). Suppression of intestinal neoplasia by DNA hypomethylation. *Cell* **81,** 197-205.

Lees-Murdock,D.J., De Felici,M., and Walsh,C.P. (2003). Methylation dynamics of repetitive DNA elements in the mouse germ cell lineage. *Genomics* **82,** 230-237.

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

Lei,H., Oh,S.P., Okano,M., Juttermann,R., Goss,K.A., Jaenisch,R., and Li,E. (1996). De novo DNA cytosine methyltransferase activities in mouse embryonic stem cells. *Development* **122,** 3195-3205.

Li,E., Beard,C., and Jaenisch,R. (1993). Role for DNA methylation in genomic imprinting. *Nature* **366,** 362-365.

Li,E., Bestor,T.H., and Jaenisch,R. (1992). Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell* **69,** 915-926.

Li,J.Y., Lees-Murdock,D.J., Xu,G.L., and Walsh,C.P. (2004). Timing of establishment of paternal methylation imprints in the mouse. *Genomics* **84,** 952-960.

Li,Q., Lee,B.T., and Zhang,L. (2005). Genome-scale analysis of positional clustering of mouse testis-specific genes. *BMC. Genomics* **6,** 7.

Lippman,Z., Gendrel,A.V., Black,M., Vaughn,M.W., Dedhia,N., McCombie,W.R., Lavine,K., Mittal,V., May,B., Kasschau,K.D. et al. (2004). Role of transposable elements in heterochromatin and epigenetic control. *Nature* **430,** 471-476.

Lippman,Z. and Martienssen,R. (2004). The role of RNA interference in heterochromatic silencing. *Nature* **431,** 364-370.

Liu,W.M., Maraia,R.J., Rubin,C.M., and Schmid,C.W. (1994). Alu transcripts: cytoplasmic localisation and regulation by DNA methylation. *Nucleic Acids Res.* **22,** 1087-1095.

Lo,Y.M., Wong,I.H., Zhang,J., Tein,M.S., Ng,M.H., and Hjelm,N.M. (1999). Quantitative analysis of aberrant p16 methylation using real-time quantitative methylation-specific polymerase chain reaction. *Cancer Res.* **59,** 3899-3903.

Lucifero,D., Mann,M.R., Bartolomei,M.S., and Trasler,J.M. (2004). Gene-specific timing and epigenetic memory in oocyte imprinting. *Hum. Mol. Genet.* **13,** 839-849.

Maatouk,D.M., Kellam,L.D., Mann,M.R., Lei,H., Li,E., Bartolomei,M.S., and Resnick,J.L. (2006). DNA methylation is a primary mechanism for silencing postmigratory primordial germ cell genes in both germ cell and somatic cell lineages. *Development* **133,** 3411-3418.

MacLean,J.A. and Wilkinson,M.F. (2005). Gene regulation in spermatogenesis. *Curr. Top. Dev. Biol.* **71,** 131-197.

Maloisel,L. and Rossignol,J.L. (1998). Suppression of crossing-over by DNA methylation in Ascobolus. *Genes Dev.* **12,** 1381-1389.

Marahrens,Y., Panning,B., Dausman,J., Strauss,W., and Jaenisch,R. (1997). Xist-deficient mice are defective in dosage compensation but not spermatogenesis. *Genes Dev.* **11,** 156-166.

McGrath,J. and Solter,D. (1984). Completion of mouse embryogenesis requires both the maternal and paternal genomes. *Cell* **37,** 179-183.

McIntosh,L.P., Grieger,I., Eckstein,F., Zarling,D.A., van de Sande,J.H., and Jovin,T.M. (1983). Left-handed helical conformation of poly[d(A-m5C).d(G-T)]. *Nature* **304,** 83-86.

McLaren,A. (1995). Germ cells and germ cell sex. *Philos. Trans. R. Soc. Lond B Biol. Sci.* **350,** 229-233.

Mertineit,C., Yoder,J.A., Taketo,T., Laird,D.W., Trasler,J.M., and Bestor,T.H. (1998). Sex-specific exons control DNA methyltransferase in mammalian germ cells. *Development* **125,** 889-897.

Michaud,E.J., van Vugt,M.J., Bultman,S.J., Sweet,H.O., Davisson,M.T., and Woychik,R.P. (1994). Differential expression of a new dominant agouti allele (Aiapy) is correlated with methylation state and is influenced by parental lineage. *Genes Dev.* **8,** 1463-1472.

Moens,P.B., Pearlman,R.E., Heng,H.H., and Traut,W. (1998). Chromosome cores and chromatin at meiotic prophase. *Curr. Top. Dev. Biol.* **37,** 241-262.

Mondal,S. and Heidelberger,C. (1980). Inhibition of induced differentiation of C3H/10T 1/2 clone 8 mouse embryo cells by tumor promoters. *Cancer Res.* **40,** 334-338.

Monk,M., Boubelik,M., and Lehnert,S. (1987). Temporal and regional changes in DNA methylation in the embryonic, extraembryonic and germ cell lineages during mouse embryo development. *Development* **99,** 371-382.

Moore,T. (2001). Genetic conflict, genomic imprinting and establishment of the epigenotype in relation to growth. *Reproduction.* **122,** 185-193.

Morgan,H.D., Sutherland,H.G., Martin,D.I., and Whitelaw,E. (1999). Epigenetic inheritance at the agouti locus in the mouse. *Nat. Genet.* **23,** 314-318.

Mruk,D.D. and Cheng,C.Y. (2004). Sertoli-Sertoli and Sertoli-germ cell interactions and their significance in germ cell movement in the seminiferous epithelium during spermatogenesis. *Endocr. Rev.* **25,** 747-806.

Muller-Tidow,C., Bornemann,C., Diederichs,S., Westermann,A., Klumpen,S., Zuo,P., Wang,W., Berdel,W.E., and Serve,H. (2001). Analyses of the genomic methylation status of the human cyclin A1 promoter by a novel real-time PCR-based methodology. *FEBS Lett.* **490,** 75-78.

Nagano,R., Tabata,S., Nakanishi,Y., Ohsako,S., Kurohmaru,M., and Hayashi,Y. (2000). Reproliferation and relocation of mouse male germ cells (gonocytes) during prespermatogenesis. *Anat. Rec.* **258,** 210-220.

Namekawa,S.H., Park,P.J., Zhang,L.F., Shima,J.E., McCarrey,J.R., Griswold,M.D., and Lee,J.T. (2006). Postmeiotic sex chromatin in the male germline of mice. *Curr. Biol.* **16,** 660-667.

Neddermann,P., Gallinari,P., Lettieri,T., Schmid,D., Truong,O., Hsuan,J.J., Wiebauer,K., and Jiricny,J. (1996). Cloning and expression of human G/T mismatch-specific thymine-DNA glycosylase. *J. Biol. Chem.* **271,** 12767-12774.

Nimura,K., Ishida,C., Koriyama,H., Hata,K., Yamanaka,S., Li,E., Ura,K., and Kaneda,Y. (2006). Dnmt3a2 targets endogenous Dnmt3L to ES cell chromatin and induces regional DNA methylation. *Genes Cells* **11,** 1225-1237.

Numata,M., Ono,T., and Iseki,S. (1994). Expression and localization of the mRNA for DNA (cytosine-5)- methyltransferase in mouse seminiferous tubules. *J. Histochem. Cytochem.* **42,** 1271-1276.

O'Neill,R.J., O'Neill,M.J., and Graves,J.A. (1998). Undermethylation associated with retroelement activation and chromosome remodelling in an interspecific mammalian hybrid. *Nature* **393,** 68-72.

Oakberg,E.F. (1971). Spermatogonial stem-cell renewal in the mouse. *Anat. Rec.* **169,** 515-531.

Oakes,C.C., La Salle,S., Robaire,B., and Trasler,J.M. (2006). Evaluation of a Quantitative DNA Methylation Analysis Technique using Methylation-Sensitive/Dependent Restriction Enzymes and Real-Time PCR. *Epigenetics* **1,** 146-152.

Oakes,C.C., La Salle,S., Smiraglia,D.J., Robaire,B., and Trasler,J.M. (2007a). Developmental Acquisition of Genome-Wide DNA Methylation Occurs Prior to Meiosis in Male Germ Cells. *Developmental Biology, in review.*

Oakes,C.C., La,S.S., Smiraglia,D.J., Robaire,B., and Trasler,J.M. (2007b). A unique configuration of genome-wide DNA methylation patterns in the testis. *Proc. Natl. Acad. Sci. U. S. A* **104,** 228-233.

Oakes,C.C., Smiraglia,D.J., Plass,C., Trasler,J.M., and Robaire,B. (2003). Aging results in hypermethylation of ribosomal DNA in sperm and liver of male rats. *Proc. Natl. Acad. Sci. U. S. A* **100,** 1775-1780.

Oka,M., Meacham,A.M., Hamazaki,T., Rodic,N., Chang,L.J., and Terada,N. (2005). De novo DNA methyltransferases Dnmt3a and Dnmt3b primarily mediate the cytotoxic effect of 5-aza-2'-deoxycytidine. *Oncogene* **24,** 3091-3099.

Okano,M., Bell,D.W., Haber,D.A., and Li,E. (1999). DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell* **99,** 247-257.

Okano,M., Xie,S., and Li,E. (1998). Dnmt2 is not required for de novo and maintenance methylation of viral DNA in embryonic stem cells. *Nucleic Acids Res.* **26,** 2536-2540.

Okazaki,Y., Okuizumi,H., Sasaki,N., Ohsumi,T., Kuromitsu,J., Hirota,N., Muramatsu,M., and Hayashizaki,Y. (1995). An expanded system of restriction landmark genomic scanning (RLGS Ver. 1.8). *Electrophoresis* **16,** 197-202.

Ordway,J.M., Bedell,J.A., Citek,R.W., Nunberg,A., Garrido,A., Kendall,R., Stevens,J.R., Cao,D., Doerge,R.W., Korshunova,Y. et al. (2006). Comprehensive DNA methylation profiling in a human cancer genome identifies novel epigenetic targets. *Carcinogenesis* **27,** 2409-2423.

Orwig,K.E., Ryu,B.Y., Avarbock,M.R., and Brinster,R.L. (2002). Male germ-line stem cell potential is predicted by morphology of cells in neonatal rat testes. *Proc. Natl. Acad. Sci. U. S. A* **99,** 11706-11711.

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

Pan,H., O'Brien,M.J., Wigglesworth,K., Eppig,J.J., and Schultz,R.M. (2005). Transcript profiling during mouse oocyte development and the effect of gonadotropin priming and development in vitro. *Dev. Biol.* **286,** 493-506.

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

Park,J.G. and Chapman,V.M. (1994). CpG island promoter region methylation patterns of the inactive-X-chromosome hypoxanthine phosphoribosyltransferase (Hprt) gene. *Mol. Cell Biol.* **14,** 7975-7983.

Peaston,A.E., Evsikov,A.V., Graber,J.H., De Vries,W.N., Holbrook,A.E., Solter,D., and Knowles,B.B. (2004). Retrotransposons regulate host genes in mouse oocytes and preimplantation embryos. *Dev. Cell* **7,** 597-606.

Rakyan,V.K., Chong,S., Champ,M.E., Cuthbert,P.C., Morgan,H.D., Luu,K.V., and Whitelaw,E. (2003). Transgenerational inheritance of epigenetic states at the murine Axin(Fu) allele occurs after maternal and paternal transmission. *Proc. Natl. Acad. Sci. U. S. A* **100,** 2538-2543.

Ramsahoye,B.H., Biniszkiewicz,D., Lyko,F., Clark,V., Bird,A.P., and Jaenisch,R. (2000). Non-CpG methylation is prevalent in embryonic stem cells and may be mediated by DNA methyltransferase 3a. *Proc. Natl. Acad. Sci. U. S. A* **97,** 5237-5242.

Reik,W., Dean,W., and Walter,J. (2001). Epigenetic reprogramming in mammalian development. *Science* **293,** 1089-1093.

Reik,W. and Walter,J. (2001). Genomic imprinting: parental influence on the genome. *Nat. Rev. Genet.* **2,** 21-32.

Riggs,A.D. (1975). X inactivation, differentiation, and DNA methylation. *Cytogenet. Cell Genet.* **14,** 9-25.

Robertson,K.D. (2005). DNA methylation and human disease. *Nat. Rev. Genet.* **6,** 597-610.

Robertson,K.D., Keyomarsi,K., Gonzales,F.A., Velicescu,M., and Jones,P.A. (2000). Differential mRNA expression of the human DNA methyltransferases (DNMTs) 1, 3a and 3b during the G(0)/G(1) to S phase transition in normal and tumor cells. *Nucleic Acids Res.* **28,** 2108-2113.

Rollins,R.A., Haghighi,F., Edwards,J.R., Das,R., Zhang,M.Q., Ju,J., and Bestor,T.H. (2006). Large-scale structure of genomic methylation patterns. *Genome Res.* **16,** 157-163.

Rossant,J., Sanford,J.P., Chapman,V.M., and Andrews,G.K. (1986). Undermethylation of structural gene sequences in extraembryonic lineages of the mouse. *Dev. Biol.* **117,** 567-573.

Rountree,M.R., Bachman,K.E., and Baylin,S.B. (2000). DNMT1 binds HDAC2 and a new co-repressor, DMAP1, to form a complex at replication foci. *Nat. Genet.* **25,** 269-277.

Rush,L.J. and Plass,C. (2002). Restriction landmark genomic scanning for DNA methylation in cancer: past, present, and future applications. *Anal. Biochem.* **307,** 191-201.

Russell,L.D., Ettlin,R., Sinha-Hikim,A.P., and Clegg,E. (1990). *Histological and histopathological evaluation of the testis.* Clearwater, FL: Cache River Press.

Saccone,S., De,S.A., Wiegant,J., Raap,A.K., Della,V.G., and Bernardi,G. (1993). Correlations between isochores and chromosomal bands in the human genome. *Proc. Natl. Acad. Sci. U. S. A* **90,** 11929-11933.

Sado,T. and Ferguson-Smith,A.C. (2005). Imprinted X inactivation and reprogramming in the preimplantation mouse embryo. *Hum. Mol. Genet.* **14 Spec No 1,** R59-R64.

Sado,T., Okano,M., Li,E., and Sasaki,H. (2004). De novo DNA methylation is dispensable for the initiation and propagation of X chromosome inactivation. *Development* **131,** 975-982.

Sanford,J., Forrester,L., Chapman,V., Chandley,A., and Hastie,N. (1984). Methylation patterns of repetitive DNA sequences in germ cells of Mus musculus. *Nucleic Acids Res.* **12,** 2823-2836.

Sanford,J.P., Clark,H.J., Chapman,V.M., and Rossant,J. (1987). Differences in DNA methylation during oogenesis and spermatogenesis and their persistence during early embryogenesis in the mouse. *Genes Dev.* **1,** 1039-1046.

Sansom,O.J., Berger,J., Bishop,S.M., Hendrich,B., Bird,A., and Clarke,A.R. (2003). Deficiency of Mbd2 suppresses intestinal tumorigenesis. *Nat. Genet.* **34,** 145-147.

Santi,D.V., Norment,A., and Garrett,C.E. (1984). Covalent bond formation between a DNA-cytosine methyltransferase and DNA containing 5-azacytosine. *Proc. Natl. Acad. Sci. U. S. A* **81,** 6993-6997.

Santoro,R. and Grummt,I. (2001). Molecular mechanisms mediating methylation-dependent silencing of ribosomal gene transcription. *Mol. Cell* **8,** 719-725.

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

Schneider-Stock,R., Diab-Assef,M., Rohrbeck,A., Foltzer-Jourdainne,C., Boltze,C., Hartig,R., Schonfeld,P., Roessner,A., and Gali-Muhtasib,H. (2005). 5-Aza-cytidine is a potent inhibitor of DNA methyltransferase 3a and induces apoptosis in HCT-116 colon cancer cells via Gadd45- and p53-dependent mechanisms. *J. Pharmacol. Exp. Ther.* **312,** 525-536.

Schumacher,A., Kapranov,P., Kaminsky,Z., Flanagan,J., Assadzadeh,A., Yau,P., Virtanen,C., Winegarden,N., Cheng,J., Gingeras,T. et al. (2006). Microarray-based DNA methylation profiling: technology and applications. *Nucleic Acids Res.* **34,** 528-542.

Shapiro,J.A. (2005). Retrotransposons and regulatory suites. *Bioessays* **27,** 122-125.

Sharpe,R. (1994). Regulation of Spermatogenesis. In *Physiology of Reproduction.* (ed. Knobil,E. and Neill,J.D.), New York: Raven Press Ltd.

Sheehy,A.M., Gaddis,N.C., Choi,J.D., and Malim,M.H. (2002). Isolation of a human gene that inhibits HIV-1 infection and is suppressed by the viral Vif protein. *Nature* **418,** 646-650.

Shemer,R., Birger,Y., Riggs,A.D., and Razin,A. (1997). Structure of the imprinted mouse Snrpn gene and establishment of its parental-specific methylation pattern. *Proc. Natl. Acad. Sci. U. S. A* **94,** 10267-10272.

Shen,J.C., Rideout,W.M., III, and Jones,P.A. (1994). The rate of hydrolytic deamination of 5-methylcytosine in double-stranded DNA. *Nucleic Acids Res.* **22,** 972-976.

Shibata,H., Yoshino,K., Sunahara,S., Gondo,Y., Katsuki,M., Ueda,T., Kamiya,M., Muramatsu,M., Murakami,Y., Kalcheva,I. et al. (1996). Inactive allele-specific methylation and chromatin structure of the imprinted gene U2af1-rs1 on mouse chromosome 11. *Genomics* **35,** 248-252.

Shima,J.E., McLean,D.J., McCarrey,J.R., and Griswold,M.D. (2004). The murine testicular transcriptome: characterizing gene expression in the testis during the progression of spermatogenesis. *Biol. Reprod.* **71,** 319-330.

Shiota,K., Kogo,Y., Ohgane,J., Imamura,T., Urano,A., Nishino,K., Tanaka,S., and Hattori,N. (2002). Epigenetic marks by DNA methylation specific to stem, germ and somatic cells in mice. *Genes Cells* **7,** 961-969.

Shovlin,T.C., Bourc'his,D., La Salle,S., O'doherty,A., Trasler,J.M., Bestor,T.H., and Walsh,C.P. (2006). Sex-specific promoters regulate Dnmt3L expression in mouse germ cells. *Hum. Reprod.*

Simpson,A.J., Caballero,O.L., Jungbluth,A., Chen,Y.T., and Old,L.J. (2005). Cancer/testis antigens, gametogenesis and cancer. *Nat. Rev. Cancer* **5,** 615-625.

Singer-Sam,J., Grant,M., LeBon,J.M., Okuyama,K., Chapman,V., Monk,M., and Riggs,A.D. (1990). Use of a HpaII-polymerase chain reaction assay to study DNA methylation in the Pgk-1 CpG island of mouse embryos at the time of X-chromosome inactivation. *Mol. Cell Biol.* **10,** 4987-4989.

Siu,M.K. and Cheng,C.Y. (2004). Dynamic cross-talk between cells and the extracellular matrix in the testis. *Bioessays* **26,** 978-992.

Smiraglia,D.J., Fruhwald,M.C., Costello,J.F., McCormick,S.P., Dai,Z., Peltomaki,P., O'dorisio,M.S., Cavenee,W.K., and Plass,C. (1999). A new tool for the rapid cloning of amplified and hypermethylated human DNA sequences from restriction landmark genome scanning gels. *Genomics* **58,** 254-262.

Smiraglia,D.J., Kazhiyur-Mannar,R., Oakes,C.C., Trasler,J.M., Wenger,R., and Plass,C. (2007). RLGS spot identification by second generation virtual RLGS in multiple genomes with multiple enzyme combinations. *in preparation.*

Song,F., Smith,J.F., Kimura,M.T., Morrow,A.D., Matsuyama,T., Nagase,H., and Held,W.A. (2005). Association of tissue-specific differentially methylated regions (TDMs) with differential gene expression. *Proc. Natl. Acad. Sci. U. S. A* **102,** 3336-3341.

Steinberg,R.A. and Gorman,K.B. (1992). Linked spontaneous CG----TA mutations at CpG sites in the gene for protein kinase regulatory subunit. *Mol. Cell Biol.* **12,** 767-772.

Stewart,F.J., Panne,D., Bickle,T.A., and Raleigh,E.A. (2000). Methyl-specific DNA binding by McrBC, a modification-dependent restriction enzyme. *J. Mol. Biol.* **298,** 611-622.

Su,A.I., Wiltshire,T., Batalov,S., Lapp,H., Ching,K.A., Block,D., Zhang,J., Soden,R., Hayakawa,M., Kreiman,G. et al. (2004). A gene atlas of the mouse and human protein-encoding transcriptomes. *Proc. Natl. Acad. Sci. U. S. A* **101,** 6062-6067.

Suetake,I., Morimoto,Y., Fuchikami,T., Abe,K., and Tajima,S. (2006). Stimulation effect of Dnmt3L on the DNA methylation activity of Dnmt3a2. *J. Biochem. (Tokyo)* **140,** 553-559.

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

Szyf,M., Bozovic,V., and Tanigawa,G. (1991). Growth regulation of mouse DNA methyltransferase gene expression. *J. Biol. Chem.* **266,** 10027-10030.

Tada,M., Tada,T., Lefebvre,L., Barton,S.C., and Surani,M.A. (1997). Embryonic germ cells induce epigenetic reprogramming of somatic nucleus in hybrid cells. *EMBO J.* **16,** 6510-6520.

Takada,S., Paulsen,M., Tevendale,M., Tsai,C.E., Kelsey,G., Cattanach,B.M., and Ferguson-Smith,A.C. (2002). Epigenetic analysis of the Dlk1-Gtl2 imprinted domain on mouse chromosome 12: implications for imprinting control from comparison with Igf2-H19. *Hum. Mol. Genet.* **11,** 77-86.

Takai,D. and Jones,P.A. (2002). Comprehensive analysis of CpG islands in human chromosomes 21 and 22. *Proc. Natl. Acad. Sci. U. S. A* **99,** 3740-3745.

Talbert,P.B. and Henikoff,S. (2006). Spreading of silent chromatin: inaction at a distance. *Nat. Rev. Genet.* **7,** 793-803.

Tam,P.P. and Snow,M.H. (1981). Proliferation and migration of primordial germ cells during compensatory growth in mouse embryos. *J. Embryol. Exp. Morphol.* **64,** 133-147.

Taylor,S.M. and Jones,P.A. (1982). Mechanism of action of eukaryotic DNA methyltransferase. Use of 5-azacytosine-containing DNA. *J. Mol. Biol.* **162,** 679-692.

Thomas,C.A., Jr. (1971). The genetic organization of chromosomes. *Annu. Rev. Genet.* **5,** 237-256.

Tollefsbol,T.O. (2004). Methods of epigenetic analysis. *Methods Mol. Biol.* **287,** 1-8.

Trasler,J.M., Alcivar,A.A., Hake,L.E., Bestor,T., and Hecht,N.B. (1992). DNA methyltransferase is developmentally expressed in replicating and non-replicating male germ cells. *Nucleic Acids Res.* **20,** 2541-2545.

Trasler,J.M., Hake,L.E., Johnson,P.A., Alcivar,A.A., Millette,C.F., and Hecht,N.B. (1990). DNA methylation and demethylation events during meiotic prophase in the mouse testis. *Mol. Cell Biol.* **10,** 1828-1834.

Tremblay,K.D., Saam,J.R., Ingram,R.S., Tilghman,S.M., and Bartolomei,M.S. (1995). A paternal-specific methylation imprint marks the alleles of the mouse H19 gene. *Nat. Genet.* **9,** 407-413.

Tweedie,S., Charlton,J., Clark,V., and Bird,A. (1997). Methylation of genomes and genes at the invertebrate-vertebrate boundary. *Mol. Cell Biol.* **17,** 1469-1475.

Ueda,T., Abe,K., Miura,A., Yuzuriha,M., Zubair,M., Noguchi,M., Niwa,K., Kawase,Y., Kono,T., Matsuda,Y. et al. (2000). The paternal methylation imprint of the mouse H19 locus is acquired in the gonocyte stage during foetal testis development. *Genes Cells* **5,** 649-659.

Vanyushin,B.F. (2006). DNA methylation in plants. *Curr. Top. Microbiol. Immunol.* **301,** 67-122.

Vergouwen,R.P., Jacobs,S.G., Huiskamp,R., Davids,J.A., and de Rooij,D.G. (1991). Proliferative activity of gonocytes, Sertoli cells and interstitial cells during testicular development in mice. *J. Reprod. Fertil.* **93,** 233-243.

Vermaak,D., Ahmad,K., and Henikoff,S. (2003). Maintenance of chromatin states: an open-and-shut case. *Curr. Opin. Cell Biol.* **15,** 266-274.

Viegas-Pequignot,E. and Dutrillaux,B. (1976). Segmentation of human chromosomes induced by 5-ACR (5-azacytidine). *Hum. Genet.* **34,** 247-254.

Walsh,C.P., Chaillet,J.R., and Bestor,T.H. (1998). Transcription of IAP endogenous retroviruses is constrained by cytosine methylation. *Nat. Genet.* **20,** 116-117.

Walsh,C.P. and Xu,G.L. (2006). Cytosine methylation and DNA repair. *Curr. Top. Microbiol. Immunol.* **301,** 283-315.

Warnecke,P.M., Mann,J.R., Frommer,M., and Clark,S.J. (1998). Bisulfite sequencing in preimplantation embryos: DNA methylation profile of the upstream region of the mouse imprinted H19 gene. *Genomics* **51,** 182-190.

Wasserman,P.M. and Albertini,D.F. (1994). The Mammalian Ovum. In *Physiology of Reproduction.* (ed. Knobil,E. and Neill,J.D.), New York: Raven Press, Ltd.

Webster,K.E., O'Bryan,M.K., Fletcher,S., Crewther,P.E., Aapola,U., Craig,J., Harrison,D.K., Aung,H., Phutikanit,N., Lyle,R. et al. (2005). Meiotic and epigenetic defects in Dnmt3L-knockout mouse spermatogenesis. *Proc. Natl. Acad. Sci. U. S. A* **102,** 4068-4073.

Weisenberger,D.J., Velicescu,M., Cheng,J.C., Gonzales,F.A., Liang,G., and Jones,P.A. (2004). Role of the DNA methyltransferase variant DNMT3b3 in DNA methylation. *Mol. Cancer Res.* **2,** 62-72.

Weisenberger,D.J., Velicescu,M., Preciado-Lopez,M.A., Gonzales,F.A., Tsai,Y.C., Liang,G., and Jones,P.A. (2002). Identification and characterization of alternatively spliced variants of DNA methyltransferase 3a in mammalian cells. *Gene* **298,** 91-99.

Weitzel,J.M., Buhrmester,H., and Stratling,W.H. (1997). Chicken MAR-binding protein ARBP is homologous to rat methyl-CpG-binding protein MeCP2. *Mol. Cell Biol.* **17,** 5656-5666.

Xu,G.L., Bestor,T.H., Bourc'his,D., Hsieh,C.L., Tommerup,N., Bugge,M., Hulten,M., Qu,X., Russo,J.J., and Viegas-Pequignot,E. (1999). Chromosome instability and immunodeficiency syndrome caused by mutations in a DNA methyltransferase gene. *Nature* **402,** 187-191.

Yamada,Y., Watanabe,H., Miura,F., Soejima,H., Uchiyama,M., Iwasaka,T., Mukai,T., Sakaki,Y., and Ito,T. (2004). A comprehensive analysis of allelic methylation status of CpG islands on human chromosome 21q. *Genome Res.* **14,** 247-266.

Yoder,J.A., Soman,N.S., Verdine,G.L., and Bestor,T.H. (1997a). DNA (cytosine-5)-methyltransferases in mouse cells and tissues. Studies with a mechanism-based probe. *J. Mol. Biol.* **270,** 385-395.

Yoder,J.A., Walsh,C.P., and Bestor,T.H. (1997b). Cytosine methylation and the ecology of intragenomic parasites. *Trends Genet.* **13,** 335-340.

Yoon,B.J., Herman,H., Sikora,A., Smith,L.T., Plass,C., and Soloway,P.D. (2002). Regulation of DNA methylation of Rasgrf1. *Nat. Genet.* **30,** 92-96.

Yoshida,S., Sukeno,M., Nakagawa,T., Ohbo,K., Nagamatsu,G., Suda,T., and Nabeshima,Y. (2006). The first round of mouse spermatogenesis is a distinctive program that lacks the self-renewing spermatogonia stage. *Development* **133,** 1495-1505.

Yoshimizu,T., Sugiyama,N., De Felice,M., Yeom,Y.I., Ohbo,K., Masuko,K., Obinata,M., Abe,K., Scholer,H.R., and Matsui,Y. (1999). Germline-specific expression of the Oct-4/green fluorescent protein (GFP) transgene in mice. *Dev. Growth Differ.* **41,** 675-684.

Yu,L., Liu,C., Bennett,K., Wu,Y.Z., Dai,Z., Vandeusen,J., Opavsky,R., Raval,A., Trikha,P., Rodriguez,B. et al. (2004). A NotI-EcoRV promoter library for studies of genetic and epigenetic alterations in mouse models of human malignancies. *Genomics* **84,** 647-660.

Zeschnigk,M., Bohringer,S., Price,E.A., Onadim,Z., Masshofer,L., and Lohmann,D.R. (2004). A novel real-time PCR assay for quantitative analysis of methylated alleles (QAMA): analysis of the retinoblastoma locus. *Nucleic Acids Res.* **32,** e125.

Zhang,Z., Joh,K., Yatsuki,H., Wang,Y., Arai,Y., Soejima,H., Higashimoto,K., Iwasaka,T., and Mukai,T. (2006). Comparative analyses of genomic imprinting and CpG island-methylation in mouse Murr1 and human MURR1 loci revealed a putative imprinting control region in mice. *Gene* **366,** 77-86.

FORMS & CERTIFICATES

Animal and Radioactive Use Compliance Forms

Reprint Permissions

Requests for Permission to Reprint

PNAS authors need not obtain permission for the following cases: (1) to use their original figures or tables in their future works; (2) to make copies of their papers for their own personal use, including classroom use, or for the personal use of colleagues, provided those copies are not for sale and are not distributed in a systematic way; (3) to include their papers as part of their dissertations; or (4) to use all or part of their article in a printed compilation of their own works. Citation of the original source must be included and copies must include the copyright notice of the original report.

Anyone may, without requesting permission, use original figures or tables published in PNAS for noncommercial and educational use (i.e., in a review article, in a book that is not for sale) provided that the original source and copyright notice are cited.

For permission to reprint material in Vols. 90 (1993) to present, requests must be sent via email, fax, or mail and include the following information about the original material:

1. Your full name, affiliation, and title.
2. Your complete mailing address, phone number, fax number, and e-mail address.
3. PNAS volume number, issue number, and issue date.
4. PNAS article title.
5. PNAS authors' names.
6. Page numbers of items to be reprinted.
7. Figure/table number or portion of text to be reprinted.

Also include the following information about the intended use of the material:

1. Title of work in which PNAS material will appear.
2. Authors/editors of work.
3. Publisher of work.
4. Retail price of work.
5. Number of copies of work to be produced.
6. Intended audience.
7. Whether work is for nonprofit or commercial use.

Authors whose work will be reused should be notified. PNAS cannot supply original artwork. Use of PNAS material must not imply any endorsement by PNAS or the National Academy of Sciences. The full journal reference must be cited and "Copyright YEAR National Academy of Sciences, U.S.A."

Requests for permission should be made in writing and addressed to: PNAS Permissions Editor, 500 Fifth Street NW, NAS 340, Washington, DC 20001 USA; phone 1-202-334-2679, fax 1-202-334-2739, e-mail PNASpermissions@nas.edu .

For permission to reprint material in Vols. 1-89 (1915-1992), requests should be addressed to the original authors, who hold the copyright. The full journal reference must be cited.

Hi Dr. Oakes,

Absolutely. Permission is granted. Actually, permission is granted for all
authors to reproduce their article or parts of it (as long as the original
article is properly cited) in the signed copyright transfer form.

Best, Kim

Kimberly Mitchell
Journal Publications Director
Landes Bioscience
1002 West Ave, 2nd Floor
Austin TX 78701
tel 512.637.6050
fax 512.637.6079
kmitchell@landesbioscience.com
http://www.landesbioscience.com/

From: "Christopher Charles Oakes, Mr" <christopher.oakes@mcgill.ca>
Date: Tue, 23 Jan 2007 11:41:12 -0500
To: <kmitchell@landesbioscience.com>
Conversation: reprint permission
Subject: reprint permission

Dear Kim Mitchell,
I am writing to request permission to reprint an article in my PhD dissertation that I have recently published in Epigenetics. I am the first author of the article and I am required by my university to obtain permission to reproduce the article within my PhD thesis:

Evaluation of a Quantitative DNA Methylation Analysis Technique using Methylation-Sensitive/Dependent Restriction Enzymes and Real-Time PCR
Christopher C Oakes, Sophie La Salle, Bernard Robaire and Jacquetta M Trasler
Volume: 1 | Issue: 3 | Pages: 146 - 152

Please let me know what needs to be done so that I can obtain this permission.

Sincerely,
Christopher Oakes
Graduate Student,
McGill University