University of the Witwatersrand
National Health Laboratory Service
Division of Human Genetics
School of Pathology

Effect of alcohol on global and locus specific DNA methylation in spermatozoa: implications for fetal alcohol spectrum disorders (FASD)

Sanam Patel
2012

A dissertation submitted to the Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, in fulfillment of the requirements for the degree in Master of Science (Medicine) in the Division of Human Genetics

Declaration

I, Sanam Patel, hereby declare this dissertation to be my own work. It is being submitted for the degree of Master of Science (Medicine) at the University of the Witwatersrand, Johannesburg. It has not been previously submitted for any degree or examination at this, or any other University.

Signed this 12th day of the month of July 2012.

_____

Acknowledgements

Acknowledgement goes to the following persons:

- Prof. Michèle Ramsay and Dr. Zané Lombard for their supervision.

- To the Cryobank, andrology laboratory and sperm bank, in Parktown, for sample collection.

- To the NHLS research trust and the NRF for funding and bursaries.

- To all participants for donating samples.

Abstract

Fetal alcohol spectrum disorders (FASD) is an umbrella term that describes a range of symptoms associated with prenatal alcohol exposure. Fetal Alcohol Syndrome (FAS) is the most severe disorder in the spectrum and is a major health problem in South Africa, with a prevalence rate of 68.0-89.2 per 1000 children of school-going age. The primary cause of FAS is *in utero* alcohol exposure. However, secondary factors that contribute to the syndrome include various genetic, epigenetic and additional environmental factors. The proposal that paternal preconception alcohol exposure has adverse effects on offspring development is supported by children born with FASD-like characteristics whose mothers did not drink but whose fathers were alcoholics. Mouse models further support these findings. One of the main epigenetic factors that have been shown to be affected by alcohol is DNA methylation. This chemical modification of DNA is associated with developmentally important genes known as imprinted genes. Imprinted genes are expressed in a parent of origin specific manner. Methylation occurs at specific regions in these genes known as differentially methylated regions (DMRs) or imprinting control regions (ICRs). Alcohol's ability to alter DNA methylation at imprinted genes raises the possibility that epigenetic disruption could contribute to the clinical features seen in FASD.

The main aim of this research was to examine global DNA methylation and locus specific *H19* ICR DNA methylation in spermatozoa, related to alcohol exposure. This was done using the luminometric methylation assay (LUMA) and bisulfite based quantitative pyrosequencing, respectively. In this study there was no significant correlation between alcohol exposure and global DNA methylation ($p = 0.17$), nor was there a significant correlation with drinking frequency ($p = 0.31$). Although not significant, a slight trend towards decreased global DNA methylation in alcohol-exposed spermatozoa was observed. This suggests that either alcohol does not affect global sperm DNA methylation or that the technique used in this

study was not sensitive enough to detect minor changes in global DNA methylation percentage.

There was also no significant correlation between alcohol exposure and average *H19* ICR DNA methylation ($p = 0.051$), nor was there a significant correlation with drinking frequency ($p = 0.47$). There was no significant correlation between alcohol exposure and DNA methylation at individual CpG sites except for CpG 3, where there was a significant increase in DNA methylation in the drinking group ($p = 0.03$).

The findings of this study together with the findings of significant selective demethylation at individual CpG sites within the *IG-DMR* from another study on the same sperm samples, suggest that alcohol may have the ability to affect DNA methylation levels in spermatozoa at certain loci within the sperm genome. However, these loci-specific effects are not reflected in global DNA methylation levels. These findings do not disprove the hypothesis that there is an epigenetic mechanism responsible for the paternal effects seen in FASD. Instead they suggest that the techniques used in this study were not sensitive enough to detect these changes in DNA methylation or alternatively, alcohol may be exerting its effects through other epigenetic mechanisms.

Table of Contents

Table of figures

List of tables

Glossary of terms

µl	Microlitre
A	Adenine
ARBD	Alcohol-related birth defects
ARND	Alcohol-related neurodevelopmental disorder
ART	Assisted reproductive technology
ATP	Adenosine triphosphate
au	Arbitrary units
bp	Base pairs
BWS	Beckwith-Wiedemann syndrome
C	Cytosine
cm	Centimetre
CNS	Central nervous system
CTCF	CCCTC-binding factor
ddH_2O	Deionised distilled water
dH_2O	Distilled water
dNTP	Deoxynucleotide triphosphate
DMR	Differentially methylated region
DNA	Deoxyribonucleic acid
DNMT	DNA methyltransferase
E	Enzyme
EDTA	Ethylene diamine tetraacetic acid
EtBr	Ethidium bromide
EtOH	Ethanol

FAE	Fetal alcohol effects
FAS	Fetal alcohol syndrome
FASD	Fetal alcohol spectrum disorders
g	Gram
G	Guanine
HDAC	Histone deacetylase protein
ICR	Imprinting control region
IVF	In vitro fertilization
L	Litre
LUMA	Luminometric methylation assay
M	Molar
mg	Milligram
$MgCl_2$	Magnesium chloride
min	Minutes
ml	Millilitre
mM	Millimolar
mRNA	Messenger ribonucleic acid
miRNA	Micro ribonucleic acid
NaOH	Sodium hydroxide
ng	Nanogram
°C	Degrees Celsius
NIAAA	National institute on alcohol abuse and alcoholism
PAE	Prenatal alcohol effects
PCR	Polymerase chain reaction
PSQ	Pyrosequencing

RNA	Ribonucleic acid
rpm	Revolutions per minute
RSA	Republic of South Africa
S	Substrate
SNP	Single nucleotide polymorphism
T	Thymine
Taq	*Thermus aquaticus polymerase*
TBE	Tris borate ethylene diamine tetraacetic acid
TRDMT1	Transfer RNA aspartic acid methyltransferase
tRNA	Transfer ribonucleic acid
UPD	Uniparental disomy
U	Units
V	Volts

Effect of alcohol on global and locus specific DNA methylation in spermatozoa: implications for fetal alcohol spectrum disorders (FASD)

1. Introduction

1.1. Fetal Alcohol Syndrome (FAS)

Alcohol is a teratogen, and consumption during pregnancy may have harmful effects on the developing fetus, an association that has been recorded since biblical times (Abel, 1997). Further recognition of the adverse effects of alcohol on fetal outcome occurred during the early 18th century gin epidemic in England (Sokol, 1981). Fetal Alcohol Spectrum Disorders (FASD) is a comprehensive term that covers a range of symptoms associated with all levels of prenatal alcohol exposure (National Institute on Alcohol Abuse and Alcoholism [NIAAA], 2006). It includes five conditions: Prenatal Alcohol Effects (PAE), Fetal Alcohol Effects (FAE), Alcohol-Related Neurodevelopmental Disorder (ARND), Alcohol-Related Birth Defects (ARBD) and Fetal Alcohol Syndrome (FAS) (NIAAA, 2006). Amongst these, FAS is the only condition which has been defined by the International Statistical Classification of Diseases and Related Health Problems and given an ICD-9 diagnosis. This emphasizes the universal importance and severity of the condition.

FAS was first systematically described as a syndrome in 1973 (Jones, Smith, Ulleland, & Streissguth, 1973) and to date its disease etiology has not yet been fully explored or understood. There are three phenotypes associated with a FAS diagnosis: growth

deficiency, specific FAS facial morphologies and central nervous system (CNS) damage. These can occur in the presence of confirmed or unknown prenatal alcohol exposure. The three most specific FAS facial morphologies include a smooth philtrum, a thin upper lip and short palpebral fissures (*Figure 1.1*). These features are distinctive of FAS individuals and are unlikely to present simultaneously in conditions which share overlapping features with FAS (Sokol, Delaney-Black, & Nordstrom, 2003). There is great variability in both the severity and range of symptoms presenting in individuals with FASD. This relates to the stage of development at which the fetus was exposed to alcohol as well as to the amount and frequency of alcohol exposure. The severity of individual symptoms also depends on the rate of alcohol metabolism in both the mother and the fetus (NIAAA, 2006). As irreversible damage occurs during development, a cure for FAS is unlikely and treatment is a life-long process. Maternal risk factors include an interaction of many factors, which include social, biological, historical, familial and psychological factors, and thus, prevention strategies need to be multidisciplinary (May, et al., 2000).

Figure 1.1: Baby showing the three distinct FAS facial morphologies: smooth philtrum, a thin upper lip and short palpebral fissures (http://www.come-over.to/FAS/JohnGrowsUp.htm).

1.1.1. FAS in South Africa

FAS has been studied in numerous countries and populations. From these studies it has been found that it is most prevalent in minority groups with low socio-economic status (May, et al., 2000; Viljoen, et al., 2005). In the United States of America the rate of FAS is 0.05-2.0 per 1000 live births (May, et al., 2000). Rural regions in the Western Cape, South Africa, have been reported to have the highest documented prevalence of FAS in the world (May, et al., 2000). Prevalence rates in these predominantly mixed ancestry communities were reported to be 46 per 1000 children of school-going age (May, et al., 2000). In a second study, this increased to 65.2-74.2 per 1000 (Viljoen, et al., 2005), and subsequently rose again to 68.0-89.2 FAS affected children per 1000 of school-going age (May, et al., 2007).

There are many proposed reasons for the high rates of FAS in the Republic of South Africa (RSA). These reasons fall within an historical, social and psychological context. The first of these is the “dop” system. This system was originally introduced under colonial agriculture and was a form of payment to farm workers where alcohol was used as partial wage payment to labourers. Although outlawed about 16 years ago, the effects of this system still continue (May, et al., 2000). Individuals with low income and moderate living conditions value alcohol as a commodity (May, et al., 2000). Alcohol is cheap, and as a consequence weekend drinking is a popular and common recreational activity (Viljoen, et al., 2005). Another reason commonly given by mothers of FAS children for drinking, is the stress of everyday living due to poor living conditions and lack of adequate income, and drinking is considered reprieve from these dire circumstances (May, et al., 2000).

1.1.2 Factors contributing to FAS

Although prenatal alcohol exposure has been recognized as the cause of FAS, secondary factors that contribute to the syndrome include various genetic, environmental, and epigenetic factors. It has also long been proposed that paternal drinking may have adverse effects on developing offspring and contribute to FAS (Abel & Bilitzke, 1990).

It has been shown that there is a differential sensitivity to prenatal alcohol exposure in dizygotic twins whereas monozygotic twins have a high concordance rate for FAS (reviewed by Abel, 1995). A key study by Streissguth & Dehaene (1993) looked at sixteen pairs of twins born to alcohol-abusing mothers. These twin pairs were made up of both monozygotic and dizygotic twins. The twin pairs were clinically examined for FAS diagnosis and given standardized IQ tests as well as mental development tests. It was found that monozygotic twin pairs had a higher concordance for disease and a more comparable IQ level than dizygotic twins. This evidence shows a clear genetic component of the condition, as siblings with higher genetic similarity also have higher concordance rates for disease. Every individual has a different genetic susceptibility to FAS. For example, polymorphisms in enzymes responsible for alcohol metabolism could contribute to the likelihood of developing some FAS symptoms (Warren & Li, 2005). Consequently, if there are some harmful intermediates of alcohol metabolism that cannot be broken down due to a non-functional or impaired enzyme, or a reduced amount of enzyme, they could cause damage.

An example of an environmental factor that contributes to FAS is where and into what circumstances you are born. For example, race is a cultural and biological concept; however, it is often linked with a social context, with certain races falling into lower socio-economic groupings (reviewed by Abel, 1998). This results in races where the majority of individuals

fall within lower socio-economic status groups being predominantly affected by FAS (reviewed by Abel, 1995).

The first cases of children born with characteristics of FAS whose mothers did not drink but whose fathers were alcoholics were reported in 1968 (Lemoine, 1968 as cited in Jamerson, Wulser, & Kimler, 2004). It is a well-established phenomenon that women who drink heavily associate with men who have similar behavior. In addition, a very high proportion of children born with FASD have fathers who are alcoholics. Therefore, some of the anomalies of FASD attributed solely to maternal drinking may actually be a result of paternal alcohol consumption preconception (Abel & Bilitzke, 1990).

The precise mechanism by which alcohol exerts its effects on the developing fetus has not yet been fully elucidated, and at present FAS disease etiology is poorly understood (Kaminen-Ahola, et al., 2010). It is known that alcohol has the ability to cause cell death, which can explain some of the symptoms seen in FAS such as neurodegeneration leading to central nervous system damage (Nayak & Murthy, 2008). In addition, evidence shows that alcohol is also able to alter epigenetic states by disrupting DNA methyltransferase levels resulting in altered DNA methylation (Garro, McBeth, Lima, & Lieber, 1991), which is responsible for genomic imprinting and is critical in development.

1.2 Development

1.2.1 Epigenetics and DNA methylation

The term epigenetics was coined by Conrad Waddington in 1942. He devised the concept of an epigenetic landscape, which describes how cells are able to choose individual paths during development and ultimately differentiate into specific cell types that make up different

tissues (Slack, 2002). According to Waddington, epigenetics was the term broadly used to describe how genes change their expression during development (Holliday, 2006). Today, epigenetics has been re-defined as the study of mechanisms that cause stable and heritable changes in gene expression without altering the DNA sequence (Rodenhiser & Mann, 2006). These mechanisms include DNA methylation, histone modification and RNA directed gene silencing (Egger, Liang, Aparicio, & Jones, 2004).

DNA methylation is the most commonly studied epigenetic mechanism and involves the addition of a methyl group to the 5^{th} position of cytosine residues within CpG dinucleotide pairs (Rodenhiser & Mann, 2006). It commonly occurs in areas of the genome known as CpG islands, which are clusters of CpG pairs that are approximately 500 base pairs in size and are usually located in promoter regions of genes (Egger, et al., 2004). DNA methylation regulates gene expression as well as chromosomal stability and is typically associated with gene silencing (Liu, Li, & Tollefsbol, 2008). DNA methylation can regulate gene expression in various ways: Firstly, methylation allows methylation dependent proteins to bind to DNA and induce gene repression (Jeltsch, 2002); secondly, methylation can directly block transcription factor binding and thus transcription initiation (Jeltsch, 2002); and lastly, hypermethylated DNA is able to induce chromatin condensation, preventing transcription factor access and repression of gene activity (Jeltsch, 2002). It is important to note that these methods of gene regulation are not solely attributed to DNA methylation. Another epigenetic modification known as histone modification works together with DNA methylation to regulate gene expression (reviewed by Cedar & Bergman, 2009).

The fundamental building blocks of chromosomes are nucleosomes. Nucleosomes are made up of 146 base pairs of DNA wrapped around a core of eight histone proteins (Peterson & Laniel, 2004). Repeating units of nucleosomes are joined by DNA to form chromatin fibers (Kornberg, 1977), which are then packaged into chromosomes. Each histone has an N-terminal tail which is the substrate for histone modification (Liu, et al., 2008). The different

types of histone modifications include methylation, acetylation, phosphorylation and ubiquitylation and determine active and silenced chromatin states (Liu, et al., 2008). There is increasing evidence that DNA methylation and histone acetylation work together in gene regulation in many cases (Liu, et al., 2008). DNA methylation and histone modification share a biological relationship and interact intricately to control gene expression (reviewed by Cedar & Bergman, 2009). DNA methylation together with various combinations of histone modifications produces unique and reversible epigenetic signatures, which regulate gene expression through chromatin modification and chromatin remodeling (Rodenhiser & Mann, 2006), resulting in tissue specific gene expression. Changes in the pattern of DNA methylation and histone modification can cause disease by dysregulating gene expression (Rodenhiser & Mann, 2006). Studies have shown that many of the proteins involved in DNA methylation interact with proteins involved in histone modification (reviewed by Sims, Nishioka, & Reinberg, 2003). One of these protein families is the DNA methyltransferases (DNMTs).

DNMTs are the enzymes responsible for establishing and maintaining DNA methylation (Jones & Takai, 2001). There are three known DNMTs that function in mammals: DNMT1, DNMT3a and DNMT3b (Jeltsch, 2002). DNMT1 is a maintenance methyltransferase and functions in maintaining existing methylation patterns through cell division (Jeltsch, 2002; Leonhardt, Page, Weier, & Bestor, 1992). DNMT3a and DNMT3b are known as *de novo* methyltransferases and are essential for establishing new methylation patterns. A fourth DNMT, DNMT2 (also known as transfer RNA aspartic acid methyltransferase 1-TRDMT1) has been identified in mammals (Goll, et al., 2006; Jeltsch, 2002). This DNMT has been shown to methylate a small RNA, known as aspartic acid transfer RNA (Goll, et al., 2006).

Studies have shown that DNMTs are vital for development, as significantly lower levels of methyltransferase activity result in lethality (Li, Bestor, & Jaenisch, 1992). Okano, Bell, Haber, & Li (1999) showed that Dnmt3-deficient mice have decreased global DNA

methylation which results in lethality. They showed this by creating mice from multiple Embryonic Stem (ES) cell lines heterozygous for either *Dnmt3a* or *Dnmt3b* mutations. They showed that *Dnmt3a*$^{-/-}$ mice die shortly after birth whereas *Dnmt3b*$^{-\backslash-}$ mice die in the embryonic stages. They also showed that double knockout mice (*Dnmt3a*$^{-/-}$*; Dnmt3b*$^{-/-}$) also die before birth.

Thus, DNMTs are vital for development and function in establishing genomic stability as well as in maintaining gene expression by ensuring maintenance of specific methylation patterns through cellular replication, including those found at imprinted regions.

1.2.2 Genomic imprinting and development

Genomic imprinting was discovered in the early 1980s when scientists questioned why diploid parthenogenetic embryos, having the normal number of chromosomes (although originating only from the mother), did not develop to full term (Solter, 2006). This question gave rise to the hypothesis that sperm makes a vital non-genetic contribution to development. Further experimentation led to the hypothesis that maternal and paternal genomes are functionally different from one another and that certain genes are inherited in an active form from one parent and repressive form from the other. These two sets of chromosomes and genes, one from each parent, are essential for normal development (Solter, 2006). This phenomenon was given the name imprinting (Surani et al., 1984 as cited in Solter, 2006). Genomic imprinting is a form of epigenetic regulation involving the reversible modification of gene expression depending on which parent the gene is inherited from (Jiang, Bressler, & Beaudet, 2004). Allele specific methylation in the two germ lines results in methylation of one parental allele (imprinted allele) and hypomethylation of the other parental allele (Lucifero, Chaillet, & Trasler, 2004). Therefore, if a gene is maternally

imprinted it will be paternally expressed and if it is paternally imprinted it will be maternally expressed (reviewed by Bartolomei & Tilghman, 1997).

There are many proposed hypotheses for the evolution of imprinting. Two of the most established are the conflict hypothesis and the ovarian time bomb hypothesis (Weisstein, Feldman, & Spencer, 2002). The conflict hypothesis states that maternal genetic interests are maximized by conserving resources (inactivation of growth enhancing alleles), whereas paternal genetic interests are maximized by expending resources (activation of growth enhancing alleles). In mammals, both parents want to ensure survival of their offspring. In females, during gestation there is a constant need for nutrients by the embryo. In order to conserve nutrients and share them equally among all offspring within a litter, mothers inhibit growth of the embryo by inactivating certain growth enhancing genes. Conserving nutrients is a way to ensure that she has many offspring. In males, fitness and survival of offspring is optimized by promoting growth, and activating growth-enhancing genes. Thus, there is a conflict of contributions from the parental alleles (reviewed by Morison, Ramsay, & Spencer, 2005).

The ovarian time-bomb hypothesis states that imprinting evolved as a defensive mechanism against trophoblastic disease in females. The spontaneous development of an unfertilized egg in the ovaries could lead to this cancerous condition. Therefore, the maternal inactivation of growth enhancing genes ensures that spontaneous embryogenesis is unlikely to occur. A paternal contribution with active growth enhancing genes is necessary for embryogenesis to take place.

When fertilization occurs, the egg and the sperm contain very dissimilar epigenetic marks. These germ cell epigenetic marks are established in the germ line prior to fertilization. Primordial germ cells are initially highly methylated in both males and females, after which they undergo a genome wide active demethylation (*Figure 1.2*). In both sexes the cells then

enter mitotic (male) and meiotic (female) arrest. Several days later *de-novo* re-methylation takes place and imprints are re-established (Reik, Dean, & Walter, 2001). Upon fertilization the paternal genome undergoes a dramatic change whereby the protamines present around the DNA are replaced by histones. This is accompanied by a wave of genome wide demethylation (Reik, et al., 2001). Only imprinted genes resist this demethylation and remain methylated (Santos, Hendrich, Reik, & Dean, 2002). Once the embryo starts to divide there is more passive demethylation, which is again escaped by imprinted genes. Thus, these imprints are maintained throughout development in a parent of origin specific manner. This epigenetic reprogramming is essential in establishing global DNA methylation patterns.

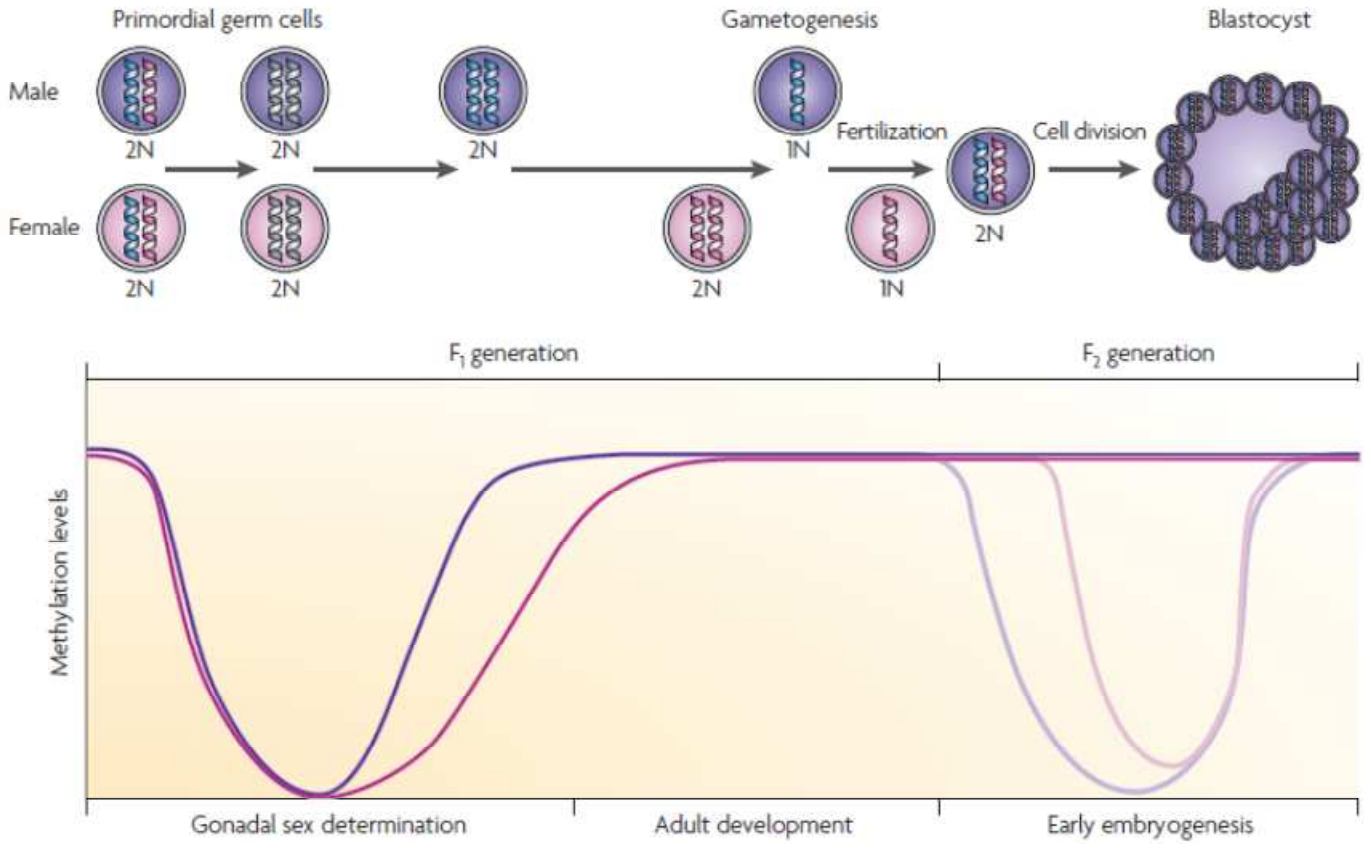

Figure 1.2: Reprogramming of DNA methylation during mammalian development (obtained from Jirtle & Skinner, 2007).

Genes which are subjected to genomic imprinting and show different methylation patterns in the embryo are known as imprinted genes (Lucifero, et al., 2004). The first three imprinted genes were discovered in the early 1990s (Reik, et al., 2001). These are the *insulin-like growth factor 2 receptor* (*Igf2r)*, *the insulin-like growth factor 2* (*Igf2)* and the *H19* genes. Today, over two hundred imprinted genes have been identified. Imprinted genes are critical in growth and development. Imprinting is a reversible epigenetic modification that must be established and maintained correctly with each reproductive cycle (Lucifero, et al., 2004). The proteins of several imprinted genes regulate embryonic growth, and monoallelic expression of these genes is essential for fetal development (Constancia, Kelsey, & Reik, 2004). One such gene is the *IGF2* gene, whose aberrant expression, resulting from altered imprinting patterns, has been linked to developmental disorders such as Beckwith-Wiedemann Syndrome (BWS) (Lucifero, et al., 2004), as well as cancer (Gabory, Ripoche, Yoshimizu, & Dandolo, 2006).

The dysregulation of imprinted genes has been implicated in several sporadic, inherited and environmentally induced growth disorders. Beckwith-Wiedemann Syndrome (BWS) is a model example of an overgrowth syndrome which is 85% sporadic and 15% inherited (Choufani, Shuman, & Weksberg, 2010). The alteration of the imprinted region on chromosome 11p15.5 is involved in this syndrome. Within this region there are two imprinting domains, the *H19-IGF2* domain and the *CDKN1C-KCNQ1OT2* domain (Choufani, et al., 2010). One common cause for this disorder is uniparental disomy (UPD) of chromosome 11p15.5 (Miozzo & Simoni, 2002). Uniparental disomy occurs when homologous chromosomes are inherited form a single parent (Miozzo & Simoni, 2002). In the case of BWS paternal UPD results in the *IGF2* gene, which is a growth factor, being expressed in double the amount than normal and thus the overgrowth phenotype (Choufani, et al., 2010). Environmental agents such as toxins have been shown to be able to alter epigenetic states (Anway, Cupp, Uzumcu, & Skinner, 2005). By doing so, they have a vast

effect on a number of different cellular processes including cell growth. Uncontrollable cell growth leads to cancerous conditions (Esteller, 2006).

1.2.3 *H19/IGF2* locus

An example of an imprinted locus is the *H19/IGF2* locus. The imprint at this locus is set up during spermatogenesis (Davis, Yang, McCarrey, & Bartolomei, 2000). The *H19* gene is found on chromosome 11p15.5 (Leibovitch, et al., 1991). This imprinted gene is significant in development and is associated with BWS. It is 2.7kb long and includes 4 small introns (Leibovitch, et al., 1991). *H19* is transcribed into a non-coding mRNA, which is processed into a micro RNA (miRNA-miR-675) (Cai & Cullen, 2007). The exact function of *H19* is unknown (Ulaner, et al., 2003). Approximately 90 kb 5' to the *H19* is the *IGF2* (Insulin-like growth factor) gene. Murine models show that this gene plays a role in fetal growth (DeChiara, Robertson, & Efstratiadis, 1991).

The expression of imprinted genes depends on the methylation status of CpG islands located in regions known as differentially methylated regions (DMRs). DMRs can be separated into two groups, namely primary DMRs and secondary DMRs (Kobayashi, et al., 2006). Primary DMRs gain their methylation patterns through the germline and retain this pattern throughout development whereas secondary DMRs gain their methylation patterns after fertilization (Kobayashi, et al., 2006). Promoters, silencers and activators are examples of DMRs (Lopes, et al., 2003). Primary DMRs are referred to as Imprinting Control Regions (ICRs) when the loss of the DMR results in the aberrant expression of its associated imprinted genes. DMRs associated with imprinted gene expression control gene expression in both maternal and paternal alleles (Kobayashi, et al., 2006).

H19 has a primary DMR/ICR that is located upstream. This ICR is situated in between the *IGF2* and *H19* genes, and is responsible for the expression of both these genes. There is reciprocal expression of the *H19* and *IGF2* genes (Gabory, et al., 2006). The paternal allele of the *H19* gene is imprinted which results in *H19* repression and concurrent *IGF2* expression (Park, Sellars, Grinberg, Huang, & Pfeifer, 2004). The opposite is true on the maternal allele where *H19* is expressed and *IGF2* is repressed. This reciprocal expression of *H19* and *IGF2* is the result of two different regulatory mechanisms resulting from methylation (*Figure 1.3*). The methylation on the paternal allele maintains *H19* repression by preventing transcription factors from binding to their target sites and initiating transcription. Whereas it promotes *IGF2* expression by preventing binding of a protein, CCCTC-binding factor (CTCF) which functions in *IGF2* repression. DNA from sperm contains a hypermethylated *H19* ICR whereas oocytes contain a hypomethylated *H19* ICR (Park, et al., 2004). Thus, when the male and female copies of the gene are brought together in the embryo they show different methylation patterns, which result in differential gene expression.

CTCF is a nuclear protein which is able to bind to a variety of sequences (Ohlsson, Renkawitz, & Lobanenkov, 2001). It can act as a transcription factor or a chromatin insulator protein (Ohlsson, et al., 2001). As a transcription factor, CTCF represses gene expression by interacting with co-repressors and recruiting histone deacetylases (HDACs) (Ohlsson, et al., 2001). Histone deacetylases are enzymes that remove acetylation marks from histones and in so doing repress gene expression (Liu, et al., 2008). As a chromatin insulator, CTCF mediates gene repression by preventing enhancers from acting on silent genes (Kurukuti, et al., 2006).

A key locus where CTCF functions as an insulator is the *H19* ICR (Ulaner, et al., 2003). This ICR has been shown to be a chromatin boundary, where it prevents the modification of chromatin into a transcriptionally active state. It also prevents enhancers from interacting with unsuitable promoters (Hark, et al., 2000). As this ICR lies between enhancers and the

IGF2 promoter, it is able to exert its effect (Hark, et al., 2000). CTCF binding mediates repression of the maternal *IGF2* gene by preventing enhancers from interacting with the gene (Pant, et al., 2004). Many previous studies have shown that specific sites in the *H19* ICR are critical for CTCF binding and insulator function. Both Bell & Felsenfeld (2000) and Hark et al. (2000) showed that DNA methylation controls *H19/IGF2* gene expression by preventing CTCF binding to its target sites. On the paternal allele, methylation of the *H19* ICR prevents CTCF binding and allows distal enhancers to interact with *IGF2* promoters resulting in *IGF2* expression (Kurukuti, et al., 2006). At the same time this methylation prevents *H19* transcription (Kurukuti, et al., 2006) *(Figure 1.3a)*. On the maternal allele the *H19* ICR is unmethylated allowing CTCF binding to its target sites and consequent *IGF2* silencing (Kurukuti, et al., 2006). *H19* expression occurs due to *H19* ICR hypomethylation (Kurukuti, et al., 2006) *(Figure 1.3b)*.

In humans there are seven CTCF binding sites within the *H19* ICR (Ulaner, et al., 2003). Of these sites, only the sixth site has been shown to have allele-specific differential methylation (Ulaner, et al., 2003). It has been shown that methylation of this site always corresponds with *H19* repression (Ulaner, et al., 2003). Although this site is central to the regulation of the *IGF2/H19* locus, CTCF binding alone is insufficient to regulate *IGF2/H19* imprinting in humans (Ulaner, et al., 2003). Many other factors such as histone modifications also play a role and collectively regulate *IGF2/H19* gene expression.

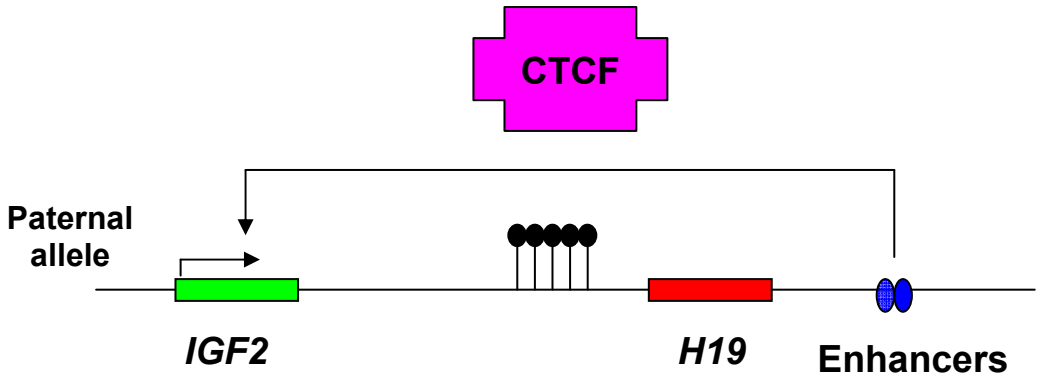

Figure 1.3a: Paternal *H19* ICR locus. The paternal ICR is hypermethylated (shown by the filled lollipops). This prevents CTCF binding and allows enhancers (blue) to interact with *IGF2* resulting in its expression. At the same time the hypermethylation directly prevents *H19* expression (figure based on Ohlsson, et al., 2001).

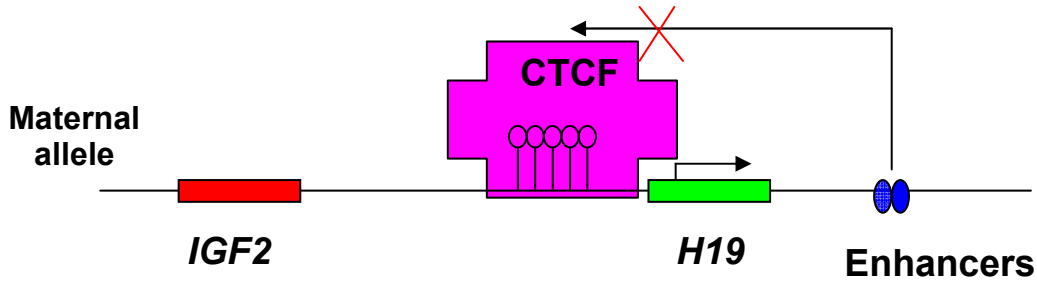

Figure 1.3b: Maternal *H19* ICR locus. The maternal ICR is hypomethylated (shown by the unfilled lollipops). This allows CTCF binding and prevents enhancers (blue) from interacting with *IGF2* resulting in its repression. At the same time the hypomethylation directly allows *H19* expression (figure based on Ohlsson, et al., 2001).

1.3 Epigenetics and alcohol

It is well established that alcohol is a teratogen and is able to cause an array of developmental abnormalities (Becker, Diaz-Granados, & Randall, 1996; Cudd, 2005). Human studies and mouse models have both demonstrated this (Streissguth, Landesman-Dwyer, Martin, & Smith, 1980). Prenatal alcohol exposure is the most common preventable cause of mental retardation worldwide, and can result in growth anomalies, central nervous system damage and structural malformations (Becker, et al., 1996; Cudd, 2005). The primary cytotoxic effect of alcohol is demonstrated in FAS where the fetus is directly exposed to alcohol *in utero*. This includes the effect of alcohol metabolism, resulting in the production of harmful metabolites, which then have a range of physiological effects (Mascord, Smith, Starmer, & Whitfield, 1992). To date many studies have been done on elucidating the molecular mechanisms underlying FAS, however, these have focused on the effects of alcohol metabolism and cell biology (Kotch & Sulik, 1992; Menegola, Broccia, Di Renzo, & Giavini, 2001). Studies on possible epigenetic mechanisms underlying alcohol teratogenesis have been conducted. However, the testing of the theory that alcohol teratogenesis could occur via epigenetic mechanisms, is still in its infancy (Haycock, 2009; Kaminen-Ahola, et al., 2010).

1.3.1 Maternal ethanol consumption and epigenetic effects

Two key studies looking at the epigenetic effects of maternal ethanol consumption were recently conducted (Haycock & Ramsay, 2009; Kaminen-Ahola, et al., 2010).

Kaminen-Ahola et al. (2010) developed a mouse model for chronic alcohol exposure, which produces measurable phenotypes in adulthood. They used the Agouti locus in order to test the epigenetic effects of both pre- and post-conception alcohol exposure. In both cases, they

found that ethanol treated mice had a higher methylation level at this locus, which corresponded to changes in coat colour. Further experimentation showed that ethanol exposure caused a consistent decrease in the expression of three genes all of which are associated with growth. When they looked at phenotypic measurements, they found postnatal growth retardation as well as craniofacial dysmorphism in offspring sired from alcohol exposed mothers. These features are consistent with features seen in individuals born with FASD. Taken together these results postulate a potential role for epigenetics in the etiology of FASD.

Haycock & Ramsay (2009) exposed preimplantation embryos to ethanol in order to test the hypothesis that a change in *H19/IGF2* methylation may play a role in growth retardation seen in FASD. This was done in a case-control manner, where mice were either administered ethanol or a saline solution. Placentae and embryos were then harvested in the prenatal period (at day 10.5 of gestation). It was found that both the placentae and the embryos were severely growth retarded (measured by weight) in the ethanol-exposed group compared with controls. However, DNA methylation at the parental alleles in the fetuses was not significantly affected. They did note that paternal alleles in the placentae had significantly less methylation in ethanol treated mice compared to controls. Despite the growth retardation observed in the embryos there was no change in *H19* ICR methylation.

Taken together, both these studies provide evidence for the ability of alcohol to either increase or decrease DNA methylation at different loci within the genome. In addition, candidate gene and microarray analyses have detected both up- and down-regulation of numerous genes following alcohol exposure (reviewed by Kaminen-Ahola et al. 2010), providing further support for this argument.

1.3.2 Paternal ethanol consumption and epigenetic effects

The effect of paternal alcohol consumption on the developing fetus would be a secondary effect. There is no direct contact between the paternally consumed alcohol and the fetus. Alcohol circulates throughout the body, including to the testes (Bielawski, Zaher, Svinarich, & Abel, 2002), and both animal models and human studies have shown that sperm is adversely affected by alcohol exposure (Abel & Moore, 1987; Donnelly, McClure, Kennedy, & Lewis, 1999). Sperm motility and morphology were both compromised in humans after varying degrees of alcohol exposure (Donnelly, et al., 1999), and in mice, alcohol consumption led to a significant increase in the percentage of abnormal sperm (Abel & Moore, 1987). As defective sperm containing anomalies or disrupted gene regulation are still capable of fertilization, embryos could be negatively affected (reviewed by Emery & Carrell, 2006). In addition to these physiological effects, alcohol has been shown to be able to alter DNA methylation in various parts of the body including the liver, uterus and sperm (Bielawski, et al., 2002; Garro, et al., 1991). In mice it has been shown that abnormal spermatozoa can still penetrate oocytes (Burruel, Yanagimachi, & Whitten, 1996). A proportion of abnormal spermatozoa are produced by males across all mammalian species and a high proportion of these lowers fertility but does not render males infertile (Burruel, et al., 1996). Infact, Seunanez, Carothers, Martin, & Short (1977) noted that fertile men also have a proportion of abnormal spermatozoa. Thus, the direct effects of alcohol or its metabolites occur in the sperm. These effects are then mediated through the sperm to the fetus.

Bielawski et al. (2002) studied the effects of paternal alcohol consumption on cytosine methyltransferase mRNA levels in sperm. They used mRNA levels as an indicator of DNMT levels and therefore DNA methylation levels, and hypothesized that alcohol is able to alter DNA methylation during spermatogenesis, and in doing so disrupt offspring development. Using a mouse model, they were able to show that cytosine methyltransferase mRNA levels

were significantly lower in ethanol treated males. In addition, they were able to show that offspring of alcohol-exposed males had decreased fetal weight. Their results were the first to demonstrate a possible mechanism for paternal alcohol effects on offspring development, namely altered genomic imprinting.

Genetic contributions of the sperm are vital for development and some imprinting errors are paternally inherited (reviewed by Emery & Carrell, 2006; Kobayashi, et al., 2009). These imprinting errors are the result of alcohol altering DNA methylation and having various effects on histone modifications (reviewed by Shukla, et al., 2008). For example, ethanol has been shown to decrease DNA methylation as well as histone methylation and increase histone acetylation (reviewed by Shukla, et al., 2008). These changes are both associated with the activation of genes. Animal models have shown that alcohol decreases global DNA methylation by lowering methylase activity (Garro, et al., 1991). Li, Beard, & Jaenisch (1993) showed that in mice without methylase activity the normally silent *H19* paternal allele was activated and the normally active *IGF2* allele was silenced. In addition, the normally silent *IGF2* maternal gene was activated (Li, et al., 1993).

Ouko et al. (2009) studied male gametes in order to establish a link between DNA methylation and alcohol consumption. Two loci were studied, the *H19* ICR and the *IG*-DMR. Despite great inter-individual variation in methylation at the *H19* ICR, a trend of DNA hypomethylation was observed with chronic alcohol use. Due to the small sample size however, no statistically significant results were obtained. A study by Knezovich & Ramsay (2012), showed that the *H19* ICR, in murine sperm, had no significant decrease in methylation with chronic alcohol use. However, the offspring sired from these alcohol exposed males and non-alcoholic females showed hypomethylation of the *H19* ICR and an age specific reduction in growth. One of the proposed reasons for this observation is that alcohol may be exerting its effect on another epigenetic mechanism, histone modification. As histone modification and DNA methylation are intricately linked an alteration in histone

modifications could result in DNA hypomethylation in the offspring. Another proposed mechanism for this is RNA mediated effects. It was always thought that sperm was the vehicle for the paternal haploid genome only; however, recent findings suggest that along with this haploid genome, sperm also carries miRNAs to the unfertilized ovum (Rassoulzadegan, et al., 2006). Rassoulzadegan et al. (2006) showed that miRNAs inherited from sperm are sufficient to cause a change in gene expression and consequently phenotype. This provides support that miRNAs inherited from sperm could be responsible for the epigenetic changes observed in the offspring sired from alcohol-exposed males.

1.3.3 Global DNA methylation and alcohol

Many studies have been conducted on the effects of various environmental exposures on global DNA methylation (Baccarelli & Bollati, 2009). These studies have found that certain metals such as arsenic, nickel and cadmium can change global DNA methylation. In addition, endocrine-disrupting chemicals such as benzene, as well as reproductive toxicants, such as diethylstilbestrol, have also been found to affect global DNA methylation (Baccarelli & Bollati, 2009). Another environmental exposure, which has been shown to change global DNA methylation in various tissues, is air pollution (Baccarelli, et al., 2009; Yauk, et al., 2008). Yauk et al. (2008) found that particulate air pollution increased global germline DNA methylation in male mice. They also noted that these changes persisted after the exposure was removed. Baccarelli et al. (2009) showed global hypomethylation in human blood after exposure to particulate emissions from vehicles in traffic.

Not much work has been done on linking alcohol and its effect on global DNA methylation. However, one key study was done by Garro et al. (1991). This study used a mouse model to see what effect maternal ethanol consumption would have on fetal DNA methylation as well as methyltransferase activity. They found that ethanol treated females gave birth to fetuses with global DNA hypomethylation resulting from lower methyltransferase activity.

1.4 DNA methylation analysis techniques

There are numerous techniques available to study DNA methylation: These include techniques used to analyze global as well as locus specific DNA methylation (Fouse, Nagarajan, & Costello, 2010). These include typing and profiling platforms (Beck & Rakyan, 2008). DNA methylation typing platforms are used when only a few loci are under investigation whereas DNA methylation profiling platforms are used when looking at multiple loci in many samples (Beck & Rakyan, 2008). Typically the type of DNA methylation technique used in a study depends on what is being investigated. *Figure 1.4* shows some DNA methylation techniques and what they would be used for. For the purposes of this study the Luminometric Methylation Assay (LUMA) and bisulfite-based pyrosequencing were used for analysis and will be discussed further.

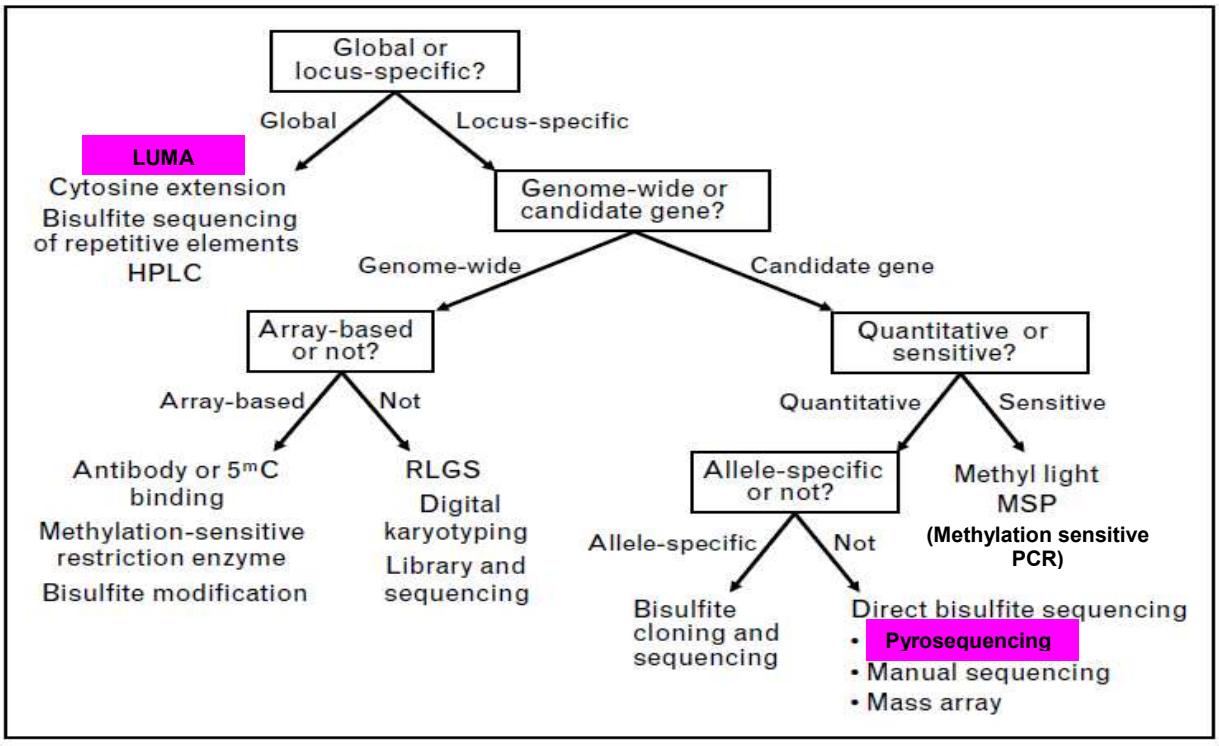

Figure 1.4: Appropriate DNA methylation analysis methods to use (figure modified from Shen & Waterland, 2007).

1.4.1 Luminometric Methylation Assay (LUMA)

Many assays have been developed to measure global DNA methylation; however, a majority of them are labour intensive and involve the use of radioactive isotopes (Karimi Arzenani, 2009). In order to minimize these shortcomings, a novel assay known as LUMA was more recently developed by Karimi Arzenani (2009). LUMA is a method used to quantitate global genomic DNA methylation based on restriction digests, followed by a polymerase extension assay using the Pyrosequencing^{TM} platform (Karimi, et al., 2006). It is a highly reproducible and rapid technique, which requires no prior modification of genomic DNA. In addition, it contains an internal control for DNA input and can be used for high throughput applications (Karimi, et al., 2006).

1.4.2. Pyrosequencing

One of the techniques used to study a biological system, which arguably gives the most detailed data, is DNA sequencing. Dideoxy chain termination (Sanger sequencing) is the most commonly used first generation sequencing methodology and has been used for the sequencing of numerous prokaryotic genomes (Sanger, Nicklen, & Coulson, 1977). In the 1990s, a new method called pyrosequencing was developed (Ronaghi, Karamohamed, Pettersson, Uhlen, & Nyren, 1996). Since then, this method has been successfully used in *de novo* as well as confirmatory sequencing. The advantages of this technique include speed, accuracy and automation. In addition, there is no need for labeled nucleotides or gel electrophoresis (Ronaghi, 2001). Pyrosequencing can radically reduce the time taken to genotype SNPs within a 200-300 base pair sequence. This would take several hours using conventional sequencing methods but approximately an hour using pyrosequencing (Ahmadian, Ehn, & Hober, 2006). With the advantages however, come the disadvantages, including cost and read length. As pyrosequencing can only sequence 300-500 base pairs at a time, conventional sequencing is more beneficial when creating genome assemblies, as it

can sequence 800-1000 base pairs at a time. Thus, depending on application, different sequencing techniques are advantageous over others (Ahmadian, et al., 2006).

Common applications of pyrosequencing include: SNP genotyping, microbial typing, resequencing and tag sequencing (Ahmadian, et al., 2006). SNP genotyping is the most popular application and has recently been extended to DNA methylation analysis (England & Pettersson, 2005). In this study, pyrosequencing was used for methylation analysis and this application will be discussed further. Please note that in this instance pyrosequencing does not refer to its use in next generation sequencing technologies. It will be used for locus specific sequencing.

1.5 Study rationale

FAS is a multifactorial disease arising from the interaction of various genetic, epigenetic and environmental factors, and is a major health problem particularly in South Africa. FAS disease etiology is poorly understood and the possible epigenetic changes resulting from alcohol consumption by pregnant women have yet to be explored. In addition, the possible epigenetic changes resulting from preconception alcohol consumption by men has not been widely explored. Incorrect gene expression resulting from imprinting defects can result in various diseases such as cancer and BWS. Thus, if preconception alcohol consumption by men can alter methylation patterns, there is a possibility that disruption of the methylation pattern at imprinted loci could be partly responsible for clinical features seen in FASD patients. The *H19/IGF2* locus is an imprinted locus in mammals, the deregulation of which may contribute to FAS (Ouko et al., 2009). In both human and animal models, the most common effect of prenatal alcohol exposure is growth retardation (Simpson, Duggal, & Keiver, 2005). It has been suggested that *IGF2* promotes fetal growth by increasing nutrient transport from the placenta to the developing fetus (reviewed by Constancia, et al., 2004). As insufficient placental nutrient transport may be associated with retarded fetal growth

(Livy, Maier, & West, 2004), in conditions such as FAS where fetal growth is compromised, it is important to understand the dynamics between nutrient supply and growth.(deleted) Murine models have shown that decreased methylation results in paternal *H19* activation and *IGF2* silencing. As such, alcohol may affect the *IGF2* gene by decreasing methylation and altering gene expression. If paternal expression of this gene is altered, the inactive *IGF2* allele could influence placental nutrient transport and in turn can cause retarded fetal growth during development, and in that way contribute to the growth phenotype observed in FAS.

1.5.1 Hypothesis

There is a difference in the DNA methylation patterns in spermatozoa exposed to alcohol and unexposed spermatozoa and this difference develops in a dosage-dependent manner.

1.5.2 Aims and objectives

Main Aim:

To examine global DNA methylation and locus specific *H19* ICR DNA methylation in spermatozoa, related to alcohol exposure.

Objectives:

To optimize the LUMA technique in blood and sperm DNA samples.

To assess differences between global DNA methylation in alcohol exposed spermatozoa and non-alcohol exposed spermatozoa, and to determine the effects of alcohol on global DNA methylation in a dosage-dependent manner.

To characterize differences in DNA methylation at the sixth CTCF binding site in the *H19* ICR between alcohol exposed spermatozoa and non-alcohol exposed spermatozoa, and to determine the dosage-dependent effects of alcohol on DNA methylation at this locus.

2. Participants and methods

All protocols presented below were followed according to the manufacturer's specifications unless otherwise stated. In addition, the full protocols can be found in *Appendix A*. All preparations of solutions will be described in detail within the text. If they do not appear in the text, they can be found in *Appendix B*.

2.1 Sample collection

After ethics approval (*M090555*) was obtained from the Human Research Ethics Committee (Medical), University of the Witwatersrand, to conduct this research (*Appendix C*), a total of 149 human semen samples were collected. These semen samples were collected from individuals visiting the Cryobank, andrology laboratory and sperm bank, in Parktown. The details and nature of the study were first thoroughly explained to participants, and any questions and concerns that they raised addressed. Following this, an information sheet (*Appendix D*) with a consent form attached to a short questionnaire was administered and completed by each participant (*Appendix E*). This questionnaire assessed personal history, individual drinking patterns, smoking and other drug use, as well as family drinking history. Of these 149 samples, 29 were disregarded due to insufficient information and 7 were azoospermic (contained no sperm cells). Thus a total of 113 sperm samples were used in this study. After assessing each individual's drinking behavior, participants were separated

into 2 groups: the control group (non-drinkers) and the drinkers. These groups were formed based on whether or not the participants consumed any alcohol. Those individuals who are complete teetotalers and those that have stopped drinking for at least 6 months were classified as controls. This is because the average length of one spermatogenesis cycle is 42-76 days (Chan, 2006), and we are focusing on alcohol affecting the current sperm cycle of the individual. Individuals who consume 1 or more drinks per week were classified as drinkers. In order to calculate the effect of alcohol in a dosage-dependent manner, the drinking frequency per month was calculated based on how often the individual consumes alcohol and how many drinks they consume per session. Once the questionnaire was completed and informed consent obtained, the participants were asked to donate semen and all samples were anonymised and coded. The samples together with information on sperm parameters, age, ethnicity, drinking habits, drug use and smoking can be found in *Appendix F*.

Sperm parameters included sperm count, sperm morphology, and sperm motility. These parameters were routinely analyzed at the fertility clinic. *Table 2.1* shows the values considered normal for these parameters according to the World Health Organization (Cooper, et al., 2010). If sperm count is <15 million/ml the individual is said to be oligozoospermic. If sperm morphology is <4% the individual is said to be teratozoospermic and if sperm motility is below 40% the individual is said to be asthenozoospermic. In our study, all samples were used irrespective of how many abnormal sperm parameters they had.

Table 2.1: Values considered normal for the three parameters analyzed

	Sperm Parameters		
	Sperm Count (million/ml)	**Sperm Morphology (%)**	**Sperm Motility (%)**
Reference Values	>15	>4	>40

2.2 Sperm DNA extraction

DNA extraction from sperm samples was done using the QIAamp® DNA Micro Kit (QIAGEN). Sperm samples stored at -40˚C were thawed after which the supplementary protocol for purification of DNA from epithelial cells mixed with sperm cells was used (*Appendix A*). This protocol includes several wash steps, which help to eliminate any contamination from epithelial cells. Steps 1-21 of the QIAGEN Supplementary Protocol, Purification of DNA from epithelial cells mixed with sperm cells using the QIAamp® DNA Micro Kit, were followed according to the manufacturer's protocol. Only three small changes were made to this protocol. At step 1, 1 ml of semen was added to 20 µl proteinase K and 500 µl Buffer ATL instead of a swab/piece of fabric. This is because the sample was collected in a liquid form and not on a swab. At step 8, an additional wash step was added in order to minimize epithelial cell contamination. At step 10, samples were incubated at 56˚C overnight instead of just for one hour. This helps facilitate complete digestion of the sperm head. The addition of carrier RNA to samples was stated as optional in the protocol. We added carrier RNA to samples as it enhances the binding of DNA to the column membrane during the extraction procedure. This is beneficial when working with very small amounts of DNA. Sperm DNA was then eluted in 35µl of buffer AE. The final concentrations of the sperm DNA ranged from 17ng/µl to 237.3ng/µl, and samples were stored at -20˚C.

2.3 Bisulfite modification

Bisulfite modification of DNA samples was performed using the EZ DNA Methylation-Gold™ Kit (Zymo Research). Bisulfite modification involves the treatment of DNA with sodium bisulfite, which deaminates all unmethylated cytosine molecules converting them to uracil and leaves methylated cytosines unchanged (*Figure 2.1*). The treated DNA sequence will differ from its original composition at unmethylated cytosines. Thus, after PCR and

sequencing one can differentiate between cytosines and methylated cytosines as cytosines would be converted to thymine and methylated cytosines would remain unchanged. PCR primers are designed according to the chemically modified sequence and the amplified DNA is then analyzed by sequencing. DNA samples were bisulfite modified as per the manufacturer's specifications. 300ng of DNA was used for conversion of sperm DNA and this was eluted in 12µl.

	Original sequence (5' to 3')	**After bisulfite treatment (5' to 3')**
Unmethylated DNA	G-T-T-G-**C**-G-**C**-T-**C**-A-**C**	G-T-T-G-**U**-G-**U**-T-**U**-A-**U**
Methylated DNA	G-T-T-G-**C**-G-**C**-T-**C**-A-**C**	G-T-T-G-**C**-G-**C**-T-**C**-A-**C**

Figure 2.1: Bisulfite modification of DNA (Adapted from Zymo Research, EZ DNA Methylation-GoldTM Kit Instruction Manual).

2.4 Polymerase Chain Reaction (PCR)

2.4.1 *H19* PCR

PCR amplification was optimized and performed using semi nested PCR. This PCR involves two rounds of amplification and is used when amplifying a target sequence when the number of initial DNA copies is very small. The first round of amplification uses 'external' primers, which amplify the region of interest, whereas the second round uses one 'internal' primer. This primer anneals to a sequence just internal to the PCR product from the first round of

amplification. This procedure is used as it increases amplification and specificity of the target region (*Figure 2.2*).

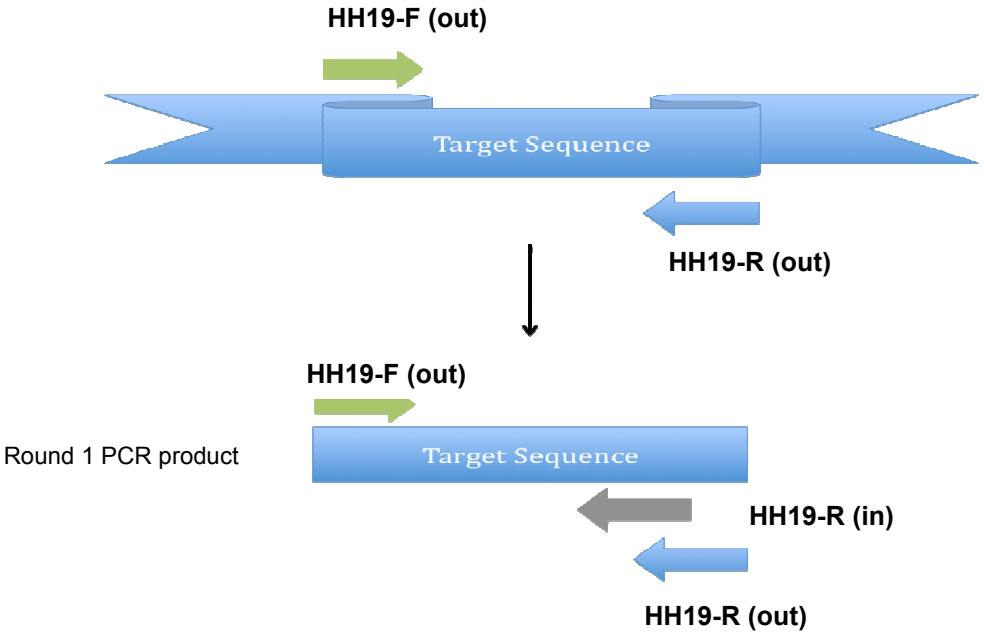

Figure 2.2: Principle of semi nested PCR for *H19*: First round of amplification used an outer reverse primer (blue) and an outer forward primer (green). Second round of amplification involved the addition of an inner/nested reverse primer (grey) to first round PCR product.

Primers and PCR conditions for the amplification of the *H19* ICR had been previously designed and optimized (M. Masemola, personal communication). As the second round of PCR used here is prior to pyrosequencing it requires three primers: a normal primer, a tagged primer and a universal biotinylated primer. The tagged primer and the universal primer share a short tag sequence. The universal primer is used, as it is needed for binding of PCR products to the sepharose beads in the pyrosequencing reaction. As a lower concentration of tagged primer is added compared to universal primer, more amplicons should have universal primer incorporated than tagged primer. This is important for binding of PCR products in the next step. Betaine is used in round 2 PCR, in order to prevent secondary structure formation of GC-rich DNA sequences, and thus improve their

amplification (Henke, Herdel, Jung, Schnorr, & Loening, 1997). PCR conditions prior to pyrosequencing can be found in *Table 2.2 a) and b)* and primer sequences for this reaction can be found in *Table 2.3*.

Table 2.2 a) and b): *H19* ICR PCR conditions (M. Masemola, personal communication).

a)

PCR reagents – Round 1		**PCR Cycle conditions**		
Reagents				
ddH_2O	29.8µl	94°C	5 min	
Buffer (10x)	5µl	94°C	30 sec	
$MgCl_2$ (25 mM)	4µl	61.5°C	30 sec	x28 cycles
dNTP Mix (5 mM)	5µl	72°C	30 sec	
Primer F/R (10 mM)	2µl	72°C	5 min	
Taq pol (5 U/µl)	0.2µl	4°C	∞	
DNA	2µl*			
	50µl			

* 2 µl bisulfite modified DNA

b)

PCR reagents - Round2		**PCR Cycle conditions**		
Reagents				
ddH2O	25.1µl	95°C	5 min	
Buffer (10x)	5µl	95°C	20 sec	
$MgCl_2$ (25 mM)	4µl	53°C	30 sec	x45 cycles
dNTP mix (5 mM)	2.5µl	72°C	20 sec	
Normal primer (10 mM)	1µl	72°C	5 min	
Tagged primer (1mM)	1µl	4°C	∞	
Universal primer (10 mM)	1µl			
Taq pol (5 U/µl)	0.4µl			
DNA	3µl*			
Betaine	7µl			
	50µl			

* PCR product from round 1

Table 2.3: *H19* ICR PCR primer sequences (M. Masemola, personal communication).

Nested PCR	Primer	Sequence	Amplicon size
Round 1	*H19*-F (out)	5' - GTATATGGGTATTTTTGGAGGT - 3'	317 bp
	H19-R (out)	5' – CTTAAATCCCAAACCATAACACTA - 3'	
Round 2	*H19*-F (out)	5' - GTATATGGGTATTTTTGGAGGT - 3'	240 bp
	H19-R (inlabel)	5'- GACGGGACACCGCTGATCGTTTAATAT CCTATTCCCAAATAAC - 3'	
	Universal Primer	5' - GGGACACCGCTGATCGTTTA - 3'	
Sequencing Primer		5' – TTGGTTGTAGTTGTGGAAT – 3'	

The *H19* ICR contains 32 CpG dinucleotides. These comprise seven CTCF binding sites of which the sixth has been shown to have allele specific differential methylation (Ulaner, et al., 2003). This CTCF site contains 5 CpG dinucleotides (*Figure 2.3*), which were assessed in this study for differential methylation. Within these sites, CpG site 4 was excluded from the analysis, as it is the site of a single nucleotide polymorphism (SNP - rs2107425). This SNP was excluded from the analysis as it corresponds to a C/T polymorphism. This polymorphism results in individuals with varying methylation profiles, based on whether they are homozygous C/C, heterozygous C/T or homozygous TT (Tost, et al., 2007). In addition, the CpG site immediately following the sixth CTCF binding site was included in the analysis. After nested PCR, the *H19* ICR PCR fragments were run on 3% agarose gels at 5 V/cm for approximately 45 minutes using a 50bp DNA size marker.

a *H19* ICR

gaaacatcccaggtcatccaagc**cg**gg**cg**ccacagggttcacaggggt**cg**tgaggtataggacactcatgggag
ccatat**cg**ggcta**cg**tgtctgattcaccccagggtgcactgttgaaggttggggagatgggaggagatactagggga
acaatgaggtgtcccagttccatggatgatggggatct**cg**gccctagtgtgaaacccttct**cg**cagggtctctggcag
gcacagagcc**cg**ggggctcttgcatagcacatgggtatttctggaggcttctccttcggtctcac**cg**cctggatggca
cggaattggttgtagttgtggaat**cg**gaagtggc**cg**c**g**c**g**g**cg**gcagtgcaggctcacacatcacagcc**cg**agc
c**cg**ccccaactggggtt**cg**ccc**g**tggaaa**cg**tccc**g**ggtcacccaagcca**cg**c**g**t**cg**cagggttca**cg**ggggt
catctgggaataggacactcatgggagc**cg**caccagatcttcaggt**cg**ggcattatccacagccc**cg**tggccc**cg**
ggtcacactc**cg**agggcttcagtgtcatggcctgggactcaagtca**cg**cctacttatgtgatgatcacagt

b Bisulfite modified *H19* ICR

gaaatattttaggttatttaagt**cg**gg**cg**ttatagggtttataggggt**cg**tgaggtataggatatttatgggagttatat**cg**
ggtta**cg**tgtttgatttatttttagggtgtattgttgaaggttggggagatgggaggagatattaggggaataatgaggtgttt
tagttttatggatgatggggatttt**cg**gttttagtgtgaaatttttttt**cg**tagggtttttggtaggtatagagtt**cg**gggggttttgt
atagtatatgggtatttttggaggttttttttt**cg**gttttat**cg**tttggatggta**cg**gaattggttgtagttgtggaat**CG**GAA
GTGGT**CGCGCG**G**CG**gtagtgtaggtttatatattatagtt**cg**agtt**cg**ttttaattggggtt**cg**tt**cg**tggaaa**c**
gttt**cg**ggttatttaagtta**cgcg**t**cg**tagggttta**cg**ggggttatttgggaataggatattttatgggagt**cg**tattagattt
ttaggt**cg**ggtattatttatagttt**cg**tggttt**cg**ggttatatttt**cg**agggtttt<u>agtgttatggtttgggatttaag</u>tta**cg**tttatt
tatgtgatgattatagt

Legend

- CpG dinucleotides are shown in blue
- External reverse primer is shown in green
- Internal primers are shown in pink
- Sequencing primer is shown in orange
- Sequence in uppercase red is the sixth CTCF binding site

Figure 2.3: a The *H19* ICR with CpG sites shown in green; **b** The *H19* ICR after bisulfite modification with CpG sites shown in blue.

2.4.2 *SNRPN* PCR

In addition to the *H19* ICR PCR, PCR was performed on another gene known as the small nuclear ribonucleoprotein polypeptide N (*SNRPN*). This gene is maternally imprinted and as such is expected to be completely hypomethylated in male sperm. This gene was used as a control for somatic cell contamination within the extracted sperm DNA. Published PCR conditions and primers for the amplification of the promoter region of the *SNRPN* gene encompassing 4 CpG sites were used (White, Durston, Harvey, & Cross, 2006). These PCR conditions and primers can be found in *Table 2.4* and *Table 2.5*. After the *SNRPN* PCR, products were run on 3% agarose gels at 5 V/cm for approximately 45 minutes using a 50bp DNA size marker. This was followed by pyrosequencing.

Table 2.4: *SNRPN* PCR conditions

PCR reagents *SNRPN*		**PCR Cycle conditions**		
Reagents				
ddH2O	16.8µl	94°C	7 min	
Buffer (10x)	5µl	94°C	30 sec	
$MgCl_2$ (25 mM)	5µl	58°C	30 sec	x45 cycles
dNTP mix (5 mM)	8µl	72°C	30 sec	
Normal primer (10 mM)	1µl	72°C	7 min	
Tagged primer (1mM)	1µl	15°C	∞	
Universal primer (10 mM)	1µl			
Taq pol (5 U/µl)	0.2µl			
DNA	2µl*			
Betaine	10µl			
	50µl			

* 2 µl bisulfite modified DNA

Table 2.5: *SNRPN* PCR primer sequences

Primer	Sequence	Amplicon Size
SNRPN-F primer	5'- AGGGAGTTGGGATTTTTGTATT **-** 3'	282 bp
SNRPN-R primer (tagged)	5' – CCCCAAACTATCTCTTAAAAAAAAC – 3'	
Universal Primer	5' - GGGACACCGCTGATCGTTTA - 3'	
Sequencing Primer	5'- GTAGAGGTAGGTTGG - 3'	

2.5. Pyrosequencing

2.5.1 Introduction to pyrosequencing methodology

Pyrosequencing is based on the sequencing-by-synthesis principle first described by Melamede & Wallace (1985). This principle involves the sequential addition of nucleotides to a primed template. The sequence of the template is then deduced by checking the order in which nucleotides are incorporated into the growing DNA chain (reviewed by Ahmadian, et al., 2006). Pyrosequencing uses the same principle and the sequencing reaction involves four very specific enzymes. A series of enzymatic reactions, upon the addition of a complementary nucleotide to a template strand of interest, produces light. This light production is central to the method of pyrosequencing.

When analyzing a sequence for DNA methylation, genomic DNA has to be treated with bisulfite (bisulfite modification). The sequence of interest is then PCR amplified and pyrosequenced. As a result of bisulfite modification, unmethylated cytosines are converted to thymine and methylated cytosines remain unchanged. The pyrosequencer can then

discriminate between the two as a C/T SNP. All amplicons within one sample are analyzed, after which methylation percentage is shown above each CpG site. Thus, a quantitative methylation analysis is achieved.

Before sequencing can take place, an assay must be designed using the PSQ assay design software (Biotage AB, Sweden). This software allows one to design an appropriate assay based on the gene sequence being analyzed. Following input of the gene sequence with all of the SNP positions clearly marked, the software generates possible primer sets to use for PCR as well as for the sequencing reaction. The program also generates a dispensation order for nucleotides based on the input sequence. The software picks suitable primer sets and scores them form 0 to 100 based on how well they are expected to work in the assay. Appropriate primers can then be chosen and the complete assay then saved.

As with other sequencing methods, PCR has to be carried out before pyrosequencing. The PCR reaction used here requires three primers: a normal primer, a tagged primer and a universal biotinylated primer. The tagged primer and the universal primer share a short tag sequence. As a lower concentration of tagged primer is added compared to universal primer, more amplicons should have universal primer incorporated than tagged primer. This is important for binding of PCR products in the next step.

Once PCR is complete, sepharose beads and binding buffer are added to the PCR product in a 96 well plate. The plate is then placed on an orbital shaker, which facilitates the binding of PCR products to the sepharose beads. Amplicons with the universal biotinylated primer incorporated bind to the sepharose beads, because the biotin molecule attached to the primer is attracted to the magnetic sepharose beads.

Following this, a vacuum prep tool containing 96 needles is applied to the samples. One needle is inserted into each well and the contents of each well sucked up. All the liquid from

each well is transferred to a waste tank, and the sepharose beads with bound PCR product form each well are caught on the ends of the needles.

The needles are then sequentially dipped into 70% ethanol, denaturation solution, wash buffer and ddH_2O. This removes any remaining components of the preceeding PCR reaction and denatures double-stranded DNA to single-stranded DNA. Thus, pure single-stranded DNA is obtained which can be used for pyrosequencing.

With this complete, the vacuum prep tool is switched off, and the beads released into a 96 well Pyrosequencing (PSQ) plate containing sequencing primer and annealing buffer. The plate is then left at 80˚C for a few minutes. This allows the sequencing primer to anneal to PCR products.

Once this is complete, a pyrosequencing cartridge is set up. This cartridge has separate wells to which a substrate mix, enzyme mix and each of the four dNTPs are added (*Figure* 2.4). The enzyme mix contains: Klenow DNA polymerase, ATP Sulfurylase, luciferase and apyrase. The polymerase catalyzes DNA synthesis. It is used for primer extension with subsequent release of pyrophosphate. ATP Sulfurylase catalyzes ATP release from pyrophosphate. Luciferase catalyzes light production form ATP, and apyrase degrades unincorporated nucleotides and ATP between base additions.

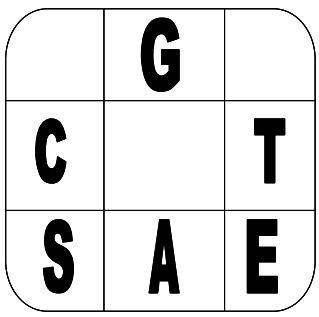

Figure 2.4: Pyrosequencing cartridge set up. S-substrate; E-enzyme; A,C,G,T-4 nucleotides.

The cartridge and the PSQ plate are then both placed into the pyrosequencing machine.

As the reaction commences, substrate mix and enzyme mix are added to each well. This is followed by the addition of separate solutions containing the four dNTPs. Successful incorporation of a nucleotide into the growing DNA chain results in the release of pyrophosphate. This is then converted, via the various enzymes, to drive light emission (*Figure 2.5*). Light emission is proportional to the number of nucleotide incorporations: Light emissions are recorded and analyzed by the pyrosequencing machine and a pyrogram is generated. A pyrogram is similar to an electropherogram, in which peak heights are proportional to the number of nucleotides incorporated.

Pyrosequencing gives a quantitative methylation analysis of individual CpG sites in a sample. Before pyrosequencing, an assay is set up in which the sequence being analyzed with all CpG positions clearly marked is entered into the assay design software. The machine then creates a dispensation order of nucleotides based on the input sequence. As sequencing commences each CpG site is treated as a C/T SNP due to the bisulfite modification of DNA. The pyrogram created shows a methylation percentage above each CpG site. The entire population of amplicons within a sample is analyzed and a quantitative result obtained. *Figure 2.6* shows how methylation analysis is achieved using pyrosequencing.

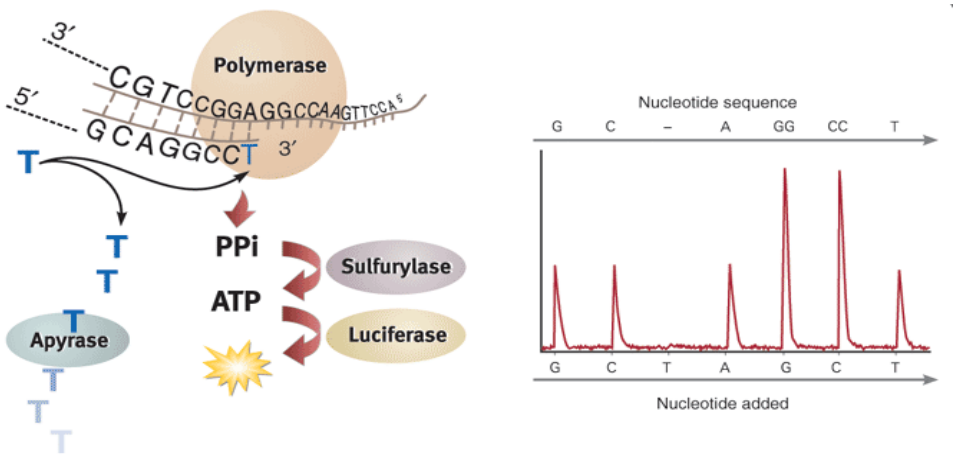

Figure 2.5: Pyrosequencing reaction with accompanying pyrogram (figure obtained from England & Pettersson, 2005).

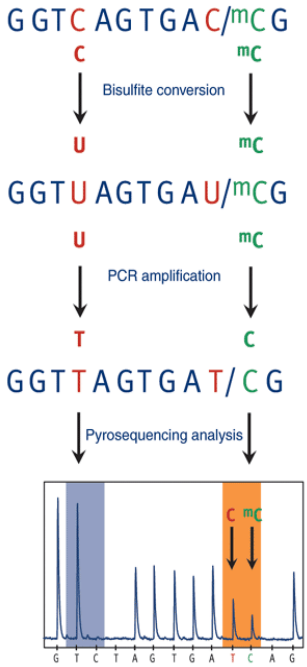

Figure 2.6: Methylation analysis using pyrosequencing (figure obtained from England & Pettersson, 2005).

2.5.2 Pyrosequencing protocol:

1. The vacuum prep tool was checked by filling a 96 well injection plate with 40µl of water and sucking this water up.
2. The four pyrosequencing troughs were filled with ddH_2O, 70% ethanol, wash buffer and denaturing buffer in preparation for the pyrosequencing run.
3. 40µl of PCR product was added to wells in a 96 well injection plate.
4. These wells were then filled with sepharose beads and binding buffer. 6µl sepharose beads and 40µl of binding buffer were added to each sample.
5. This plate was then covered and left on an orbital shaker for 10 minutes at 300 r.p.m.
6. In the interim a PSQ plate was filled with sequencing primer and annealing buffer. 1.6µl of sequencing primer and 38.4µl of annealing buffer were added to each well.
7. Once the MicroAmp plate was removed from the orbital shaker, samples were resuspended using a pipette, after which the vacuum prep tool was applied to the samples.

8. Following this, the vacuum prep tool now containing the sepharose beads plus PCR product attached is sequentially dipped into the troughs containing denaturing buffer, ethanol, wash buffer and water for approximately 12 seconds each.
9. The vacuum pump is then turned off and the beads released into the PSQ plate containing sequencing primer and annealing buffer.
10. The PSQ plate was then placed on a heating block at 80°C for 3 minutes.
11. While the samples are denaturing the pyrosequencing cartridge is set up. Substrate mix, enzyme mix, and the four dNTPs are placed into specific wells in the cartridge. The amounts of each liquid to be added is obtained from the pyrosequencing software and differs according to how many samples are being analyzed at a time.
12. The cartridge and the PSQ plate were then placed into their respective positions in the pyrosequencer after which the run was started.

2.5.3 Pyrosequencing data analysis

Pyrosequencing data was analyzed using the Pyro Q-CpG Software (Biotage). The output given after analysis is in the form of quantitative percentages of DNA methylation per CpG site within the sequence being analyzed. The analysis also includes an inbuilt control for detecting incomplete bisulfite modification. This is important as the incomplete conversion of cytosines to thymines may give rise to false positive methylation levels. If this was detected, then the data from the sample being analyzed was discarded.

2.6 Luminometric Methylation Assay (LUMA)

LUMA has 3 steps:

- First, genomic DNA samples are incubated at 37˚C for 4 hours in two reaction mixtures. One containing the methylation sensitive *HpaII* restriction enzyme together with the control enzyme *EcoRI* and the other containing *MspI* together with *EcoRI*. All enzymes used were at a concentration of 10U/µl. *HpaII* and *MspI* both recognize the sequence CCGG, however, *HpaII* cannot cut if the internal C is methylated. Therefore, *MspI* will cut all CG sites whereas *HpaII* will cut only unmethylated CG sites. *HpaII* and *MspI* both leave 5'-CG overhangs and *EcoRI* leaves a 5'-AATT overhang.

- Second, after complete DNA digestion the overhangs are filled in using a polymerase extension assay using the Pyrosequencing^{TM} platform. The pyrosequencing reaction nucleotides are added sequentially: 1) dATP (fills in the *EcoRI* overhangs); 2) mixture of dCTP and dGTP (fills in the *HpaII* and *MspI* overhangs); 3) dTTP (fills in the *EcoRI* overhangs) and 4) mixture of dCTP and dGTP (used as a control step and no peak is expected as all overhangs should have been filled in at step 2) (*Figure 2.7*).

- Third, peak heights are analyzed and a *HpaII/MspI* ratio calculated. The peak heights are analyzed on the pyrosequencing machine and then used to calculate a ratio which can be translated into a global DNA methylation percentage. The ratio is calculated as follows: 1- ($\frac{HpaII}{MspI}$) x 100.

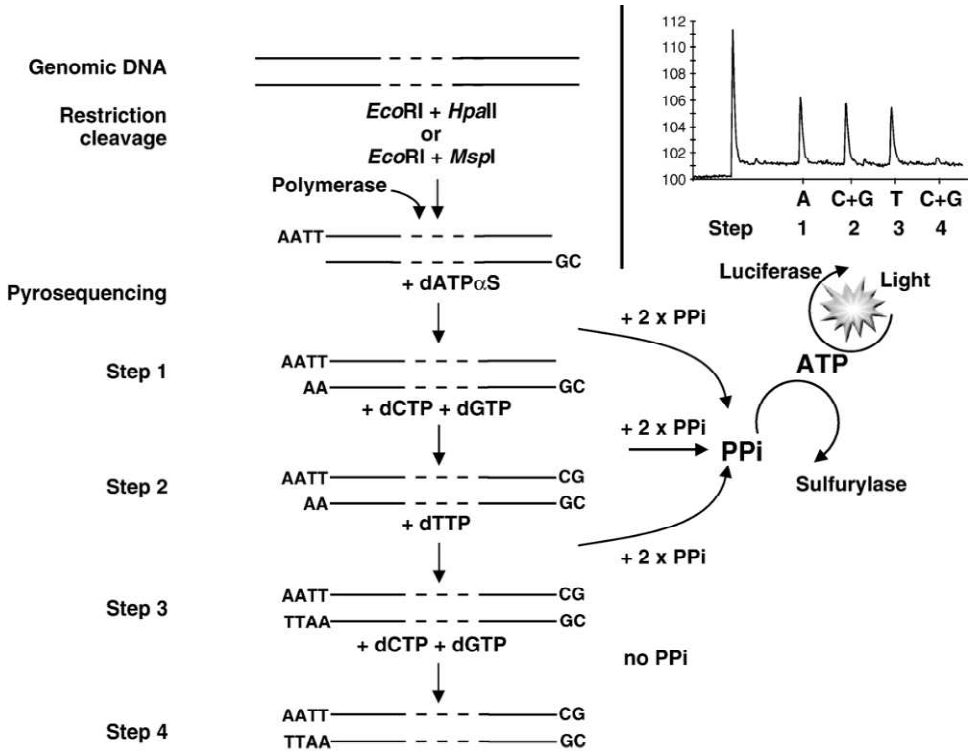

Figure 2.7: Diagram showing the LUMA pyrosequencing reaction (figure obtained from Karimi, et al., 2006)

2.6.1 LUMA data analysis

LUMA data was analyzed using the PyroMark Q96MA application software (Biotage AB, Sweden). The output is in the form of peak heights from which an appropriate ratio can be calculated and a global DNA methylation percentage obtained.

2.7. Statistical analysis of data

All statistical analyses were performed using Statistica, version 9.1 (StatSoft Inc., UK).

2.7.1 LUMA optimization on blood samples

LUMA requires a very high starting concentration of genomic DNA. For analysis of one sample, 2500ng of genomic DNA is required.

When collecting sperm samples and extracting DNA, very small amounts of DNA are obtained. As the sperm samples collected were needed for more than just one analysis, the LUMA technique was first optimized using blood DNA. One sample was used in 3 independent experimental runs, each time across five different concentrations ranging from 50ng/µl – 250ng/µl. It was tested for both reproducibility as well as reproducibility across lower concentrations. All the results from these runs are reported.

2.7.2 Global DNA methylation and *H19* ICR DNA methylation

2.7.2.1 Descriptive statistics

Descriptive statistics for both global DNA methylation and *H19* ICR DNA methylation are reported in terms of central tendency (mean and median levels) and variation (standard deviation and range), and graphically represented as box and whisker plots. The DNA methylation per CpG site within the *H19* ICR is graphically represented as a bar graph. Included in the descriptive statistics on DNA methylation are the descriptive statistics of all possible confounding variables.

In addition to the above, descriptive statistics for the *SNRPN* locus are also presented. This locus is maternally imprinted and was used to assess the level of somatic contamination in our samples.

2.7.2.2 Hypothesis testing

The main aim of this study is to examine global and locus specific *H19* ICR DNA methylation in spermatozoa, related to alcohol exposure. This includes alcohol exposure as a whole and alcohol exposure in a dosage-dependent manner.

In order to test the hypothesis that alcohol use decreases DNA methylation both globally and locus specifically, both Mann-Whitney U and T-tests for independent groups were used. The Mann-Whitney U test is the non-parametric equivalent of the independent samples T-test, and is used to asses differences between two groups when the data is not normally distributed.

For global DNA methylation analysis, the T-test for independent groups was used. This test is used when data is split into two groups, is continuous and is normally distributed. P-values of <0.05 were considered significant.

For average *H19* ICR DNA methylation as well as methylation per CpG site, the Mann-Whitney U test was used. This is because the data in these cases are split into two groups, and are not normally distributed. P-values of <0.05 were considered significant.

In order to test the hypothesis that alcohol use decreases DNA methylation both globally and locus specifically in a dosage-dependent manner, Spearman's rank correlations were used. Spearman's rank correlations are the non-parametric equivalent of Pearson's correlations, and indicate both the strength and direction of linear relationships between variables. This correlation assesses how well the relationship between DNA methylation and alcohol use can be described. This correlation was used because the data are not normally distributed and are continuous in nature. P-values of <0.05 were considered significant.

2.7.2.3 Correlations and multiple regression

There are many confounding variables apart from alcohol that have been shown to influence DNA methylation. These include: age, drug use, smoking, sperm concentration, sperm morphology and sperm motility. Therefore, in order to test whether or not these confounders were significantly associated with DNA methylation in our samples, the non-parametric Spearman's rank correlations were calculated between each variable and DNA methylation. Those that were found to be significantly associated with DNA methylation at the $p < 0.05$ level, are presented.

Those confounding variables that were found to be significantly associated with DNA methylation were then put into multiple regression models together with alcohol exposure. Multiple regression, allows us to predict the quantitative effect of several variables on an outcome variable. For example, in our study it would allow us to predict the quantitative effect of alcohol and smoking on DNA methylation. In addition, it will also allow us to see how much each variable on its own contributes to the observed DNA methylation. Multiple regression is expressed in the form of an equation, shown below.

$$Y = a + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3$$

In this equation, Y is the dependent variable, in our case DNA methylation. 'a' is a constant or the intercept. The intercept is defined as the mean of the dependent variable when all predictor variables are equal to zero. β_1 is the beta value of X_1 or variable one. The beta value tells us how strongly X_1 influences the dependent variable. Therefore, together X_1, X_2 and X_3 can tell us how much they each contribute to the dependent variable and how much they contribute together in predicting the dependant variable. If the regression model has

only one variable then the β value would be equal to the correlation value between that variable and the dependent variable.

For regression models in which the predictor variables are strongly correlated with each other, a forward stepwise method of regression is used. This method allows us to add one variable into the model at a time to see how significantly it correlates with the dependent variable. Once the variable with the largest effect is added, it takes the other variables and assesses their effect. If no significant effect is seen in the presence of the variable with the largest effect, the variable is then removed.

All regression models give us an adjusted R^2 value. This value indicates the proportion of variance in the dependent variable that is accounted for by our model.

For each significant correlation found, two regression models are presented. One model, containing treatment and the significant cofounding variables, and the other model containing drinking frequency and the appropriate confounding variables. In this way, we can see whether or not alcohol has a significant effect on DNA methylation as a whole or in a dosage-dependent manner.

2.7.3 Sperm parameters and DNA methylation

In order to test the effect of individual sperm parameters on DNA methylation independently from alcohol use, the samples were divided into abnormal and normal sperm parameters and then tested for differences in methylation using Mann-Whitney U tests. Tests that revealed significance at the $p < 0.05$ level are presented accompanied by box and whisker plots.

3. Results

The main aim of this study was to examine global DNA methylation and locus specific *H19* ICR DNA methylation in spermatozoa, related to alcohol exposure. To examine global DNA methylation, the LUMA technique had to be optimized in both blood and sperm derived DNA samples. LUMA was then used to assess differences in global DNA methylation levels in alcohol exposed spermatozoa and non-alcohol exposed spermatozoa, and the dosage-dependent effects of alcohol on global DNA methylation was also assessed. *H19* ICR DNA methylation levels in alcohol exposed spermatozoa and non-alcohol exposed spermatozoa were compared using pyrosequencing and the dosage-dependent effects of alcohol on *H19* ICR DNA methylation was also assessed.

3.1 LUMA

3.1.1 Optimization on blood samples

The DNA concentration required for LUMA is high; therefore we attempted to try to lower the concentration of input DNA as much as possible without compromising the quality of the results. The LUMA technique was optimized using blood derived DNA. The technique was tested for reproducibility as well as reproducibility across different concentrations. One blood sample was used for this analysis in three independent experiments, each time across five different concentrations. These were 50ng/µl, 100ng/µl, 150ng/µl, 200ng/µl and 250ng/µl. The lowest concentration found to work reliably was 100ng/µl. The values of the global DNA methylation at each concentration across the three runs are shown below (*Table 3.1*).

Table 3.1: Values of global DNA methylation at each concentration across the three experimental runs

Concentration (ng/µl)	% Methylation		
250	72	71	72
200	71	73	73
150	70	73	73
100	70	72	73
50	66	69	66

As can be seen, 100-250ng/µl showed good reproducibility of results whereas 50ng/µl was inconsistent. Therefore 100ng/µl was used for all subsequent analyses.

3.1.2 Global DNA methylation analysis of sperm samples

LUMA was performed on a total of 58 sperm DNA samples. These consisted of 18 non-drinkers/controls and 40 drinkers. The global DNA methylation results for these samples together with the ratios used to calculate them are shown in *Table 3.2*.

Table 3.2: *HpaII/EcoRI* and *HpaII/MspI* ratios and global DNA methylation results for sperm samples

Sample	Msp I/EcoRI	Hpa II/EcoRI	Hpa II/ Msp I	(1 - Hpa II/Msp I)*100
				Global Methylation (%)
Controls				
HSS 027	2.173	0.652	0.300	70
HSS 031	1.661	0.524	0.315	69
HSS 033	1.864	0.607	0.326	67
HSS035	3.508	0.811	0.231	77
HSS 037	2.508	0.706	0.281	72
HSS 038	1.762	0.785	0.446	55
HSS 046	2.384	0.838	0.352	65
HSS 056	1.488	0.548	0.368	63
HSS 070	1.684	0.467	0.277	72
HSS 074	2.531	0.799	0.316	68
HSS 077	1.534	0.534	0.348	65
HSS 082	2.166	0.763	0.352	65
HSS 085	2.076	0.698	0.336	66
HSS 088	1.459	0.469	0.321	68
HSS 089	2.056	0.694	0.338	66
HSS 094	2.481	0.713	0.287	71
HSS 096	2.461	0.861	0.35	65
HSS 131	2.355	0.741	0.315	69
Drinkers				
HSS 003	1.885	0.725	0.385	62
HSS 023	2.048	0.753	0.368	63
HSS 024	2.159	0.675	0.313	69
HSS 028	1.54	0.56	0.364	64
HSS 036	3.604	1.459	0.405	60
HSS 043	1.916	0.723	0.377	62
HSS 053	1.871	0.536	0.286	71
HSS 061	1.731	0.474	0.274	73
HSS 062	2.086	0.654	0.314	69
HSS 064	2.188	0.677	0.309	69
HSS 068	1.827	0.565	0.309	69
HSS 069	2.016	0.652	0.323	68
HSS 072	2.705	0.775	0.287	71
HSS 073	1.622	0.481	0.297	70
HSS 076	2.097	0.703	0.335	67
HSS 078	1.817	0.562	0.309	69
HSS 081	1.937	0.632	0.326	67
HSS 084	1.967	0.674	0.343	66
HSS 086	1.921	0.671	0.349	65
HSS 090	1.621	0.542	0.334	67
HSS 092	1.987	0.664	0.334	67
HSS 095	2.196	0.97	0.442	56
HSS 103	1.796	0.745	0.415	59
HSS 105	2.18	0.736	0.338	66
HSS 111	2.191	0.715	0.326	67
HSS 114	2.582	0.816	0.316	68

HSS 117	2.515	0.809	0.322	68
HSS 119	2.508	0.942	0.376	62
HSS 122	1.772	0.637	0.359	64
HSS 123	2.012	0.916	0.455	55
HSS 126	1.828	0.572	0.313	69
HSS 127	2.156	0.658	0.305	70
HSS 128	2.093	0.818	0.391	61
HSS 130	1.755	0.721	0.411	59
HSS 132	2.17	0.723	0.333	67
HSS 133	2.449	0.903	0.369	63
HSS 134	2.92	0.735	0.252	75
HSS 135	1.912	0.657	0.344	66
HSS 137	1.934	0.757	0.391	61
HSS 140	1.314	0.435	0.331	67

An example of how global methylation was calculated for sample HSS 027 is as follows:

$$1 - \frac{(HpaII)}{MspI} \times 100$$

$$= 1 - \frac{0.652}{2.173} \times 100$$

$$= 69.996$$

$$= 70\%$$

3.1.2.1 LUMA descriptive statistics and hypothesis testing

Table 3.3 shows the descriptive statistics of global DNA methylation as well as possible confounding factors of the LUMA samples. The majority of the samples are of African and Caucasian descent and the age range between the controls and the drinkers is similar ($p = 0.54$). Sperm parameters are variable in both groups and the range of the drinking frequency in the drinkers is extremely large. Cigarette use was high in the drinking group compared to the control group ($p = 0.08$), suggesting possible co-occurrence of smoking and drinking. None of the confounding variables (sperm parameters, smoking and drug use) were found to

correlate significantly with global DNA methylation. *Figure 3.1* shows a box plot for global DNA methylation of the two treatment groups: controls and drinkers.

Table 3.3: Global DNA methylation descriptive statistics (58 samples)

	Treatment Groups				
	Controls		Drinkers		Controls vs. Drinkers
Sample Size (N)	18		40		
Statistic	**Mean ± SD**	**Median (Range)**	**Mean ± SD**	**Median (Range)**	
Global DNA Methylation (%)	67.5 ± 4.5	67.7 (21.5)	65.7 ± 4.5	66.6 (20.3)	p = 0.17
Age (Years)	32.4 ± 4.4	33.0 (16)	33.9 ± 5.1	33.0 (24)	p = 0.54
Drinking Frequency (drinks/month)	-	-	38.0 ± 64.5	12.0 (359)	
Sperm concentration (million/ml)	54.6 ± 30.9	55.0 (119)	73.8 ± 69.3	12.0 (305.8)	
Sperm Morphology (% Normal)	10.9 ± 4.3	10.0 (15)	12.1 ± 6.3	57.0 (24)	
Sperm Motility (%)	54.4 ± 10.1	50.0 (30)	57.6 ± 10.9	60.0 (55)	
Ethnicity (%)	**Proportion of N**		**Proportion of N**		
White	0.0 (0/18)		47.5 (19/40)		
Black	61.1 (11/18)		37.5 (15/40)		
Coloured	5.6 (1/18)		7.5 (3/40)		
Indian	33.3 (6/18)		7.5 (3/40)		
Cigarette Use (%)	16.7 (3/18)		40.0 (16/40)		p = 0.08
Drug Use (%)	0.0 (0/18)		7.5 (3/40)		p = 0.23

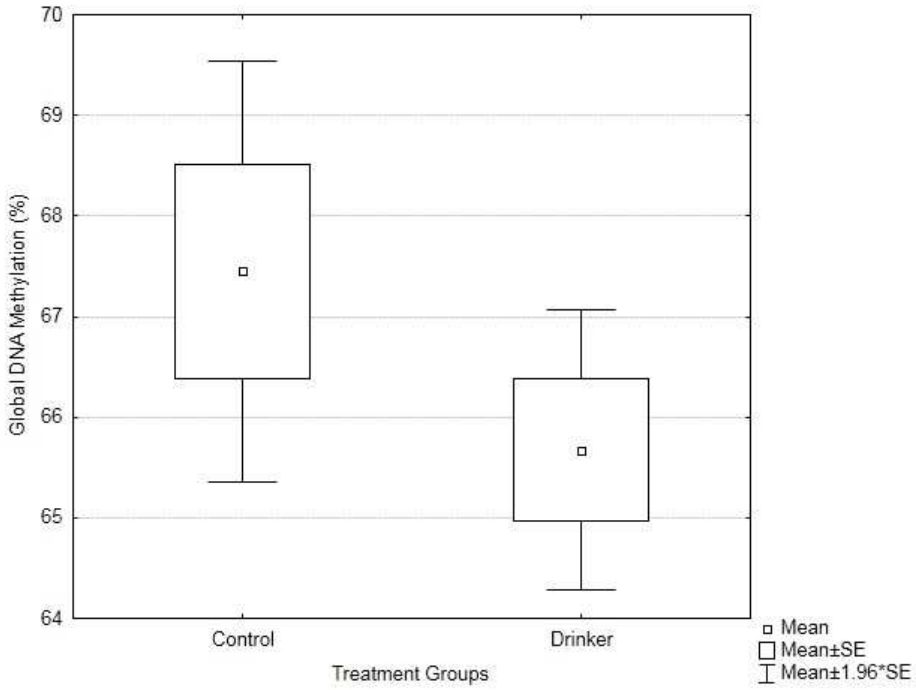

Figure 3.1: Box and whisker plot showing the spread of global DNA methylation in the two treatment groups on a scale from 64-70%.

The effect of alcohol on global DNA methylation was assessed using the T-test for independent groups and a p-value of 0.17 was obtained. A trend towards global demethylation was observed in the drinkers. To test whether or not alcohol had an effect on global DNA methylation in a dosage-dependent manner, global DNA methylation and drinking frequency were correlated using the Spearman's rank correlation. The Spearman's rho value was -0.14 ($p = 0.31$), which means there is a decrease in global DNA methylation with increasing alcohol consumption although it is not significant. In this study there is no significant difference in global DNA methylation between the drinking group and the control group, nor is there a correlation with drinking frequency.

3.2 *SNRPN*

Table 3.4 shows the descriptive statistics of average *SNRPN* DNA methylation. A representative sample of 20% of the total sample was analyzed for contamination. This included both controls and drinkers. As can be seen the mean methylation of *SNRPN* in these samples is 2.4%, indicating that there is no appreciable contamination of somatic cells in the sperm samples. Raw data of *SNRPN* results are shown in *Appendix G*.

Table 3.4: *SNRPN* descriptive statistics

Sample Size (N)	22	
Statistic	**Mean ± SD**	**Median (Range)**
Average *SNRPN*DNA Methylation (%)	2.4 ± 1.4	2.5 (4.9)

3.3 *H19*

Results for a total of 110 samples in triplicate were obtained. These consisted of 32 controls and 78 drinkers. The missing three results are due to insufficient amounts of DNA for analysis. *Table 3.5* shows the descriptive statistics of the DNA methylation at the *H19* ICR as well as descriptive statistics of possible confounding factors. CpG site 4 was excluded from all analyses as it is the site of a single nucleotide polymorphism (SNP) and would skew the data. The full table of results including CpG site 4 can be found in *Appendix H*.

Table 3.5: Descriptive statistics of the *H19* ICR (110 samples). Average DNA methylation as well as DNA methylation per CpG site is shown. The descriptive statistics of confounding factors are also shown.

	Treatment Groups				
	Controls		**Drinkers**		**Controls vs. Drinkers**
Sample Size (N)	32		78		
Statistic	**Mean ± SD**	**Median (Range)**	**Mean ± SD**	**Median (Range)**	
Average *H19* ICR DNA Methylation (%)	90.2 ± 2.0	90.7 (8.5)	91.4 ± 3.0	92.0 (19.7)	P = 0.0514
***H19* CpG 1 Methylation (%)**	91.8 ± 2.6	92.5 (10.3)	91.5 ± 4.6	92.6 (28.7)	P = 0.73
***H19* CpG 2 Methylation (%)**	99.8 ± 0.6	100.0 (2.7)	99.7 ± 1.2	100.0 (9.7)	P = 0.69
***H19* CpG 3 Methylation (%)**	75.7 ± 7.0	75.0 (24)	78.7 ± 8.5	80.5 (44)	P = 0.03
***H19* CpG 5 Methylation (%)**	91.2 ± 2.5	91.3 (15.3)	90.6 ± 3.4	90.6 (24.3)	P = 0.23
***H19* CpG 6 Methylation (%)**	96.1 ± 2.0	96.0 (7.7)	96.4 ± 3.0	96.6 (22)	P = 0.30
Age (Years)	34.4 ± 6.0	35.0 (28)	33.8 ± 5.5	34.0 (29)	P = 0.70
Drinking Frequency (drinks/month)	-	-	36.1 ± 51.4	15.0 (359)	
Sperm concentration (million/ml)	52.2 ± 35.9	43.0 (156)	58.4 ± 58.4	41.5 (305.8)	
Sperm Morphology (% Normal)	11.2 ± 5.7	11.0 (20)	11.0 ± 6.4	12.0 (24)	
Sperm Motility (%)	55.3 ± 10.2	55.0 (40)	55.7 ± 11.2	60.0 (55)	
Ethnicity (%)	**Proportion of N**		**Proportion of N**		
White	6.3 (2/32)		41.0 (32/78)		
Black	56.3 (18/32)		46.2 (36/78)		
Coloured	9.4 (3/32)		5.1 (4/78)		
Indian	28.1 (9/32)		7.7 (6/78)		
Cigarette Use (%)	18.8 (6/32)		41.0 (32/78)		P = 0.26
Drug Use (%)	3.1 (1/32)		6.4 (5/78)		P = 0.49

The majority of the sample are of African and Caucasian descent and the age range between the two groups is similar (p = 0.70). Sperm parameters are similar in both groups and the range of the drinking frequency in the drinkers is extremely large. Cigarette use was high in the drinking group compared to the control group, but not statistically significant (p = 0.26), suggesting possible co-occurrence of smoking and drinking. Average *H19* ICR DNA methylation as well as DNA methylation per CpG site (excluding CpG site 4) is similar in both groups. All methylation levels are high (>90%) except for CpG site 3 where levels are lower (75-80%). The standard deviation and range for CpG site 3 was larger than for any of the other sites. *Figure 3.2* shows a box plot for average *H19* ICR DNA methylation of the two treatment groups: controls and drinkers. The effect of alcohol on average *H19* ICR DNA methylation was assessed using the Mann-Whitney U test (p = 0.0514), and was borderline significant. As can be seen the control group has a smaller range (whiskers) compared to the drinking group, however, the clustering of half of the drinking group is much tighter than that of the control group (box).

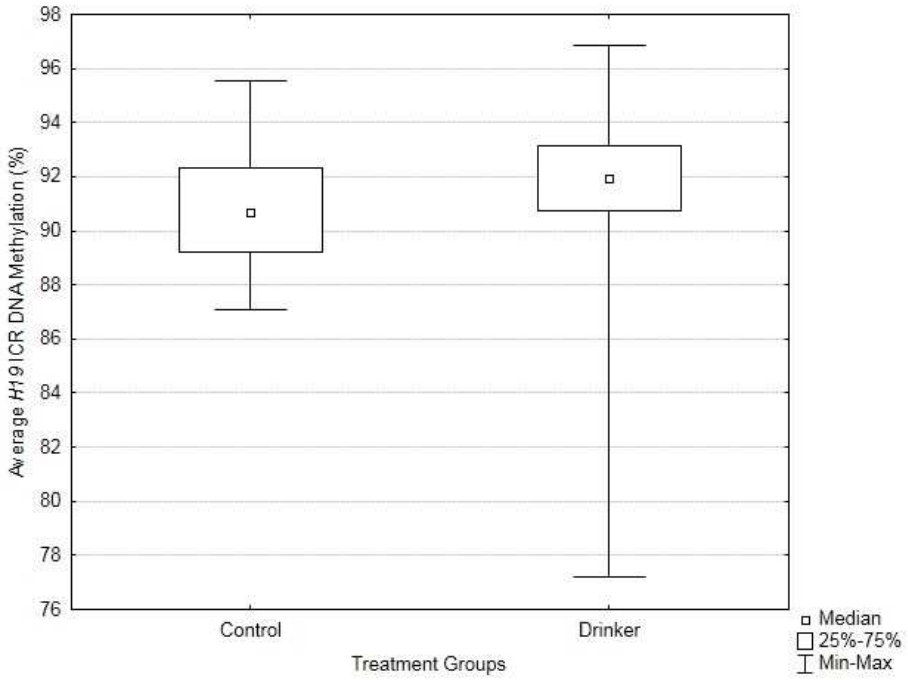

Figure 3.2: Box and whisker plot showing the spread of average *H19* ICR DNA methylation in the two treatment groups on a scale from 76-98%.

To test whether or not alcohol had an effect on average *H19* ICR DNA methylation in a dosage-dependent manner, average *H19* ICR DNA methylation and drinking frequency were correlated using the Spearman's rank correlation. The Spearman's rho value was 0.07 (p = 0.47), which means there is an increase in *H19* ICR DNA methylation with increasing alcohol consumption although it is not significant

The effect of alcohol on individual CpG sites (excluding CpG site 4) within the *H19* ICR was assessed using the Mann-Whitney U test. As can be seen in *Figure 3.3*, there was only a significant difference in methylation at CpG site 3 within the *H19* ICR (p = 0.03). In this study there is no significant correlation between alcohol exposure and average *H19* ICR DNA methylation, nor is there a correlation with drinking frequency. There is also no significant correlation between alcohol exposure and DNA methylation at individual CpG sites except for CpG 3, where there is an increase in DNA methylation in the drinking group.

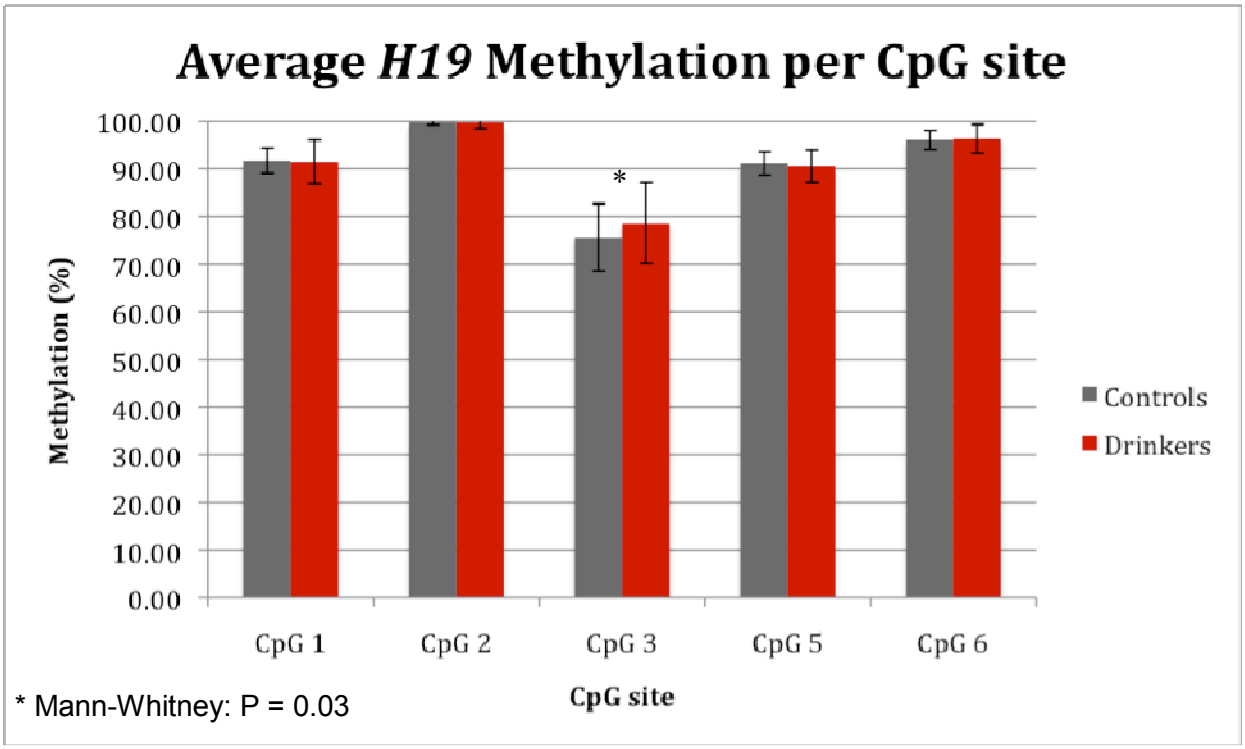

Figure 3.3: Bar graph showing the DNA methylation per CpG site in the *H19* ICR in the two treatment groups.

To test whether or not alcohol had an effect, on DNA methylation at individual CpGs within the *H19* ICR, in a dosage-dependent manner, the DNA methylation at each CpG site was correlated with drinking frequency using the Spearman's rank correlation. The values obtained from these analyses are shown in *Table 3.6*. There were no significant correlations between DNA methylation per CpG site and drinking frequency.

Table 3.6: Spearman's correlation values for methylation per CpG site and drinking frequency

CpG site	Spearman's rho	p-value
1	-0.08	0.39
2	0.01	0.89
3	0.10	0.26
5	-0.13	0.18
6	0.08	0.42

3.3.1 *H19* correlations

Table 3.7 shows confounding variables that were found to be significantly correlated with *H19* ICR DNA methylation ($p < 0.05$). These values were obtained by looking at the drinking group only. This was done because one of the aims of this study is to look at factors in addition to alcohol, which are predicting DNA methylation. In the control group, alcohol use is not a predicting variable, and therefore, there cannot be any variables that confound the effect of alcohol. Age significantly affects DNA methylation at CpG sites 1, 3 and 5 as well as overall average *H19* ICR DNA methylation. In all instances it causes an increase in DNA methylation. Sperm motility is associated with a significant decrease in DNA methylation at CpG site 3.

Table 3.7: Spearman's correlation values for confounding variables that were significantly associated with DNA methylation

Confounding Variables	*H19* DNA Methylation	Spearman's rho	p-value
Age	Average *H19* ICR DNA Methylation	0.30	0.001
Age	CpG site 1	0.30	0.007
Age	CpG site 3	0.37	0.0009
Sperm Motility	CpG site 3	-0.22	0.049
Age	CpG site 5	0.23	0.047

3.3.2 *H19* multiple regression models

Multiple regression allows us to predict the quantitative effect of several variables on an outcome variable. In addition, it will also allow us to see how much each variable on its own contributes to the observed DNA methylation. The confounding variables found to be significantly associated with DNA methylation at the *H19* ICR (*Table 3.7*) were put into regression models together with alcohol use and drinking frequency in order to assess how much each variable on its own contributes to the observed DNA methylation and what the quantitative effects of the variables are on the observed DNA methylation.

Table 3.8 shows the multiple regression models obtained for *H19* ICR DNA methylation. As can be seen, average *H19* ICR DNA methylation can be expressed by the following equations:

Average *H19* ICR DNA Methylation = 84.50127 + 0.18640 Age + 0.57869 Treatment

The first equation, tells us that average *H19* ICR DNA methylation was shown to significantly increase by 0.19% for every 1 year increase in age ($p < 0.001$) and increase by 0.58% with alcohol consumption. This increase however, was not found to be significant ($p= 0.288$). Average *H19* ICR DNA methylation would be 84.50% if the regression coefficients of age

and treatment were zero and this was shown to be significant ($p < 0.001$). The adjusted R^2 value for this model was 0.13. Therefore 0.13% of the variation in average DNA methylation at the *H19* ICR can be explained by age and treatment.

Average *H19* ICR DNA Methylation = 85.10311 + 0.18382 Age - 0.00323 Drinking Frequency

Likewise, the second equation tells us that average *H19* ICR DNA methylation was shown to significantly increase by 0.18% for every 1 year increase in age ($p < 0.001$) and decrease by 0.003% with increasing levels of alcohol consumption. This decrease however, was not found to be significant ($p=0.522$). Average *H19* ICR DNA methylation would be 85.1% if the regression coefficients of age and drinking frequency were zero; this was shown to be significant ($p < 0.001$). The adjusted R^2 value for this model was 0.13. Therefore 0.13% of the variation in average DNA methylation at the *H19* ICR can be explained by age and drinking frequency. The same regression models were applied to CpG 1, 3, 5 and 6, as shown in *Table 3.8*.

In this study, all regression models gave and R^2 value between 0.002-0.13 indicating that the variance in DNA methylation that our models account for is extremely small. For CpG site 3, age and sperm motility were both significantly correlated with DNA methylation. These two variables were also significantly correlated with one another and therefore a forward stepwise method of regression was used. As age has the largest effect on DNA methylation at this site, it was added to the regression model first. Sperm motility does not have a significant effect on DNA methylation in the presence of age and therefore only age and treatment/drinking frequency were left in the model.

Table 3.8: Multiple regression models for *H19* ICR DNA methylation

	Variable	%	SE	P-value
Average *H19* ICR DNA Methylation	Intercept	84.5	1.59	0.000
	Treatment	0.58	0.54	0.288
	Age	0.19	0.04	0.000
	Adjusted R^2 = 0.13			
Average *H19* ICR DNA Methylation	Intercept	85.1	1.53	0.000
	Drinking Frequency	-0.003	0.01	0.522
	Age	0.18	0.04	0.000
	Adjusted R^2 = 0.13			
CpG 1	Intercept	82.71	2.37	0.000
	Treatment	-0.13	0.81	0.874
	Age	0.26	0.07	0.000
	Adjusted R^2 = 0.11			
CpG 1	Intercept	82.95	2.28	0.000
	Drinking Frequency	-0.008	0.01	0.316
	Age	0.26	0.07	0.000
	Adjusted R^2 = 0.12			
CpG 3	Intercept	60	4.75	0.000
	Treatment	3.32	1.62	0.043
	Age	0.46	0.13	0.001
	Adjusted R^2 = 0.11			
CpG 3	Intercept	62.82	4.61	0.000
	Age	0.44	0.13	0.001
	Adjusted R^2 = 0.03			
CpG 5	Intercept	87.17	1.9	0.000
	Treatment	-0.53	0.64	0.412
	Age	0.11	0.05	0.0271
	Adjusted R^2 = 0.03			
CpG 5	Intercept	86.94	1.84	0.000
	Drinking Frequency	-0.005	0.01	0.429
	Age	0.12	0.05	0.028
	Adjusted R^2 = 0.03			
CpG 6	Intercept	94.41	1.28	0.000
	Treatment	0.16	0.57	0.78
	Drug Use	1.67	1.15	0.147
	Adjusted R^2 = 0.002			
CpG 6	Intercept	94.55	1.24	0.000
	Drinking Frequency	-0.002	0.01	0.694
	Drug Use	1.7	1.14	0.14
	Adjusted R^2 = 0.003			

3.3.4 Sperm parameters and *H19* DNA methylation

Table 3.9 shows the breakdown of the 110 samples into their sperm abnormalities without taking into account alcohol consumption.

Table 3.9: Sperm samples categorized by abnormalities

***H19* Sperm Samples**	
Abnormality	**Number of samples**
Oligozoospermic	10
Teratozoospermic	6
Asthenozoospermic	1
Oligo-astheno-teratozoospermic	2
Oligo-teratozoospermic	7
Astheno-teratozoospermic	1
Normozoospermic	83
Total	**110**

When looking at the difference in *H19* average DNA methylation between normozoospermic individuals and the above categories, only one significant difference was found. This was between normozoospermic and teratozoospermic individuals. A Mann-Whitney p-value of 0.007, indicating a significant increase in average *H19* DNA methylation in teratozoospermic individuals, was obtained. *Figure 3.4* is a box plot showing this result. It must be noted that the teratozoospermic individuals were only a sample of 6.

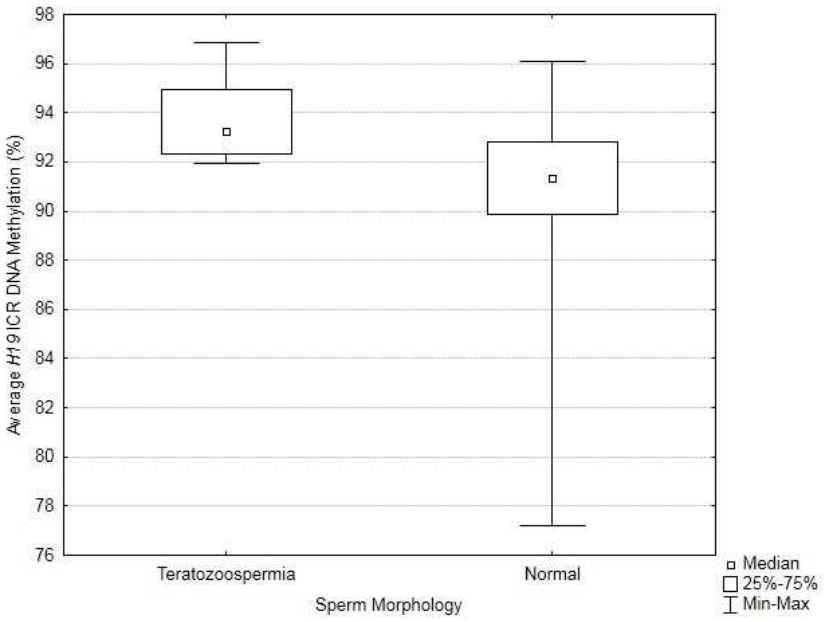

Figure 3.4: Box and whisker plot showing the spread of average *H19* ICR DNA methylation in morphologically normal (normozoospermic, n = 83) and morphologically abnormal (teratozoospermic, n = 6) sperm on a scale from 76-98%.

In summary, the results of this study indicate no significant difference in global DNA methylation between the drinking group and the control group (p = 0.17), and no significant correlation with drinking frequency (p = 0.31). There is no significant difference in average *H19* ICR DNA methylation between the drinking group and the control group (p = 0.0514), nor is there a correlation with drinking frequency (p = 0.47). No significant differences occur at individual CpG sites between the drinking group and the control group within the *H19* ICR, except for CpG site 3 (p = 0.03), which showed an increase in DNA methylation. Also, there are no significant correlations between DNA methylation per CpG site and drinking frequency.

4. Discussion

Paternal preconception alcohol exposure has been shown to result in phenotypic abnormalities (Jamerson, et al., 2004). Epigenetic changes resulting from preconception alcohol consumption by men has not been widely explored. If preconception alcohol consumption by men can alter DNA methylation patterns, there is a possibility that disruption of the methylation pattern at imprinted loci could be partly responsible for clinical features seen in FASD patients. The *H19/IG2* locus is a paternally imprinted locus and contains an ICR, which controls the reciprocal expression of these two genes. *IGF2* is a significant contributor to fetal growth. One of the key features of FAS is growth retardation and an alcohol induced decrease in *H19* ICR DNA methylation in the paternal germline, would be expected to result in biallelic *IGF2* repression in the fetus resulting in a growth retardation phenotype.

The overall objective of this study was to examine global and locus specific *H19* ICR DNA methylation in spermatozoa, related to alcohol exposure. The findings of this study show a trend towards decreased global DNA methylation in alcohol exposed spermatozoa. This was not a significant result and suggests that either alcohol does not affect global sperm DNA methylation or that the technique used in this study was not sensitive enough to detect small, but significant changes in global DNA methylation. Average *H19* ICR DNA methylation was not significantly affected by alcohol exposure, although a significant site specific change (at CpG 3) was observed. These findings suggest that we cannot implicate paternally induced DNA methylation changes at the *H19* ICR in contributing to the FASD phenotype of growth retardation. The discussion that follows further explains these findings and suggests possible reasons for, and implications of, these findings.

The mean methylation level obtained for *SNRPN* in our study was 2.4%. *SNRPN* is a maternally imprinted gene and as such was used to check for somatic cell contamination. In sperm samples *SNRPN* should theoretically be completely unmethylated. Geuns, De Rycke, Van Steirteghem, & Liebaers (2003) looked at *SNRPN* methylation in sperm using cloning and a sequencing method and found it to be 0-5% methylated. Elsewhere the DNA methylation level of a maternally imprinted gene in paternal uniparental disomy was found to be less than 25% (Choufani, et al., 2011). As can be seen when put into practice, unmethylated DNA is never 0% or fully unmethylated. The average *SNRPN* methylation levels for individual samples in this study were between 0 and 5% and as the overall mean *SNRPN* methylation level was <5%, this indicates that there was no significant somatic cell contamination in our sperm DNA samples. This is important because somatic cell contamination would greatly affect the DNA methylation levels observed in this study. Somatic cells have different DNA methylation patterns at imprinted loci as well as globally. Pyrosequencing is a quantitative method and analyzes the entire cell population in a given sample. If there was somatic cell contamination in our samples the difference in methylation patterns between the two cell populations would be reflected in the methylation pattern, and result in an inaccurate representation of DNA methylation levels in sperm.

4.1 Global DNA methylation

4.1.1 LUMA as a measure for global DNA methylation

In order to examine global DNA methylation the LUMA technique needed to be optimized. During this optimization the reproducibility of the technique was examined. In addition, as the starting quantity of genomic DNA needed for the technique is so high, the lowest amount of DNA that could be used without compromising the quality of the results was determined.

500ng (5µl of 100ng/µl) of DNA was found to be the lowest quantity that did not compromise results.

Most studies looking at global DNA methylation are conducted using repetitive elements such as long interspersed nuclear element-1 (*LINE-1*), as surrogate markers for estimating global DNA methylation (Bollati, et al., 2007; Zhu, et al., 2010). This is because *LINE-1* is one of the most common and well-characterized repetitive elements, and more than 90% of CpG methylation within the genome occurs in such repetitive elements (Baccarelli & Bollati, 2009; Zhu, et al., 2010). These studies show blood DNA methylation levels to range from 57-85%, with the majority of individuals falling into the range of 72-78% (Bollati, et al., 2007; Hsiung, et al., 2007; Zhu, et al., 2010). These methylation levels are from individuals who are not diseased but have a range of characteristics with respect to age, gender and lifestyle habits.

In this study the average methylation level obtained for the single blood sample used in LUMA optimization was approximately 72%, which falls neatly within the range observed in the above studies.

4.1.2 Global DNA methylation and alcohol use

Many studies have looked at global sperm DNA methylation in both mice and humans (Benchaib, et al., 2005; Tunc & Tremellen, 2009; Yauk, et al., 2008). Benchaib et al. (2005) looked at the relationship between global sperm DNA methylation levels and fertilization rates and Tunc & Tremellen (2009) looked at global sperm DNA methylation levels in infertile men and the implications for fertilization rates. Yauk et al. (2008) looked at a mouse model in order to understand the effect of air pollution on sperm DNA integrity, mutation rates and DNA methylation. These studies have measured global sperm DNA methylation it terms of

arbitrary units (au) using immunohistochemical-staining methods (Benchaib, et al., 2005; Tunc & Tremellen, 2009).

Benchaib et al. (2005) analyzed semen samples left over from in vitro fertilization (IVF) procedures. They correlated global DNA methylation levels in sperm with fertilization and pregnancy rates. They found that pregnancy rates were significantly lower when global DNA methylation levels are lower than a threshold of 555 au, and concluded that altered global sperm DNA methylation levels affect embryo development.

Oxidative DNA damage and hyperhomocysteinaemia have been shown to affect DNA methylation in somatic tissues. As infertile men show impaired sperm DNA methylation, Tunc & Tremellen (2009) looked at oxidative DNA damage and hyperhomocysteinaemia as possible mechanisms underlying germ line hypomethylation. Correlations between global sperm DNA methylation levels, reactive oxygen species levels (ROS), and homocysteine levels were assesed. A significant negative correlation was found between global DNA methylation and ROS levels. This hypomethylation decreased with 3 months of antioxidant therapy. They concluded that oxidative stress does interfere with germline global DNA methytlation and that antioxidant supplementation can result in significant improvement of both DNA methylation and integrity.

Yauk et al. (2008) using a mouse model, aimed to confirm findings that chemical pollutants can cause heritable paternal germ line mutations. They also looked at the relationships of these mutations to other types of DNA modifications and the consequences thereof. They measured global sperm DNA methylation using the cytosine extension assay and the methyl-acceptance assay, at three time points after particulate air pollution exposure. These were after 3, 10 and 16 weeks of exposure respectively. No change in global DNA methylation was found after 3 weeks of exposure but global DNA hypermethylation was found at both 10 and 16 weeks of exposure, and this hypermethylation persisted into mature

sperm even after the exposure was removed. They concluded that this global hypermethylation may affect progeny, but it is uncertain if it will lead to reprogramming of methylation post fertilization. They mention that it is important to note that the global sperm DNA methylation level does not reflect which individual genes are up or down regulated and thus the consequences of the observed hypermethylation in this study is unknown.

My study is the first to examine global DNA methylation levels in human sperm using the LUMA technique and obtaining a global sperm DNA methylation percentage. The average global sperm DNA methylation levels in this study are lower than that seen in somatic tissues. The level of DNA methylation seen in the control sperm samples, ranges from 63-72%, with the majority of individuals falling into the range of 65-68.5%. Mouse studies using restriction enzyme based analyses have also reported that germ line global DNA methylation levels are lower than those found in somatic tissues (Monk, Boubelik, & Lehnert, 1987; Razin, et al., 1984).

When comparing global sperm DNA methylation levels between males who consume alcohol and those who do not, there was no significant difference ($p=0.17$). However, when looking at the box and whisker plot (*Figure 3.1*), the drinking group does show a trend towards lower global sperm DNA methylation levels compared to the control group. In addition, the Spearman's correlation assessing global sperm DNA methylation levels and drinking frequency showed a negative, but not significant, correlation of -0.14 ($p = 0.31$). This suggests that global sperm DNA methylation may be decreasing with increasing alcohol consumption, but this would need to be confirmed or refuted in a larger study. This trend agrees with findings in animal models that chronic alcohol exposure decreases global DNA methylation in sperm (Garro, et al., 1991). The exact mechanism by which this occurs is unknown. In animal models the decrease in global DNA methylation was found to be the result of lower methylase activity as a result of alcohol exposure (Garro, et al., 1991). It is

therefore plausible that this also occurs in human sperm. Yauk et al. (2008) found that air pollution causes global hypermethylation of sperm DNA in mice, which persists after exposure is removed. If alcohol does cause a decrease in global sperm DNA methylation, and this decrease persists after removal of the substance, it could have serious implications for offspring born to men who consume large quantities of alcohol. It is however, uncertain whether a very modest global methylation difference would have any biological effect.

Global DNA methylation levels do not accurately reflect the status of specific genes and whether they are activated or silenced (Carrell & Hammoud, 2010; Kobayashi, et al., 2007; Marques, et al., 2008; Yauk, et al., 2008). If some genes are significantly up regulated and others significantly down regulated, they would balance each other out, and the overall global DNA methylation percentage would not reflect this. These locus specific effects may influence phenotypes in important ways (Benchaib, et al., 2005; Yauk, et al., 2008), and this would not be picked up with methods that examine global DNA methylation.

As an alternate approach to the LUMA technique, whole methylome analysis using high throughput techniques could yield more specific results. Several techniques have been developed in the past few years to learn more about the human methylome, particularly for cancer studies (Estecio & Issa, 2009; Jacinto, Ballestar, & Esteller, 2008). Many assays using various restriction enzymes and microarray platforms have been developed. An inherent limitation of this approach is that it is not truly a genome wide analysis. Restriction enzymes rely on certain sequences to be present in the genome in order to work. This results in a bias, as only certain parts of the genome end up being analyzed (Estecio & Issa, 2009; Jacinto, et al., 2008). In order to circumvent this problem a technique known as methyl-DNA immunoprecipitation (MeDIP) was developed. This technique uses an antibody targeted against 5-methyl-cytosine and isolates the entire portion of methylated CpG sequences within a sample. This isolated sample can then be further analyzed either by PCR amplification or hybridization to a microarray or DNA sequencing (Jacinto, et al., 2008).

This will indicate which genes are present in the methylated portion of the sample and which are not. Alternatively, LUMA could be used in conjunction with a gene expression microarray. This would result in an overall methylation level (net effect) together with an indication as to which genes are up or down regulated giving rise to the observed net effect.

Zhu et al. (2010) looked at global DNA methylation levels in blood to see if certain lifestyle factors, including alcohol use had a significant effect on global DNA methylation. They reported no association between alcohol use and *LINE-1* DNA methylation. They do mention however, that this result is observed from a sample made up of mostly social drinkers and studies need to focus on alcohol abusers.

4.1.3 Global DNA methylation and confounding factors

Global DNA methylation levels vary among healthy individuals and also in response to age; gender and lifestyle factors, and these must therefore be taken into account when investigating associations between global DNA methylation levels and various outcomes (Zhu, et al., 2010).

When looking at the effect of possible confounding factors (age, smoking and drug use) on global sperm DNA methylation, none were found to be significantly correlated with global sperm DNA methylation. Several studies have looked at an association between global DNA methylation levels and age (Fuke, et al., 2004; Moore, et al., 2008). There is conflicting evidence concerning a true association between age and global DNA methylation. Zhu et al. (2010) found that age did not exert a global effect on DNA methylation but rather acts at some loci and not others. This age-related alteration found in DNA methylation was very small relative to inter-individual variations in DNA methylation. When looking at smoking and global DNA methylation no significant effect was found (Zhu, et al., 2010).

My study looked at a population of men with varying sperm parameters, according to WHO standards. Other studies assessing global sperm DNA methylation and sperm parameters focused on only infertile men (Benchaib, et al., 2005; Tunc & Tremellen, 2009). These studies yielded conflicting results with one showing no association when looking at sperm parameters and global sperm DNA methylation (Benchaib, et al., 2005), and the other showing an association between global sperm DNA methylation and both sperm concentration as well as sperm morphology (Tunc & Tremellen, 2009). The small sample size in my study could have been insufficiently powered to observe the effects of sperm parameters on global sperm DNA methylation.

4.2 *H19* ICR DNA methylation

4.2.1 *H19* ICR DNA methylation and alcohol use

This study showed no significant difference in average *H19* ICR sperm DNA methylation levels between men who consume alcohol and those who do not. Boissonnas et al. (2010) used pyrosequencing to assess the differences in *IGF2* DMR and *H19* ICR DNA methylation, in spermatozoa from men with normal and abnormal sperm parameters. In the normozoospermic samples they found the average methylation of the loci studied to be 86.8±6.9%. Looking at inter- assay variation they suggest that approximately 10% of the spermatozoa in human sperm is completely unmethylated. Pyrosequencing looks at the entire population of cells within a sample and gives an overall methylation level. The presence of spermatozoa with different methylation patterns within one sperm sample may mask the effects of alcohol within the *H19* locus to a certain extent.

Heijmans, Kremer, Tobi, Boomsma, & Slagboom (2007) looked at the contribution of both heritable and environmental factors to variation in DNA methylation at the *H19* locus. They assessed 327 adolescent and middle aged twins for methylation at various CpG sites throughout the locus. They noted great inter-individual variation in methylation for each CpG site. This finding is supported in our study by the large ranges of DNA methylation seen at individual CpG sites within the *H19* locus (*Table 3.5*). They also found that epigenetic variation at this locus is primarily heritable and resistant to age-related changes (Heijmans, et al., 2007). As the *H19* locus is more resistant to environmental influence, the absence of a change in DNA methylation in response to alcohol may be explained. Loci, which show greater epigenetic stability during aging, could have more serious phenotypic consequences if altered (Heijmans, et al., 2007). If a germ line alteration at this locus were to occur due to alcohol use it would be stably inherited and possibly have a more serious phenotypic consequence.

In the present study, one site-specific difference in *H19* ICR DNA methylation was displayed. Pyrosequencing data showed that out of the 5 CpG sites analyzed, alcohol only had a significant effect on CpG 3 ($p=0.03$). This was shown to be an increase in DNA methylation, contrary to our hypothesis. This finding was also shown by Knezovich & Ramsay (2012), in murine sperm, where a significant increase in DNA methylation at an individual *H19* CpG site in alcohol exposed murine sperm was reported.

Renda et al. (2007) showed through a series of experiments involving deleting certain zinc fingers and mutating individual CpG sites within the CTCF protein, that two CpG sites within a conserved 12bp core sequence of CTCF binding sites is critical for CTCF binding. With the first of these two sites being the most crucial site for protein binding. These two sites correspond with CpG sites 2 and 3 within the 6th CTCF binding site of the *H19* ICR in this study. Even though CpG site 3 showed an increase in DNA methylation in alcohol

consuming males in our study, on comparison to the other CpG sites within the sequence, it displays the lowest methylation levels in both groups (*Table 3.5*). As can be seen, on average the methylation level at this site is almost 20% lower than that of the other sites. In our study CpG 3 was also the most amenable to fluctuations in DNA methylation as both sperm motility (p=0.05) and age (p=0.002) were significantly correlated with DNA methylation at this site. When comparing the range of methylation percentage at CpG site 3 between the drinkers and the controls we see that the range in the drinkers is almost twice as large as that of the control group (*Table 3.5*). This could reflect the ability of alcohol to affect methylation at this site and the large extent of inter-individual variation at this CpG site.

Alcohol has been shown to have a varying effect on DNA methylation (Liu, Balaraman, Wang, Nephew, & Zhou, 2009; Zhou, et al., 2011). It is able to exert site-specific increases and decreases in DNA methylation. Another study done in our research group (P.N. Pitamber, personal communication), investigated the methylation patterns of the *IG-DMR* in the same sperm samples as the ones used in this study. Selective demethylation at individual CpG sites in response to alcohol consumption was shown to be significant. This further suggests that alcohol may be exerting site-specific increases/decreases within the sperm epigenome.

4.2.2 *H19* ICR DNA methylation and confounding factors

When testing all possible confounding factors with *H19* ICR DNA methylation per CpG site and with average DNA methylation at the *H19* ICR, age came up consistently as being associated with levels of DNA methylation. Although significant, the Spearman's rho values were all small indicating that age does not have a substantial effect on DNA methylation. Age on its own could explain 0.21-0.30% of the variation seen in DNA methylation. Heijmans et al. (2007) looked specifically at the effect of age on DNA methylation at the *H19* locus and

found an extremely limited role for age and environmental related changes in comparison to heritable factors. This substantiates our findings of almost no change in DNA methylation at the *H19* ICR in response to age and other confounding variables. Other sources have found age-related changes in DNA methylation, however, these were changes in global DNA methylation levels (Ahuja & Issa, 2000; Fuke, et al., 2004). The parallel study conducted in our department on these same sperm samples found age to be significantly associated with an increase in the overall methylation at *IG*-DMR and at individual sites within *IG*-DMR. This suggests that age does have an effect on DNA methylation; however, it is likely to be site/locus specific.

One factor that was found to be significantly associated with CpG 3 DNA methylation was sperm motility. There was a negative correlation indicating that DNA methylation decreases with lowered sperm motility ($r = -0.19$; $p = 0.05$). However, when put into a regression model together with alcohol consumption, this effect was eliminated. This significant correlation is small and would not likely have a phenotypic effect. Pacheco et al. (2011) noted a mix of both hypermethylated and hypomethylated DNA at the *H19* locus with lowered sperm motility.

All regression models accounted for less than 1% of the variation in DNA methylation seen at each individual CpG site and overall at the *H19* ICR. A relevant question to raise would be how biologically relevant is this variation? Is a 1% change in DNA methylation at this locus enough to make a phenotypic difference? Renda et al. (2007) did find that CpG 2 and 3 are critical for CTCF binding, but it has not been specified how big a change needs to occur for it to have a significant effect.

Even if this change in methylation did occur in the sperm and were significant enough to have a phenotypic effect, is this change heritable? Kobayashi et al. (2007) showed that sperm containing epigenetic defects can be used in assisted reproductive technology (ART)

procedures to produce both a phenotypically and genotypically healthy child. This suggests that the epigenetic defect in the sperm was somehow corrected and not inherited. Smallwood et al. (2011), in trying to elucidate the mechanisms and requirements for *de novo* DNA methylation in germ cells, discovered a more complicated process of DNA methylation in the embryo than previously reported. They propose that after fertilization not all imprinted loci have absolute maintenance of DNA methylation, and not all non-imprinted loci lack this maintenance. They suggest that there is a much greater diversity of methylation choices. This flexibility in maintenance of DNA methylation could be a mechanism by which errors in DNA methylation are corrected, or alternatively this could be the point where errors in DNA methylation are made. Ibala-Romdhane et al. (2011) studied the heritability of imprinting defects by observing them in embryos that failed to develop normally and in the gametes used to create these embryos. They found that despite the variability in abnormal sperm parameters, all sperm samples showed a normal high methylation percentage at the *H19* ICR. This did not match some of the imprinting defects seen in the embryos produced by these sperm samples, leading them to conclude that the imprinting defects seen in the embryos did not originate in the father's germ line. This led to the hypothesis that the observed hypomethylation at the *H19* ICR is due to an alteration of DNA methylation maintenance in the embryo. All of these studies illustrate the complexity of epigenetic mechanisms and their inheritance.

Despite the complexity of epigenetic inheritance, it is important to try and understand how a paternally mediated epigenetic change could occur. As previously stated DNA methylation and histone modifications are intricately linked, and together control imprinted gene expression (reviewed by Cedar & Bergman, 2009; Liu, et al., 2008). It has been shown that although most sperm DNA is coated in protamines, approximately 15% of the sperm genome retains histones and that these are retained at loci critical in early development, including at the *IGF2 locus* (Ooi & Henikoff, 2007). Alcohol has the ability to disrupt histone modifications (reviewed by Shukla, et al., 2008). If this is what is happening in the sperm

genome, it is possible that this histone modification is transmitted to offspring and upon re-setting of imprints affects DNA methylation (reviewed by Filipponi & Feil, 2009). Another possible mechanism is through miRNAs. It was long believed that the sperm delivers only the haploid paternal genome to offspring. However, it has been found that along with this, the sperm head actually contains numerous miRNAs that are transmitted to the egg along with the DNA (Ostermeier, Miller, Huntriss, Diamond, & Krawetz, 2004). More intriguing is the fact that expression of these miRNAs is affected by ethanol exposure (Sathyan, Golden, & Miranda, 2007; Wang, et al., 2009). This together with the observation that miRNAs inherited from sperm are sufficient to cause a change in gene expression and consequently phenotype (Rassoulzadegan, et al., 2006), provides support that miRNAs inherited from sperm could be responsible for the epigenetic changes observed in offspring sired from alcohol exposed males.

4.2.3 *H19* ICR DNA methylation and sperm parameters

When looking at the relationship between abnormal sperm parameters and *H19* ICR DNA methylation, not taking alcohol use into account, a significant increase in average DNA methylation at the *H19* ICR in teratozoospermic individuals was observed ($p = 0.007$). Boissonnas et al. (2010) showed that the 6th CTCF binding site in the *H19* ICR showed site specific decreases in DNA methylation in teratozoospermic individuals. This was only noted in approximately 60% of the samples while the rest of the samples displayed normal levels of DNA methylation. The average DNA methylation levels observed in their study did not differ significantly from the normozoospermic group. It must be noted that they did analyze 16 CpGs encompassing the *H19* ICR whereas this study focused on only 5 CpG sites within the CTCF binding site. In addition, our study only had 6 teratozoospermic individuals compared to 83 normozoospermic individuals (*Table 3.9*). This gave less power to detect small effects.

Houshdaran et al. (2007) looked at several genes, including *DIRAS3*, *MEST* and *PLAGL1* throughout the sperm genome and found DNA methylation at these genes to be elevated in abnormal sperm. They suggest a broad epigenetic defect associated with abnormal sperm parameters, which could possibly arise from improper erasure of imprints during spermatogenesis. It has also been found that men with normal sperm counts, can still have abnormal DNA methylation patterns which arise from incorrect chromatin packaging (Carrell & Hammoud, 2010). As can be seen there are many complex mechanisms underlying the possible relationships between abnormal sperm parameters and DNA methylation. More studies need to be done on the interaction between various epigenetic mechanisms and the outcomes on sperm parameters.

4.3 Limitations

Sample size, particularly for global DNA methylation analysis was relatively small, resulting in this study being insufficiently powered to detect modest changes in DNA methylation levels. An experiment would be meaningful if the statistical power is 80% or greater. With a sample size of 58, there would have to be approximately a 32% change in global DNA methylation levels in order for the results to have an 80% power to detect a statistically significant difference. The LUMA method was not the most appropriate method to detect changes in global DNA methylation. LUMA gives and overall percentage of the DNA methylation of a sample and does not look at individual genes which could be up or down regulated within the sample.

The self-report questionnaires used in this study required participant honesty. This could have led to the under-reporting of socially unacceptable behavior, such as alcohol and drug use.

Alcohol use in this study had a wide range, from 1 to 360 drinks per month, with the majority falling well below 100 drinks per month. In order to get a clearer indication of the effects of alcohol on DNA methylation more individuals with heavy drinking patterns should be used.

The dichotomous nature of data obtained for smoking and drug use (yes/no answer) does not give an accurate representation of the range of substance abuse and the possible effects it could have on DNA methylation. More detailed information concerning the type and quantity of smoking and drug use should be collected. It is crucial to control more rigorously for these factors as they have been linked with changes in DNA methylation.

4.4 Future studies

As highlighted by the limitations above, several strategies can be implemented to improve this study. Future studies would benefit from assessing the functional impact of alcohol mediated paternally induced epigenetic changes on offspring phenotypes. This needs to be done in a longitudinal way so that alcohol exposed fathers as well as their offspring can be analyzed. Genome wide loci could be looked at using a next generation sequencing approach, and gene expression data of these loci could be integrated with histone modification data and DNA methylation data resulting in a more holistic picture. The ethical issues associated with such a study may imply that it is not a feasible approach. However, mouse models dealing with these aspects may offer a good starting point.

5. Conclusion

The findings of this study together with the findings of significant selective demethylation at individual CpG sites within the *IG-DMR* in the same sperm samples, suggest that alcohol may have the ability to affect DNA methylation levels in spermatozoa at certain loci within the sperm genome. However, these loci specific effects are not reflected in global DNA methylation levels. These findings do not disprove the hypothesis that there is an epigenetic mechanism responsible for the paternal effects seen in FASD.

The sample size used in this study to analyze global sperm DNA methylation changes in response to alcohol exposure was relatively small. The trend of a decrease in global DNA methylation with alcohol exposure that was noted may prove to be important in the future, should a large scale study demonstrate a significant result. The challenge would then be to see if the global change persists in offspring and what effect it may have on development. It is also important to keep in mind that global DNA methylation levels represent the net effect of increases and decreases at certain loci within the genome. It is thus important to locate these loci and test the effects of such site specific changes. Alternate methylome-wide high throughput techniques could be used in future to assess global methylation levels and at the same time reveal specific loci that have been altered by alcohol use.

Renda et al. (2007) have shown that CpG sites 2 and 3 at the *H19* ICR are critical for CTCF binding. CpG site 3 in our study showed the lowest methylation percentage among all CpG sites. It was also the most amenable to changes by environmental factors and sperm parameters. Contradictory to our hypothesis there was a significant increase in DNA methylation at this CpG site in alcohol exposed spermatozoa. Even though this was the case, the average methylation at this site was still approximately 20% lower than at other CpG sites within the *H19* ICR, in both the control and drinking groups. Could this have an effect on CTCF binding? Although CpG sites critical for CTCF binding have been identified,

the question of how much of a difference in DNA methylation would cause a significant biological change still needs to be addressed.

In conclusion, this study has shown a trend towards decreased global DNA methylation in alcohol-exposed spermatozoa. Alcohol did not have an effect on average DNA methylation at the *H19* ICR but did have a site-specific significant increase. This study suggests that we cannot implicate paternally induced DNA methylation changes in the epigenetic deregulation of the *H19/IGF2* locus in contributing to the FAS phenotype of growth retardation. It is possible, however, that other epigenetic mechanisms may be altered at this locus and this is what contributes to the phenotype. The *H19/IGF2* locus was an attractive candidate to look at in this study for several reasons. Firstly, it is paternally imprinted and is theoretically 100% methylated in the male germ line. Secondly, the *H19* ICR is situated in the centre of the *H19* and *IGF2* gene region and controls their reciprocal expression. Thirdly, *IGF2* is a significant contributor to fetal growth, and one of the key features of FAS is growth retardation. Therefore, if an alcohol induced decrease in *H19* ICR DNA methylation in the paternal germline occurs, CTCF binding would be prevented and result in biallelic *IGF2* repression in the fetus resulting in a growth retardation phenotype.

Although this study could not shed light on the possible paternal mechanisms underlying FASD, future studies using different techniques, looking at more loci and integrating other epigenetic mechanisms may be able to do so.

6. Reference list

6.1 Internet references

Chan, N. (2006). *New findings on sperm life cycle could impact fertility treatments*. Retrieved November 23, 2011, from http://www.ucsf.edu/news/2006/03/6621/new-findings-sperm-life-cycle-could-impact-fertility-treatments

National Institute on Alcohol Abuse and Alcoholism. (2006). *About Fetal Alcohol Syndrome and Fetal Alcohol Spectrum Disorders.* Retrieves March 27, 2008, from http://www.niaaa.nih.gov/

http://come-over.to/FAS/JohnGrowsUp.htm [Accessed August 2008]

6.2 Software references

PSQ Assay Design Software version 1.0.6 (Biotage AB, Kungsgatan, Uppsala, Sweden)

PyroMark Q96MA application software (Biotage AB, Kungsgatan, Uppsala, Sweden)

Pyro Q-CpG version 1.0.9 (Biotage AB, Kungsgatan, Uppsala, Sweden)

Statistica v19.1 (StatSoft Inc. Tulsa, UK)

6.3 Literary references

Abel, E. L. (1995). An update on incidence of FAS: FAS is not an equal opportunity birth defect. *Neurotoxicology and Teratology, 17*(4), 437-443.

Abel, E. L. (1997). Was the fetal alcohol syndrome recognized in the ancient Near East? *Alcohol Alcohol, 32*(1), 3-7.

Abel, E. L. (1998). Fetal alcohol syndrome: the 'American Paradox'. *Alcohol Alcohol, 33*(3), 195-201.

Abel, E. L., & Bilitzke, P. (1990). Paternal alcohol exposure: paradoxical effect in mice and rats. *Psychopharmacology (Berl), 100*(2), 159-164.

Abel, E. L., & Moore, C. (1987). Effects of paternal alcohol consumption in mice. *Alcohol Clinical and Experimental Research, 11*(6), 533-535.

Ahmadian, A., Ehn, M., & Hober, S. (2006). Pyrosequencing: history, biochemistry and future. *Clinica Chimica Acta, 363*(1-2), 83-94.

Ahuja, N., & Issa, J. P. (2000). Aging, methylation and cancer. *Histology and Histopathology, 15*(3), 835-842.

Anway, M. D., Cupp, A. S., Uzumcu, M., & Skinner, M. K. (2005). Epigenetic transgenerational actions of endocrine disruptors and male fertility. *Science, 308*(5727), 1466-1469.

Baccarelli, A., & Bollati, V. (2009). Epigenetics and environmental chemicals. *Current Opinion in Pediatrics, 21*(2), 243-251.

Baccarelli, A., Wright, R. O., Bollati, V., Tarantini, L., Litonjua, A. A., Suh, H. H., et al. (2009). Rapid DNA methylation changes after exposure to traffic particles. *American Journal of Respiratory and Critical Care Medicine, 179*(7), 572-578.

Bartolomei, M. S., & Tilghman, S. M. (1997). Genomic imprinting in mammals. *Annual Review of Genetics, 31*(1), 493-525.

Beck, S., & Rakyan, V. K. (2008). The methylome: approaches for global DNA methylation profiling. *Trends in Genetics, 24*(5), 231-237.

Becker, H. C., Diaz-Granados, J. L., & Randall, C. L. (1996). Teratogenic actions of ethanol in the mouse: a minireview. *Pharmacology Biochemistry and Behavior, 55*(4), 501-513.

Bell, A. C., & Felsenfeld, G. (2000). Methylation of a CTCF-dependent boundary controls imprinted expression of the Igf2 gene. *Nature, 405*(6785), 482-485.

Benchaib, M., Braun, V., Ressnikof, D., Lornage, J., Durand, P., Niveleau, A., et al. (2005). Influence of global sperm DNA methylation on IVF results. *Human Reproduction, 20*(3), 768-773.

Bielawski, D. M., Zaher, F. M., Svinarich, D. M., & Abel, E. L. (2002). Paternal alcohol exposure affects sperm cytosine methyltransferase messenger RNA levels. *Alcohol Clinical and Experimental Research, 26*(3), 347-351.

Boissonnas, C. C., Abdalaoui, H. E., Haelewyn, V., Fauque, P., Dupont, J. M., Gut, I., et al. (2010). Specific epigenetic alterations of IGF2-H19 locus in spermatozoa from infertile men. *European Journal of Human Genetics, 18*(1), 73-80.

Bollati, V., Baccarelli, A., Hou, L., Bonzini, M., Fustinoni, S., Cavallo, D., et al. (2007). Changes in DNA methylation patterns in subjects exposed to low-dose benzene. *Cancer Research, 67*(3), 876-880.

Burruel, V. R., Yanagimachi, R., & Whitten, W. K. (1996). Normal mice develop from oocytes injected with spermatozoa with grossly misshapen heads. *Biology of Reproduction, 55*(3), 709-714.

Cai, X., & Cullen, B. R. (2007). The imprinted H19 noncoding RNA is a primary microRNA precursor. *RNA, 13*, 313-316.

Carrell, D. T., & Hammoud, S. S. (2010). The human sperm epigenome and its potential role in embryonic development. *Molecular Human Reproduction, 16*(1), 37-47.

Cedar, H., & Bergman, Y. (2009). Linking DNA methylation and histone modification: patterns and paradigms. *Nature Reviews Genetics, 10*(5), 295-304.

Choufani, S., Shapiro, J. S., Susiarjo, M., Butcher, D. T., Grafodatskaya, D., Lou, Y., et al. (2011). A novel approach identifies new differentially methylated regions (DMRs) associated with imprinted genes. *Genome Research, 21*(3), 465-476.

Choufani, S., Shuman, C., & Weksberg, R. (2010). Beckwith-Wiedemann syndrome. *American Journal of Medical Genetics Part C Seminars in Medical Genetics, 154C*(3), 343-354.

Constancia, M., Kelsey, G., & Reik, W. (2004). Resourceful imprinting. *Nature, 432*(7013), 53-57.

Cooper, T. G., Noonan, E., von Eckardstein, S., Auger, J., Baker, H. W., Behre, H. M., et al. (2010). World Health Organization reference values for human semen characteristics. *Hum Reprod Update, 16*(3), 231-245.

Cudd, T. A. (2005). Animal model systems for the study of alcohol teratology. *Experimental Biology and Medicine (Maywood), 230*(6), 389-393.

Davis, T. L., Yang, G. J., McCarrey, J. R., & Bartolomei, M. S. (2000). The H19 methylation imprint is erased and re-established differentially on the parental alleles during male germ cell development. *Human Molecular Genetics, 9*(19), 2885-2894.

DeChiara, T. M., Robertson, E. J., & Efstratiadis, A. (1991). Parental imprinting of the mouse insulin-like growth factor II gene. *Cell, 64*(4), 849-859.

Donnelly, G. P., McClure, N., Kennedy, M. S., & Lewis, S. E. (1999). Direct effect of alcohol on the motility and morphology of human spermatozoa. *Andrologia, 31*(1), 43-47.

Egger, G., Liang, G., Aparicio, A., & Jones, P. A. (2004). Epigenetics in human disease and prospects for epigenetic therapy. *Nature, 429*(6990), 457-463.

Emery, B. R., & Carrell, D. T. (2006). The effect of epigenetic sperm abnormalities on early embryogenesis. *Asian Journal of Andrology, 8*(2), 131-142.

England, R., & Pettersson, M. (2005). PyroQ-CpG: quantitatve analysis of methylation in multiple CpG sites by pyrosequencing *Nature Methods*.

Estecio, M. R., & Issa, J. P. (2009). Tackling the methylome: recent methodological advances in genome-wide methylation profiling. *Genome Medicine, 1*(11), 106.

Esteller, M. (2006). Epigenetics provides a new generation of oncogenes and tumour-suppressor genes. *British Journal of Cancer, 94*(2), 179-183.

Filipponi, D., & Feil, R. (2009). Perturbation of genomic imprinting in oligozoospermia. *Epigenetics, 4*(1), 27-30.

Fouse, S. D., Nagarajan, R. P., & Costello, J. F. (2010). Genome-scale DNA methylation analysis. *Epigenomics, 2*(1), 105-117.

Fuke, C., Shimabukuro, M., Petronis, A., Sugimoto, J., Oda, T., Miura, K., et al. (2004). Age related changes in 5-methylcytosine content in human peripheral leukocytes and placentas: an HPLC-based study. *Annals of Human Genetics, 68*(Pt 3), 196-204.

Gabory, A., Ripoche, M. A., Yoshimizu, T., & Dandolo, L. (2006). The H19 gene: regulation and function of a non-coding RNA. *Cytogenetic and Genome Research, 113*(1-4), 188-193.

Garro, A. J., McBeth, D. L., Lima, V., & Lieber, C. S. (1991). Ethanol consumption inhibits fetal DNA methylation in mice: implications for the fetal alcohol syndrome. *Alcohol Clinical and Experimental Research, 15*(3), 395-398.

Geuns, E., De Rycke, M., Van Steirteghem, A., & Liebaers, I. (2003). Methylation imprints of the imprint control region of the SNRPN-gene in human gametes and preimplantation embryos. *Human Molecular Genetics, 12*(22), 2873-2879.

Goll, M. G., Kirpekar, F., Maggert, K. A., Yoder, J. A., Hsieh, C.-L., Zhang, X., et al. (2006). Methylation of tRNA Asp by the DNA Methyltransferase Homolog Dnmt2. *Science, 311*, 395-398.

Hark, A. T., Schoenherr, C. J., Katz, D. J., Ingram, R. S., Levorse, J. M., & Tilghman, S. M. (2000). CTCF mediates methylation-sensitive enhancer-blocking activity at the H19/Igf2 locus. *Nature, 405*(6785), 486-489.

Haycock, P. C. (2009). Fetal alcohol spectrum disorders: the epigenetic perspective. *Biology of Reproduction, 81*(4), 607-617.

Haycock, P. C., & Ramsay, M. (2009). Exposure of mouse embryos to ethanol during preimplantation development: effect on DNA methylation in the h19 imprinting control region. *Biology of Reproduction, 81*(4), 618-627.

Heijmans, B. T., Kremer, D., Tobi, E. W., Boomsma, D. I., & Slagboom, P. E. (2007). Heritable rather than age-related environmental and stochastic factors dominate variation in DNA methylation of the human IGF2/H19 locus. *Human Molecular Genetics, 16*(5), 547-554.

Henke, W., Herdel, K., Jung, K., Schnorr, D., & Loening, S. A. (1997). Betaine improves the PCR amplification of GC-rich DNA sequences. *Nucleic Acids Research, 25*(19), 3957-3958.

Holliday, R. (2006). Epigenetics: a historical overview. *Epigenetics, 1*(2), 76-80.

Houshdaran, S., Cortessis, V. K., Siegmund, K., Yang, A., Laird, P. W., & Sokol, R. Z. (2007). Widespread epigenetic abnormalities suggest a broad DNA methylation erasure defect in abnormal human sperm. *PLoS One, 2*(12), e1289.

Hsiung, D. T., Marsit, C. J., Houseman, E. A., Eddy, K., Furniss, C. S., McClean, M. D., et al. (2007). Global DNA methylation level in whole blood as a biomarker in head and neck squamous cell carcinoma. *Cancer Epidemiology, Biomarkers & Prevention, 16*(1), 108-114.

Ibala-Romdhane, S., Al-Khtib, M., Khoueiry, R., Blachere, T., Guerin, J. F., & Lefevre, A. (2011). Analysis of H19 methylation in control and abnormal human embryos, sperm and oocytes. *European Journal of Human Genetics, 19*(11), 1138-1143.

Jacinto, F. V., Ballestar, E., & Esteller, M. (2008). Methyl-DNA immunoprecipitation (MeDIP): hunting down the DNA methylome. *Biotechniques, 44*(1), 35, 37, 39 passim.

Jamerson, P. A., Wulser, M. J., & Kimler, B. F. (2004). Neurobehavioral effects in rat pups whose sires were exposed to alcohol. *Brain Res Dev Brain Res, 149*(2), 103-111.

Jeltsch, A. (2002). Beyond Watson and Crick: DNA methylation and molecular enzymology of DNA methyltransferases. *ChemBioChem, 3*(4), 274-293.

Jiang, Y. H., Bressler, J., & Beaudet, A. L. (2004). Epigenetics and human disease. *Annual Review of Genomics and Human Genetics, 5*, 479-510.

Jirtle, R. L., & Skinner, M. K. (2007). Environmental epigenomics and disease susceptibility. *Nature Reviews Genetics, 8*(4), 253-262.

Jones, K. L., Smith, D. W., Ulleland, C. N., & Streissguth, P. (1973). Pattern of malformation in offspring of chronic alcoholic mothers. *Lancet, 1*(7815), 1267-1271.

Jones, P. A., & Takai, D. (2001). The role of DNA methylation in mammalian epigenetics. *Science, 293*(5532), 1068-1070.

Kaminen-Ahola, N., Ahola, A., Maga, M., Mallitt, K. A., Fahey, P., Cox, T. C., et al. (2010). Maternal ethanol consumption alters the epigenotype and the phenotype of offspring in a mouse model. *PLoS Genetics, 6*(1), e1000811.

Karimi Arzenani, M. (2009). *Genomic DNA methylation in health and disease.* Karolinska Institutet, Solna.

Karimi, M., Johansson, S., Stach, D., Corcoran, M., Grander, D., Schalling, M., et al. (2006). LUMA (LUminometric Methylation Assay)--a high throughput method to the analysis of genomic DNA methylation. *Experimental Cell Research, 312*(11), 1989-1995.

Knezovich, J. G., & Ramsay, M. (2012). The Effect of Preconception Paternal Alcohol Exposure on Epigenetic Remodeling of the H19 and Rasgrf1 Imprinting Control Regions in Mouse Offspring. *Frontiers in Genetics, 2*(10), 1-10.

Kobayashi, H., Hiura, H., John, R. M., Sato, A., Otsu, E., Kobayashi, N., et al. (2009). DNA methylation errors at imprinted loci after assisted conception originate in the parental sperm. *European Journal of Human Genetics, 17*(12), 1582-1591.

Kobayashi, H., Sato, A., Otsu, E., Hiura, H., Tomatsu, C., Utsunomiya, T., et al. (2007). Aberrant DNA methylation of imprinted loci in sperm from oligospermic patients. *Human Molecular Genetics, 16*(21), 2542-2551.

Kobayashi, H., Suda, C., Abe, T., Kohara, Y., Ikemura, T., & Sasaki, H. (2006). Bisulfite sequencing and dinucleotide content analysis of 15 imprinted mouse differentially methylated regions (DMRs): paternally methylated DMRs contain less CpGs than maternally methylated DMRs. *Cytogenetic and Genome Research, 113*(1-4), 130-137.

Kornberg, R. D. (1977). Structure of chromatin. *Annual Review of Biochemistry, 46*, 931-954.

Kotch, L. E., & Sulik, K. K. (1992). Experimental fetal alcohol syndrome: proposed pathogenic basis for a variety of associated facial and brain anomalies. *American Journal of Medical Genetics, 44*(2), 168-176.

Kurukuti, S., Tiwari, V. K., Tavoosidana, G., Pugacheva, E., Murrell, A., Zhao, Z., et al. (2006). CTCF binding at the H19 imprinting control region mediates maternally inherited higher-order chromatin conformation to restrict enhancer access to Igf2. *Proceedings of the National Academy of Sciences USA, 103*(28), 10684-10689.

Leibovitch, M. P., Nguyen, V. C., Gross, M. S., Solhonne, B., Leibovitch, S. A., & Bernheim, A. (1991). The human ASM (adult skeletal muscle) gene: expression and chromosomal assignment to 11p15. *Biochemical and Biophysical Research Communications, 180*(3), 1241-1250.

Leonhardt, H., Page, A. W., Weier, H. U., & Bestor, T. H. (1992). A targeting sequence directs DNA methyltransferase to sites of DNA replication in mammalian nuclei. *Cell, 71*(5), 865-873.

Li, E., Beard, C., & Jaenisch, R. (1993). Role for DNA methylation in genomic imprinting. *Nature, 366*(6453), 362-365.

Li, E., Bestor, T. H., & Jaenisch, R. (1992). Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell, 69*(6), 915-926.

Liu, L., Li, Y., & Tollefsbol, T. O. (2008). Gene-environment interactions and epigenetic basis of human diseases. *Current Issues in Molecular Biology, 10*(1-2), 25-36.

Liu, Y., Balaraman, Y., Wang, G., Nephew, K. P., & Zhou, F. C. (2009). Alcohol exposure alters DNA methylation profiles in mouse embryos at early neurulation. *Epigenetics, 4*(7), 500-511.

Livy, D. J., Maier, S. E., & West, J. R. (2004). Long-term alcohol exposure prior to conception results in lower fetal body weights. *Birth Defects Research Part B: Developmental and Reproductive Toxicology, 71*(3), 135-141.

Lopes, S., Lewis, A., Hajkova, P., Dean, W., Oswald, J., Forne, T., et al. (2003). Epigenetic modifications in an imprinting cluster are controlled by a hierarchy of DMRs suggesting long-range chromatin interactions. *Human Molecular Genetics, 12*(3), 295-305.

Lucifero, D., Chaillet, J. R., & Trasler, J. M. (2004). Potential significance of genomic imprinting defects for reproduction and assisted reproductive technology. *Human Reproduction Update, 10*(1), 3-18.

Marques, C. J., Costa, P., Vaz, B., Carvalho, F., Fernandes, S., Barros, A., et al. (2008). Abnormal methylation of imprinted genes in human sperm is associated with oligozoospermia. *Molecular Human Reproduction, 14*(2), 67-74.

Mascord, D., Smith, J., Starmer, G. A., & Whitfield, J. B. (1992). Effects of increasing the rate of alcohol metabolism on plasma acetate concentration. *Alcohol Alcohol, 27*(1), 25-28.

May, P. A., Brooke, L., Gossage, J. P., Croxford, J., Adnams, C., Jones, K. L., et al. (2000). Epidemiology of fetal alcohol syndrome in a South African community in the Western Cape Province. *American Journal of Public Health, 90*(12), 1905-1912.

May, P. A., Gossage, J. P., Marais, A. S., Adnams, C. M., Hoyme, H. E., Jones, K. L., et al. (2007). The epidemiology of fetal alcohol syndrome and partial FAS in a South African community. *Drug and Alcohol Dependence, 88*(2-3), 259-271.

Melamede, R. J., & Wallace, S. S. (1985). A possible secondary role for thymine-containing DNA precursors. *Basic Life Sciences, 31*, 67-102.

Menegola, E., Broccia, M. L., Di Renzo, F., & Giavini, E. (2001). Acetaldehyde in vitro exposure and apoptosis: a possible mechanism of teratogenesis. *Alcohol, 23*(1), 35-39.

Miozzo, M., & Simoni, G. (2002). The role of imprinted genes in fetal growth. *Biology of the Neonate, 81*(4), 217-228.

Monk, M., Boubelik, M., & Lehnert, S. (1987). Temporal and regional changes in DNA methylation in the embryonic, extraembryonic and germ cell lineages during mouse embryo development. *Development, 99*(3), 371-382.

Moore, L. E., Pfeiffer, R. M., Poscablo, C., Real, F. X., Kogevinas, M., Silverman, D., et al. (2008). Genomic DNA hypomethylation as a biomarker for bladder cancer susceptibility in the Spanish Bladder Cancer Study: a case-control study. *The Lancet Oncology, 9*(4), 359-366.

Morison, I. M., Ramsay, J. P., & Spencer, H. G. (2005). A census of mammalian imprinting. *Trends in Genetics, 21*(8), 457-465.

Nayak, R. B., & Murthy, P. (2008). Fetal alcohol spectrum disorder. *Indian Pediatrics, 45*(12), 977-983.

Ohlsson, R., Renkawitz, R., & Lobanenkov, V. (2001). CTCF is a uniquely versatile transcription regulator linked to epigenetics and disease. *Trends in Genetics, 17*(9), 520-527.

Okano, M., Bell, D. W., Haber, D. A., & Li, E. (1999). DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell, 99*(3), 247-257.

Ooi, S. L., & Henikoff, S. (2007). Germline histone dynamics and epigenetics. *Current Opinion in Cell Biology, 19*(3), 257-265.

Ostermeier, G. C., Miller, D., Huntriss, J. D., Diamond, M. P., & Krawetz, S. A. (2004). Reproductive biology: delivering spermatozoan RNA to the oocyte. *Nature, 429*(6988), 154.

Ouko, L. A., Shantikumar, K., Knezovich, J., Haycock, P., Schnugh, D. J., & Ramsay, M. (2009). Effect of alcohol consumption on CpG methylation in the differentially methylated regions of H19 and IG-DMR in male gametes: implications for fetal alcohol spectrum disorders. *Alcohol Clinical and Experimental Research, 33*(9), 1615-1627.

Pacheco, S. E., Houseman, E. A., Christensen, B. C., Marsit, C. J., Kelsey, K. T., Sigman, M., et al. (2011). Integrative DNA methylation and gene expression analyses identify DNA packaging and epigenetic regulatory genes associated with low motility sperm. *PLoS One, 6*(6), e20280.

Pant, V., Kurukuti, S., Pugacheva, E., Shamsuddin, S., Mariano, P., Renkawitz, R., et al. (2004). Mutation of a single CTCF target site within the H19 imprinting control region leads to loss of Igf2 imprinting and complex patterns of de novo methylation upon maternal inheritance. *Molecular and Cellular Biology, 24*(8), 3497-3504.

Park, K. Y., Sellars, E. A., Grinberg, A., Huang, S. P., & Pfeifer, K. (2004). The H19 differentially methylated region marks the parental origin of a heterologous locus without gametic DNA methylation. *Molecular and Cellular Biology, 24*(9), 3588-3595.

Peterson, C. L., & Laniel, M. A. (2004). Histones and histone modifications. *Current Biology, 14*(14), R546-551.

Rassoulzadegan, M., Grandjean, V., Gounon, P., Vincent, S., Gillot, I., & Cuzin, F. (2006). RNA-mediated non-mendelian inheritance of an epigenetic change in the mouse. *Nature, 441*(7092), 469-474.

Razin, A., Webb, C., Szyf, M., Yisraeli, J., Rosenthal, A., Naveh-Many, T., et al. (1984). Variations in DNA methylation during mouse cell differentiation in vivo and in vitro. *Proceedings of the National Academy of Sciences USA, 81*(8), 2275-2279.

Reik, W., Dean, W., & Walter, J. (2001). Epigenetic reprogramming in mammalian development. *Science, 293*(5532), 1089-1093.

Renda, M., Baglivo, I., Burgess-Beusse, B., Esposito, S., Fattorusso, R., Felsenfeld, G., et al. (2007). Critical DNA binding interactions of the insulator protein CTCF: a small number of zinc fingers mediate strong binding, and a single finger-DNA interaction controls binding at imprinted loci. *Journal of Biological Chemistry, 282*(46), 33336-33345.

Rodenhiser, D., & Mann, M. (2006). Epigenetics and human disease: translating basic biology into clinical applications. *Canadian Medical Association Journal, 174*(3), 341-348.

Ronaghi, M. (2001). Pyrosequencing sheds light on DNA sequencing. *Genome Research, 11*(1), 3-11.

Ronaghi, M., Karamohamed, S., Pettersson, B., Uhlen, M., & Nyren, P. (1996). Real-time DNA sequencing using detection of pyrophosphate release. *Analytical Biochemistry, 242*(1), 84-89.

Sanger, F., Nicklen, S., & Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences USA, 74*(12), 5463-5467.

Santos, F., Hendrich, B., Reik, W., & Dean, W. (2002). Dynamic reprogramming of DNA methylation in the early mouse embryo. *Developmental Biology, 241*(1), 172-182.

Sathyan, P., Golden, H. B., & Miranda, R. C. (2007). Competing Interactions between Micro-RNAs Determine Neural Progenitor Survival and Proliferation after Ethanol Exposure: Evidence from an Ex Vivo Model of the Fetal Cerebral Cortical Neuroepithelium. *The Journal of Neuroscience, 27*(32), 8546-8557.

Seuanez, H. N., Carothers, A. D., Martin, D. E., & Short, R. V. (1977). Morphological abnormalities in spermatozoa of man and great apes. *Nature, 270*(5635), 345-347.

Shen, L., & Waterland, R. A. (2007). Methods of DNA methylation analysis. *Current Opinion in Clinical Nutrition & Metabolic Care, 10*(5), 576-581.

Shukla, S. D., Velazquez, J., French, S. W., Lu, S. C., Ticku, M. K., & Zakhari, S. (2008). Emerging role of epigenetics in the actions of alcohol. *Alcohol Clinical and Experimental Research, 32*(9), 1525-1534.

Simpson, M. E., Duggal, S., & Keiver, K. (2005). Prenatal ethanol exposure has differential effects on fetal growth and skeletal ossification. *Bone, 36*(3), 521-532.

Sims, R. J., 3rd, Nishioka, K., & Reinberg, D. (2003). Histone lysine methylation: a signature for chromatin function. *Trends in Genetics, 19*(11), 629-639.

Slack, J. M. (2002). Conrad Hal Waddington: the last Renaissance biologist? *Nature Reviews Genetics, 3*(11), 889-895.

Smallwood, S. A., Tomizawa, S., Krueger, F., Ruf, N., Carli, N., Segonds-Pichon, A., et al. (2011). Dynamic CpG island methylation landscape in oocytes and preimplantation embryos. *Nature Genetics, 43*(8), 811-814.

Sokol, R. J. (1981). Alcohol and abnormal outcomes of pregnancy. *Canadian Medical Association Journal, 125*(2), 143-148.

Sokol, R. J., Delaney-Black, V., & Nordstrom, B. (2003). Fetal alcohol spectrum disorder. *Journal of the American Medical Association, 290*(22), 2996-2999.

Solter, D. (2006). Imprinting today: end of the beginning or beginning of the end? *Cytogenetic and Genome Research, 113*(1-4), 12-16.

Streissguth, A. P., & Dehaene, P. (1993). Fetal alcohol syndrome in twins of alcoholic mothers: concordance of diagnosis and IQ. *American Journal of Medical Genetics, 47*(6), 857-861.

Streissguth, A. P., Landesman-Dwyer, S., Martin, J. C., & Smith, D. W. (1980). Teratogenic effects of alcohol in humans and laboratory animals. *Science, 209*(4454), 353-361.

Tost, J., Jammes, H., Dupont, J.-M., Buffat, C., Robert, B., Mignot, T.-M., et al. (2007). Retraction: Non-random, individual-specific methylation profiles are present at the sixth CTCF binding site in the human H19/IGF2 imprinting control region. *Nucleic Acids Research, 35*(2), 701.

Tunc, O., & Tremellen, K. (2009). Oxidative DNA damage impairs global sperm DNA methylation in infertile men. *Journal of Assisted Reproduction and Genetics, 26*(9-10), 537-544.

Ulaner, G. A., Yang, Y., Hu, J. F., Li, T., Vu, T. H., & Hoffman, A. R. (2003). CTCF binding at the insulin-like growth factor-II (IGF2)/H19 imprinting control region is insufficient to regulate IGF2/H19 expression in human tissues. *Endocrinology, 144*(10), 4420-4426.

Viljoen, D. L., Gossage, J. P., Brooke, L., Adnams, C. M., Jones, K. L., Robinson, L. K., et al. (2005). Fetal alcohol syndrome epidemiology in a South African community: a second study of a very high prevalence area. *Journal of Studies on Alcohol and Drugs, 66*(5), 593-604.

Wang, L. L., Zhang, Z., Li, Q., Yang, R., Pei, X., Xu, Y., et al. (2009). Ethanol exposure induces differential microRNA and target gene expression and teratogenic effects which can be suppressed by folic acid supplementation. *Human Reproduction, 24*(3), 562-579.

Warren, K. R., & Li, T. K. (2005). Genetic polymorphisms: impact on the risk of fetal alcohol spectrum disorders. *Birth Defects Research. Part A, Clinical and Molecular Teratology, 73*(4), 195-203.

Weisstein, A. E., Feldman, M. W., & Spencer, H. G. (2002). Evolutionary genetic models of the ovarian time bomb hypothesis for the evolution of genomic imprinting. *Genetics, 162*(1), 425-439.

White, H. E., Durston, V. J., Harvey, J. F., & Cross, N. C. (2006). Quantitative analysis of SNRPN(correction of SRNPN) gene methylation by pyrosequencing as a diagnostic test for Prader-Willi syndrome and Angelman syndrome. *Clinical Chemistry online, 52*(6), 1005-1013.

Yauk, C., Polyzos, A., Rowan-Carroll, A., Somers, C. M., Godschalk, R. W., Van Schooten, F. J., et al. (2008). Germ-line mutations, DNA damage, and global hypermethylation in mice exposed to particulate air pollution in an urban/industrial location. *Proceedings of the National Academy of Sciences USA, 105*(2), 605-610.

Zhou, F. C., Zhao, Q., Liu, Y., Goodlett, C. R., Liang, T., McClintick, J. N., et al. (2011). Alteration of gene expression by alcohol exposure at early neurulation. *BMC Genomics, 12*, 124.

Zhu, Z. Z., Hou, L., Bollati, V., Tarantini, L., Marinelli, B., Cantone, L., et al. (2010). Predictors of global methylation levels in blood DNA of healthy subjects: a combined analysis. *International Journal of Epidemiology*.

Appendices

Appendix A: Experimental protocols

QIAGEN Supplementary Protocol:

Purification of DNA from epithelial cells mixed with sperm cells using the QIAamp® DNA Micro Kit

This protocol is designed for purification of total (genomic and mitochondrial) DNA from fabrics or swabs containing epithelial cells mixed with sperm cells using the QIAamp DNA Micro Kit.

IMPORTANT: Please read the *QIAamp DNA Micro Handbook*, paying careful attention to the "Safety Information" and "Important Notes" sections, before beginning this procedure. For safety information on the additional chemicals mentioned in this protocol, please consult the appropriate material safety data sheets (MSDSs), available from the product supplier.

Equipment and reagents to be supplied by the user

When working with chemicals, always wear a suitable lab coat, disposable gloves, and protective goggles. For more information, consult the appropriate material safety data sheets (MSDSs), available from the product supplier.

- Ethanol (96–100%)
- 1.5 ml or 2 ml microcentrifuge tubes (for lysis steps)
- 1.5 ml microcentrifuge tubes (for elution steps) (available from Brinkmann [Safe-Lock, cat. no. 022363204], Eppendorf [Safe-Lock, cat. no. 0030 120.086], or Sarstedt [Safety Cap, cat. no. 72.690])*
- Pipet tips (to avoid cross contamination, we recommend pipet tips with aerosol barriers)
- Thermomixer, heated orbital incubator, heating block, or water bath
- Microcentrifuge with rotor for 2 ml tubes
- Vortexer
- Scissors
- 1 M dithiothreitol (DTT). This solution can be made in advance, and stored at –20°C in appropriate aliquots. A thawed aliquot should be discarded after use.
- **Optional**: QIAshredder spin columns (cat. no. 79654 or 79656; to harvest lysate remaining in the swab or fabric)

Important points before starting

- Perform all centrifugation steps at room temperature (15–25°C).
- Check whether carrier RNA is required (see the *QIAamp DNA Micro Handbook*).

* This is not a complete list of suppliers and does not include many important vendors of biological supplies.

Things to do before starting

- Equilibrate Buffer AE or distilled water for elution to room temperature (15–25°C).
- Set a thermomixer or heated orbital incubator to 56°C for use in steps 3 and 10, and a second thermomixer or heated orbital incubator to 70°C for use in step 13. If thermomixers or heated orbital incubators are not available, heating blocks or water baths can be used instead.
- If Buffer AL or Buffer ATL contains precipitates, dissolve by heating to 70°C with gentle agitation.
- Ensure that Buffers AW1 and AW2 have been prepared according to the instructions in the *QIAamp DNA Micro Handbook*.

Procedure

1. **Place the swab or a piece of fabrics (≤0.5cm²) in a 2 ml microcentrifuge tube (not provided).**

   Separate the cotton or DACRON® swab from its shaft by hand or using scissors.

2. **Add 20 µl proteinase K and 500µl Buffer ATL to the sample. Close the cap and mix by pulse-vortexing for 10 s.**
3. **Place the 2 ml tube in a thermomixer or heated orbital incubator, and incubate at 56°C with shaking at 900 rpm for at least 1 h.**

   If using a heating block or water bath, vortex the tube for 10 s every 10 min to improve lysis.

4. **Briefly centrifuge the 2 ml tube to remove drops from the inside of the lid.**
5. **Remove the solid material from the tube.**

   Note: Up to 200 µl of lysate remains in the swab or fabric. To harvest this remaining lysate, place the swab or fabric in a QIAshredder spin column (not supplied), place the QIAshredder spin column containing the solid material in the 2 ml tube containing the lysate, and centrifuge at full speed (20,000 x g; 14,000 rpm) for 2 min. Remove and discard the QIAshredder spin column containing the solid material.

6. **Centrifuge the tube for 5 min at full speed. Carefully transfer all but 30 µl of the supernatant to a new tube without disturbing the pellet.**

   Note: For isolation of DNA from epithelial cells, transfer 300 µl of the supernatant into a 2 ml microcentrifuge tube and continue with step 12.

7. **Resuspend the pellet in 500 µl Buffer ATL. Close the lid and mix by pulse-vortexing for 10 s. Centrifuge the tube for 5 min at full speed. Carefully aspirate and discard all but 30 µl of the supernatant without disturbing the pellet.**
8. **Repeat step 7 at least three times.**

   Note: The ratio of epithelial cells to sperm cells influences the number of repeats needed for purification of sperm nuclei.

9. **Add 300 µl Buffer ATL, 10 µl proteinase K, and 10 µl 1 M DTT to the pellet. Close the lid and mix by pulse-vortexing for 10 s.**

10. **Place the 2 ml tube in a thermomixer or heated orbital incubator, and incubate at 56°C with shaking at 900 rpm for at least 1 hour.**

    If using a thermoblock or waterbath, vortex the tube for 10 s every 10 min to improve lysis.

11. **Briefly centrifuge the tube to remove drops from the inside of the lid.**

12. **Add 300 µl Buffer AL, close the lid, and mix by pulse-vortexing for 10 s.**

    To ensure efficient lysis, it is essential that the sample and Buffer AL are thoroughly mixed to yield a homogeneous solution.

    A white precipitate may form when Buffer AL is added to Buffer ATL. The precipitate does not interfere with the QIAamp procedure and will dissolve during incubation in step 13.

    Note: If carrier RNA is required (see the *QIAamp DNA Micro Handbook*), add 1 µg dissolved carrier RNA to 300 µl Buffer AL. Note that carrier RNA does not dissolve in Buffer AL. It must first be dissolved in Buffer AE and then added to Buffer AL.

13. **Place the tube in the thermomixer or heated orbital incubator, and incubate at 70°C with shaking at 900 rpm for 10 min.**

    If using a heating block or water bath, vortex the tube for 10 s every 3 min to improve lysis.

14. **Centrifuge the tube at full speed (20,000 x *g*; 14,000 rpm) for 1 min.**

15. **Carefully transfer the supernatant from step 14 to the QIAamp MinElute® column without wetting the rim.**

16. **Close the lid, and centrifuge at 6000 x *g* (8000 rpm) for 1 min. Place the QIAamp MinElute column in a clean 2 ml collection tube, and discard the collection tube containing the flow-through.**

    If the lysate has not completely passed through the membrane after centrifugation, centrifuge again at a higher speed until the QIAamp MinElute Column is empty.

17. **Carefully open the QIAamp MinElute column and add 500 µl Buffer AW1 without wetting the rim. Close the lid and centrifuge at 6000 x *g* (8000 rpm) for 1 min. Place the QIAamp MinElute column in a clean 2 ml collection tube, and discard the collection tube containing the flow-through.**

18. **Carefully open the QIAamp MinElute column and add 500 µl Buffer AW2 without wetting the rim. Close the lid and centrifuge at 6000 x *g* (8000 rpm) for 1 min. Place the QIAamp MinElute column in a clean 2 ml collection tube, and discard the collection tube containing the flow-through.**

    Contact between the QIAamp MinElute column and the flow-through should be avoided. Some centrifuge rotors may vibrate upon deceleration, resulting in the flow-through, which contains ethanol, coming into contact with the QIAamp MinElute column. Take care when removing the QIAamp MinElute column and collection tube from the rotor, so that flow-through does not come into contact with the QIAamp MinElute column.

19. **Centrifuge at full speed (20,000 x *g*; 14,000 rpm) for 3 min to dry the membrane completely.**

    This step is necessary, since ethanol carryover into the eluate may interfere with some downstream applications.

20. **Place the QIAamp MinElute column in a clean 1.5 ml microcentrifuge tube (not provided) and discard the collection tube containing the flow-through. Carefully open the lid of the QIAamp MinElute column and apply 20–50 µl Buffer AE or distilled water to the center of the membrane.**

    If high pH or EDTA affects sensitive downstream applications, use water for elution (see the *QIAamp DNA Micro Handbook*).

    Important: Ensure that Buffer AE or distilled water is equilibrated to room temperature (15–25°C). Dispense Buffer AE or distilled water onto the center of the membrane to ensure complete elution of bound DNA.

    QIAamp MinElute columns provide flexibility in the choice of elution volume. Choose a volume according to the requirements of the downstream application. Remember that the volume of eluate will be up to 5 µl less than the volume of elution solution applied to the column.

21. **Close the lid and incubate at room temperature (15–25°C) for 1 min. Centrifuge at full speed (20,000 x *g*; 14,000 rpm) for 1 min.**

    Incubating the QIAamp MinElute column loaded with Buffer AE or water for 5 min at room temperature before centrifugation generally increases DNA yield.

ZYMO RESEARCH

The Beauty of Science is to Make Things Simple

INSTRUCTION MANUAL

EZ DNA Methylation-Gold™ Kit

Catalog Nos. **D5005 & D5006**

Highlights

- Complete bisulfite conversion of GC-rich DNA in less than 3 hours.
- A coupled heat denaturation/conversion reaction step streamlines the conversion of unmethylated cytosines into uracil.
- DNA precipitations are omitted. Instead, DNA is cleaned and desulphonated in a single step using state-of-the-art spin columns.
- Eluted, ultra-pure DNA is ideal for use in subsequent molecular-based analyses.

Contents

Product Contents:

EZ DNA Methylation-Gold™ Kit	D5005 50 rxns.	D5006 200 rxns.	Storage Temperature
CT Conversion Reagent*	5 Tubes	20 Tubes	Room Temp.
M-Dilution Buffer	1.5 ml	7 ml	Room Temp.
M-Dissolving Buffer	500 µl	1.2 ml	Room Temp.
M-Binding Buffer	30 ml	125 ml	Room Temp.
M-Wash Buffer**	6 ml	24 ml	Room Temp.
M-Desulphonation Buffer	10 ml	40 ml	Room Temp.
M-Elution Buffer	1 ml	4 ml	Room Temp.
Zymo-Spin™ IC Columns	50 ct.	200 ct.	Room Temp.
Collection Tubes	50 ct.	200 ct.	Room Temp.
Instruction Manual	1	1	-

Note - Integrity of kit components is guaranteed for one year from date of purchase. Reagents are routinely tested on a lot-to-lot basis to ensure they provide maximal performance and reliability.

* 900 µl water, 300 µl **M-Dilution Buffer**, and 50 µl **M-Dissolving Buffer** must be added per tube of **CT Conversion Reagent** prior to use.

** Add 24 ml of 100% ethanol to the 6 ml **M-Wash Buffer** concentrate (D5005) or 96 ml of 100% ethanol to the 24 ml **M-Wash Buffer** concentrate (D5006) before use.

Specifications:

- **DNA Input:** Samples containing 500 pg - 2 µg of DNA. For optimal results, the amount of input DNA should be from 200 to 500 ng.
- **Conversion Efficiency:** > 99% of non-methylated C residues are converted to U; > 99% protection of methylated cytosines.
- **DNA Recovery:** > 75%

Reagent Preparation:

- **Preparation of CT Conversion Reagent**

  The **CT Conversion Reagent** supplied within this kit is a solid mixture and must be prepared prior to first use. Prepare as follows:

  1. Add 900 µl water, 300 µl of **M-Dilution Buffer**, and 50 µl **M-Dissolving Buffer** to a tube of **CT Conversion Reagent**.
  2. Mix at room temperature with frequent vortexing or shaking for 10 minutes.

  Note: It is normal to see trace amounts of undissolved reagent in the **CT Conversion Reagent**. Each tube of **CT Conversion Reagent** is designed for 10 separate DNA treatments.

  Storage: The **CT Conversion Reagent** is light sensitive, so minimize its exposure to light. For best results, the **CT Conversion Reagent** should be used immediately following preparation. If not used immediately, the **CT Conversion Reagent** solution can be stored overnight at room temperature, one week at 4°C, or up to one month at -20°C. Stored **CT Conversion Reagent** solution must be warmed to 37°C, then vortexed prior to use.

- **Preparation of M-Wash Buffer**

  Add 24 ml of 100% ethanol to the 6 ml **M-Wash Buffer** concentrate (D5005) or 96 ml of 100% ethanol to the 24 ml **M-Wash Buffer** concentrate (D5006) before use.

Protocol:

Note: For DNA volumes >20 µl, an adjustment needs to be made during the preparation of the **CT Conversion Reagent**. The amount of water is decreased 100 µl for each 10 µl increase in DNA sample volume. For example, for a 40 µl DNA sample, 700 µl of water is added to make the **CT Conversion Reagent**. The maximum DNA sample volume to be used for each conversion reaction is 50 µl. Do not adjust the volumes of either the **M-Dissolving Buffer** or **M-Dilution Buffer**.

1. Add 130 µl of the **CT Conversion Reagent** to 20 µl of your DNA sample in a PCR tube. If the volume of the DNA sample is less than 20 µl, make up the difference with water. Mix the sample by flicking the tube or pipetting the sample up and down, then centrifuge the liquid to the bottom of the tube.

2. Place the sample tube in a thermal cycler and perform the following steps*:

    1. 98°C for 10 minutes
    2. 64°C for 2.5 hours
    3. 4°C storage up to 20 hours.

    *For some samples, alternative parameters may yield improved results (see Appendix). If you have been using this kit with good results using different reaction conditions than described above, you can continue using those same conditions.

3. Add 600 µl of **M-Binding Buffer** to a **Zymo-Spin™ IC Column** and place the column into a provided **Collection Tube**.

4. Load the sample (from Step 2) into the **Zymo-Spin™ IC Column** containing the **M-Binding Buffer**. Close the cap and mix by inverting the column several times.

The capacity of the collection tube with the column inserted is 800 µl. Empty the collection tube whenever necessary to prevent contamination of the column contents by the flow-through.

5. Centrifuge at full speed (≥10,000 x *g*) for 30 seconds. Discard the flow-through.

6. Add 100 µl of **M-Wash Buffer** to the column. Centrifuge at full speed for 30 seconds.

7. Add 200 µl of **M-Desulphonation Buffer** to the column and let stand at room temperature (20°C – 30°C) for 15 - 20 minutes. After the incubation, centrifuge at full speed for 30 seconds.

8. Add 200 µl of **M-Wash Buffer** to the column. Centrifuge at full speed for 30 seconds. Add another 200 µl of **M-Wash Buffer** and centrifuge for an additional 30 seconds.

Alternatively, water or TE (pH ≥ 6.0) can be used for elution if required for your experiments.

9. Place the column into a 1.5 ml microcentrifuge tube. Add 10 µl of **M-Elution Buffer** directly to the column matrix. Centrifuge for 30 seconds at full speed to elute the DNA.

    The DNA is ready for immediate analysis or can be stored at or below -20°C for later use. For long term storage, store at or below -70°C. We recommend using 1 - 4 µl of eluted DNA for each PCR, however, up to 10 µl can be used if necessary. The elution volume can be > 10 µl depending on the requirements of your experiments, but small elution volumes will yield more concentrated DNA.

Appendix B: Solution preparations

dNTP Mix

dATP (10mM)	10µl
dTTP (10mM)	10µl
dCTP (10mM)	10µl
dGTP (10mM)	10µl
ddH_2O	760µl
	800µl

dNTPs are commercially available from Roche Diagnostics

3% Agarose Gel

Agarose	12g
1 x TBE*	400ml
EtBr* (10mg/ml)	12µl

EtBr (Ethidium Bromide: 3µl EtBr/100ml 1 x TBE)

10 x TBE

Tris	216g
Boric acid	110g
EDTA	14.88g

Make up to 2 litres with dH2O

Ethidium Bromide (EtBr)

EtBr powder	10mg
ddH20	1ml

EtBr is light sensitive and must be stored in a dark bottle.

This reagent is also commercially available from Sigma Aldrich.

Ficoll-bromophenol Blue Loading Dye (Ficoll)

50% Sucrose crystals	50g
0.5M EDTA (pH 7.0)	0.1ml
0.1% Bromophenol blue dye	0.1g
10% Ficoll powder	10g
Final volume (ddH2O)	100ml

70% Ethanol (EtOH)

100% Ethanol	70ml
ddH2O	30ml
	100ml

Binding buffer pH 7.6

10mM Tris-HCl	1.21g
2M NaCl	117g
1mM EDTA	0.292g
0.1% Tween 20	1ml
	1L

Made up with Millipore water and stored at 4°C

1X Annealing buffer, pH 7.6

20mM Tris-Acetate	2.42g
2mM Mg-Acetate	0.43g
	1L

Made up with Millipore water and stored at 4°C

Denaturation solution, 0.2M NaOH

Sodium Hydroxide (NaOH)	8g
	1L

Made up with Millipore water and stored at room temperature

Washing Buffer, pH 7.6
10mM Tris-Acetate

	1.21g
	1L

Made up with Millipore water and stored at room temperature

50 bp DNA Length Marker (100ng/µl)

Ficoll	1µl
50 bp DNA stock (1µg/µl)	1µl
ddH_20	8µl

Appendix C: Ethics approval certificate

UNIVERSITY OF THE WITWATERSRAND, JOHANNESBURG

Division of the Deputy Registrar (Research)

HUMAN RESEARCH ETHICS COMMITTEE (MEDICAL)
R14/49 Misses P Pitamber/S Patel

CLEARANCE CERTIFICATE	**M090555**
PROJECT	P: methylation Profiling of Two Paternally Imprinted loci, RASGRF and IG-DMR, in Male Gametes following Alcohol Exposure/S: Profiling the Epigenetic Signatures at the H19 DMR in Sperm of Alcoholic Males-Implications for Fetal...
INVESTIGATORS	Misses P Pitamber/S Patel.
DEPARTMENT	School of Pathology/Division of Human Genetics
DATE CONSIDERED	09.05.29
DECISION OF THE COMMITTEE*	Approved unconditionally

Unless otherwise specified this ethical clearance is valid for 5 years and may be renewed upon application.

DATE 2009/09/14 **CHAIRPERSON** .....
(Professor PE Cleaton-Jones)

*Guidelines for written 'informed consent' attached where applicable

cc: Supervisor : Prof M Ramsay

UNIVERSITY OF THE WITWATERSRAND
PROF AJ WOODWISS
Co - Chair
2009 -09- 14
HREC (Medical)
JOHANNESBURG

DECLARATION OF INVESTIGATOR(S)

To be completed in duplicate and **ONE COPY** returned to the Secretary at Room 10004, 10th Floor, Senate House, University.
I/We fully understand the conditions under which I am/we are authorized to carry out the abovementioned research and I/we guarantee to ensure compliance with these conditions. Should any departure to be contemplated from the research procedure as approved I/we undertake to resubmit the protocol to the Committee. **I agree to a completion of a yearly progress report.**

PLEASE QUOTE THE PROTOCOL NUMBER IN ALL ENQUIRIES...

Appendix D: Information sheet

STUDY TITLE: Methylation and histone profiling of two paternally imprinted loci, in male gametes, following alcohol exposure.

INVESTIGATORS: Miss Punita Navnitlal Pitamber and Miss Sanam HarishKumar Patel
(MSc [Med] Research Students)

Good Day,

We, Punita Pitamber and Sanam Patel, master's students from the Division of Human Genetics at the NHLS would like to invite you to participate in a research study. This research focuses on the effect of alcohol consumption on sperm DNA and your participation in this study is entirely voluntary. Before agreeing to participate, it is important that you read the following document to understand the purpose of the study. If you have any questions about the study or the terms used, please do not hesitate to ask one of the study co-ordinators.

The aim of the study is to understand the effects of alcohol consumption on sperm and the effect that this may have on the development of subsequent children. We already know that alcohol can affect the normal development of a child when a mother drinks during pregnancy causing a condition called Fetal Alcohol Syndrome (FAS). Recent research in mice has shown that some of the symptoms of FAS can develop in a baby born to a mother who does not drink but whose father consumed alcohol before conception.

DNA forms part of chromosomes and we all inherit half of our chromosomes from our fathers and half from our mothers. The father's sperm DNA has unique patterns, which are different to the mother's DNA patterns. These DNA patterns are very important for the normal development of the child during pregnancy. Our study seeks to examine whether alcohol can change the parameters of sperm (e.g. motility and morphology) and the patterns present in sperm DNA in such a way that may affect the normal development of a child.

If you volunteer to participate in this study we would request that you complete a questionnaire, which will take about 30-45minutes to complete. You will be asked to answer questions about your family and ethnic history, sexual health, drinking, smoking and drug use patterns. A few questions will be asked about your parents' drinking habits. Some of the questions may be sensitive and if you feel you cannot answer them, you may leave them unanswered. You may participate in this study without giving your name. If you reveal information that may have serious legal implications we may be obligated to disclose it if compelled by a court of law. You can withdraw from the study at any point without consequence to you by contacting the project investigators.

After completing the questionnaire, we will ask you to provide us with a semen sample. A private room will be made available for this purpose. Semen should only be given if you have not ejaculated in the past two days. In the laboratory, specific tests will be performed on the sperm DNA to determine if any changes have occurred in the DNA patterns. We shall not test for any general health conditions (e.g. sexually transmitted diseases) and your samples will not be used for any other studies. The samples will be given a unique identification code to ensure anonymity. The results from the test are for this research only and will not be returned to you. However, should you wish to discuss further any issues related to or raised in the course of this research, you could approach the investigators with regard to appropriate referral for counselling.

Thank you for your time.

Sincerely,

_____
Punita Pitamber

_____
Sanam Patel

Tel: (011) 489 9225

Appendix E: Consent form and questionnaire

Study Titles:

1) Methylation Profiling of Two Paternally Imprinted loci, *RASGRF* and *IG*-DMR, in Male Gametes following Alcohol Exposure

2) Profiling the Epigenetic Signatures at the *H19* DMR in Sperm of Alcoholic Males-Implications for Fetal Alcohol Spectrum Disorders (FASD)

Investigators: Miss Punita Navnitlal Pitamber and Miss Sanam Harishkumar Patel

Site of Investigation: ……
Date: …….
Code: ……

Consent Form

The information that you divulge in the questionnaire will be kept confidential

Do you understand the information related to this study?	Yes	No
Do you have any questions about the research?	Yes	No
Do we have your permission to administer the questionnaire?	Yes	No
Are you willing to contribute a semen sample?	Yes	No
Do we have your permission to conduct the described research on your semen sample?	Yes	No
Do you give permission for storage of your sample, to be used for future research on the effect of alcohol on sperm?	Yes	No

Full Name of Participant
(Participation in these studies can be anonymous): _____

Signature: _____

Date: _____

I have answered all the questions from the participant to the best of my ability.

Name of Study Co-ordinator: _____

Signature: _____

Study Titles:

1) Methylation Profiling of Two Paternally Imprinted loci, *RASGRF* and *IG*-DMR, in Male Gametes following Alcohol Exposure

2) Profiling the Epigenetic Signatures at the *H19* DMR in Sperm of Alcoholic Males-Implications for Fetal Alcohol Spectrum Disorders (FASD)

Investigators: Miss Punita Navnitlal Pitamber and Miss Sanam Harishkumar Patel

Site of Investigation: ……
Date: …….
Code: ……

Questionnaire

Please answer the following questions:

A. Personal History

1. Age: ___
2. Your first language: _____
3. Your place of birth: _____
4. Mother's first language: _____
5. Mother's place of birth: _____
6. Father's first language: _____
7. Father's place of birth: _____
8. Have you ever conceived a child? Yes No
9. Have your genitals ever been seriously injured? Yes No
10. Have you ever contracted any sexually transmitted diseases? Yes No
11. Have you ejaculated in the past 2 days? Yes No

B. sMAST Questions

Please circle the appropriate answer

1. Do you feel you are a normal drinker of alcohol? Yes No
2. Do you ever feel bad about drinking? Yes No
3. Do friends or relatives think you are a normal drinker? Yes No
4. Do your spouse/partner or parents worry of complain about your drinking? Yes No
5. Are you always able to stop drinking when you want to? Yes No
6. Have you ever attended a meeting of Alcoholics Anonymous or any similar group? Yes No
7. Has drinking ever created problems between you and your spouse? Yes No
8. Have you ever gotten into trouble at work because of drinking? Yes No
9. Have you ever neglected your obligations, your family, or your work for 2 days or more days in a row because you were drinking? Yes No
10. Have you ever gone to anyone for help about your drinking? Yes No
11. Have you ever been in hospital because of drinking? Yes No
12. Have you ever been arrested, even for a few hours, because of drinking? Yes No
13. Have you ever been arrested for drunk driving or driving after drinking? Yes No

C. Drinking Frequency

1. How old were you when you started drinking regularly? _____
2. How often do you drink? (Tick the one that fits you best)

i. Never	vi. 1-2 days a week
ii. Once a year at most	vii. Almost every day
iii. Several times a year	viii. Every day
iv. Once a month	ix. Not sure
v. 2-3 times a month	

3. Have you stopped drinking? Yes No

   If *yes* at what age? _____

4. What is your usual alcoholic beverage or beverages? _____

5. On days that you drink, how many drinks do you have? _____

D. Family Drinking History

	Mother		**Father**	
1. Is he / she alive?	Yes	No	Yes	No
2. Cause of death?	_____		_____	
3. Current age of parent or age at death?	_____		_____	
4. Does he / she drink alcohol regularly?	Yes	No	Yes	No
5. If not currently a drinker did they drink when you were growing up?	Yes	No	Yes	No

E. Smoking

1. When did you last use tobacco (smoke or chew) or snuff?
   - Never
   - More than a year ago
   - In the past year
   - In the past 30 days
   - Don't recall

2. How old were you when you first smoked or used tobacco? _____
3. Are you still smoking or using tobacco? Yes No
4. How old where you when you stopped using tobacco? _____
5. How many cigarettes a day do/did you smoke? _____

F. Other Drug Use

1. Have you ever used any drugs regularly? Yes No
2. What drugs have you used? _____

   (e.g. dagga, glue, ecstasy (XTC, E, Adam), marijuana, cocaine, heroin, sugars, mandrax, amphetamines (bennies, dexies), Ice (meth, crystal), CAT)

3. How old were you when you started using drugs regularly? _____

4. How often do/did you use drugs?

i. Never
ii. Once a year at most
iii. Several times a year
iv. Once a month
v. 2-3 times a month
vi. 1-2 days a week
vii. Almost every day
viii. Every day
ix. Not sure

5. When did you last use drugs?

- Never
- More than a year ago
- In the past year
- 1-6 months ago
- In the past 30 days
- Don't recall

Thank you very much for your co-operation ☺

Appendix F: Samples used in the study

[1] Volunteer Groups	Study ID	# Years drinking	Drinking frequency (drinks/month)	Age	Ethnicity	Drug use	Smoking	Total Sperm Count (million)	Sperm concentration (million/ml)	Sperm morphology (% normal)	Sperm motility (%)
Controls	HSS 019	0	0	25	B	no	no	275	110	15	60
	HSS 025	0	0	38	B	no	no	268.4	44	5	60
	HSS 027	0	0	28	I	no	no	205.2	54	20	60
	HSS 031	0	0	32	B	no	no	180	30	7	45
	HSS 032		Stopped: 4yrs	29	W	no	no	13	13	11	50
	HSS 033		Stopped: 6yrs	36	I	no	yes	255	85	15	50
	HSS 035	0	0	29	B	no	no	129.6	18	10	40
	HSS 037		Stopped: 4yrs	33	B	no	no	383.6	137	7	60
	HSS 038	0	0	37	B	no	no	138	23	8	60
	HSS 042		Stopped: 9yrs	40	B	no	no	36	18	12	50
	HSS 046	0	0	33	B	no	no	385	77	5	50
	HSS 050		Stopped: 20yrs	50	I	no	no	126	42	4	40
	HSS 055		Stopped: 2yrs	29	C	no	no	72.6	22	6	60
	HSS 056	0	0	30	B	no	yes	84	21	11	50
	HSS 070	0	0	34	B	no	no	160.6	73	14	70
	HSS 074	0	0	30	B	no	no	180	72	6	40
	HSS 075	0	0	40	B	no	no	25.2	36	16	70
	HSS 077	0	0	28	I	no	no	167.7	39	5	50
	HSS 079	0	0	39	B	no	no	21.5	167	20	50
	HSS 082	0	0	22	I	no	no	201.6	56	15	65
	HSS 085	0	0	38	B	no	no	148.2	57	10	50

[1] This table shows the study IDs of the volunteers together with information on drinking frequency, age, ethnicity [European descent (W), African descent (B), Mixed Ancestry (C), Asian descent (I)], drug use, and smoking. Also included are sperm parameters. The numbers in red represent those that fall below the normal threshold determined by the World Health Organization in 2010. These are ≥ 15 million sperm per ml concentration, at least 4% normal sperm morphology and > 50% sperm motility.

Volunteer Groups	Study ID	# Years drinking	Drinking frequency (drinks/month)	Age	Ethnicity	Drug use	Smoking	Total Sperm Count (million)	Sperm concentration (million/ml)	Sperm morphology (% normal)	Sperm motility (%)
Controls	HSS 088	0	0	38	I	no	yes	53.2	28	13	65
	HSS 089	0	0	38	B	no	no	151.2	36	15	70
	HSS 094	0	0	28	I	no	no	166	83	16	65
	HSS 096	0	0	36	B	no	no	119.6	23	10	40
	HSS097	0	0	36	B	no	no	19.8	11	1	40
	HSS 107	0	0	43	I	no	yes	115.6	68	16	55
	HSS 116	0	0	45	B	no	no	228.8	52	19	55
	HSS 121	0	0	28	I	no	no	188.5	53	18	70
	HSS 131	0	0	33	C	no	no	329	70	10	50
	HSS 136	0	0	36	B	no	no	118.4	37	21	50
	HSS 141	0	0	37	W	no	yes	57.6	32	1	30
	HSS 142		Stopped: 6yrs	36	C	no	yes	137.6	65.5	3	70
	HSS 149	0	0	33	B	yes	no	42	14	9	60
Drinkers	HSS 001	6	30	23	B	no	yes	152	76	15	60
	HSS 003	12	15	31	W	no	yes	255	102	18	50
	HSS 004	9	72	27	W	no	yes	480	96	14	60
	HSS 005	4	6	28	W	no	no	217	70	19	65
	HSS 006	20	3	38	W	no	no	30	10	13	60
	HSS 008	6	60	22	W	no	yes	88.8	24	15	60
	HSS 010	10	6	35	B	no	no	84.8	16	9	60
	HSS 011	2	75	20	B	no	no	80	40	20	60
	HSS 016	7	6	26	W	no	no	155	62	12	60
	HSS 017	21	15	31	B	no	yes	297	110	11	60
	HSS 018	4	15	34	B	no	no	78.4	28	14	60
	HSS 020	16	120	34	B	no	no	23.4	18	18	60
	HSS 021	16	72	34	W	no	yes	40.5	**8.8**	2	60
	HSS 022	11	30	29	W	no	no	17.6	4.4	10	50
	HSS 023	12	6	33	W	yes	yes	1624	280	17	40
	HSS 024	15	1	33	I	no	no	409.6	128	12	60

Volunteer Groups	Study ID	# Years drinking	Drinking frequency (drinks/month)	Age	Ethnicity	Drug use	Smoking	Total Sperm Count (million)	Sperm concentration (million/ml)	Sperm morphology (% normal)	Sperm motility (%)
Drinkers	HSS 028	18	1	36	C	no	no	162	81	19	60
	HSS 036	22	36	40	B	no	no	74.8	44	5	45
	HSS 040	22	18	38	B	yes	no	21	**14**	10	50
	HSS 043	10	9	29	I	no	no	163.4	86	16	50
	HSS 044	13	2	34	B	no	yes	19	**10**	**1**	45
	HSS 048	17	60	37	B	no	no	55	**11**	**1**	40
	HSS 051	15	4	35	I	no	yes	110	110	9	45
	HSS 052	21	60	39	B	no	no	11	**10**	5	60
	HSS 053	15	84	33	C	yes	no	238	70	16	70
	HSS 054	7	14	30	I	no	yes	176	80	5	65
	HSS 057	19	150	40	W	no	yes	40.5	27	4	50
	HSS 058	17	36	34	W	no	no	178.2	99	12	60
	HSS 061	19	8	35	W	no	no	364.8	76	15	70
	HSS 062	12	2	29	W	no	yes	729	162	20	75
	HSS 064	17	72	35	C	no	no	13.9	**6.3**	4	**20**
	HSS 067	18	1	36	B	no	no	171	45	20	45
	HSS 068	22	30	39	W	no	yes	189.6	38	14	55
	HSS 069	16	12	35	W	no	no	70	20	15	50
	HSS 071	7	24	34	B	no	yes	17.5	**7**	12	40
	HSS 072	15	6	35	B	no	yes	750	25	15	70
	HSS 073	28	6	46	B	no	no	48	24	10	40
	HSS 076	8	36	28	B	no	yes	176	176	20	65
	HSS 078	15	15	45	W	no	yes	501	43	17	40
	HSS 081	13	30	34	B	no	no	172.2	123	18	65
	HSS 083	14	5	39	B	no	no	124.2	54	19	70

Volunteer Groups	Study ID	# Years drinking	Drinking frequency (drinks/month)	Age	Ethnicity	Drug use	Smoking	Total Sperm Count (million)	Sperm concentration (million/ml)	Sperm morphology (% normal)	Sperm motility (%)
Drinkers	HSS 084	18	24	37	B	no	no	63.8	29	16	60
	HSS 086	21	6	39	W	no	no	315.4	83	20	60
	HSS 090	9	45	25	W	no	yes	47.6	28	9	60
	HSS 092	12	108	27	W	no	yes	139.2	58	8	60
	HSS 095	5	144	28	B	no	no	15.1	4.2	5	60
	HSS 098	20	48	35	B	no	no	172.8	48	19	70
	HSS102	13	60	41	B	no	yes	19.8	11	2	40
	HSS 103	15	2	31	W	yes	no	151.2	151.2	2	70
	HSS 105	18	1	36	W	no	no	310	310	14	60
	HSS 110	16	9	35	W	no	yes	24	16	1	40
	HSS 111	12	23	33	W	no	yes	246.4	28	6	60
	HSS112	22	42	42	B	no	no	12	6	8	45
	HSS 114	24	60	49	B	no	yes	17.1	9	2	50
	HSS 117	15	7	31	W	no	yes	26.6	5.9	6	50
	HSS118	13	5	48	B	no	no	20.8	8	1	40
	HSS 119	11	4	30	W	no	yes	72.8	28	15	50
	HSS 120	18	15	35	B	no	no	147	49	17	30
	HSS 122	13	6	31	W	no	no	186	155	25	65
	HSS 123	16	144	36	B	no	no	186	62	20	60
	HSS 124	10	42	29	W	no	no	126	36	11	60
	HSS 125	16	15	34	W	no	yes	66.3	17	1	60
	HSS 126	8	3	29	I	no	no	264.6	63	4	70
	HSS 127	5	36	30	B	no	no	128.8	56	10	70
	HSS 128	11	6	29	W	no	no	30	12	1	50
	HSS 129	16	36	41	B	no	yes	59.4	66	16	60

Volunteer Groups	Study ID	# Years drinking	Drinking frequency (drinks/month)	Age	Ethnicity	Drug use	Smoking	Total Sperm Count (million)	Sperm concentration (million/ml)	Sperm morphology (% normal)	Sperm motility (%)
Drinkers	HSS 130	17	360	32	B	no	yes	208.8	58	7	60
	HSS 133	18	96	33	B	no	no	22.8	5.3	3	60
	HSS 132	18	42	39	B	no	no	237.5	95	15	60
	HSS 134	15	12	33	B	no	yes	285.5	47	15	65
	HSS 135	9	10	33	B	no	no	121	55	8	60
	HSS 137	16	1	34	W	no	no	172.5	23	4	50
	HSS 140	14	9	32	W	no	yes	181.8	101	16	70
	HSS 143	16	10	33	C	no	yes	72	18	2	50
	HSS 144	6	15	31	B	no	no	210.9	37	10	50
	HSS 145	17	30	38	W	no	yes	135	27	4	60
	HSS 146	14	60	34	I	yes	no	198.4	124	15	60
	HSS 147	10	18	26	B	no	no	70	28	17	60
	HSS 148	18	10	39	B	no	no	132.6	162	3	60

Appendix G: *SNRPN* results

Sample ID	CpG Sites Analyzed					Mean Methylation (%) per Sample
		1	**2**	**3**	**4**	
Controls						
HSS 032	PCR1	4	4	4	5	
	PCR2	4	0	4	4	
	PCR3	0	0	0	0	
	Mean Methylation (%)	**3**	**1**	**3**	**3**	**2**
HSS 037	PCR1	0	3	0	4	
	PCR2	5	6	5	6	
	PCR3	0	0	0	7	
	Mean Methylation (%)	**2**	**3**	**2**	**6**	**3**
HSS 038	PCR1	6	0	0	5	
	PCR2	0	0	0	0	
	PCR3	5	5	4	5	
	Mean Methylation (%)	**4**	**2**	**1**	**3**	**3**
HSS 085	PCR1	0	0	0	3	
	PCR2	0	5	4	4	
	PCR3	0	0	0	5	
	Mean Methylation (%)	**0**	**2**	**1**	**4**	**2**
HSS 088	PCR1	0	0	0	0	
	PCR2	5	0	0	0	
	PCR3	5	0	4	0	
	Mean Methylation (%)	**3**	**0**	**1**	**0**	**1**
Drinkers						
HSS 005	PCR1	4	0	0	0	
	PCR2	0	0	0	0	
	PCR3	0	0	0	0	
	Mean Methylation (%)	**1**	**0**	**0**	**0**	**0**
HSS 006	PCR1	6	5	6	6	
	PCR2	0	0	0	8	
	PCR3	4	5	5	7	
	Mean Methylation (%)	**3**	**3**	**4**	**7**	**4**
HSS 017	PCR1	0	5	0	7	
	PCR2	4	4	0	7	
	PCR3	0	0	2	4	
	Mean Methylation (%)	**1**	**3**	**1**	**6**	**3**
HSS 018	PCR1	0	0	0	0	
	PCR2	0	0	0	0	
	PCR3	0	4	5	5	
	Mean Methylation (%)	**0**	**1**	**2**	**2**	**1**
HSS 020	PCR1	0	0	0	5	

	PCR2	0	6	5	7	
	PCR3	4	5	5	6	
	Mean Methylation (%)	**1**	**4**	**3**	**6**	**4**
HSS 022	PCR1	0	0	0	0	
	PCR2	0	0	0	0	
	PCR3	0	0	0	0	
	Mean Methylation (%)	**0**	**0**	**0**	**0**	**0**
HSS 023	PCR1	6	6	4	4	
	PCR2	3	3	3	4	
	PCR3	6	7	6	7	
	Mean Methylation (%)	**5**	**5**	**4**	**5**	**5**
HSS 024	PCR1	3	0	4	4	
	PCR2	0	0	0	9	
	PCR3	4	4	5	6	
	Mean Methylation (%)	**2**	**1**	**3**	**6**	**3**
HSS 043	PCR1	0	0	0	0	
	PCR2	4	0	0	0	
	PCR3	5	3	4	5	
	Mean Methylation (%)	**3**	**1**	**1**	**2**	**2**
HSS 048	PCR1	4	5	5	5	
	PCR2	0	6	0	9	
	PCR3	5	4	0	5	
	Mean Methylation (%)	**3**	**5**	**2**	**6**	**4**
HSS 054	PCR1	3	0	0	5	
	PCR2	4	0	0	5	
	PCR3	6	5	4	7	
	Mean Methylation (%)	**4**	**2**	**1**	**6**	**3**
HSS 067	PCR1	6	5	0	7	
	PCR2	6	3	4	4	
	PCR3	3	5	0	5	
	Mean Methylation (%)	**5**	**4**	**1**	**5**	**4**
HSS 073	PCR1	4	4	4	4	
	PCR2	0	0	0	0	
	PCR3	3	0	3	4	
	Mean Methylation (%)	**2**	**1**	**2**	**3**	**2**
HSS 081	PCR1	0	0	0	0	
	PCR2	0	4	0	6	
	PCR3	0	0	0	0	
	Mean Methylation (%)	**0**	**1**	**0**	**2**	**1**
HSS 122	PCR1	0	0	0	4	
	PCR2	0	0	0	0	
	PCR3	5	0	0	0	
	Mean Methylation (%)	**2**	**0**	**0**	**1**	**1**
HSS 128	PCR1	0	0	0	0	
	PCR2	4	3	5	4	
	PCR3	0	0	5	6	
	Mean Methylation (%)	**1**	**1**	**3**	**3**	**2**
HSS 129	PCR1	5	0	6	7	

PCR2	0	0	0	6	
PCR3	0	4	5	6	
Mean Methylation (%)	**2**	**1**	**4**	**6**	**3**

Appendix H: *H19* results

[2]

Sample ID		CpG Sites Analyzed: Sperm HH19 ICR (CTCF binding site)						Mean Methylation (%) per Sample
		1	2	3	4	5	6	
Controls								
HSS 019	PCR1	89	100	68	89	91	99	
	PCR2	89	100	69	88	89	99	
	PCR3	89	100	69	88	90	97	
	Mean Methylation (%)	**89**	**100**	**69**	**88**	**90**	**98**	**89**
HSS 025	PCR1	90	100	66	88	93	96	
	PCR2	92	100	71	85	91	95	
	PCR3	89	100	68	85	88	99	
	Mean Methylation (%)	**90**	**100**	**68**	**86**	**91**	**97**	**89**
HSS 027	PCR1	92	100	74	44	90	95	
	PCR2	93	100	77	45	89	96	
	PCR3	92	100	74	44	90	97	
	Mean Methylation (%)	**92**	**100**	**75**	**44**	**90**	**96**	**91**
HSS 031	PCR1	94	100	77	91	91	98	
	PCR2	94	100	77	89	92	98	
	PCR3	93	100	78	90	93	99	
	Mean Methylation (%)	**94**	**100**	**77**	**90**	**92**	**98**	**92**
HSS 032	PCR1	94	100	82	42	87	97	
	PCR2	92	100	83	43	87	96	
	PCR3	93	100	83	43	88	97	
	Mean Methylation (%)	**93**	**100**	**83**	**43**	**87**	**97**	**92**
HSS 033	PCR1	95	100	87	49	100	96	
	PCR2	95	100	88	47	100	96	
	PCR3	95	100	88	47	98	95	
	Mean Methylation (%)	**95**	**100**	**88**	**48**	**99**	**96**	**96**
HSS 035	PCR1	86	100	79	87	86	97	
	PCR2	85	100	80	84	88	98	
	PCR3	83	100	75	84	78	95	
	Mean Methylation (%)	**85**	**100**	**78**	**85**	**84**	**97**	**89**
HSS 037	PCR1	92	100	76	88	92	100	
	PCR2	93	100	78	88	91	99	
	PCR3	94	100	78	89	94	97	
	Mean Methylation (%)	**93**	**100**	**77**	**88**	**92**	**99**	**92**
HSS 038	PCR1	89	100	71	88	93	97	
	PCR2	89	100	72	86	92	99	
	PCR3	90	100	72	86	93	94	
	Mean Methylation (%)	**89**	**100**	**72**	**87**	**93**	**97**	**90**

[2] Note: CpG site 4 results are included in the table for interest but values have not been used for any analyses or for calculating the mean methylation (%) per sample.

HSS 042	PCR1	94	100	78	89	93	91	
	PCR2	90	100	80	88	91	94	
	PCR3	92	100	78	88	91	93	
	Mean Methylation (%)	**92**	**100**	**79**	**88**	**92**	**93**	**91**
HSS 046	PCR1	90	100	75	87	92	94	
	PCR2	91	100	76	89	93	98	
	PCR3	90	100	75	87	92	94	
	Mean Methylation (%)	**90**	**100**	**75**	**88**	**92**	**95**	**91**
HSS 050	PCR1	94	100	82	87	90	94	
	PCR2	93	100	82	88	91	96	
	PCR3	94	100	83	87	91	95	
	Mean Methylation (%)	**94**	**100**	**82**	**87**	**91**	**95**	**92**
HSS 055	PCR1	91	100	67	86	92	92	
	PCR2	90	100	66	88	91	94	
	PCR3	90	100	66	87	91	93	
	Mean Methylation (%)	**90**	**100**	**66**	**87**	**91**	**93**	**88**
HSS 056	PCR1	94	100	83	87	91	93	
	PCR2	95	100	83	88	94	95	
	PCR3	94	100	83	87	94	96	
	Mean Methylation (%)	**94**	**100**	**83**	**87**	**93**	**95**	**93**
HSS 070	PCR1	89	98	64	86	89	99	
	PCR2	88	100	64	84	90	93	
	PCR3	89	100	64	84	89	96	
	Mean Methylation (%)	**89**	**99**	**64**	**85**	**89**	**96**	**87**
HSS 075	PCR1	92	96	71	89	88	94	
	PCR2	93	97	72	90	92	98	
	PCR3	94	99	74	90	93	100	
	Mean Methylation (%)	**93**	**97**	**72**	**90**	**91**	**97**	**90**
HSS 077	PCR1	93	100	74	89	91	97	
	PCR2	93	100	75	89	92	92	
	PCR3	92	100	75	87	90	94	
	Mean Methylation (%)	**93**	**100**	**75**	**88**	**91**	**94**	**91**
HSS 079	PCR1	94	100	81	87	89	94	
	PCR2	95	100	83	89	91	94	
	PCR3	95	100	79	88	93	93	
	Mean Methylation (%)	**95**	**100**	**81**	**88**	**91**	**94**	**92**
HSS 082	PCR1	86	98	66	50	92	94	
	PCR2	85	97	67	50	92	93	
	PCR3	86	99	67	49	93	91	
	Mean Methylation (%)	**86**	**98**	**67**	**50**	**92**	**93**	**87**
HSS 085	PCR1	87	99	65	85	90	96	
	PCR2	89	100	65	88	92	96	
	PCR3	90	100	66	89	93	96	
	Mean Methylation (%)	**89**	**100**	**65**	**87**	**92**	**96**	**88**
HSS 088	PCR1	95	100	86	43	91	97	
	PCR2	94	100	87	41	93	95	
	PCR3	94	100	86	41	93	96	
	Mean Methylation (%)	**94**	**100**	**86**	**42**	**92**	**96**	**94**

HSS 089	PCR1	90	99	69	81	89	97	
	PCR2	91	100	69	87	93	96	
	PCR3	89	100	67	87	92	94	
	Mean Methylation (%)	**90**	**100**	**68**	**85**	**91**	**96**	**89**
HSS 094	PCR1	88	100	70	85	90	95	
	PCR2	90	100	72	86	93	95	
	PCR3	90	100	71	88	91	94	
	Mean Methylation (%)	**89**	**100**	**71**	**86**	**91**	**95**	**89**
HSS 096	PCR1	94	100	81	90	91	98	
	PCR2	93	100	81	88	91	97	
	PCR3	94	100	81	88	91	97	
	Mean Methylation (%)	**94**	**100**	**81**	**89**	**91**	**97**	**93**
HSS097	PCR1	91	100	72	90	95	100	
	PCR2	91	100	71	91	94	96	
	PCR3	91	100	71	91	93	96	
	Mean Methylation (%)	**91**	**100**	**71**	**91**	**94**	**97**	**91**
HSS 107	PCR1	94	100	72	91	94	99	
	PCR2	94	100	73	90	93	99	
	PCR3	95	100	72	92	93	100	
	Mean Methylation (%)	**94**	**100**	**72**	**91**	**93**	**99**	**92**
HSS 116	PCR1	92	100	74	87	89	90	
	PCR2	93	99	75	90	91	93	
	PCR3	92	99	76	90	91	94	
	Mean Methylation (%)	**92**	**99**	**75**	**89**	**90**	**92**	**90**
HSS 121	PCR1	95	100	84	46	93	97	
	PCR2	94	100	84	45	91	99	
	PCR3	95	100	85	45	93	99	
	Mean Methylation (%)	**95**	**100**	**84**	**45**	**92**	**98**	**94**
HSS 131	PCR1	90	100	69	87	94	100	
	PCR2	90	100	69	87	93	97	
	PCR3	93	98	70	87	93	100	
	Mean Methylation (%)	**91**	**99**	**69**	**87**	**93**	**99**	**90**
HSS 136	PCR1	94	100	72	88	89	100	
	PCR2	93	100	74	87	89	100	
	PCR3	94	100	72	87	93	100	
	Mean Methylation (%)	**94**	**100**	**73**	**87**	**90**	**100**	**91**
HSS 141	PCR1	95	99	88	5	90	96	
	PCR2	95	99	88	5	88	97	
	PCR3	95	100	88	7	89	94	
	Mean Methylation (%)	**95**	**99**	**88**	**6**	**89**	**96**	**93**
HSS 142	PCR1	93	100	84	46	87	97	
	PCR2	93	100	85	47	89	95	
	PCR3	93	100	84	47	89	96	
	Mean Methylation (%)	**93**	**100**	**84**	**47**	**88**	**96**	**92**
HSS 149	PCR1	95	100	83	89	90	98	
	PCR2	93	100	83	90	90	99	
	PCR3	91	100	83	93	91	96	
	Mean Methylation (%)	**93**	**100**	**83**	**91**	**90**	**98**	**93**

Drinkers

HSS 001	PCR1	92	100	86	41	90	96	
	PCR2	90	100	86	40	90	96	
	PCR3	90	100	85	39	90	97	
	Mean Methylation (%)	**91**	**100**	**86**	**40**	**90**	**96**	**93**
HSS 003	PCR1	78	100	63	44	89	95	
	PCR2	79	100	63	43	92	96	
	PCR3	75	100	65	47	89	100	
	Mean Methylation (%)	**77**	**100**	**64**	**45**	**90**	**97**	**86**
HSS 004	PCR1	92	100	83	91	91	100	
	PCR2	89	100	83	91	90	100	
	PCR3	92	100	82	90	91	100	
	Mean Methylation (%)	**91**	**100**	**83**	**91**	**91**	**100**	**93**
HSS 005	PCR1	100	100	88	8	100	96	
	PCR2	94	100	86	8	100	94	
	PCR3	100	100	86	8	100	97	
	Mean Methylation (%)	**98**	**100**	**87**	**8**	**100**	**96**	**96**
HSS 006	PCR1	93	100	84	46	92	99	
	PCR2	90	100	86	41	91	98	
	PCR3	93	100	82	45	93	100	
	Mean Methylation (%)	**92**	**100**	**84**	**44**	**92**	**99**	
HSS 008	PCR1	87	100	82	6	87	94	
	PCR2	89	100	83	6	87	98	
	PCR3	86	100	82	6	91	96	
	Mean Methylation (%)	**87**	**100**	**82**	**6**	**88**	**96**	**91**
								93
HSS 010	PCR1	94	100	81	88	94	97	
	PCR2	92	100	78	86	94	99	
	PCR3	94	100	79	89	93	98	
	Mean Methylation (%)	**93**	**100**	**79**	**88**	**94**	**98**	**93**
HSS 011	PCR1	76	93	63	70	72	79	
	PCR2	77	89	63	74	78	78	
	PCR3	78	89	69	75	77	77	
	Mean Methylation (%)	**77**	**90**	**65**	**73**	**76**	**78**	**77**
HSS 016	PCR1	69	99	47	79	87	96	
	PCR2	70	100	50	80	89	100	
	PCR3	69	100	47	76	87	95	
	Mean Methylation (%)	**69**	**100**	**48**	**78**	**88**	**97**	**80**
HSS 017	PCR1	92	100	69	85	86	97	
	PCR2	93	100	71	86	86	96	
	PCR3	94	100	71	86	82	96	
	Mean Methylation (%)	**93**	**100**	**70**	**86**	**85**	**96**	**89**
HSS 018	PCR1	100	100	72	87	90	100	
	PCR2	96	100	74	87	89	100	
	PCR3	97	100	72	88	90	100	
	Mean Methylation (%)	**98**	**100**	**73**	**87**	**90**	**100**	**92**

HSS 020	PCR1	94	100	81	90	82	97	
	PCR2	93	100	81	89	86	99	
	PCR3	93	100	84	87	81	96	
	Mean Methylation (%)	**93**	**100**	**82**	**89**	**83**	**97**	**91**
HSS 021	PCR1	83	98	67	40	87	91	
	PCR2	83	98	65	41	87	89	
	PCR3	81	98	66	41	87	93	
	Mean Methylation (%)	**82**	**98**	**66**	**41**	**87**	**91**	**85**
HSS 022	PCR1	90	100	67	88	100	99	
	PCR2	89	100	67	88	100	97	
	PCR3	88	100	67	89	100	96	
	Mean Methylation (%)	**89**	**100**	**67**	**88**	**100**	**97**	**91**
HSS 023	PCR1	91	100	87	43	90	97	
	PCR2	91	100	82	44	90	100	
	PCR3	91	100	83	43	89	98	
	Mean Methylation (%)	**91**	**100**	**84**	**43**	**90**	**98**	**93**
HSS 024	PCR1	90	100	80	43	92	96	
	PCR2	90	100	79	41	90	97	
	PCR3	90	100	79	41	89	95	
	Mean Methylation (%)	**90**	**100**	**79**	**42**	**90**	**96**	**91**
HSS 036	PCR1	93	100	79	90	92	95	
	PCR2	93	100	78	89	91	95	
	PCR3	93	100	79	92	93	96	
	Mean Methylation (%)	**93**	**100**	**79**	**90**	**92**	**95**	**92**
HSS 040	PCR1	94	100	78	92	94	98	
	PCR2	90	100	79	90	92	98	
	PCR3	94	100	78	92	94	98	
	Mean Methylation (%)	**93**	**100**	**78**	**91**	**93**	**98**	**92**
HSS 043	PCR1	85	99	67	85	92	96	
	PCR2	86	100	66	81	89	98	
	PCR3	86	100	68	83	90	96	
	Mean Methylation (%)	**86**	**100**	**67**	**83**	**90**	**97**	**88**
HSS 044	PCR1	89	100	69	85	91	93	
	PCR2	90	100	71	87	92	96	
	PCR3	86	100	70	86	91	92	
	Mean Methylation (%)	**88**	**100**	**70**	**86**	**91**	**94**	**89**
HSS 048	PCR1	89	100	76	87	91	96	
	PCR2	90	100	74	85	90	97	
	PCR3	91	100	76	88	91	95	
	Mean Methylation (%)	**90**	**100**	**75**	**87**	**91**	**96**	**90**
HSS 051	PCR1	95	100	84	86	89	97	
	PCR2	93	100	83	87	91	97	
	PCR3	95	100	82	88	91	98	
	Mean Methylation (%)	**94**	**100**	**83**	**87**	**90**	**97**	**93**
HSS 052	PCR1	87	100	92	89	89	96	
	PCR2	88	100	90	86	92	94	
	PCR3	91	100	89	87	87	94	
	Mean Methylation (%)	**89**	**100**	**90**	**87**	**89**	**95**	**93**

HSS 053	PCR1	93	100	79	87	89	97	
	PCR2	93	100	81	87	91	98	
	PCR3	94	100	80	87	91	97	
	Mean Methylation (%)	**93**	**100**	**80**	**87**	**90**	**97**	**92**
HSS 054	PCR1	82	100	65	43	89	93	
	PCR2	80	98	67	45	88	94	
	PCR3	82	96	65	44	89	95	
	Mean Methylation (%)	**81**	**98**	**66**	**44**	**89**	**94**	**86**
HSS 057	PCR1	95	100	90	6	88	91	
	PCR2	93	100	94	7	88	92	
	PCR3	92	100	92	4	91	91	
	Mean Methylation (%)	**93**	**100**	**92**	**6**	**89**	**91**	**93**
HSS 058	PCR1	93	100	89	40	90	94	
	PCR2	91	100	90	41	86	96	
	PCR3	94	100	88	38	90	93	
	Mean Methylation (%)	**93**	**100**	**89**	**40**	**89**	**94**	**93**
HSS 061	PCR1	92	100	76	40	91	97	
	PCR2	93	100	77	43	94	95	
	PCR3	93	100	79	41	92	97	
	Mean Methylation (%)	**93**	**100**	**77**	**41**	**92**	**96**	**92**
HSS 062	PCR1	91	100	78	44	86	95	
	PCR2	90	100	76	47	82	98	
	PCR3	92	100	79	44	89	97	
	Mean Methylation (%)	**91**	**100**	**78**	**45**	**86**	**97**	**90**
HSS 064	PCR1	91	100	83	44	91	94	
	PCR2	87	99	83	47	89	94	
	PCR3	93	99	84	49	92	97	
	Mean Methylation (%)	**90**	**99**	**83**	**47**	**91**	**95**	**92**
HSS 067	PCR1	93	100	85	88	91	92	
	PCR2	90	100	85	88	89	96	
	PCR3	92	100	86	89	92	95	
	Mean Methylation (%)	**92**	**100**	**85**	**88**	**91**	**94**	**92**
HSS 068	PCR1	93	100	84	90	93	94	
	PCR2	92	100	83	87	88	98	
	PCR3	93	100	86	89	91	97	
	Mean Methylation (%)	**93**	**100**	**84**	**89**	**91**	**96**	**93**
HSS 069	PCR1	93	100	87	48	89	90	
	PCR2	93	100	86	48	88	94	
	PCR3	95	100	88	48	91	92	
	Mean Methylation (%)	**94**	**100**	**87**	**48**	**89**	**92**	**92**
HSS 071	PCR1	90	100	69	85	92	96	
	PCR2	88	100	68	86	90	94	
	PCR3	92	100	70	85	92	94	
	Mean Methylation (%)	**90**	**100**	**69**	**85**	**91**	**95**	**89**
HSS 072	PCR1	92	100	79	43	91	96	
	PCR2	92	100	80	44	89	95	
	PCR3	92	100	81	43	90	90	
	Mean Methylation (%)	**92**	**100**	**80**	**43**	**90**	**94**	**91**

HSS 073	PCR1	95	100	88	86	90	93	
	PCR2	93	100	87	87	90	95	
	PCR3	94	100	89	87	92	95	
	Mean Methylation (%)	**94**	**100**	**88**	**87**	**91**	**94**	**93**
HSS 076	PCR1	91	100	73	87	91	99	
	PCR2	93	100	72	87	90	97	
	PCR3	93	100	73	87	92	97	
	Mean Methylation (%)	**92**	**100**	**73**	**87**	**91**	**98**	**91**
HSS 078	PCR1	90	100	87	16	83	95	
	PCR2	91	100	88	11	85	92	
	PCR3	93	100	85	10	89	92	
	Mean Methylation (%)	**91**	**100**	**87**	**12**	**86**	**93**	**91**
HSS 084	PCR1							
	PCR2	91	95	77	89	90	94	
	PCR3	95	99	76	91	93	97	
	Mean Methylation (%)	95	99	78	90	93	94	
		94	**98**	**77**	**90**	**92**	**95**	**91**
HSS 086	PCR1	94	100	91	6	87	94	
	PCR2	93	100	90	5	88	94	
	PCR3	95	99	90	5	90	93	
	Mean Methylation (%)	**94**	**100**	**90**	**5**	**88**	**94**	**93**
HSS 090	PCR1	93	100	81	45	89	96	
	PCR2	93	100	82	46	85	92	
	PCR3	95	100	83	45	92	96	
	Mean Methylation (%)	**94**	**100**	**82**	**45**	**89**	**95**	**92**
HSS 092	PCR1	93	100	81	48	87	99	
	PCR2	91	100	85	48	86	95	
	PCR3	92	100	83	50	88	96	
	Mean Methylation (%)	**92**	**100**	**83**	**49**	**87**	**97**	**92**
HSS 095	PCR1	90	100	68	86	91	97	
	PCR2	91	100	66	85	90	96	
	PCR3	91	100	66	85	90	97	
	Mean Methylation (%)	**91**	**100**	**67**	**85**	**90**	**97**	**89**
HSS 098	PCR1	92	100	80	88	92	95	
	PCR2	92	100	82	90	92	97	
	PCR3	92	100	80	89	92	96	
	Mean Methylation (%)	**92**	**100**	**81**	**89**	**92**	**96**	**92**
HSS102	PCR1	95	100	83	92	93	100	
	PCR2	96	100	83	91	93	100	
	PCR3	95	100	84	92	92	97	
	Mean Methylation (%)	**95**	**100**	**83**	**92**	**93**	**99**	**94**
HSS 103	PCR1	94	100	81	90	94	98	
	PCR2	93	100	81	91	94	100	
	PCR3	94	100	81	90	92	98	
	Mean Methylation (%)	**94**	**100**	**81**	**90**	**93**	**99**	**93**
HSS 105	PCR1	95	100	79	90	93	100	
	PCR2	95	100	79	90	94	99	
	PCR3	95	100	79	90	94	97	
	Mean Methylation (%)	**95**	**100**	**79**	**90**	**94**	**99**	**93**

HSS 110	PCR1	95	100	86	91	93	97	
	PCR2	96	100	88	90	95	100	
	PCR3	95	100	88	90	92	99	
	Mean Methylation (%)	**95**	**100**	**87**	**90**	**93**	**99**	**95**
HSS 111	PCR1	91	100	73	89	93	93	
	PCR2	94	97	75	89	93	96	
	PCR3	93	100	73	88	92	97	
	Mean Methylation (%)	**93**	**99**	**74**	**89**	**93**	**95**	**91**
HSS112	PCR1	95	100	83	90	91	100	
	PCR2	96	100	85	90	93	100	
	PCR3	94	100	86	89	92	100	
	Mean Methylation (%)	**95**	**100**	**85**	**90**	**92**	**100**	**94**
HSS 114	PCR1	94	100	85	89	92	96	
	PCR2	91	100	85	89	92	97	
	PCR3	92	100	81	92	95	97	
	Mean Methylation (%)	**92**	**100**	**84**	**90**	**93**	**97**	**93**
HSS 117	PCR1	94	100	73	89	94	99	
	PCR2	92	100	73	88	93	100	
	PCR3	94	99	74	90	94	98	
	Mean Methylation (%)	**93**	**100**	**73**	**89**	**94**	**99**	**92**
HSS118	PCR1	98	100	88	90	93	100	
	PCR2	96	100	88	89	93	100	
	PCR3	97	100	88	92	93	99	
	Mean Methylation (%)	**97**	**100**	**88**	**90**	**93**	**100**	**96**
HSS 119	PCR1	95	100	90	6	90	93	
	PCR2	96	100	91	6	90	94	
	PCR3	95	100	90	6	89	93	
	Mean Methylation (%)	**95**	**100**	**90**	**6**	**90**	**93**	**94**
HSS 120	PCR1	92	100	72	83	94	96	
	PCR2	93	100	72	85	91	99	
	PCR3	94	100	72	86	93	99	
	Mean Methylation (%)	**93**	**100**	**72**	**85**	**93**	**98**	**91**
HSS 122	PCR1	96	100	77	90	93	98	
	PCR2	96	100	79	89	93	100	
	PCR3	96	100	78	89	93	100	
	Mean Methylation (%)	**96**	**100**	**78**	**89**	**93**	**99**	**93**
HSS 123	PCR1	94	100	75	87	90	99	
	PCR2	92	100	73	88	90	99	
	PCR3	93	100	73	88	90	97	
	Mean Methylation (%)	**93**	**100**	**74**	**88**	**90**	**98**	**91**
HSS 125	PCR1	93	100	88	6	90	95	
	PCR2	94	100	88	5	91	94	
	PCR3	93	100	89	6	89	93	
	Mean Methylation (%)	**93**	**100**	**88**	**6**	**90**	**94**	**93**

HSS 126	PCR1	93	100	73	88	91	98	
	PCR2	94	100	75	88	92	99	
	PCR3	93	100	73	89	91	98	
	Mean Methylation (%)	**93**	**100**	**74**	**88**	**91**	**98**	**91**
HSS 127	PCR1	96	100	81	90	92	100	
	PCR2	97	98	82	90	93	99	
	PCR3	96	100	81	91	94	98	
	Mean Methylation (%)	**96**	**99**	**81**	**90**	**93**	**99**	**94**
HSS 128	PCR1	93	100	84	6	90	95	
	PCR2	92	100	84	5	91	96	
	PCR3	93	99	84	6	90	97	
	Mean Methylation (%)	**93**	**100**	**84**	**6**	**90**	**96**	**93**
HSS 129	PCR1	96	100	78	89	93	100	
	PCR2	95	100	76	90	93	99	
	PCR3	96	100	77	89	93	100	
	Mean Methylation (%)	**96**	**100**	**77**	**89**	**93**	**100**	**93**
HSS 130	PCR1	88	100	63	87	94	97	
	PCR2	87	100	65	86	95	98	
	PCR3	87	100	65	85	93	97	
	Mean Methylation (%)	**87**	**100**	**64**	**86**	**94**	**97**	**89**
HSS 133	PCR1	93	100	82	88	93	99	
	PCR2	94	100	80	90	92	100	
	PCR3	94	100	83	90	93	100	
	Mean Methylation (%)	**94**	**100**	**82**	**89**	**93**	**100**	**94**
HSS 132	PCR1	90	100	86	88	94	97	
	PCR2	90	100	89	88	93	99	
	PCR3	88	100	88	87	95	100	
	Mean Methylation (%)	**89**	**100**	**88**	**88**	**94**	**99**	**94**
HSS 134	PCR1	89	100	65	85	90	96	
	PCR2	89	100	65	90	92	97	
	PCR3	90	100	66	86	90	98	
	Mean Methylation (%)	**89**	**100**	**65**	**87**	**91**	**97**	**88**
HSS 135	PCR1	86	100	62	83	88	98	
	PCR2	86	100	62	86	90	98	
	PCR3	87	100	62	84	90	99	
	Mean Methylation (%)	**86**	**100**	**62**	**84**	**89**	**98**	**87**
HSS 137	PCR1	93	100	90	5	89	97	
	PCR2	93	100	89	6	88	96	
	PCR3	92	100	89	5	87	98	
	Mean Methylation (%)	**93**	**100**	**89**	**5**	**88**	**97**	**93**
HSS 140	PCR1	94	100	88	6	88	97	
	PCR2	94	100	88	6	87	94	
	PCR3	94	100	89	5	88	93	
	Mean Methylation (%)	**94**	**100**	**88**	**6**	**88**	**95**	**93**
HSS 143	PCR1	93	98	82	40	89	97	
	PCR2	94	100	81	39	88	95	
	PCR3	93	100	82	41	87	100	
	Mean Methylation (%)	**93**	**99**	**82**	**40**	**88**	**97**	**92**

HSS 144	PCR1	91	100	81	87	89	95	
	PCR2	87	100	84	86	87	92	
	PCR3	92	100	80	86	89	97	
	Mean Methylation (%)	**90**	**100**	**82**	**86**	**88**	**95**	**91**
HSS 145	PCR1	91	100	80	42	87	95	
	PCR2	90	100	80	44	88	94	
	PCR3	91	98	81	43	88	98	
	Mean Methylation (%)	**91**	**99**	**80**	**43**	**88**	**96**	**91**
HSS 146	PCR1	90	100	63	89	94	99	
	PCR2	89	100	64	90	93	96	
	PCR3	87	100	62	88	94	97	
	Mean Methylation (%)	**89**	**100**	**63**	**89**	**94**	**97**	**89**
HSS 147	PCR1	93	100	77	88	90	97	
	PCR2	92	100	80	88	90	99	
	PCR3	93	100	78	89	91	100	
	Mean Methylation (%)	**93**	**100**	**78**	**88**	**90**	**99**	**92**
HSS 148	PCR1	95	100	90	47	100	100	
	PCR2	95	100	89	48	100	100	
	PCR3	95	100	89	46	100	100	
	Mean Methylation (%)	**95**	**100**	**89**	**47**	**100**	**100**	**97**