

https://dspace.org/

Epigenetic inheritance of aberrant DNA methylation signatures as a consequence of chronic paternal alcohol exposure and the effect on embryonic gene expression in mice

Ismail, Ayesha

2015
http://hdl.handle.net/10539/21283

Downloaded from a DSpace repository. https://dspace.org/

University of the Witwatersrand

National Health Laboratory Service
School of Pathology
Division of Human Genetics

2015 MSc (Med) Dissertation







Epigenetic inheritance of aberrant DNA methylation signatures as a consequence of chronic paternal alcohol exposure and the effect on embryonic gene expression in mice

Supervisors: Prof M Ramsay, Dr J Knezovich, and Dr Z Lombard
Collaborator: Prof Anne Ferguson Smith (University of Cambridge, UK)

Ayesha Ismail
458260

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, Ayesha Ismail, hereby declare this dissertation to be my own, unaided work. It is being submitted for the degree of Master of Science (Medicine) in Human Genetics to the Faculty of Health Sciences of the University of the Witwatersrand. This work has not been submitted for any other degrees at any other university.

Ayesha Ismail

Date

“Science is not only a disciple of reason but, also, one of romance and passion”

-Stephen Hawking

This dissertation is dedicated to my parents, Suraya and Yousuf Ismail

Presentations arising from this study

1. **South African Societiy of Human Genetics Young Reserachers Forum (Pretoria, 15th August 2015).** Presented a poster entitled " The epigenetic effect on mouse offspring following preconception paternal alcohol exposure".

2. **16th South African Society of Human Genetics Biennial Congress 2015 (Pretoria, 16th-19th August 2015).** Presented a poster entitled " The epigenetic effect on mouse offspring following preconception paternal alcohol exposure".

3. **University of the Witwatersrand, 7th Cross Faculty Graduate Symposium (Johannesburg, March 2016).** Presented a poster entitled " The epigenetic effect on mouse offspring following preconception paternal alcohol exposure". Abstract submmitted for poster presentation.

Abstract

Epigenetic mechanisms regulate gene expression, a particularly important activity during foetal development. DNA methylation contained within promoter and regulatory intergenic regions influence gene activity. *In utero* alcohol exposure as a result of maternal consumption during pregnancy has been associated with disruption of foetal DNA methylation and gene expression, leading to neurological dysfunction, growth retardation and facial anomalies. While similar phenotypes in offspring have been associated with chronic preconception paternal alcohol exposure, the mechanisms underlying these effects remain largely unexplored.

This study aimed to: (1) validate significant changes in sperm DNA methylation in a list of ten candidate genes in male mice chronically exposed for ten weeks to ethanol (n=10) compared to a calorie-equivalent sucrose solution (n=10); (2) validate significant changes in gene expression in candidate genes in the brain, liver and placenta of E16.5 embryos sired by ethanol (n=24) compared to sucrose (n=24) treated male mice; (3) quantify DNA methylation changes in candidate genes in the three embryonic tissues. (4) Lastly, previously generated microarray data were reanalysed using bioinformatics tools to generate a top ranked candidate differentially expressed gene list that was used to identify and analyse biological functions or pathways significantly over represented among these genes using PANTHER and DAVID.

This study was unable to provide validation for most of the significant differences observed in the sperm DNA methylome in the original study, most likely because of the low sperm DNA concentration. Significant methylation differences were however observed at individual CpG sites in three candidate genes (*Igf1r*, *Odc1*, *Depdc1b*) in specific tissues of embryos sired by ethanol-exposed males relative to embryos sired by sucrose-treated males. There was concordance in the direction of altered gene expression between the cases and controls using the microarray and real-time PCR approaches for two genes in the brain (*Grm7* and *Zfp317*), three genes in the liver (*Igf1r*, *Vwf* and *Depdc1b*) and one gene in the placenta

(*Vwf*). However, none of the candidate genes selected for validation showed statistically significant changes. This may be a result of the modest fold changes observed in the microarray experiment that as shown in many cases, often do not replicate. The remainder of the genes showed no changes in expression in the test embryos relative to the control. The functional enrichment analysis revealed biological processes that were over represented in the brain and liver indicating that they may be more vulnerable to the effects of alcohol, compared to the placenta.

Overall, the study could not provide a statistically significant correlation between methylation changes in the sperm that were inherited by the offspring which subsequently dysregulated gene expression in the embryo. However, as trends toward significance and significant DNA methylation changes were observed in the embryonic tissues, this study supports the idea that preconception paternal alcohol exposure can induce epigenetic alterations in a locus and organ specific manner within offspring.

Acknowledgements

I would like to extend my utmost gratitude to the following people:

To my supervisor, Professor Michèle Ramsay, you have been a great source of knowledge. Thank for your patient guidance, advice and continuous encouragement.

To my supervisor, Dr Jaysen Knezovich, thank you for your invaluable assistance, support, guidance and enthusiasm over the last three years. I could not have asked for a better mentor. You always went beyond what was expected of you, and I will always be indebted to you.

To Dr Zané Lombard, my co-supervisor, thank you for your guidance, motivation and support over the last two years. You gave your time freely and I will always be grateful.

To Professor Anne Ferguson-Smith, Dr Mitsutero Ito, Dr Hui Shi and Ms Jennifer Corish at the University of Cambridge, UK. Thank you for hosting me at your research lab and for your invaluable assistance in both the laboratory and analyses of the data. This has proven instrumental in broadening my understanding of epigenetics and my growth as a scientist.

To my friends at the AFS lab, thank you for welcoming me so warmly into your research group through my 3-month stay. Thank you for the encouragement and support you provided, I will treasure the memories.

To my friend Richard Munthali, your assistance with the bioinformatics component of this study is appreciated.

To Michelle Ungerer, thanks for keeping me motivated, for always being willing to help, and for providing continuous support when the going got tough.

To Liesl, Venesa, Andrew, Thandiswa and Richard, thank you for your friendship, for keeping the mood light and for your encouragement, support and advice.

Special thanks must go to my best friends Fatima, Aboo and Shaahid for being a constant source of laughter, love and motivation.

I would like to thank the National Research Foundation and the Anne Ferguson-Smith Lab (University of Cambridge, UK) for funding this project.

Finally, I would like to express a deep sense of gratitude to my family. My parents for giving me the gift of education and allowing me to further my studies with their constant support.

My elder brother for the excellent example he sets. Thank you all for the love and strength you continuously provide.

Table of Contents

Declaration ii

Presentations arising from this study v

Abstract vi

Acknowledgements viii

Table of Figures xiii

List of Tables xv

List of Abbreviations xvi

1.0 Introduction 2

1.1 Alcohol use and abuse 3

1.2 Paternal alcohol consumption alters offspring phenotype 5

1.3 Epigenetic regulatory mechanisms 6

1.3.1 DNA methylation as a regulator of gene expression 7

1.3.2 DNA demethylation in regulating gene expression 8

1.3.3 Histone modifications and miRNAs in regulating gene expression 9

1.4 Epigenetic reprogramming 10

1.4.1 Epigenetic reprogramming in the germ cells 11

1.4.2 Epigenetic reprogramming in the preimplantation embryo 11

1.5 Epigenetic mechanisms are senstive to environmental changes 13

1.5.1 Exposures can affect DNA methylation of adult somatic tissue 14

1.5.2 Environmental exposures can alter DNA methylation in embryonic tissue (*in utero)* 15

1.5.3 Environmental exposures can alter the epigenetic marks in the germline 17

1.6 Changes in DNA methylation as a consequence of preconception paternal environmental exposures impact on embryonic gene expression 20

1.7 Background to current study 21

1.7.1 Mouse Model 21

1.7.2 Sperm DNA extraction and DNA methylation quantification 22

1.7.3 Allocation of significantly differentially methylated CpG sites 23

1.7.4 Embryo harvesting and cDNA synthesis 23

1.7.5 Expression array 23

1.7.6 Merging methylome and gene expression datasets to obtain a list of top candidates for pyrosequencing and real-time PCR validation 24

1.8 Study Rationale, Aim and Objectives 26

1.8.1 Rationale 26

1.8.2 Aim 26

1.8.3 Study Objectives 26

2.0 Materials and Methods 29

2.1 Present study design 30

2.2 DNA sample selection and quantification 32

2.3 Bisulphite modification 33

2.4 Quantitative DNA methylation analysis by pyrosequencing 35

2.4.2 Pyrosequencing assay validation 39

2.4.3 Pyrosequencing PCR 42

2.4.4 Pyrosequencing preparation and run 43
2.4.5 Pyrosequencing data analysis 45
2.4.5.1 Data quality control 47
2.5 RNA sample selection 48
2.6 RNA quantification 48
2.7 Real-time PCR Primer design 48
2.8 Reverse Transcription of RNA into cDNA 50
2.9 Selecting an endogenous control 51
2.10 Relative gene expression analysis using quantitative real-time PCR 51
2.10.1 SYBR green based method of detection 52
2.10.2 Melting curve analysis to confirm primer specificity 53
2.10.3 PCR efficiency analysis using serial dilutions 54
2.10.4 Real Time PCR preparation and run using the Roche Light Cycler 480 55
2.10.5 Standard curve method of relative quantification 57
2.10.6 Real-time PCR data quality control 59
2.11 Statistical analysis of pyrosequencing and real-time PCR data 59
2.11.1 Descriptive statistical analysis 59
2.11.2 Mann-Whitney U test 60
2.12 Functional enrichment analysis of genes that showed a significant change in embryonic gene expression in the brain, liver and placenta 60
3.0 Results 64
3.1 The effect of preconception EtOH exposure on DNA methylation in sperm of male mice 66
3.1.1 Validation of RRBS sperm DNA mean methylation data using pyrosequencing 67
3.1.2 Validation of RRBS sperm DNA methylation data at specific CpG sites within candidate genes 69
3.2 The effect of preconception EtOH exposure on gene expression in embryonic tissues 76
3.2.1 Validation of whole genome expression array data from embryonic tissues using quantitative real-time PCR 77
Comparison of microarray and real-time PCR fold change values 84
3.2.2 Sex specific differential expression of candidate genes between treatment groups. 85
3.3 The effect of preconception EtOH exposure on methylation in embryonic tissues 86
3.3.1 Gene averaged methylation analyses of embryonic DNA 87
3.3.2 CpG specific methylation analyses of embryonic DNA 90
3.4 Bioinformatics analysis to assess differential gene expression between treatment groups in three embryonic tissues 96
3.4.1 Volcano plots to identify significantly differentially expressed genes 96
3.4.2 Functional classification analysis using PANTHER 100
3.4.3 Functional enrichment analysis using DAVID 104
4.0 Discussion 109
4.1 Effect of alcohol exposure on DNA methylation in the sperm of male mice 111
4.2 Alterations in embryonic DNA methylation following preconception paternal alcohol exposure 116
4.3 Effects of chronic preconception paternal alcohol intake on candidate gene expression in embryos 121

4.3.1 Altered *Grm7* and *Zfp317* expression in the brain of embryos sired by ethanol-treated male mice 122
4.3.2 Upregulation of *Igf1r* and *Depdc1b* in the liver of embryos sired by ethanol treated males 126
4.3.3 Altered expression of *Vwf* in the liver and placenta of embryos sired by ethanol treated males 129
4.4 Do DNA methylation changes in the sperm align with methylation and expression alterations in the embryos? 130
4.5 Challenges when validating gene expression microarray data with real-time PCR 133
4.6 Bioinformatics approach reveals molecular functions and biological processes significantly enriched in a list of differentially expressed genes in embryonic tissues 134
4.6.1 Molecular functions and biological processes identified by gene ontology analysis of differential gene expression in embryo brains 136
4.6.2 Molecular functions and biological processes identified by functional enrichment analysis of differential gene expression in embryonic liver 142
4.7 Limitations and future studies 144
4.7.1 Future model for the transfer of epigenetic alterations as a consequence of chronic preconception paternal alcohol intake 145
4.8 Conclusion 150
5.0 References 152
6.0 Appendices 164
6.1 Appendix A: Reagents used in this study 164
6.2 Appendix B: Manufacturer/Supplier information regarding the equipment and products as well as the developers of the software used 166
6.3 Appendix C: DNA Bisulphite Modification Protocol 167
6.4 Appendix D: Recommended Protocol for Methylation of Genomic DNA 169
6.5 Appendix E: Pyrosequencing preparation protocol 170
6.6 Appendix F: Validation by dilution series 171
6.7 Appendix G: Real- Time PCR Primer efficiency analysis graphs 174
6.8 Appendix H: Thermo Scientific RevertAid H Minus First Strand cDNA synthesis Kit 181
6.9 Appendix I: R script used to generate the volcano plots 182
6.10 Appendix J: Cambridge Personal Project License 183
6.11 Appendix K: Wits Ethics clearance 184
6.12 Appendix L: Sperm DNA Pyrosequencing data 186
6.13 Appendix M: Embryonic DNA pyrosequencing data 187
6.14 Appendix N: Embryonic brain real-time PCR data 188
6.15 Appendix O: Placenta real-time PCR data 189
6.16 Appendix P: Embryonic liver real-time PCR data 190
6.17 Appendix Q: Statistical analysis of sperm pyrosequencing data 191
6.18 Appendix R: Statistical analysis of embryonic pyrosequencing data 192
6.19 Appendix S: Statistical analysis of embryonic real-time PCR data 193
6.20 Appendix T: Turnitin plagiarism detection certificate 194

Table of Figures

Figure 1.1: Nucleosome structure. 7

Figure 1.2: Diagram showing the active demethylation of 5-methylcytosine. 9

Figure 1.3: There are two phases in the in the life cycle of a mouse when epigenetic reprogramming occurs. 12

Figure 1.4: Intergenerational and. transgenerational epigenetic inheritance. 17

Figure 1.5: Flow diagram of the methodology on which the present study was designed. ... 22

Figure 1.6: Allocation of differentially methylated CpG sites to the gene closest proximity .. 23

Figure 2.1: Flow Diagram of the Methodology. 31

Figure 2.2: Flow diagram of the DNA samples included in the present study.. 32

Figure 2.3: Schematic of the bisulphite conversion for the analysis of DNA methylation.. ... 34

Figure 2.4: The principle of pyrosequencing and the output pyrogram. 36

Figure 2.5: *Vwf* original genomic sequence and its bisulphite modified sequence. 38

Figure 2.6: Pyromark MD Cartridge. 45

Figure 2.7: Results of a bisulphite pyrosequencing assay to determine the DNA methylation status within a region of interest. 46

Figure 2.8: Schematic of the SYBR green based real time assay. 53

Figure 2.9: Melting curve analysis confirming specific amplification with *Depdc1b* primers.. 54

Figure 2.10: The standard curve used to analyse the efficiency of *Actb* 55

Figure 3.1: Validation of RRBS mean methylation data for five candidate genes using pyrosequencing. 68

Figure 3.2: Validation of *Grm7* RRBS CpG site-specific sperm DNA methylation data of male mice exposed to EtOH or sucrose using pyrosequencing.. 71

Figure 3.3: Validation of *Tet1* RRBS CpG site-specific sperm DNA methylation data of male mice exposed to EtOH or sucrose using pyrosequencing. 72

Figure 3.4: Validation of *Vwf* RRBS CpG site-specific sperm DNA methylation data of male mice exposed to EtOH or sucrose using pyrosequencing.. 73

Figure 3.5: Validation of *Odc1* RRBS CpG site-specific sperm DNA methylation data of male mice exposed to EtOH or sucrose using pyrosequencing. 74

Figure 3.6: Validation of *Depdc1b* RRBS CpG site-specific sperm DNA methylation data of male mice exposed to EtOH or sucrose using pyrosequencing. 75

Figure 3.7: Real-time PCR validation of embryonic brain microarray results. 79

Figure 3.8: Real-time PCR validation of embryonic liver microarray results. 81

Figure 3.9: Real-time PCR validation of placental microarray results.. 83

Figure 3.10: Differential *Grm7* expression in male and female embryos sired by male mice exposed to EtOH or sucrose.. 85

Figure 3.11: Column scatter plots of mean methylation levels of the candidates gene in different embryonic tissues. 90

Figure 3.12: Column scatter plots of DNA methylation at individual CpG sites of the embryonic brain.. 93

Figure 3.13: Column scatter plots of DNA methylation at individual CpG sites of the embryonic liver. 94

Figure 3.14: Column scatter plots of DNA methylation at individual CpG sites of the placenta. 95

Figure 3.15: Volcano plots of significance against fold change of embryonic brain (A), liver (B) and placental (C) gene expression.. 97

Figure 3.16: Pie charts of the molecular function of genes significantly ($p<0.001$ & $FC>1.2$) downregulated and upregulated in the brain and liver of embryos sired by males treated with EtOH.. 101

Figure 3.17: Pie charts of the biological processes of genes significantly ($p<0.001$ & $FC>1.2$) downregulated and upregulated in the brain and liver of embryos sired by males treated with EtOH. 103

Figure 4.1: The effect of bisulfite treatment and DNA concentration on the quantification of DNA methylation, as determined by pyrosequencing..... 113

Figure 4.2: A simplistic future model to investigate the effect of preconception paternal alcohol exposure on epigenetic mechanisms in mouse offspring. 148

List of Tables

Table 1.1: The ten candidate genes analysed in the present study 25

Table 2.1: Pyrosequencing primer properties and the CpG sites analysed 39

Table 2.2: Pyrosequencing PCR assay validation and pre-PCR reference mixes 40

Table 2.3: Reagents and conditions used for the PCR amplification of the candidate genes before pyrosequencing 43

Table 2.4: The probe ID and NM acession number used to identify the microarray probe sequence 49

Table 2.5: The microarray probe sequence used to identify the region of interest to be amplified by real-time PCR 50

Table 2.6: The primer sequences and amplicon length of the candidate genes that underwent real-time PCR analysis 50

Table 2.7: The general real-time PCR reagents and thermal cycling conditions for the amplification of the candidate genes 57

Table 3.1: Mean methylation levels across the CpG sites of the candiate genes between the treatment groups 67

Table 3.2: CpG site methylation analysis within the candidate genes between the treatment groups 70

Table 3.3: Relative gene expression analyis of candidate genes in the tissues of embryos sired by male mice exposed to EtOH or sucrose 78

Table 3.4: A comparison of the expression array and the real-time PCR data 84

Table 3.5: Mean methylation data of the candidate genes in the tissues of embryos sired by males that were exposed to EtOH or sucrose 88

Table 3.6: DNA methylation data of individual CpG sites in the candidate genes of embryos sired by male mice exposed to EtOH or sucrose 91

Table 3.7: Genes that showed a significant difference ($p<0.001$) in expression with a fold change of greater than 1.2 in the brain or liver of embryos sired by EtOH-treated males as per the microarray data 99

Table 3.8: Biological functions or pathways overrepresented among the top ranked candidate genes expressed in the brain. 106

Table 3.9: Biological functions or pathways overrepresented among the top ranked candidate genes expressed in the liver. 107

List of Abbreviations

%	Percentage
°C	Degrees Celsius
18S rRNA	*18S Ribosomal RNA*
5-caC	5-carboxylcytosine
5fC	5-formlycytosine
5hmC	5-hydroxymethylcytosine
5mC	5-methylcytosine
A	Adenine
Actb	*Beta Actin*
Aebp2	*Adipocyte Enhance Binding Protein 2*
ARBD	Alcohol Related Brain Damage
ARND	Alcohol Related Neurodevelopmental disorder
As	Arsenic
ATP	Adenosine Triphosphate
A^{vy}	Viable yellow agouti locus
Bdnf	*Brain-derived neurotrophic factor*
Bp	Base pair
C	Cytosine
CCD	Charged Coupled Device
Ccnd3	*Cyclin D3*
cDNA	Complementary Deoxyribose nucleic acid
CGIs	CpG Islands
CNS	Central Nervous System
CNV	Copy Number Variation
CpG	CpG site
C_T	Cycle Threshold
CTCF	CCCTC-binding factor
dATP	Deoxyadenosine Triphosphate
dATP S	Deoxyadenosine Alpha-Thio Triphosphate
DAVID	Database for Annotation Visualization and Integrated Discovery
ddH_2O	Double distilled water
DE	Differentially Expressed
Depdc1b	*DEP Domain Containing 1B*
dHand	*Heart and Neural Crest Derivatives Expressed*
DMR	Differentially methylated region
DNA	Deoxyribose nucleic acid

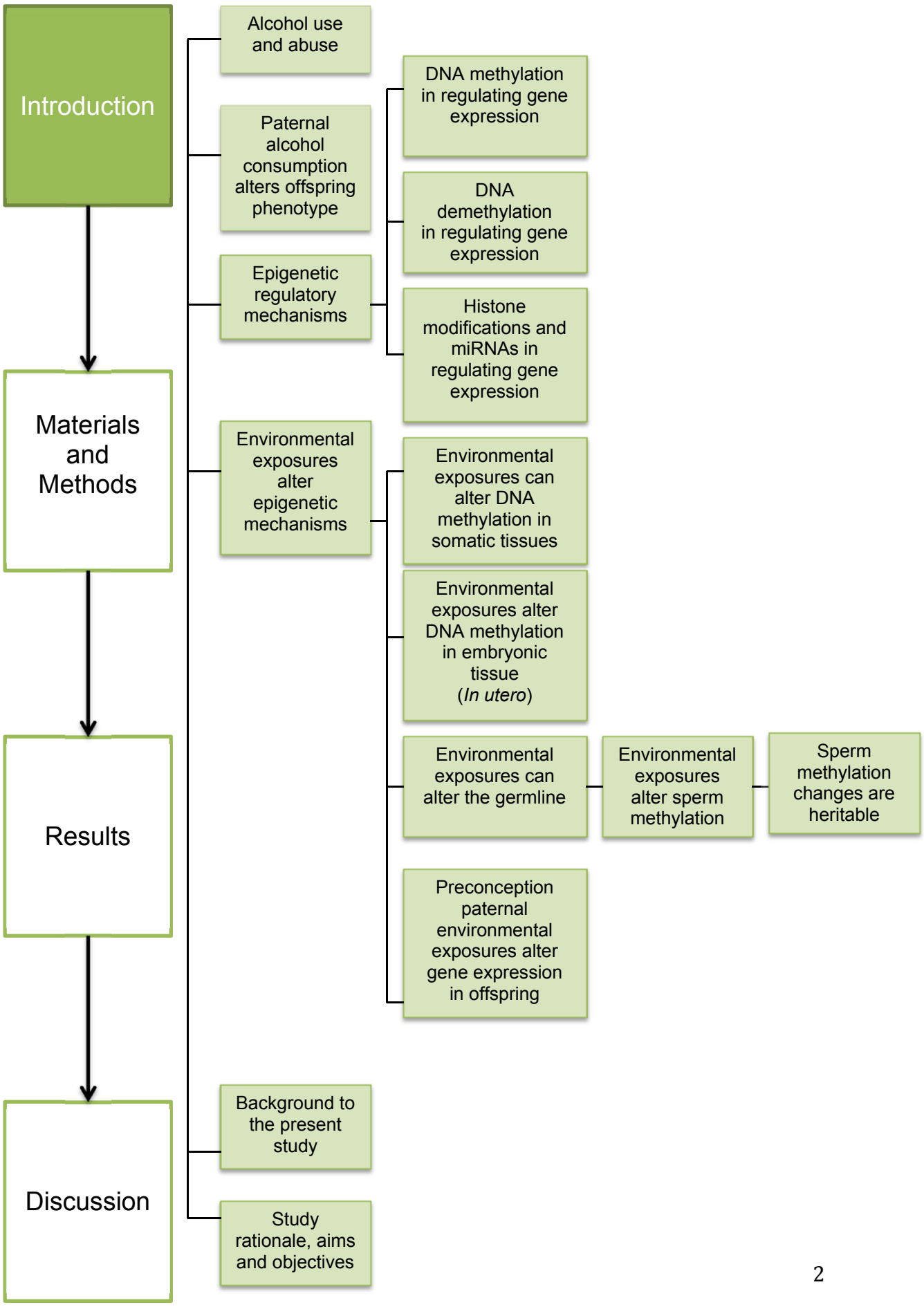
DNMT	DNA Methyl Transferase
dNTP	Deoxyribonucleic acid triphosphate
E16.5	Embryonic day 16.5
EtOH	Ethanol
F_0	Generation 0
F_1	Generation 1
F_2	Generation 2
F_3	Generation 3
FAE	Foetal Alcohol Effects
FAS	Foetal Alcohol Syndrome
FASD	Foetal Alcohol Spectrum Disorders
FC	Fold Change
FRET	Förster Resonance Energy Transfer
g	Gram
G	Guanine
Gapdh	*Glyceraldehyde 3-phosphate dehydrogenase*
Gli3	*GLI family Zinc Finger 3*
GO	Gene Ontology
Grm7	*Glutamate Receptor Metabotropic*
H19	*H19,* imprinted maternally expressed transcript
HERP	*Homocysteine-induced Endoplasmic Reticulum Protein*
IAP	Intracisternal A-Particle
IUPAC	International Union of Pure and Applied Chemistry
ICRs	Imprinting Control Regions
Igf1	*Insulin like growth factor 1*
Igf1r	*Insulin like growth factor receptor 1*
Igf2	*Insulin like growth factor 2*
IUE	*In utero* Electroporation
KEGG	Kyoto Encyclopedia of Genes and Genomes
Kg	Kilogram
MeDIP Seq	Methylated DNA Immunoprecipitation Sequencing
MEEBO	Mouse Exonic Evidence Based Oligonucleotide
mg	Milligram
min	Minute
miRNAs	micro ribonucleic acid
ml	Millilitre
mRNA	messenger RNA
mtDNA	mitochondrial DNA

NCBI Refseq	National Centre for Biotechnology Reference Sequence
ncRNAs	non-coding RNAs
ng	Nanogram
Nnat	*Neuronatin*
Odc1	*Ornithine Decarboxylase*
OMIM	Online Mendelian Inheritance in Man
PANTHER	Protein Annotation Through Evolutionary Relationships
Pcbp3	Poly (rC)-binding protein 3
PCR	Polymerase Chain Reaction
Peg3	*Paternally expressed 3*
pFAS	partial foetal alcohol syndrome
PFAS	Phosphoribosylformylglycinamidine synthase
PGC	Primordial Germ Cells
PPARα	Peroxisome proliferator-activated receptor
PPi	Pyrophosphate
PRC2	*PolyComb Repressive Complex 2*
Psmg2	*Proteasome Assembly Chaperone 2*
Pyro	Pyrosequencing
R	R programming language
Rasgrf1	*Ras Protein-Specific Guanine Nucleotide-Releasing Factor 1*
RISC	RNA-induced silencing complex
RNA	Ribonucleic acid
ROS	Reactive Oxidative Species
RRBS	Reduced Representation Bisulphite Sequencing
S	Sucrose
SAM	S-Adenosyl methionine
Sec	Seconds
Shh	*Sonic Hedgehog*
SNP	Single Nucleotide Polymorphism
SNRPN	*Small Nuclear Ribonucleoprotein Polypeptide*
T	Thymine
Tet	Ten-Eleven Translocation
Tet1	*Tet methylcytosine dioxygenase 1*
U	Uracil
Vwf	*Von Willebrand Factor*
Zfp317	*Zinc Finger protein 317*
µl	Microliter

Chapter 1

Introduction

1.0 Introduction



Introduction

It is evident that genetic factors are not solely responsible for the heritable processes that shape development and cause disease. Epigenetic mechanisms including DNA methylation, histone modifications and RNA interference regulate gene expression in a spatial and temporal manner. These mechanisms are particularly important during foetal development. Environmental exposures can influence epigenetic mechanisms, which can impact on gene expression and phenotype. *In utero* alcohol exposure as a result of maternal consumption during pregnancy has been associated with disruption of foetal DNA methylation, gene expression, and the manifestation of growth retardation phenotypes. While similar phenotypes in offspring have been associated with chronic preconception paternal alcohol exposure, the mechanisms underlying these effects remain largely unknown. This study aimed to validate the effect of chronic alcohol exposure on sperm DNA methylation, and investigated whether possible aberrations were transmitted to sired offspring and whether a change in embryonic methylation and gene expression was observed.

1.1 Alcohol use and abuse

Alcohol (ethanol) is an intoxicating ingredient produced by the fermentation of yeast, sugars and starches that can be found in beer, wine and liquor. It is a psychoactive substance with dependency producing properties (National Institute on Drug Abuse, 2015). Alcoholic beverages are widely consumed throughout the world, however, drinking behaviour differs considerably between countries. Alcohol-related harm is determined by environmental factors and genetics, drinking patterns, the volume of alcohol consumed and in some instances the quality of the alcohol (World Health Organisation 2011).

Problems associated with the consumption of alcoholic beverages have been recognised for centuries. The ancient Greeks demonstrated insight into the possible risks associated

with alcohol consumption as Plato proposed that wine consumption in people under 40 years should be limited (as reviewed by Abel, 1984). In addition to the adverse effects of alcohol on health, societies have recognised the harmful effects of alcohol on the developing foetus. Societies such as Carthage and Sparta prohibited the use of alcohol by newlyweds (as reviewed by Friedler, 1996). Biblical and Talmudic references suggest that the ancient Hebrews might have recognised that male and female preconception alcohol use could have adverse effects on their offspring (Abel, 1984). The *Anatomy of Melancholy* published in 1621, stated that babies born to drunken fathers had impaired brains and the gin epidemic in the early eighteenth century in England evoked concern about the consequences of alcohol abuse by either parent (as reviewed by Friedler, 1996).

By the mid-1900s, reports on the harmful effects of alcohol on offspring development began to peak, with a large number of publications describing the characteristics of children born to alcoholics (Lemoine et al., 1968). Descriptions of Foetal Alcohol Syndrome (FAS) - the most severe manifestation of the Foetal Alcohol Spectrum Disorders (FASD), first appeared in medical literature in 1973 (May and Gossage, 2011). Following this, research into FAS began to emerge as the research community devoted its attention toward understanding FASD. FASD is an umbrella term used for a spectrum of conditions caused by maternal alcohol abuse during pregnancy, which generally involve structural anomalies, behavioral defects and neurocognitive disabilities. (Hoyme et al., 2005). The severity of these phenotypes vary, ranging from those that are that are moderately observable (attention difficulties, reasoning and memory deficits) to spontaneous abortion (Clarke and Gibbard, 2003). Four diagnoses are included in the FASD spectrum: FAS, partial Foetal Alcohol Syndrome (pFAS), Foetal Alcohol Effects (FAE), Alcohol-Related Neurodevelopment Disorder (ARND) and Alcohol-Related Birth Defects (ARBD - the least severe of the disorders). Although the primary impact of FAS is on the brain, it is also associated with pre- and postnatal growth retardation, damage to the heart and kidney, impaired cognitive

ability, reduced stress tolerance and behavioral abnormalities affecting memory (Abel, 2004, Gearing et al., 2005).

Although FASD is primarily attributed to maternal alcohol exposure during pregnancy, human and animal studies have suggested that paternal alcohol consumption may contribute to developmental abnormalities in offspring that resemble a FASD phenotype.

1.2 Paternal alcohol consumption alters offspring phenotype

Various studies have been directed toward understanding the paternal effects of preconception alcohol exposure. As early as 1984, it had been demonstrated that children born to alcoholic fathers, whose mothers did not consume alcohol, exhibited hyperactivity and reduced cognitive performance (Hegedus et al., 1984). However, preconception paternal alcohol exposure has primarily been investigated in rat and mouse models. These studies observed reduced birth weight, compromised immunity, neurobehavioral abnormalities behavioral deficits such as anxiety and aggressive like behavior, reduced litter size, developmental retardation, increase susceptibility to infections and increased mortality in offspring sired by alcohol-exposed males (as reviewed by Curley and Mashoodh, 2010). Furthermore, it was noted that 75% of children with FAS had biological fathers who are heavy drinkers or alcoholics (Friedler, 1996, Jamerson, 2004, Abel, 2004). Unlike maternal alcohol exposure, paternal alcohol exposure does not involve the direct exposure of alcohol to the developing embryo. This evidence suggested the possibility that FAS cases that were previously thought to be the consequence of *in utero* alcohol exposure could be attributed to, or partially be a consequence of, excessive paternal preconception alcohol intake. At this level, mechanisms other than the DNA code itself could be affected by ethanol exposure.

This alludes to epigenetic mechanisms and this idea is supported by studies performed on both maternal and paternal mouse models that have shown that there is an epigenetic

contribution to a FAS-like phenotype (section 1.5.2) (Kaminen-Ahola et al., 2010, Haycock and Ramsay, 2009).

1.3 Epigenetic regulatory mechanisms

Almost all cells in a multicellular organism are genetically identical but functionally distinct, with the exception of cells that belong to the immune system and the germ cells. In order to understand how functional diversity is generated, it is necessary to understand that heritable differences in gene expression arise during the development of different cell types, which is regulated by mechanisms external to the DNA. Epigenetics is defined as mitotically heritable changes in gene expression that do not involve an alteration of the DNA sequence (Chuang and Jones, 2007).

The human genome is packaged into chromatin, a combination of genetic material and proteins, which package the genome into distinct conformations that delicately balance gene expression and repression (Davis and Brackmann, 2003). The fundamental building blocks of chromatin are the nucleosomes, which are composed of DNA wrapped around an octamer of histones and a separate linker histone, which aids further folding of the nucleoprotein complex as displayed in **Figure 1.1.** Histone proteins are rich in basic amino acids that carry a net positive charge, imparting a strong affinity for the negatively charged DNA phosphodiester backbone. Chromatin can be divided between two subtypes: partially unfolded active euchromatin and tightly coiled inactive heterochromatin. A sophisticated system called chromatin remodeling facilitates access of the DNA to regulatory machinery thereby controlling gene expression. This mechanism relies on histone variants, ATP-dependent chromatin remodeling complexes (SWI/SNF, ISWI), DNA methylation and specific enzymes that carry out covalent histone modifications to histone N- terminal tails (Wang et al., 2007). The two subtypes of chromatin have distinct classes of epigenetic modifications, which interact and regulate gene expression in a coordinated manner.

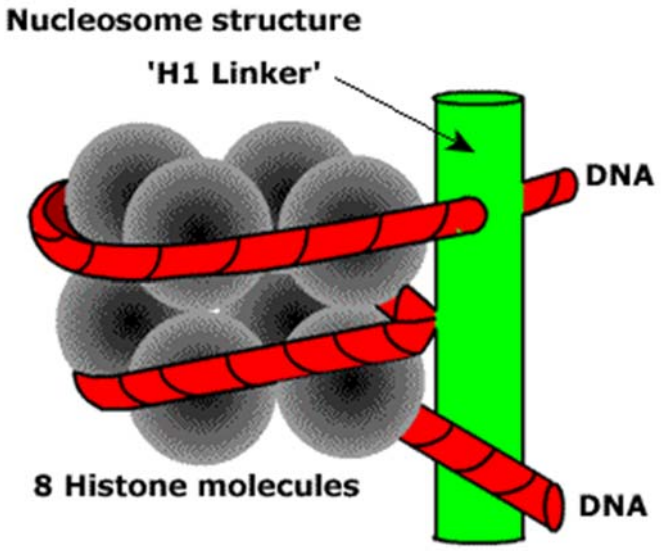


Figure 1.1: Nucleosome structure. Chromatin is comprised of histones and DNA: a 147bp chain of DNA wrapped around 8 core histone proteins forms the basic chromatic unit, the nucleosome (Sarkar, 2012).

1.3.1 DNA methylation as a regulator of gene expression

DNA methylation is arguably the best-characterised epigenetic mechanism, which involves the chemical modification of a cytosine that precedes guanine to yield 5-methylcytosine (5-mC dinucleotides). These dinucleotides are referred to as CpG sites. The cytosine base is modified by DNA methyltransferase enzymes (DNMT), which catalyse the addition of a methyl moiety from the methyl donor S-adenosyl methionine (SAM) to the fifth carbon of the cytosine residue. There are four members of the DNMT family, including DNMT1, DNMT3a, DNMT3b and DNMT3L. *DNMT1* encodes the maintenance methyltransferase, which methylate's hemi-methylated DNA during DNA replication. *DNMT3A/DNMT3B* encodes the *de novo* methyltransferases, required to establish genomic methylation. CpG sites are commonly clustered into CpG islands (CGIs), which are characterised by a CpG density of >50%, and a minimum of 500bp in length (Takai and Jones, 2002).

DNA methylation patterns vary from tissue to tissue, and even within tissues, thus conferring upon DNA a particular cellular identity. DNA methylation can alter the state of gene expression though several mechanisms. High levels of DNA methylation

(hypermethylation) in promoter regions of genes are generally associated with gene repression, whereas DNA methylation acts to condense the chromatin through recruitment of methylation binding factors that in turn attract chromatin-inactivation complexes including histone deacetylases and histone methyltransferase. Alternatively, DNA methylation prevents transcription factors from interacting with the underlying genome. Conversely, absence or low levels of DNA methylation (hypomethylation) are generally associated with gene expression. Almost all genes contain CGIs positioned in their respective promoter regions, whose state of methylation specifies silencing or expression of its associated gene (Liu et al., 2009). Cytosine methylation has a significant role in genomic imprinting and X-inactivation, with a fundamental contribution to the early stages of embryogenesis as it influences cell fate, pattern formation and terminal differentiation (Liu et al., 2009).

1.3.2 DNA demethylation in regulating gene expression

DNA methylation was originally thought to be irreversible however it is now known that DNA demethylation is an important process that involves the removal of the methyl group from the cytosine, facilitating epigenetic reprogramming (see section 1.4) and regulating gene transcription by the reactivation of silenced genes (Yamaguchi et al., 2013). Following the identification of the Ten-eleven translocation (Tet) family of enzymes, an understanding of the molecular mechanisms underlying the removal of the methyl group are beginning to unravel (Wu and Zhang, 2011). Demethylation can be achieved either actively, passively or a by combination of the two. Passive demethylation is mediated by DNMT1 and usually occurs on the newly synthesized DNA strands during replication rounds (Neidhart, 2015). Whereas active DNA demethylation primarily involves the removal of the 5-methylcytosine via the sequential modification of cytosine bases that have been converted by Tet enzyme-mediated oxidation (**Figure 1.2**) (Wu and Zhang, 2011). The Tet family of 5-mC hydroxylases consists of Tet1, Tet2 and Tet3. Together, the Tet proteins promote DNA demethylation by binding to CpG rich regions preventing the DNMT enzymes from binding, and by converting 5-mC to 5-hmC, 5-hmC to 5-fC (5-formylcytosine), and 5-fC to 5-caC (5-

carboxylcytosine) through hydroxylase activity (Wu and Zhang, 2011). While the role of 5hmC remains unknown, Ficz et al. (2011) showed that the presence of 5-hmC in the promoter regions of genes were associated with high levels of transcription, suggesting a role for short or long time regulation of gene expression. Overall, Tet proteins appear to be involved in transcriptional activation and repression, tumour suppression, and DNA methylation reprogramming processes (Wu and Zhang, 2011).

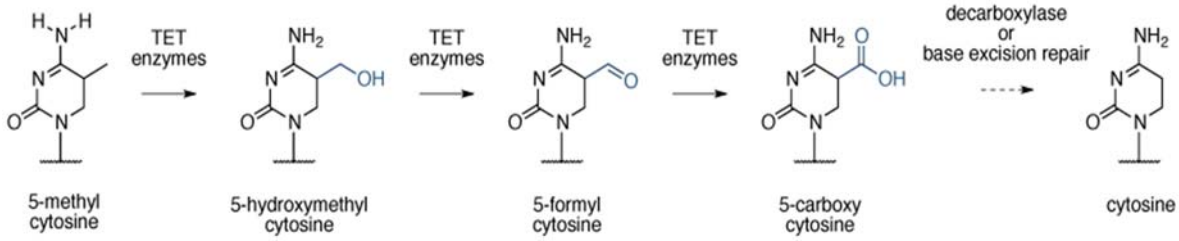


Figure 1.2: Diagram showing the active demethylation of 5-methylcytosine. 5-mC is actively oxidized by Tet proteins to 5-hmC, 5-fC, 5ac-C followed by decarboxylation/ base excision (Wu and Zhang, 2011).

DNA methylation is the focus of this study and by far the most extensively measured epigenetic mark because of its fundamental biological interest, its mitotic stability, the availability of methods to quantify global and targeted regions and its stability through the DNA extraction and purification processes. However, epigenetic modifications are not independent and it has become increasingly clear that there is an interconnection and interdependency between them. Therefore, histone modifications and microRNAs (miRNAs) are briefly discussed in the next section.

1.3.3 Histone modifications and miRNAs in regulating gene expression

Although it is not the primary focus of the present study, it is important to bear in mind that covalent modifications of histones and miRNAs are intimately involved in regulating gene expression.

Histones are subject to various enzyme-catalysed posttranslational modifications, which serve to allocate regions of the genome as active euchromatin or compact inactive regions of heterochromatin. Enzymes primarily modify the N-terminal tails of histones. Well-understood modifications include lysine acetylation, serine/threonine/tyrosine phosphorylation, lysine and arginine methylation, and serine/threonine ubiquitylation. Although these modifications were once thought to act independently, it has become increasingly apparent that there is cross talk between modifications on the same histone, between different histones, within the same nucleosome as well as different nucleosomes (abcam, 2013).

miRNAs are a class of small single stranded non-coding RNA molecules (20-24 nucleotides) that derive from regions of RNA transcripts that fold back on themselves to form hairpin structures. The hairpin is processed into mature miRNA and forms the RNA-induced silencing complex (RISC), which contains miRNA interacting proteins such as Dicer. The miRNA will pair with complementary mRNAs targets leading to gene silencing of the target (abcam, 2013)

Together, epigenetic mechanisms are thought to stabilise gene expression in specialised cell types in order to preserve cellular identity and lineage fidelity. However, there are two stages in the mammalian life cycle where epigenetic signatures are globally altered. These events are referred to as epigenetic reprogramming.

1.4 Epigenetic reprogramming

Global erasure and re-establishment of epigenetic marks occurs naturally at two stages in the mammalian cycle. These reprogramming events serve to erase epigenetic changes that may have occurred in the gametes and restore the ability of the fertilised egg (zygote) to develop into all different cell types and tissues (Reik et al., 2001, Santos and Dean, 2004).

1.4.1 Epigenetic reprogramming in the germ cells

The first reprogramming event occurs in the primordial germ cells (PGC) - the precursors of the mature gametes, where rapid genome-wide demethylation occurs to ensure erasure of parental methylation patterns followed by reestablishment of epigenetic marks in a sex specific manner. Germline epigenetic reprogramming occurs to reset the parental epigenetic marks and prevent the propagation of altered epigenetic marks (Santos and Dean, 2004, Reik et al., 2001).

1.4.2 Epigenetic reprogramming in the preimplantation embryo

The second reprogramming phase occurs shortly after fertilisation, in the preimplantation embryo. Reprogramming in the early embryo occurs in both an active and passive manner. Soon after fertilisation there is a remodelling of the sperm chromatin involving the removal of protamine's and their replacement with histones present in the oocyte, leading to the enrichment of the histone variant H3.3 (Reik et al., 2001). This is followed by active genome wide demethylation, which is completed before DNA replication commences. Thereafter, the maternal genome undergoes a step-wise decline in methylation until the morula stage. This decline occurs as a result of the absence of the primary maintenance DNA methyl transferase, DNMT1, during DNA methylation and is termed passive DNA methylation (Reik et al., 2001). Several sequences, including differentially methylated regions (DMRs) of imprinted genes, are protected from this demethylation. Embryonic DNA re-methylation is then established in a lineage-specific manner in the inner cell mass of the blastocyst (Reik et al., 2001).

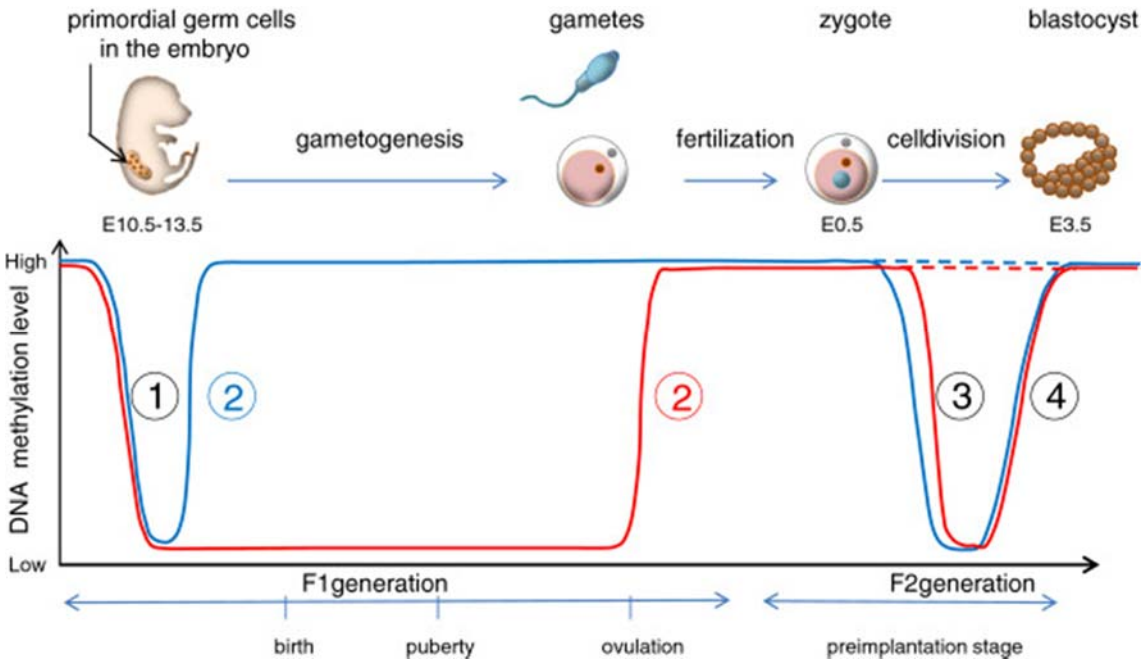


Figure 1.3: There are two phases in the in the life cycle of a mouse when epigenetic reprogramming occurs. (1) The first wave of DNA demethylation takes place in the male and female PGC of the F1 individuals; this takes place throughout the genome (including the imprinted genes) between embryonic days 10.5 to 13.5. This is followed by *de novo* methylation in the gametes. The maternal methylation marks are established at a later stage than the paternal marks. (2) The second wave of DNA demethylation occurs after fertilisation in the F2 zygote, with a more rapid demethylation in the paternal than the maternal genome. However, through preimplantation reprogramming (dotted lines) both maternal and paternal imprinted genes maintain their methylation, allowing for the inheritance of parent-specific monoallelic expression in somatic tissues of the F2 individual. Lastly, both parental genomes are remethylated at the time of implantation (Fresard et al., 2013).

Although these events serve to erase parental DNA methylation patterns and ensure cell lineage specific and sex specific placement of methylation marks, it is evident that specific genetic loci have the capacity to escape these waves of epigenetic reprogramming (Mashoodh and Champagne, 2014, Reik et al., 2001). Most notably, this has been observed in the case of retrotransposable elements and imprinted genes that appear to be both sensitive to environmental influences and have the potential to retain epigenetic marks (Lane et al., 2003).

1.5 Epigenetic mechanisms are senstive to environmental changes

During development, the effects of various signal transduction pathways on cellular function are controlled by growth factors, hormones and other signal molecules which influence the chromatin structure and consequently influences gene expression (Amey and Fisher, 2004, Cheung and Lau, 2005). A key aspect of cellular differentiation and functional variation among cells of the same individual, is the genome's ability to respond to developmental signals through hereditable alterations in chromatin structure (Amey and Fisher, 2004). The fact that epigenetic factors respond to developmental signals underlies their ability to respond to exogenous environmental stimuli. The majority of environmental factors and toxicants lack the ability to alter the DNA sequence or promote genetic mutations, but rather exert their effects through alterations of the epigenome (Skinner and Guerro-Bosanga, 2009). Furthermore, as previously stated, epigenetic marks are not always cleared through reprogramming events as certain regions have the potential to escape the reprogramming events. The modified epigenome may result in the enhancement of physiological fitness or alternatively lead to the development of complex diseases and syndromes including cancer, infertility, cardiovascular, respiratory, metabolic, and immunologic and neurodegenerative pathologies.

This study, like most other environmental epigenetic studies, chose to focus on evaluating the effect of one epigenetic mark. For this reason, the sections that follow will only provide evidence for the potential for environmental exposures to alter DNA methylation. DNA methylation can be altered in adult somatic tissue or in the tissues of offspring whose mothers were exposed to environmental toxins. Furthermore, recent studies provide evidence for an epigenetic impact of preconception maternal and paternal environmental exposures on mouse offspring. Various environmental exposures are mentioned in the sections that follow, particularly that of alcohol exposure.

1.5.1 Exposures can affect DNA methylation of adult somatic tissue

Human, animal and *in vitro* investigations have identified various environmental exposures that have the ability to modify the epigenome (as reviewed by Baccarelli and Bollati, 2009). Arsenic (As) exposure has been associated with decreased *DNMT* expression because the detoxification of As uses the universal methyl donor for methyltransferases. Evidence suggests that As exposure leads to a decrease in the methyl donor, which decreases DNMT activity due to the reduction of its substrate (Reichard et al., 2007). In addition, several *in vitro* studies have associated As exposure with global hypomethylation in a dose-dependent manner (Benbrahim-Tallaa et al., 2005, Coppin et al., 2008), including hepatic genes in rats and mice exposed to As for several weeks (as reviewed by Hou et al., 2012).

Exposure of other metals that consequently lead to alterations in epigenetic marks include: nickel, cadmium and chromium (as reviewed by Hou et al., 2012). Christensen et al. (2009) examined the association between alterations in methylation in the lung tissue of smoker's compared to non-smokers and found an alteration in the methylation status of 138 CpG loci in the lung tissue of smokers. Other environmental exposures that lead to the alteration of epigenetic mechanisms include peroxisome proliferators (trichloroethylene, dichloroacetic acid, trichloroacetic acid), air pollutants (particulate matter, black carbon, benzene), and endocrine-disrupting/reproductive toxicants (diethylstilbestrol, bisphenol A, persistent organic pollutants, dioxin) (as reviewed by Baccarelli and Bollati, 2009).

Ethanol has been shown to exert its effect on DNA methylation by disrupting the methionine-homocysteine cycle. Methionine is an essential amino acid and is a precursor of the compound, S-adenosyl methionine (SAM) (Kobor and Weinberg, 2011). Through the methionine-homocysteine cycle, methionine is converted to SAM, which in turn serves as the methyl donor in DNA methylation (Kobor and Weinberg, 2011). Alcohol affects the cycle by disturbing the enzymes necessary for methionine metabolism, therefore disrupting

the SAM dependent transmethylation reactions (Kobor and Weinberg, 2011). As a result, there is no longer a source of methyl groups for the DNA methyltransferases that use SAM as the methyl donor. Therefore, it was originally proposed that alcohol would lead to a decrease in methylation. This was supported by Choi et al. (1999) who showed that chronic exposure to ethanol leads to DNA hypomethylation throughout the genome in the colonic mucosa in rats, and that this may constitute a pathway by which carcinogenesis is enhanced. To the contrary, Bleich et al. (2006) observed that the *HERP* gene was downregulated as a consequence of hypermethylation in the promoter region in patients with alcohol dependence in comparison to non-drinker controls.

In addition to studies that have investigated the consequence of environmental exposures on DNA methylation patterns in adult somatic tissue, the effect of environmental exposures on epigenetic marks during *in utero* development has been explored.

1.5.2 Environmental exposures can alter DNA methylation in embryonic tissue (*in utero)*

Several studies have supported the association between *in utero* environmental influences on offspring DNA methylation. Holliday (1998) demonstrated that maternal use of teratogens has the ability to disrupt DNA methylation distribution in developing embryos. Garro et al. (1991) used a mouse model to evaluate the effects of *in utero* alcohol exposure from gestational day nine to eleven. Embryos exposed to alcohol demonstrated lower-than-normal global DNA methylation of foetal DNA, which was associated with a significant reduction in DNA methylase activity. A similar study by Kaminen-Ahola et al. (2010) investigated the effect of *in utero* ethanol exposure on the mouse *Agouti viable yellow* (A^{vy}) locus. A^{vy} is an epigenetically sensitive allele, as its expression is regulated by DNA methylation, which is sensitive to environmental exposures known to influence epigenetic mechanisms. Hypomethylation of the A^{vy} promoter element leads to constitutive expression of the Agouti gene, which results in a yellow coat colour. Conversely, hypermethylation is

correlated with promoter silencing and the generation of a brown pseudo-agouti coat colour. A mottled coat colour is the result of intermediate expression of A^{vy}. Prenatal exposures of both nutritional and toxic agents have been shown to alter DNA methylation at the A^{vy} locus, and dysregulate its expression in exposed offspring. Kaminen-Ahola et al. (2010) found that the proportion of pseudo-agouti offspring increased as a result of ethanol exposure. This was attributed to higher levels of transcriptional silencing which is consistent with a decreased expression of A^{vy} and promoter hypermethylation. This study highlighted the ability of prenatal alcohol exposure to alter the foetal epigenotype and, consequently the adult phenotype. In their latest study Marjonen et al. (2015) aimed to identify genes dysregulated in the hippocampi of offspring as consequence of early gestational ethanol exposure. Using a genome wide analysis approach they identified 23 genes and three miRNAs dysregulated in the ethanol exposed adolescent offspring. In addition site-specific changes in DNA methylation were observed in three CpG islands. These epigenetic changes were associated with altered brain structure in the ethanol-exposed offspring. In a study conducted by Liu et al. (2009), pregnant ethanol-treated female mice produced offspring with altered DNA methylation patterns that was associated with growth retardation and neurofacial deficits. Taken together, there is compelling evidence demonstrating the influence of environmental factors on DNA methylation and consequently dysregulation of gene expression, particularly during *in utero* development.

Subsequent to the discovery of the paramutation in maize by Brink and Coe in the 1950s (Brink, 1956, Coe, 1959), it has been well established that plants can inherit epigenetic alterations. Although it was long believed that segregation of the germ line in animal species prevented the transmission of epigenetic alterations, this view has been challenged in the past 20 years with a surge of studies providing examples of the ability of environmental factors to alter the germline epigenome and transmit these alterations to future generations.

1.5.3 Environmental exposures can alter the epigenetic marks in the germline

Rodent models provide a system for the identification of specific epigenetic changes that mediate an effect on the phenotype as well as the study of multiple generations in a relatively short time frame. Studying the transmission of these epigenetic alterations across multiple generations has led to a distinction between intergenerational and transgenerational effects in the literature (**Figure 1.4**). Intergenerational epigenetic inheritance refers to the passage of epigenetic changes from parent to offspring (F1 generation) (the premise of this study). Because developing germ cells in a foetus can also be affected by *in utero* exposures, intergenerational exposure may extend to into the F2 generation. Transgenerational inheritance refers to hereditability of an environmentally acquired phenotype into the F2 for paternal exposures or F3 generation for *in utero* exposures (Heard and Martienssen, 2014).

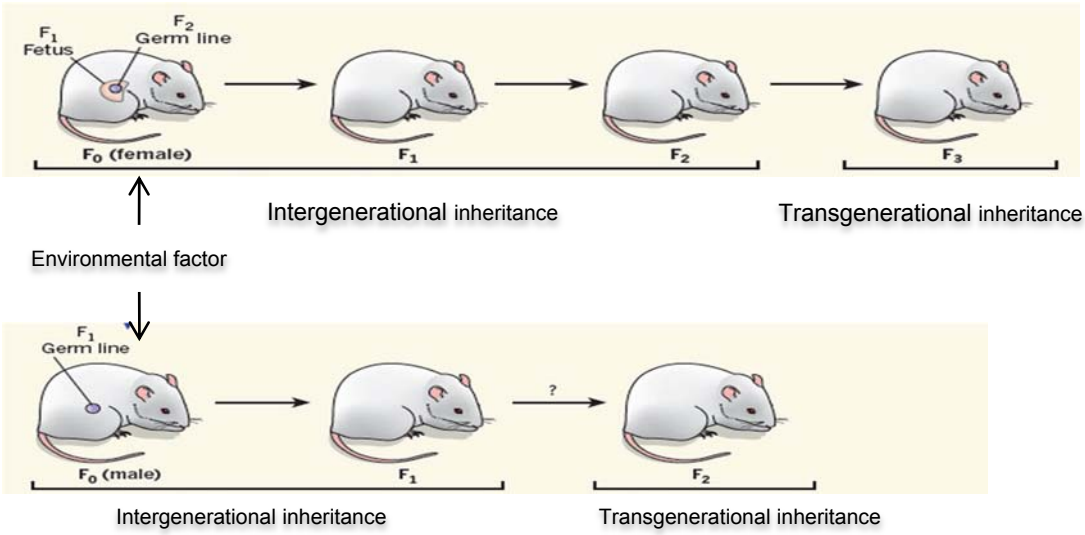


Figure 1.4: Intergenerational and transgenerational epigenetic inheritance. Epigenetic changes in mammals can arise sporadically or can be induced by the environment (toxins, nutrition, and stress). In the case of an exposed female mouse, if she is pregnant, the foetus can be affected *in utero* (F1), as can the germline of the foetus (the future F2). These are considered to be parental effects, leading to intergenerational epigenetic inheritance. Only F3 individuals can be considered as true transgenerational inheritance in the absence of exposure. In the case of males in which an epigenetic change is induced, the individual (F0) and his germline (future F1) are exposed; the F1 is thus considered as intergenerational. Only F2 and subsequent generations can be considered for evidence of transgenerational inheritance (Skinner, 2010).

Due to the mitotic heritability of DNA methylation patterns that are perpetuated through cell division and DNA replication, DNA methylation is a plausible mechanism of heritable biological transmission (as reviewed by Mashoodh and Champagne, 2014). There is however, controversy surrounding the heritability of DNA methylation patterns through meiosis and fertilisation because of epigenetic reprograming events. However, as demonstrated at IAP (Intracisternal A-particle) elements, it is becoming increasingly clear that specific genetic loci have the ability to escape these epigenetic reprogramming events (Mashoodh and Champagne, 2014, Reik et al., 2001). In such events, the altered germ line has the ability to propagate altered epigenetic marks to subsequent generations. Recent studies have begun to document the ability of environmental toxins and nutritional factors to promote epigenetic inter- and transgenerational inheritance.

Environmental exposures alter DNA methylation levels in the sperm of male mice

In a recent study, Lambrot et al. (2013) investigated the effect of a folate deficient diet on the epigenome within the sperm of male mice. Male mice were subjected to either a folate-sufficient or folate-deficient diet. Using a quantitative genome wide DNA methylation analysis, differences in methylation were observed in genes involved in development, with their function in the central nervous system, kidney, spleen and digestive tract. Recently, Radford et al. (2014) evaluated the effects of undernourished dams (pregnant female mice) on the epigenetic profiles of their offspring. 111 regions in the F1 male mice were found to be hypomethylated relative to the control sperm. The regions were enriched in nucleosomes and were low methylated regions, although did not persist in to the F2 generation. This study serves as an example of intergenerational inheritance that was described in the previous section. In line with the hypothesis that ethanol disturbs the methionine homocysteine cycle resulting in a decrease in methylation, alcohol exposure has been associated with significantly lower DNA methyltransferase mRNA levels in the sperm of male rats exposed to alcohol for nine weeks (Bielawski and Abel, 2002).

To date, there have only been three studies that investigated the effect of environmental exposures on DNA methylation levels in sperm of humans. One study analysed the effects of perfluoroalkyl substances (PFASs) on sperm and found no significant associations (Leter et al., 2014). Another focused on DNA methylation levels in the sperm of radiation-exposed workers and found that the radiation-exposed group had hypermethylated spermatozoa (Kumar et al., 2014). Ouko et al. (2009) demonstrated that chronic alcohol use in humans was associated with decreased DNA methylation of specific imprinted regions in male sperm.

Taken together, both animal and human studies suggest that environmental exposures have the potential to alter the sperm methylome. It is thus possible that these alterations can be inherited by their offspring.

Sperm methylation changes are heritable

The first known evidence of the potential multigenerational effect of alcohol exposure dates back almost a century ago, in which a guinea pig model showed that not only did chronic paternal alcohol exposure result in higher mortality rates of offspring, but the mortality rates were elevated in the grandchildren of the alcohol exposed father (Stockard, 1913). Since then the idea that paternal environmental exposures could affect offspring has been rejected. However, recently animal models have provided evidence for the persistence of environmentally induced epigenetic alterations in the sperm through generations.

In 2012, Knezovich and Ramsay investigated the DNA methylation patterns at two paternally methylated Imprinting Control Regions (ICRs) (*H19* and *Rasgrf1*) in the sperm of alcohol-exposed males and somatic DNA of sired offspring following preconception paternal alcohol intake. Imprinted regions are important for embryonic development and are regulated by parent-of-origin specific DNA methylation, which serves to silence one of the parental alleles. Knezovich and Ramsay (2012) observed significantly reduced DNA methylation at the *H19* CCCTC-binding factor (CTCF) 1 ($p=0.0027$) and CTCF 2

(p=0.0009) binding sites in the offspring of ethanol-treated males, which was correlated with a transient reduction in weight at postnatal days 35-42 ($p < 0.05$). In a study conducted by Cheng et al., (2004) male mice were exposed to chromium chloride for two weeks before mating. This exposure led to the hypomethylation of the 45S ribosomal RNA gene in their sperm, with their offspring exhibiting an elevation in both body weight and thyroxine levels.

Although there is evidence that environmental exposures have the capacity to induce epigenetic alterations and pass these variations to subsequent generations through the germline, the question now arises whether these epigenetic alterations passed onto offspring are able to produce changes in gene expression and consequently alter the offspring phenotype.

1.6 Changes in DNA methylation as a consequence of preconception paternal environmental exposures impact on embryonic gene expression

Vassoler et al. (2013) reported that male offspring sired by cocaine-treated males exhibit increased expression levels of cortical *brain-derived neurotrophic* factor (*Bdnf*), which confers a cocaine resistant phenotype. In a different study, Finegersh et al. (2015) exposed male mice to chronic amounts of vaporized ethanol or control conditions and mated them with ethanol naïve females. Adult offspring were then tested for ethanol drinking, ethanol-induced behaviours, gene expression and DNA methylation. Interestingly, they found that ethanol exposed offspring had a reduced ethanol preference and consumption, and increased *Bdnf* expression in comparison to controls.

Carone et al., (2010) exposed male mice to a control or low-protein (test) diet. The sired offspring were then sacrificed, their livers were harvested and RNA was isolated to profile global gene expression differences between the offspring of the two groups. Carone et al.

(2010) observed that offspring born to male mice that consumed a low-protein diet displayed an increase in expression of hepatic genes involved in lipid and cholesterol biosynthesis and reduced levels of cholesterol esters compared to offspring sired by males that were fed a control diet. In addition, quantitative bisulphite sequencing revealed that there were modest (~20%) alterations in cytosine methylation in the livers of offspring sired by males that were fed a low-protein diet, as well as reproducible changes in methylation over a potential enhancer for the key lipid regulator *peroxisome proliferator-activated receptor alpha* (*PPARα*).

Collectively, these studies demonstrate that environmental factors have the potential to alter DNA methylation, both in adult somatic tissue and embryonic tissues through prenatal exposure. Furthermore, evidence has been provided for the potential for environmental factors to alter both the female and male germline. However, what remains unknown and what the current study aimed to elucidate was whether paternal preconception environmental exposures (alcohol) have the ability to alter DNA methylation changes in the sperm, and whether these changes will be inherited by the offspring, thereby dysregulating gene expression in offspring.

1.7 Background to current study

The current study aimed to validate findings by a study entitled “The effects of chronic preconception paternal alcohol intake on methylation signatures and subsequent gene expression in mouse offspring” (Knezovich et al., 2014).

1.7.1 Mouse Model

Ten C57BL/6 male mice were chronically exposed to alcohol (3g/kg) by oral gavage for ten weeks, which stimulated a pharmacologically significant blood alcohol concentration (1mg/ml) and ten male mice were exposed to sucrose (**Figure 1.5)**. Sucrose was used as a control in order to equalise calorie intake in the control group and to control for physical

stress that may be caused by gavaging. Following the ten week exposure, male mice were mated with unexposed C57BL/6 females, which generated embryos.

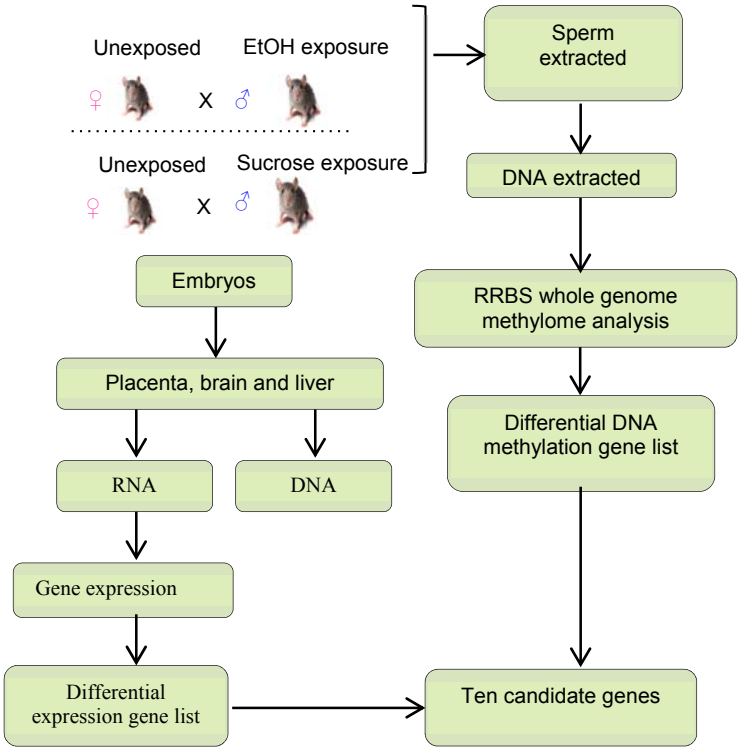


Figure 1.5: Flow diagram of the methodology on which the present study was designed. Ten C57BL/6 male mice were exposed to alcohol for ten weeks, and ten male mice were exposed to sucrose. Male mice were then mated with unexposed C57BL/6 females which generated embryos. Once female mice were inseminted, male mice were sacrificed and the sperm was extracted. RRBS was used to quantify global levels of sperm DNA methylation. Embryos were harvested at E.16.5 and whole brain, placenta and liver were harvested. RNA was extracted from each tissue. RNA was converted to cDNA and a whole genome expression array was carried out. A candidate gene list was then compiled based on genes that showed both a significant change in methylation in the sperm DNA and a significant change in gene expression in at least one of the three embryonic tissues.

1.7.2 Sperm DNA extraction and DNA methylation quantification

Once female mice were inseminated, male mice were sacrificed and mature sperm was extracted from the epididymis of each testicle using a swim-out protocol. DNA was extracted from the sperm and bisulphite modified. Reduced Representation Bisulphite Sequencing (RRBS) for whole methylome analysis was used to quantify sperm DNA methylation at CpG islands at a single-base resolution. Differences in DNA methylation between the sucrose-treated (control) and ethanol-treated (test) groups were quantified and analysed statistically for DNA methylation changes using a t-test.

1.7.3 Allocation of significantly differentially methylated CpG sites

RRBS identified CpG sites in the sperm DNA that showed differential methylation levels between the Ethanol (EtOH) and sucrose exposed male mice. The percentage methylation at each CpG site was determined and the CpG sites were allocated to the closest gene (as shown in **Figure 1.6**). The average DNA methylation percentage was calculated by averaging the quantified levels for each CpG site allocated to a specific gene then calculated per gene.

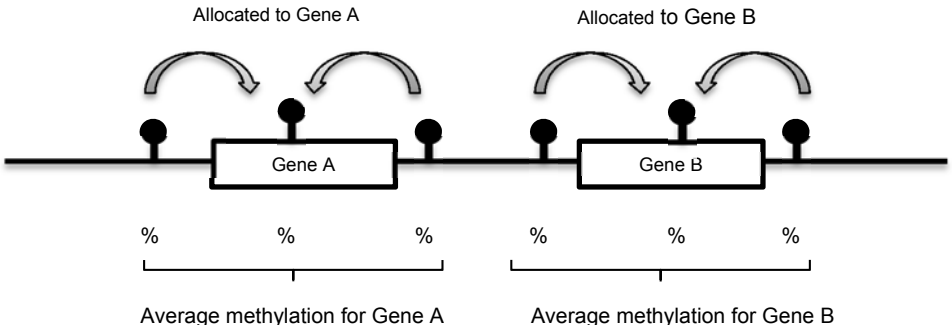


Figure 1.6: Allocation of differentially methylated CpG sites to the gene closest proximity.

1.7.4 Embryo harvesting and cDNA synthesis

Embryos were harvested at embryonic day 16.5 (E16.5). From each embryo whole placenta, brain and liver were harvested. Total RNA was extracted from each tissue, which was DNAse treated and later converted to cDNA using the Thermo Scientific Revert Aid H Minus First Strand cDNA synthesis kit (Thermo Scientific Fermentas, Leicestershire, UK). Additional RNA was stored at -80°C.

1.7.5 Expression array

A Whole-genome expression array was carried out using the MouseWG-6 v2.0 Expression BeadChip (Illumina, San Diego, CA, USA) to assess differences in gene expression levels in the brains, liver and placentae of embryos derived from sucrose-treated and EtOH-

treated fathers. The multi-sample format of the BeadChip allows more than 45,200 transcripts to be simultaneously interrogated. The BeadChip content is sourced from the National Centre for Biotechnology Information Reference Sequence (NCBI RefSeq) database (Build 36, Release 22), is supplemented with probes from the Mouse Exonic Evidence Based Oligonucliotide (MEEBO) and exemplar protein-coding sequences derived from the RIKEN FANTOM2 database (illumina, 2013).

1.7.6 Merging methylome and gene expression datasets to obtain a list of top candidates for pyrosequencing and real-time PCR validation

To compile a list of genes that showed both a significant change in DNA methylation in the sperm of ethanol-treated male mice, and a change in gene expression in embryonic placenta, brain or liver, an overlap analysis was conducted. This analysis effectively filtered for genes that demonstrated a significant change in DNA methylation ($p<0.05$, and either an increase or decrease in DNA methylation of >10% when the control group was compared to the treatment group) and a significant change in gene expression ($p<0.05$). This effectively generated the candidate gene list in **Table 1.1**. Of the nine candidate genes (**Table 1.1**), significant changes in both sperm DNA methylation (according to RRBS data) and in gene expression (according to the array data) were observed in eight genes expressed in the brain, five in the liver and two in the placenta. The percentage difference in methylation in the sperm DNA of the male mice exposed to alcohol or sucrose is described in **Table 1.1**. The measure of how much or less expressed a gene is when compared to its control is referred to as the Fold Change (FC) and each gene's respective FC value is shown in **Table 1.1**. *Tet1* only showed a significant change in methylation according to the RRBS data, however this gene was included in this study as it is a known epigenetic modulator.

Table 1.1: The ten candidate genes analysed in the present study

Gene	Function of genes	Reason for validation	Average methylation difference (%)	FC Brain	FC Liver	FC Placenta
Grm7	Metabotropic glutamate receptor 7, receptor for L-glutamate, a major excitatory neurotransmitter, essential for normal brain function	DMR in sperm; DE in Embryonic tissue	19.4	1.25		
Psmg2	Proteasome assembly chaperone 2, associated with growth retardation (arrest), expression elevated in cycling cells and downregulated in cells under growth arrest	DMR in sperm; DE in Embryonic tissue	12.4	1.16		
Pcbp3	Poly (rC)-binding protein 3, unknown function	DMR in sperm; DE in embryonic tissue	15.9		1.1	
Zfp317	Zinc finger protein 317, a KRAB-ZFP, unknown function	DMR in sperm; DE in Embryonic tissue	12.1	1.33		
Tet1	Epigenetic modulator	DMR in sperm	13.4			
Vwf	Von Willebrand factor, promotes adhesion of platelets to the sites of vascular injury, crucial to the haemostasis process	DMR in sperm; DE in Embryonic tissue	18.4	1.16	1.25	1.21
Depdc1b	DEP domain containing 1B	DMR in sperm; DE in Embryonic tissue	13.4	1.18	1.16	
Ccnd3	Cyclin D3, regulates the cell-cycle	DMR in sperm; DE in Embryonic tissue	22.8	1.12	1.14	
Igf1r	Igf1 receptor, which binds Igf2, decreased levels are associated with reduction in body mass	DMR in sperm; DE Embryonic tissue	12.4	1.18	1.23	
Odc1	Ornithine decarboxylase, regulate cell proliferation	DMR in sperm; DE in Embryonic tissue	14.3	1.14		1.16
DMR= Differentially Methylated Regions **DE= Differentially Expressed** **FC=Fold Change**						

1.8 Study Rationale, Aim and Objectives

1.8.1 Rationale

High throughput approaches are frequently used to evaluate DNA methylation and gene expression changes, as it allows for analysis on a genome-wide level. Although these approaches are valid, they provide a crude overview of both DNA methylation and expression changes, which may result in false positive and negative results, making independent validation a requirement. Based on this, the current study aimed to validate the RRBS and gene expression array findings by Knezovich et al. (2014). Therefore, candidate genes were selected and locus-specific techniques were used to determine whether specific methylation and gene expression differences observed in the original study could be replicated and validated.

1.8.2 Aim

It was hypothesised that chronic exposure of male mice to alcohol leads to DNA methylation aberrations in the sperm, which would subsequently be inherited by sired offspring and consequently dysregulate embryonic gene expression.

Therefore, this study aimed to validate the effect of alcohol on sperm DNA methylation, and investigated whether possible aberrations were transmitted to sired offspring and whether a change in embryonic DNA methylation and gene expression was observed.

1.8.3 Study Objectives

1. To validate significant changes in sperm DNA methylation (as observed by RRBS data analysis) in ten candidate genes in sires exposed to alcohol.
2. To quantify DNA methylation in the embryonic tissues (brain, liver and placenta) of the ten candidate genes.

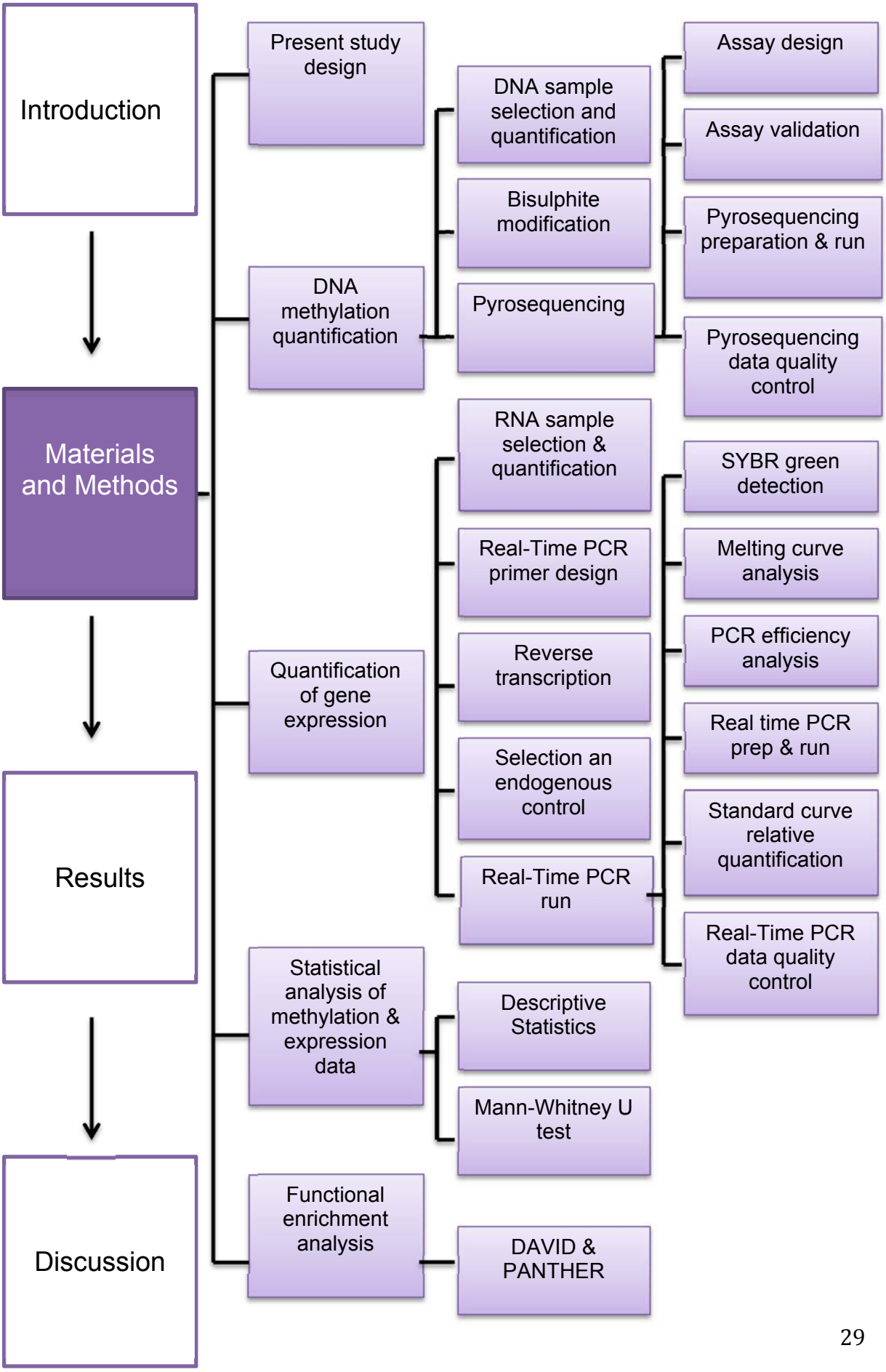
3. To validate significant changes in gene expression (as observed by whole genome expression array data analysis) in the embryos sired by males exposed to alcohol for ten weeks when compared to embryos sired by sucrose treated males.

4. To reanalyse the microarray expression data generated by Knezovich et al. (2014) to construct a list of genes that displayed a significant ($p<0.001$) increase or decrease in expression with a fold change greater than 1.2 in the embryonic, brain, liver or placenta and to identify biological pathways and functions significantly overrepresented among these genes.

Chapter 2

Materials & Methods

2.0 Materials and Methods



Materials and Methods

The sections covered in the materials and methods chapter will include the present study design, the samples and techniques used, the genes analysed as well as the statistical approaches used to analyse the data. All of the laboratory work for the present study was conducted in the Physiology, Development and Neuroscience Department at the University of Cambridge, UK.

2.1 Present study design

This study had four aspects to it (**Figure 2.1**), which included (1) validation of significant changes in sperm DNA methylation for ten genes in sires exposed to alcohol or sucrose, by bisulphite modification and quantitative pyrosequencing; (2) Quantification of embryonic tissue DNA methylation by bisulphite modification and pyrosequencing; (3) Validation of significant changes in candidate gene expression in the brain, liver or placenta of their embryos by quantitative real time PCR; (4) Reanalysis of the previously generated microarray gene expression data by using bioinformatics tools to generate a top ranked candidate gene expression list and to identify and analyse biological functions or pathways significantly over represented among these genes using the Protein Annotation Through Evolutionary Relations (PANTHER) browser (Thomas et al., 2003) and the Database for Annotation Visualization and Integrated Discovery (DAVID) (Huang et al., 2005, Huang et al., 2009).

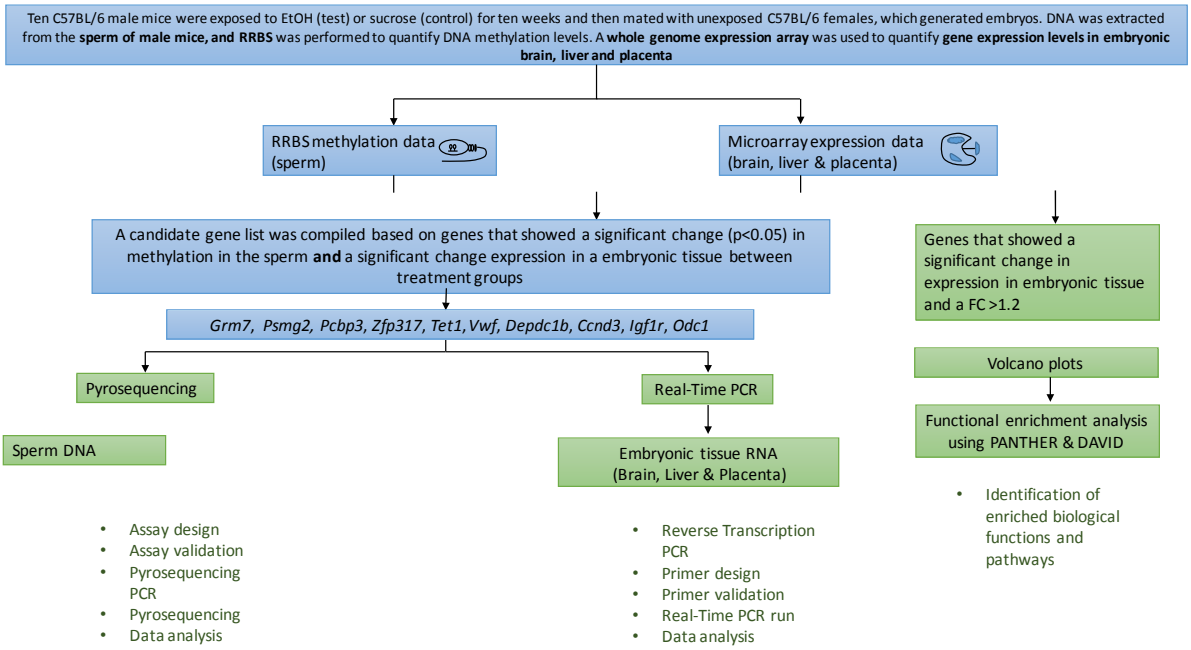


Figure 2.1: Flow Diagram of the Methodology. The previous study is depicted in blue and explained in section 1.7. The current study is shown in green. Pyrosequencing was performed on both the sperm and embryonic DNA to quantify DNA methylation levels. Real-time PCR was used to quantify gene expression levels in embryonic tissues. The previously generated microarray expression data was renalaysed to generate a new gene list and bioinformatic approaches were used to identify biological processes and functions over represented among this list. See text below for details.

Sperm and embryonic DNA methylation quantification

2.2 DNA sample selection and quantification

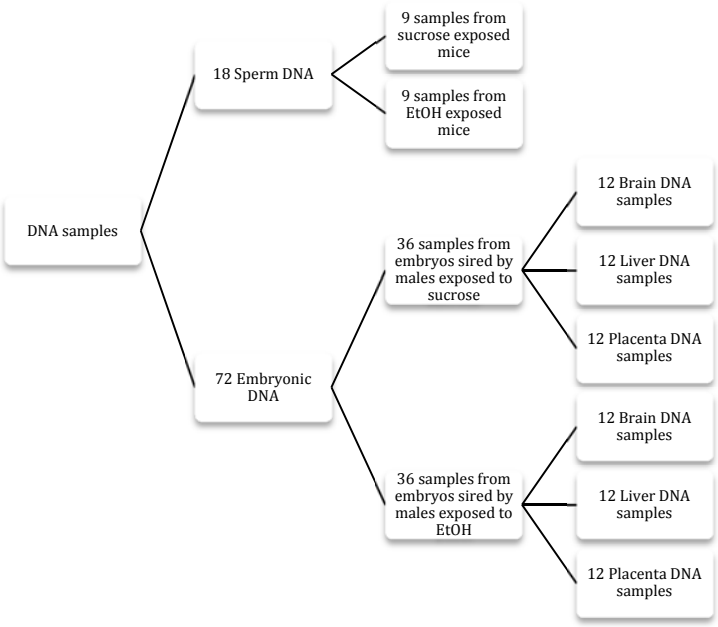


Figure 2.2: Flow diagram of the DNA samples included in the present study. Every DNA sample underwent bisulphite treatment followed by pyrosequencing to quantify DNA methylation levels.

Nine sperm DNA samples were selected from the test group (male mice chronically exposed to EtOH) and the control group (male mice exposed to sucrose) (**Figure 2.2).** The concentration range for the sperm DNA samples was low and varied between 2ng/µl to 4ng/µl.

Thirty-six DNA samples were selected from embryos sired by EtOH or sucrose treated male mice and separated according to tissue type as described in **Figure 2.2**. Embryonic DNA samples were diluted to a final concentration of 50ng/µl using double distilled water (ddH_20) (an optimum DNA concentration for pyrosequencing). The dilution volumes and DNA concentrations were determined using the equation: $C_1V_1=C_2V_2$.

DNA concentration was measured using the NanoDrop®ND-1000 Spectrophotometer and ND–1000 V3.3 software (NanoDrop Technologies, Wilmington, DE USA). All DNA samples were stored at -20°C.

2.3 Bisulphite modification

The ability to efficiently and accurately detect and quantify DNA methylation has become a vital component for the study of cancer, gene expression, genetic diseases, as well as many other aspects of biology. To date, various methods have been developed to facilitate the identification and quantification of DNA methylation, however the bisulphite conversion and quantitative pyrosequencing method remains the "gold standard" (Patterson et al., 2011).

Bisulphite modification of DNA is the most commonly used method for DNA methylation studies (Patterson et al., 2011). This technique was first described by Frommer et al. (1992) and is based on the finding that treating DNA with sodium bisulphite deaminates unmethylated cytosines converting them to uracil, (**Figure 2.3**) contrary to methylated cytosine's (5-methylcytosines) which remain unchanged. Therefore, the treated sequence of DNA will differ from the original sequence at the unmethylated cytosine sites. During the PCR amplification process, newly converted uracil residues pair with adenosine followed by complementation with thymine in the subsequent PCR amplification steps, whereas 5-methylcyctosine (immune to the uracil conversion) pair with guanine followed by complementation with cytosine. The chemical reaction therefore translates epigenetic information into differential sequence information, and the positions exhibiting potentially variable degrees of methylation can be analysed as virtual C/T polymorphisms in bisulphite treated DNA (Tost and Gut, 2007). Specific primers are designed based on the chemically converted DNA sequence (see section 2.4.1).

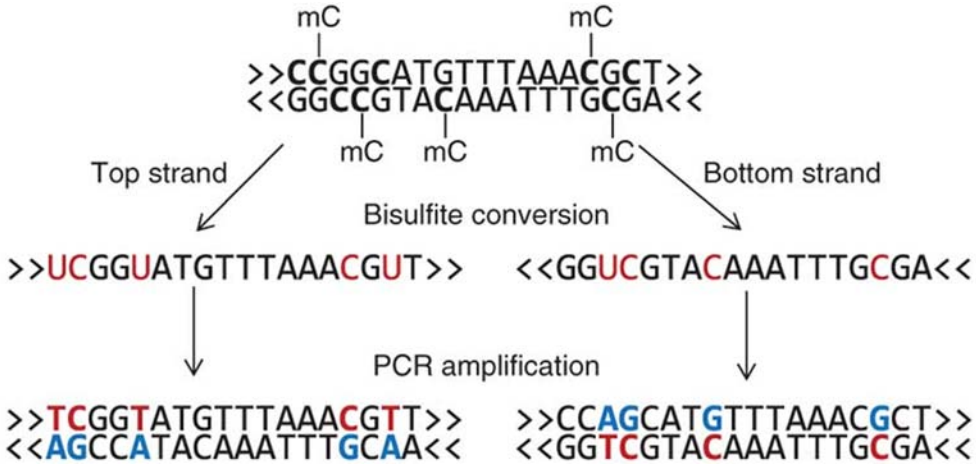


Figure 2.3: Schematic of the bisulphite conversion for the analysis of DNA methylation. Methylated cytosines (mC) remain as Cs, while unmethylated Cs are converted to uracil (U) and subsequently to thymine (T) during the PCR amplification reaction (Krueger et al., 2012).

The EZ DNA Methylation-Gold Kit (Zymo Research CA, USA) was used to bisulphite modify the DNA (sperm and embryonic DNA) according to the manufacturer's instructions (full protocol shown in Appendix C). The reaction process involves five basic steps:

(1) DNA denaturation to separate double stranded DNA

(2) Incubation with bisulphite at elevated temperatures to allow for the conversion of unmethylated cytosines to uracil. Converted single-stranded DNA is then bound to the membrane of the spin-column, which is purified through a wash step.

(3) Desalting to remove any bisulphite ions.

(4) Desulfonation completes the bisulphite conversion by facilitating the removal of the sulfonate group from uracil, as sulfonate inhibits downstream PCR reactions if not removed. This is followed by a wash step to remove the desulfonation agent.

(5) The elution step in which pure converted DNA is eluted in 10µl of elution buffer

This process results in the complete conversion (99.5-99.7%) of DNA. Bisulphite modified DNA was stored at -20°C.

2.4 Quantitative DNA methylation analysis by pyrosequencing

Pyrosequencing is a sequencing-by-synthesis technique that allows for quantitative CpG site-specific DNA methylation analysis that can be monitored in real time (Tost and Gut, 2007). This technique involves a light-emitting reaction that relies on a sequential addition and incorporation of nucleotides to a single stranded bisulphite modified DNA template in a primer-directed polymerase extension. The DNA methylation level at a certain CpG position is determined based on parallel sequencing of all possible templates (presence of a C or T at each CpG site) present in a single PCR reaction and is displayed as an average (Tost and Gut, 2007). This technique is cost effective and can be used in several applications including the analysis of genetic variation, genotyping, determination of copy number and the analysis of DNA methylation (Tost and Gut, 2007).

This technique involves isolating the template strand from its complementary strand following PCR amplification, accomplished as a result of incorporation of biotin at the 5' end of the primer used to initiate amplification of the particular strand destined to undergo pyrosequencing. Streptavidin-coated sepharose beads are then used to capture the biotinylated PCR products and the unlabelled strand is then washed away.

The single stranded DNA template is then annealed to a sequencing primer and incubated with two substrates (adenosine 5' phosphosulfate and luciferin) and four enzymes (luciferase, ATP-Sulfuryase, apyrase and polymerase). The first deoxyribonucleotide triphosphate is added to the reaction (according to the dispensation order, which is determined by the sequence being analysed) and is incorporated into the DNA strand. If it is complementary, if successful an inorganic pyrophosphate (PPi) is released in a quantity equimolar to the amount of incorporated nucleotide. The generated PPi is then converted

to ATP by ATP sulfurylase, which is used by luciferase to convert luciferin to oxyluciferin generating visible light in amounts that are proportional to the amount of ATP. The visible light produced is detected by a charge coupled device (CCD) camera and seen as a peak in the raw data output (Pyrogram) (**Figure 2.4**). The height of each peak (light signal) is proportional to the number of nucleotides incorporated (England and Pettersson, 2005, Ronaghi, 2001).

The apyrase nucleotide-degrading enzyme is responsible for the degradation of any unincorporated nucleotides and any ATP that remains in the reaction mixture. When degradation is complete another nucleotide is added. Addition of dNTPs is performed sequentially. As the process continues, the complementary DNA strand is built up and the nucleotide sequence is determined from the signal peaks in the pyrogram. The intensity of the light emitted for each reaction well is used to provide information regarding the quantity of template present and to confirm that the sequence produced matches the sequence of the template (England and Pettersson, 2005, Tost and Gut, 2007).

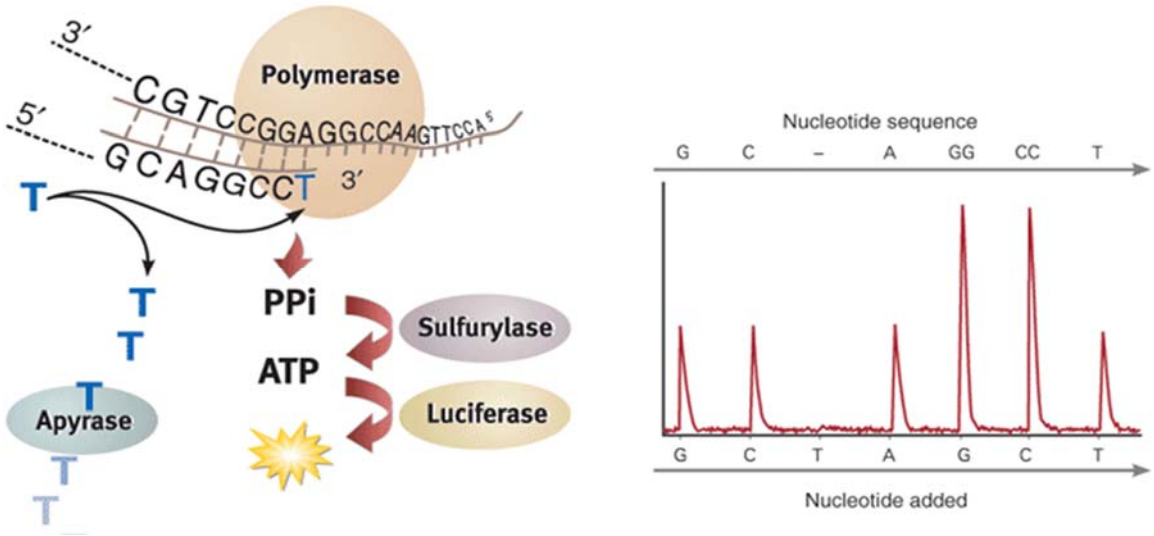


Figure 2.4: The principle of pyrosequencing and the output pyrogram. This technique involves the release of pyrophosphate (PPi) following a nucleic acid polymerisation reaction due to nucleotide incorporation on the single stranded complementary strand. Once the PPi is released it is converted to ATP by the ATP sulfurylase enzyme thereby providing energy to luciferase to mediate the conversion of luciferin to oxyluciferin, generating light. A light signal proportional the quantity of PPi released is produced and captured by a CCD camera and displayed as peaks in a pyrogram. Obtained from England and Pettersson (2005).

2.4.1 Pyrosequencing assay design

Pyrosequencing assays were designed using the Assay Design software version 2.0 (Biotage, Uppsala, Sweden). This software specifically designs an assay to amplify a region of interest that contains a number of single nucleotide polymorphisms (SNPs). In this instance, the SNP is equivalent to the cytosine nucleotide contained within a CpG dinucleotide. Following bisulphite modification and subsequent PCR, if the cytosine (C) residue was methylated it will remain cytosine, however if it was unmethylated it will be amplified at thymine (T). Thus, the "polymorphism" of interest is C/T. The sequence of interest was imported into the design software, with the CpG sites denoted as C/TG.

The region of interest was based on the CpG sites that showed significant alterations in methylation in sperm of alcohol-exposed mice in comparison to the sucrose-exposed mice as per the RRBS data. The target region encompassing the differentially methylated CpG sites was identified using the UCSC website. The pyrosequencing assay design software (Biotage, Uppsala, Sweden) was then used to design a list of potential primers based on optimum annealing temperatures, primer lengths and minimisation of non-specific binding. Using the software the user can select the primer pair that covers the CpG sites of interest. Generally, a target region of 400 base pairs or greater is selected and from this an amplicon of 80-200 base pairs is generated. The primers generated by the software include: a pair of pyrosequencing primers (forward and reverse) (**Table 2.1**) to amplify the region of interest and a single sequencing primer to sequence the region of interest. Primer sequences were sent to Sigma-Aldrich, UK, for synthesis.

The number of CpG sites analysed ranged from one to five per PCR template, depending on the target site in the gene of interest. The same CpG sites were analysed in the embryonic DNA samples as those analysed in the sperm DNA samples by RRBS. Shown below is an example of the primer design for *Vwf* (one of the candidate genes). This figure highlights the complementarity of the primer pairs to the bisulphite modified DNA. *Vwf*

contained six CpG sites that showed significant alterations in methylation according to the RRBS data, however only two CpG sites were analysed for validation as a consequence of the pyrosequencing software's ability to only analyse small regions. The CpG sites that showed differential methylation according to the RRBS data are highlighted in green below. The pyrosequencing primers captured the two differentially methylated sites and an additional CpG site (highlighted in yellow and underlined in **Figure 2.5A**. One of the primers is always biotin labelled as this facilitates immobilisation to the sephrose beads.

Vwf pyrosequencing primer

A. Original sequence

5'GCAGGCATTCCACGGGACACAGAGAGAGGGAAGGCAAAGGCGGAGCCATTGGCCCAGCCT
TTATTTTGACCACTGGCCGTTGAGATCCACATTTGGCTCATTTCCTTTTCAAAGTTTGTTTATTGA
ACACCCTCCATCCATGGGTGAAATCCCAAGAAAGAATAGAATATTGGGTCGTTTCTGGCCCGG
AAAAGACCCAGTCTAGAGCAGCGGTGGGTGTGGGAGCTTAAAGTTGACTTGTATGATAGATGC
GCTGCAGTGACCCTGTGTGGACACGGTGCTGCGACATCTCTGCTCTTGCTAGCTTCTGCTACT
CTAACCACCGTCTCCGGTCAGTGGGTCTTTGGGGAACGCACATTTACCCGTTACTGAGCCAAA
CGCAGAACTGTACATTGGCACCAGAAGCCAGAGCCAGATAGATTAATTTTGGTAGCAACAGAAA
AGGGAACTGGCTTTGTTTGTGGGGA-3'

B. Bisulphite modified sequence

5'GTAGGTATTTTACGGGATATAGAGAGAGGGAAGGTAAAGGCGGAGTTATTGGTTTAGTTTTTA
TTTTGATTATTGGTCGTTGAGATTTATATTTGGTTTATTTTTTTTTTAAAGTTTGTTTATTGAATATT
TTTTATTTATGGGTGAAATTTTAAGAAAGAATAGAATATTGGGTCGTTTTTGGTTCGGAAAAGATT
TAGTTTAGAGTAGCGGTGGGTGTGGGAGTTTAAAGTTGATTTGTATGATAGATGCGTTGTAGTG
ATTTTGTGTGGATACGGTGTTGCGATATTTTTGTTTTTGTTAGTTTTTGTTATTTTAATTATCGTTT
TCGGTTAGTGGGTTTTTGGGGAACGTATATTTATTCGTTATTGAGTTAAACGTAGAATTGTATAT
TGGTATTAGAAGTTAGAGTTAGATAGATTAATTTTGGTAGTAATAGAAAAGGGAATTGGTTTTGT
TTGTGGGGA-3'

Figure 2.5: *Vwf* original genomic sequence and its bisulphite modified sequence. A. *Mus musculus* genomic sequence spanning a region of the *Vwf* gene. The sequence shown is the *Vwf* sequence in chr6 at nucleotides: 125553577-125554043 in the mouse (July 2007) NCBI37/mm9 assembly UCSC browser (https://genome.ucsc.edu/). The CpG sites fall within exon 14 or the intronic regions spanning either side. The CpG sites within the target sequence are highlighted in yellow; the CpG sites that showed significant alterations in methylation data according to RRBS data

are highlighted in green. The CpG site highlighted in yellow and underlined did not display differential methylation in the RRBS data but was captured and analysed by the pyrosequencing primers B. The bisulphite modified sequence spanning the same region as A. The sequence that is highlighted in blue represents the reverse primer, the sequence highlighted in grey represents the sequencing primer and the sequence underlined is the biotinylated reverse primer. The pyrosequencing primers amplified a 98bp amplicon.

Table 2.1: Pyrosequencing primer properties and the CpG sites analysed

Region Primer	Primer Sequence (5'→3')	CpG sites analysed	Amplicon size (bp)	Annealing temperature (°C)
Grm7*-F** ***Grm7*-R ***Grm7*-S**	GGGTTAAGGATTATGGTGTTGTAT AACACAACCCCTACTCCAATCCAATC AC ATGGTGTTGTATTGGG	111543873	127	54
Zfp317*-F** ***Zfp317*-R ***Zfp317*-S**	AGATTGGATTGGGGTAGG ACCTAAAACACAAACCAAATCTAC TGGGTTTTTGGGGGT	19443712 19443716	150	54
Tet1*-F** ***Tet1*-R ***Tet1*-S**	GGTAGGTGTTATGTGGTGTTAG ACCCACTTAATAATCAACCAACCACT TCC TGAGAGGGTTAGGATGTTTAAT	623329009 623329050	195	54
Vwf*-F** ***Vwf*-R ***Vwf*-S**	ATGGGTGAAATTTTAAGAAAGAATAG AATA ATCAACTTTAAACTCCCACACC AGAAAGAATAGAATATTGGG	125553557 125553763 125553787	98	54
Depdc1b*-F** ***Depdc1b*-R ***Depdc1b*-S**	GGATTGTTTTTTTAATGGGTAGGAGA T ATAACTACTAAACCAAACTAAACACT ATGAGAGTTAGTTTTGTAGGGTAA	109101807 109101098 109101330	168	54
1f1*-F** ***rr1*-R ***Ifgr1*-S**	ATGGGGGGTTTTTTAAGAAAG ACAACACTACCCTACAAAATTTTTA GGTTTTATAGGAGTGAAGAGG	75274289 *75274294* 75274301 75274319 75274324 75274349	228	54
Odc1*-F** ***Odc1*-R ***Odc1*-S**	GGATAGTTTTAGTGAATTAAAGGGAT AG CCTCAACAACTCAACCCTTCATT TGATTTGAAAGTTTGTTGT	17514120 17514135 17514144 17514154	196	54

* Biotinylated primer

2.4.2 Pyrosequencing assay validation

Before the candidate gene pyrosequencing assays were implemented, validation of the primer sets was performed to determine whether preferential amplification towards methylated or unmethylated DNA would occur during the PCR. This involved the production of unmethylated and methylated controls, bisulphite modification,

pyrosequencing PCR, pyrosequencing and analysis through calibration curves and the preparation of a dilution series.

Unmethylated and methylated DNA control preparation

PCR bias can be detected using PCR-amplified DNA with a known level of methylation. The first step of the validation process involved the generation of unmethylated (0%) and methylated (100%) control samples. Genomic DNA isolated from a triple knock out mouse was used (knocked out for all methylation enzymes: DNMT1, DNMT3A, DNMT3B) as fully unmethylated DNA. Fully methylated DNA was produced using the recommended protocol (Biolabs, New England, UK) for methylation of genomic DNA (Appendix D).

Validation Mixes and PCR amplification

In order to generate a range of methylation levels 0, 25, 50, 75, 100%, the bisulphite modified 0% and 100% controls were mixed before pyrosequencing PCR according to the ratios described in **Table 2.2**. The samples underwent pyrosequencing PCR with the designed primers (**Table 2.1**) to amplify the regions of interest within candidate genes. The reaction was performed in the Biorad DNA engine PTC 2000 Thermoblock (Biorad, Berkeley, California) as per the thermal cycling conditions described in **Table 2.3**. The PCR amplification for each sample was performed in triplicate.

Table 2.2: Pyrosequencing PCR assay validation and pre-PCR reference mixes

Percentage mix (%)	Volume of methylated DNA (µl)	Volume of unmethylated DNA (µl)
100	20	0
75	15	5
50	10	10
25	5	15
0	0	20

Following PCR, agarose gel electrophoresis was used to visualise the PCR products and assess potential contamination, which is indicated by the presence of a smeared band, a

second band in any lane or a band in the no template PCR control lane. Using the mixes, all the candidate genes were analysed by pyrosequencing.

Validation Analyses

Calibration curves were constructed for each primer and assessed at each CpG site within the candidate genes. Analyses of the PCR R^2 (a statistical measure of how close the data are to the fitted regression line) values between expected and observed methylation levels were used to determine whether there was preferential amplification of methylated DNA by the candidate gene primers. Although the R^2 for each primer set was above 0.9 (an accepted value), when the methylation levels approached 60% or higher, the expected and observed percentages did not correlate.

To ensure that preferential amplification was not occurring, the same assay was used on *H19*, an imprinted PCR primer already proven to have no bias, according to the standard protocols in the laboratory (Anne Ferguson-Smith Lab in the Physiology, Development and Neuroscience department at University of Cambridge, UK). Pyrosequencing analysis of this primer and calibration curves produced similar results as the validation analyses of the candidate genes. Therefore, it was concluded that there was an error with the assay and not the primers. An alternative method (assay validation using a dilution series) was then used to confirm the absence of preferential amplification.

Assay validation using a dilution series

Due to a failure to confirm the absence of preferential amplification using varying ratios of methylated and unmethylated DNA and the construction of calibration curves, an alternative method was carried out. This method required mixing all the unmethylated and methylated samples (of varying methylation levels) that was previously produced, in a single tube. Using ddH_20, the DNA sample was used to produce a set of serial dilutions: 1:10, 1:100, 1:1000, and 1:10 000. Using the candidate gene primers in **Table 2.1**, samples were amplified by PCR pyrosequenced in technical triplicates. This assay

validation method is based on the fact that highly concentrated DNA will plateau faster than highly diluted DNA. For example, if the 1:10 (highly concentrated) sample reaches its PCR plateau at cycle 26, the 1:100 sample will plateau at cycle 29, the 1:1000 will plateau at cycle 32 and the 1: 10 000 sample will reach its plateau at cycle 38. Therefore, the 1: 10 000 samples would have undergone 14 more PCR cycles to reach its plateau in comparison to 1:10 sample. Due to an accumulative effect, if there were preferential amplification the 1:10000 methylated sample would show a strong bias and the methylated percentage would vary considerably to the 1:10 sample. If there is no bias it would be expected that the four different DNA dilutions would produce similar methylated percentages for each CpG site as seen in the graphs shown in Appendix F.

2.4.3 Pyrosequencing PCR

PCR is a sensitive technique that was first described by Kary Mullis in the 1980s, which involves the amplification of specific regions of DNA (Mullis and Faloona, 1987). Amplification relies on repeated cycles of denaturation, primer annealing and extension each with specific temperatures to synthetically replicate targeted regions of DNA, ultimately resulting in large quantities of the desired region. Conventional PCR was used to amplify the region containing the CpG sites of interest in the ten candidate genes (Chapter 1, **Table 1.1**) from bisulphite modified DNA. Amplifications were performed as per the conditions in **Table 2.3**. PCR amplification for each sample was performed in technical triplicates on the Biorad DNA engine PTC 2000 Thermoblock (Biorad, Berkeley, California). DNA amplicons were pyrosequenced to quantitatively analyse the methylation levels at specific CpG sites in the test and control samples.

Standard electrophoretic techniques were used to confirm the specificity of the PCR amplified products. Amplified products were visualised on a 1% agarose gel containing SafeView (NBS Biologicals, Huntingdon, UK) at 100 volts in 1X TBE buffer for 40 minutes.

A 100bp DNA molecular size marker at 0.1µg/µl (Gene Ruler 100bp plus, Thermo Scientific, MA, USA) was used as a size standard.

PCR Optimisation

Throughout the study, PCR optimisation was required. Optimisation included setting up temperature gradients, which tested temperatures 5-6°C above and below the initial calculated annealing temperature, varying the volume and/or concentration of the general PCR components, increasing or decreasing the number of cycles of the reaction and/or the length of time that the particular cycle occurred.

Table 2.3: Reagents and conditions used for the PCR amplification of the candidate genes before pyrosequencing

PCR Components	Concentration	Volume (µl) (1x mix)	
DNA	Variable	1	
dNTPS	2mM	1.25	
Buffer	---	1.25	
Forward Primer	100 µM	0.125	
Reverse Primer	100 µM	0.125	
HotStarTaq Polymerase (QIAGEN)	---	0.1	
ddH_20	---	8.65	
Total Volume		12.5	
Thermal Cycling conditions	**Temperature (°C)**	**Time (min: sec)**	**Cycles**
Initial heat activation	95	5:00	1
Denaturation	94	1:00	X40
Annealing	54	1:00	
Extension	72	00:15	
Final extension	72	10:00	1
Hold	10	∞	

2.4.4 Pyrosequencing preparation and run

Pyrosequencing was performed using the PyroMark MD System (Biotage, Uppsala, Sweden) (full protocol in Appendix E). Pyrosequencing preparation begun by loading each well of a 96 well plate with: 10µl of PCR product, 1µl of Streptavidin Sepharose HP beads,

30μl ddH_20 and 39μl of binding buffer. The plate was then placed on a shaker for >6min at 1200rpm. This enabled the biotin labelled PCR products to bind to the sepharose beads.

The vacuum probe block was then used to capture the PCR product attached to the sepharose beads. The vacuum probe block, with the PCR product attached to it end, was then sequentially placed into four troughs containing (1) 70% EtOH, (2) denaturation solution, (3) washing buffer and (4) ddH_20, for 8 seconds each. This process facilitated the separation of the double stranded DNA, causing the single standard biotin labelled PCR product attached to the sepharose beads, to be captured by the vacuum probe block. The probe block with the sepharose beads and PCR product attached was then gently placed into a pyrosequencing 96 well plate, in which each well had been pre-filled with 0.5μl of 10μM sequencing primer and 11.5μl of annealing buffer. The plate was covered with an adhesive film and placed on a 90°C heating block for 3 minutes and left to cool to room temperature. This step aided in annealing the single stranded PCR products to the sequencing primer.

Using the PyroMArk MD software (Biotage, Uppsala, Sweden) as a guideline, a reagent cartridge (**Figure 2.6**) was prepared with the required amounts of enzyme, substrate and dNTPs. The filled cartridge was then loaded into the pyrosequencer and a “Tip Test” was performed. If a drop was dispensed from each tip, the test was successful and the pyrosequencing plate with the appropriate samples could be loaded. Each sample in every assay was run in triplicate.



Figure 2.6: Pyromark MD Cartridge. Arrangement of the tips in the pyrosequencing MD cartridge. S- Substrate, E-Enzyme, A, C, T, G- 4 nucleotides (McKnight, 2004).

When analysing methylation data, the ratio of C (methylated): T (unmethylated) at a given CpG site within a specific DNA sample may vary, and is therefore unlike a traditional SNP where a sample is denoted as either heterozygous or homozygous. In this case, the C and T alleles are quantified and expressed as a percentage. Pyrosequencing data was analysed using the Pyro Q-CpG Software (Biotage, Uppsala, Sweden).

2.4.5 Pyrosequencing data analysis

The Biotage PyroMark Q96 MD software generates a theoretical histogram based on the nucleotide input sequence (region of interest within the candidate gene) entered by the user. As shown in **Figure 2.7a** the x-axis represents the nucleotide dispensation order, which is labelled every 5bp, whereas the y-axis indicates the number of nucleotides expected to be incorporated. The filled bars in **Figure 2.7a** represent the nucleotides outside of the CpG sites whilst the clear bars that contain a vertical double arrow represent the CpG sites. The arrows indicate the potential theoretical number of incorporated nucleotides (Tollefsbol, 2011).

The Biotage PyroMark Q96 software (Biotage, Uppsala, Sweden) creates a pyrogram using the collected data points and the theoretical histogram as an internal control (**Figure 2.7b**). The variable positions at the CpG sites are given in the appropriate International Union of Pure and Applied Chemistry (IUPAC) code (Y stands for the bases C or T) (Tollefsbol, 2011). Similar to the histogram, the nucleotide dispensation order is

represented on the x-axis whilst the y-axis describes the relative peak intensity in arbitrary units. Incorporated nucleotides can be seen as peaks in **Figure 2.7b.**

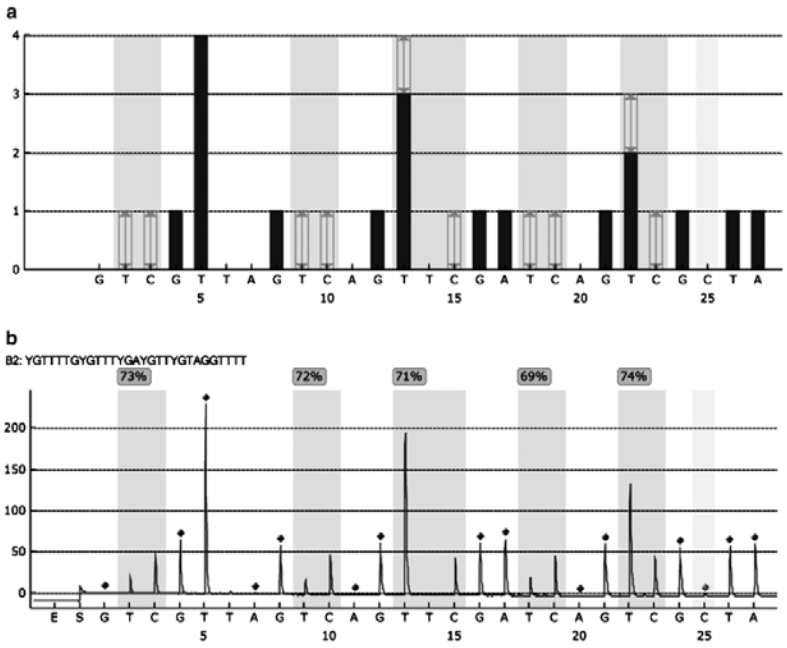


Figure 2.7: Results of a bisulphite pyrosequencing assay to determine the DNA methylation status within a region of interest. (**a**) The theoretical histogram based on the nucleotide input sequence showing the expected peak pattern for the region of interest. (**b**) The pyrogram obtained for the bisulphite sequencing describing the methylation levels at the CpG sites within the region (Tollefsbol, 2011).

The variable positions highlighted by the darker shading represent the peaks, which are used to establish the methylation ratio at a specific CpG site e.g. nucleotide positions 2 and 3 (Tollefsbol, 2011). The DNA methylation percentage calculated by the software is given above the shaded boxes. These boxes are colour coded in blue, yellow and red, which indicate the quality and reliability of the DNA methylation levels, determined by the quality control settings of the pyrosequencing software (Biotage, Uppsala, Sweden). A blue code indicates a pass, a yellow code indicates that a manual check is required, and a red code indicates a failure (Tollefsbol, 2011).

2.4.5.1 Data quality control

Following a pyrosequencing run, the quality of the run was analysed by checking if the pyrogram was blue and had passed the quality control measures set out by the software. In cases where the blocks shaded behind the program were yellow, the data was manually assessed to determine whether it was acceptable. All runs that produced a red shade were taken as unreliable and repeated.

The average methylation percentage for each CpG site was calculated by averaging the triplicate values for each CpG site for each sample. A sample was excluded from analysis if the intra-sample variation (same sample performed in triplicate) exceeded 10%. In cases where more than three separate samples were excluded due to large variation, a new plate was set up and the region was reanalysed.

As the previous study had analysed and reported the mean methylation levels per candidate gene, the mean methylation percentage per gene was then calculated to facilitate a comparison between the techniques. This was calculated by averaging the methylation percentage per CpG site across all the CpG sites analysed per gene.

To account for any bias that may occur by averaging the samples, the RRBS data was reanalysed to determine the individual CpG site methylation levels. The methylation alterations per CpG site could then be directly compared between the techniques.

Quantitative Real-Time PCR of cDNA from mouse foetal tissues

2.5 RNA sample selection

All 144 RNA samples that were analysed using the microarray were selected for validation using quantitative real-time PCR. Of the 144 samples, 72 samples were from embryos sired by males chronically exposed to EtOH (test) and 72 from embryos sired by males exposed to sucrose (control). The treatment group samples were divided evenly between the brain (n=48), liver (n=48) and placenta (n=48). Each sample had a corresponding reverse transcriptase negative control, included to verify the absence of genomic contamination. Samples were stored at -80°C. Prior to real-time PCR, samples were run on a 1% agarose gel to assess the quality of the RNA.

2.6 RNA quantification

Six RNA samples were chosen at random and quantified using the NanoDrop®ND-1000 Spectrophotometer and ND – 1000 V3.3 software (NanoDrop Technologies, Wilmington, DE USA). The concentration of the samples ranged between 1,000ng/μl and 1,300ng/μl. Therefore, an average concentration of 1,200ng/μl was chosen for the 144 RNA samples. RNA concentration was measured at 260nm and purity was assessed using the 260:280nm ratio. A ratio of 1.8 was accepted as “pure” RNA. If the ratio is substantially lower, this is indicative of phenol, protein or any other contaminants that absorb optimally at 280nm.

2.7 Real-time PCR Primer design

The real-time PCR oligonucleotide primers were manually designed for each of the candidate genes to ensure maximal efficiency and sensitivity. The forward and reverse

primers were designed in the region that was in close proximity to the probes that were used for the microarray analysis. The microarray probe sequences (**Table 2.4**) were identified using the probe IDs (**Table 2.5**) and submitted to the UCSC website, this identified the target region. The primer sequences used for the amplification of the candidate genes were designed using Primer3 (version 4.0.0) (http://primer3.ut.ee/), a computational program that enables the user to design a compatible primer pair for a specified region.

Primer3 has different input parameters that can be controlled and adjusted according to what is required by the researcher in respect to designing primers. The primers listed in **Table 2.6** were designed in such a way that they contained sequence from two adjacent exons. This way of designing primers reduces the possibility of amplifying contaminating DNA, which contains introns separating exons. One of the requirements of the real-time PCR primers is that they amplify a region smaller that 200bp.

Table 2.4: The probe ID and NM acession number used to identify the microarray probe sequence

Gene	Probe ID	NM
Vwf	7510414	NM_011708.3
Grm7	7160605	NM_177328.3
Psmg2	3850471	NM_134138.1
Zfp317	2690519	NM_172918.3
Pcbp3	1300681	NM_021568.1
Depdc1b	10161	NM_178683.4
Ccnd3	610717	NM_007632.2
Igf1r	6960678	NM_010513.2
Odc1	6840121	NM_013614.1

Table 2.5: The microarray probe sequence used to identify the region of interest to be amplified by real-time PCR

Gene	Probe sequence
Vwf	GGTCATCAACGCCATGCAGTGCAGGTGTTCTCCCCGGAACTGCAGCAAGT
Grm7	GAGGTTTTGTTCATGCCCCTATGACTGACCAATGTGATAAGGACTTTAAA
Psmg2	TTGGAAGATCCCCAGCTCTTGGCGGCTGCTCTTTGGCAGTGGGCTTCCTC
Zfp317	GCCATGACTCTGGCCGTTGACTCATAAATAGAGATAAGTGTGTGGCATTC
Pcbp3	GTGTAGCCCTCCACGCATTGGCATGCAGCTCTTTACTTTACTGCTCAACG
Depdc1b	GCAGATTTGCCTGTCGCTCTTTGATTACCCGTATGCCGGGCTGGGTCTGC
Ccnd3	CACCCCTCCACCTCATCTTCATCAGAGCAGGGTTAGGTTGGGATGGATCG
Igf1r	CTCCCCGTGCATTATTGGCCATGTCTCTCTGGTTCCACAAGATACCGTTG
Odc1	CACTTTTCAGACATTGATACTAACGTGTGCCCCTCAGCTGCTGAGCAAGC

**Tet1* not included as it did show significant differential expression on the array (see section 2.2.5)

Table 2.6: The primer sequences and amplicon length of the candidate genes that underwent real-time PCR analysis

Gene	Forward Primer	Reverse Primer	Amplicon Length (Bp)	Tissues
Vwf	CAGCCAAGGTGCAGTACATC	TCAATGGAGTACACAGCTTTGC	106	Brain, Placenta & Liver
Grm7	CCCAATGGTGAGGCAAAGAC	CCTCTGTGGTTCCATGGGAT	127	Brain
Zfp317	TGGGAAGGAAGCATCTGATGA	AGCAAGGTGAGGGCCAATAT	109	Brain
Psmg2	ACAATATCCCAGATGCCGTCA	CTTCCAAGGCAGGGCAGAT	104	Brain
Pcbp3	GCGCCAGATTACCATCACAG	AGGTTAGAGTGCACCCATCC	106	Liver
Igf1r	GGACAACTGCCCTGATATGC	CCTTGATGCTGCCGATGATC	101	Liver, Brain
Depdc 1b	TAGCCGACGACAAACTCTCC	CGCTCTCTGGGGTAGGAAAT	150	Liver, Brain
Odc1	GGTTCCAGAGGCCAAACATC	CTCTCCTGGGCACAAGACAT	139	Brain, Placenta
Tet1	CCATTCTCACAAGGACATTCACA	GCAGGACGTGGAGTTGTTCA	116	Brain

2.8 Reverse Transcription of RNA into cDNA

Prior to real-time PCR, the total RNA extracted from the embryonic tissue was reverse transcribed into cDNA using the Thermo Scientific Revert Aid H Minus First Strand cDNA synthesis Kit (Wiame, 2000) (full protocol in Appendix H) as per the manufacturers instructions. Reverse transcriptase (RT) uses an RNA template and a short primer complementary to the 3' end of the RNA to direct the synthesis of the first cDNA strand, which can be used directly as a template for accessing the expression by real-time PCR. All cDNA samples were diluted to 1:20 using ddH_20. The quality of the cDNA was

evaluated by running 4μl on a 1% agarose gel against a 1kb+ ladder. Good quality cDNA was expected to appear as a smear and degraded cDNA to form a laddering pattern.

2.9 Selecting an endogenous control

In all gene expression studies, it is essential to include an endogenous control to correct for differences in RNA sampling and to normalise the results. It is ideal for the chosen endogenous control to have a constant RNA transcription level under various experimental conditions and be significantly abundant across multiple cell types and tissues (QIAGEN, 2004).

Housekeeping genes are commonly selected as they are constitutively expressed. Genes such as Beta-actin (*Actb*), Glyceraldehyde-3-phosphate dehydrogenase (*Gapdh*) and 18S ribosomal RNA (*18S rRNA*) are widely used as housekeeping genes in gene expression studies.

Actb was chosen as the endogenous control for this study. According to literature there is a large amount of variability in the expression of *Gapdh* among different tissue types and it is very highly expressed in the brain whereas the expression of *Actb* remains constant under a wide range of physiological conditions (QIAGEN, 2004, Barber et al., 2005).

2.10 Relative gene expression analysis using quantitative real-time PCR

Real time PCR was developed due to the need to quantify differences in mRNA expression and because of the availability of only small amounts of mRNA in some procedures (Hunt, 2006). Real-time PCR is a technique for collecting data throughout the PCR process as it occurs, thus combining amplification and detection into a single step (Wong and Medrano, 2005). The real-time PCR technique, where the fluorescent signal reflecting the DNA accumulation is detected in every amplification cycle, has simplified the quantification of

gene expression levels.

2.10.1 SYBR green based method of detection

There are currently four different chemistries available for real-time PCR namely; TaqMan® (Applied Biosystems, Foster City, CA, USA), Molecular Beacons, Scorpions® and SYBR® Green (Basel, Switzerland). Each of these methods allows detection of PCR products via the generation of a fluorescent signal. The fluorescent signal is generated via the coupling of a fluorogenic dye molecule and a quencher moiety to the same or different oligonucleotide substrates. In this project, the SYBR green method was used to monitor cDNA synthesis. This is the simplest method for detecting and quantifying PCR products in real-time reactions. SYBR green is a fluorogenic dye that exhibits little fluorescence when in solution, but emits a strong fluorescent signal upon binding to double-stranded DNA (Dorak, 2012). Therefore, as the PCR product accumulates, the fluorescence increases (**Figure 2.8**). This method is sensitive and easy to use, however, SYBR green binds to any double-stranded DNA in the reaction, including primer-dimers, which can result in an overestimation of the target concentration. Therefore, this method requires extensive PCR optimisation.

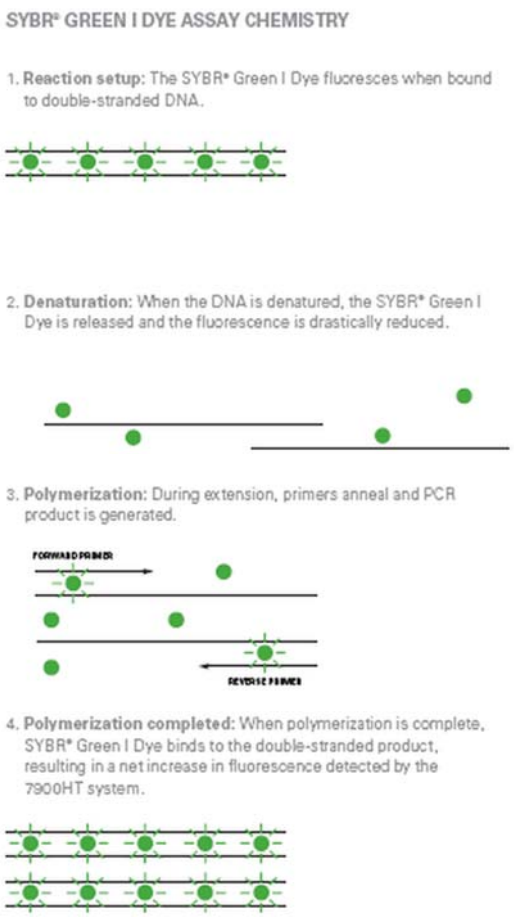


Figure 2.8: Schematic of the SYBR green based real time assay. SYBR green fluorescence is increased upon binding to double stranded DNA, therefore as the PCR amplification progresses, there is an increase in fluorescence (Thermo Scientific., 2015).

2.10.2 Melting curve analysis to confirm primer specificity

The chosen reaction conditions and primer pairs determine the specificity of the real time PCR assay. However, there is always the possibility that even well designed primers may form primer-dimers or amplify a non-specific product. There is also the possibility when performing real-time PCR that the RNA sample contains genomic DNA contamination.

The LightCycler (Roche, Basel, Switzerland) melting curve analysis was used to assess whether the real-time PCR assays for the ten candidate genes and the housekeeping gene would produce single, specific products. The single product is depicted as a single peak with no shoulders as shown in **Figure 2.9** (an example of one of the candidate genes

analysed). No primer-dimers were generated during the real-time PCR amplification cycles for any of the genes. In addition, specificity of RT-PCR products was documented with high-resolution gel electrophoresis and resulted in a single product with the desired length.

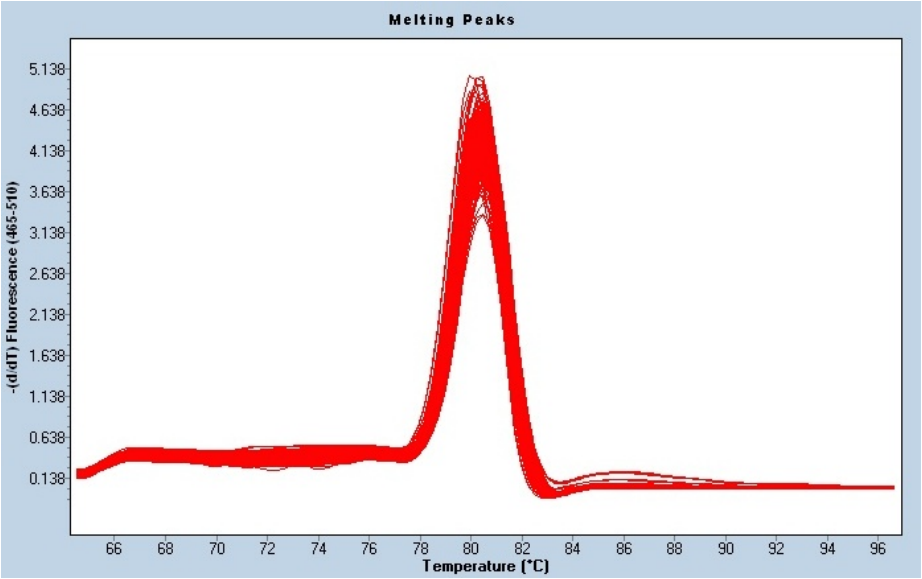
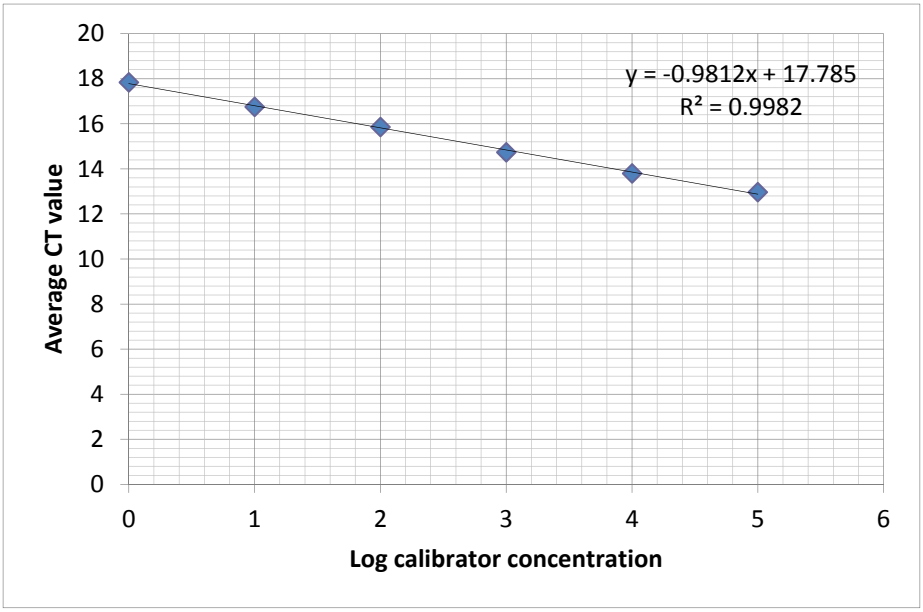


Figure 2.9: Melting curve analysis confirming specific amplification with *Depdc1b* primers. The single peak confirms the PCR amplification is specific to the region of interest.

2.10.3 PCR efficiency analysis using serial dilutions

The PCR efficiency can be determined both for the target gene and the reference (housekeeping) gene by real-time PCR analysis of serial dilutions and the construction of a standard curve. Therefore, the amplification of each primer pair was determined using a six point, two-fold serial dilution of cDNA (1:5, 1:10, 1:20, 1:40, 1:80, 1:160) that was pooled from all the samples. Dilutions were created using ddH_20. Each sample within a reaction plate was performed in triplicate. A standard curve was then constructed by plotting the log of the starting quantity of template (or the dilution factor, for unknown quantities) against the C_T (cycle threshold) value obtained during amplification of each dilution ($E = 2^{1/slope}-1$). The C_T is defined as the number of cycles required for the fluorescent signal to cross the threshold (exceed background levels). The equation of the linear regression line, along with the coefficient of determination (R^2) can then be used for evaluating whether the real-time

PCR assay has been optimised. At 100% efficiency, the template doubles after each cycle. It is unlikely that reactions will always be 100% efficient and thus a range of 85% to 105% can be considered optimised. Factors affecting efficiency include: the length of the amplicon, the G/C content of the amplicon, and secondary structure. The primer pairs for all the candidate genes were greater than 90% and therefore accepted as efficient (graphs in Appendix G).



Beta Actin amplification efficiency: $E = 2^{1/0.9812}-1$

$E = 1.02$

This is approximately 100% efficient

Figure 2.10: The standard curve used to analyse the efficiency of *Actb*

2.10.4 Real Time PCR preparation and run using the Roche Light Cycler 480

Once all the primers were optimised and the PCR efficiency was determined for all the genes, real time PCR was used to amplify the candidate genes in the cDNA from

embryonic tissues (brain, liver or placenta) of embryos sired by EtOH treated males and sucrose treated males. The real time PCR protocol and the thermal cycling conditions of the real time PCR reactions are represented in **Table 2.7**. Each run analysed a specific gene in a particular tissue (brain, liver or placenta) and included:

- The six dilution samples that formed the standard curve (1/5, 1/10, 1/20, 1/40, 1/80, 1/160).
- The 48 samples (24 EtOH and 24 sucrose) that were run in triplicate. It must be noted that the concentration of every cDNA sample fell within the dilution range of the standard curve.
- A reverse transcriptase negative sample, produced by pooling together all of the reverse transcriptase negative samples.
- A minus-reverse transcriptase control sample. This sample did not include the reverse transcriptase when converted from RNA to DNA. This sample serves as an additional control because if anything is amplified in this sample, this is indicative of DNA contamination.
- A ddH_20 sample ran in triplicate.

The real-time PCR reactions were run on the Light Cycler 480 (Roche, Basel, Switzerland) on white 384 well PCR plates covered with an optically clear film.

Table 2.7: The general real-time PCR reagents and thermal cycling conditions for the amplification of the candidate genes

PCR reagents	**Volume (μl) (1X mix)**		
ddH_20	0.5		
cDNA	3		
Roche SYBR green	4		
Primers (Forward and Reverse) 10μM	0.5		
Total volume	8		
Thermal Cycling conditions	**Temperature (°C)**	**Time (min: sec)**	**Cycles**
Pre incubation	95	10:00	1
Amplification	95	00:15	X45
	60	00:30	
Cooling		30 seconds	1

*After cooling the programme would be set to run the melting curve.

2.10.5 Standard curve method of relative quantification

Real-time PCR reactions are characterised by the cycle where the target amplification is first detected. This cycle is referred to as the cycle threshold (C_T) (Wong and Medrano, 2005). The real-time thermocycler creates an amplification profile for each sample as the amplification occurs. The success of the amplification can be determined from the amplification profile (successful amplifications are displayed as sigmoidal-shaped amplification plots). The Light Cycler 480 (Roche, Basel, Switzerland) software automatically calculated the C_T values. The software essentially selects 3 points on the linear section of the amplification curve and calculates an average. These values represent the cycle at which the first significant increase in PCR product was detected. Using the C_T values, the differential expression was calculated and normalised by the reference gene. There are two methods, both equally valid, for analysing data obtained from real time PCR: Relative Standard Curve method and Comparative C_T method. The relative standard curve method is usually used when researchers have a large number of candidate genes and a limited amount of samples, whereas the comparative C_T method is generally used when investigators have a limited number of genes but a large number of samples (RRC Core Genomics, 2003).

The real-time PCR results from this project were analysed using the relative standard curve method. This approach requires running a standard curve for both the gene of interest and the endogenous control. The template for the standard curve can be any cDNA sample that expresses both the gene of interest and endogenous control. It is important to note that the expression level of the sample being analysed should fall within the limits of your standard curve (RRC Core Genomics, 2003).

Once the threshold cycle values (C_T) for the real time experiment are obtained, the next step is to transform the sample data into units defined by the standard curve (RRC Core Genomics, 2003). This is done using the following equations:

For the natural logarithm:

- Sample (ng)= C_T value - y intercept/slope

Log input value is then converted to a real number:

- Real number = 2^{sample}

Once a sample has a value defined by the standard curve, the value needs to be normalised to the endogenous control (e.g. *Actb*):

- Sample (ng)/*Actb* (ng)

This will give the expression level of the sample relative to the expression level of the housekeeping gene.

If the fold change value is preferred, it can be calculated as follows:

- Average of the test group/ Average of the control group

This method has its advantages, as it requires the least amount of validation because the PCR efficiencies of the target and endogenous control do not have to be equivalent.

However, the method does require that each reaction plate contain standard curves, and requires more space on a reaction plate. Highly accurate quantitative results are obtained using this method as values are interpolated from the standard curve (Applied Biosystems, 2004).

2.10.6 Real-time PCR data quality control

As a consequence of experimental error, some samples may amplify insufficiently or not at all, producing C_T values that may differ significantly from their associated technical replicates. If these samples are included in the analysis, they may lead to erroneous measurements. To ensure precise relative quantification, it is important to identify replicate group outliers. Each cDNA sample was run in triplicate to rule out experimental bias or random error. If one value within the triplicate deviated from the other two resulting in a standard deviation above 0.25, the value was omitted. However, only a maximum of three values could be omitted per run. If more than three samples had to be excluded from the analysis, the plate was re-run and the gene expression levels were reanalysed.

2.11 Statistical analysis of pyrosequencing and real-time PCR data

Statistical analyses were performed using STATISTICA, version 12 (StatSoft of Inc., UK) and GraphPad Prism, version 6 (GraphPad Software Inc.La Jolla, CA).

2.11.1 Descriptive statistical analysis

The data analysed by descriptive statistics are described in terms of central tendency, specifically, mean levels, median and the mode. The box plots and histograms were generated to analyse the distribution of the data and identify outliers. The mean, median and standard deviation were calculated for each candidate gene. The Lilliefors test was used to test for normality of the data. The p-value generated by the Lilliefors test denotes either normal distribution ($p<0.05$) or non-normal distribution ($p>0.05$). The DNA

methylation and relative gene expression levels per candidate gene are graphically represented as column scatter plots (Chapter 3). The standard deviation is displayed on each graph.

2.11.2 Mann-Whitney U test

The data generated in the present study is continuous. The independent variable in this study is the treatment group: EtOH or sucrose. A t-test would typically be used to analyse continuous data. However, the t-test requires a sample size of >20 and assumes that the data follows a normal distribution. As a consequence of the small sample size and the non-normally distributed data generated in this study (determined using the Lilliefors test), the non-parametric two Mann-Whitney U test was selected for data analysis. This test was used to determine whether there was significant difference in:

(1) Sperm DNA methylation levels between male mice exposed to chronic amounts of EtOH in comparison to male mice exposed to sucrose.

(2) Expression of candidate genes in the brain, liver and placenta of E.16.5 embryos sired by EtOH and sucrose treated male mice

(3) Levels of DNA methylation in the embryonic tissues of embryos sired by EtOH or sucrose treated male mice.

2.12 Functional enrichment analysis of genes that showed a significant change in embryonic gene expression in the brain, liver and placenta

To gain insights into the biological functions and pathways of the genes differentially expressed (according to the microarray data) in the embryos sired by males exposed to alcohol, a functional classification and functional enrichment analysis was performed using

PANTHER (Thomas et al., 2003) classification system and the DAVID (Huang et al., 2005, Huang et al., 2009) respectively.

Four candidate gene lists were submitted to DAVID (Huang et al., 2005, Huang et al., 2009) and PANTHER (Thomas et al., 2003) for analysis. The gene lists were generated based on information extracted from volcano plots (Chapter 3, section 3.4.1). A volcano plot is a scatter plot often used to visualise microarray expression results. Volcano plots allow for the fold change value to be compared to the statistical significance level by arranging genes along dimensions of biological and statistical significance. The x-axis represents the fold change between the two groups (using a log scale so the up and downregulated genes appear symmetric), and the y-axis represents the p-value (usually using a negative log-scale so smaller p values appear toward the top). The x-axis (fold change) is usually indicative of the biological impact of the change, whereas the y-axis (p value) signifies the statistical significance, or the reliability of the change. The most promising candidate genes may then be selected for further analyses.

Three volcano plots were generated using R (R Core Team, 2013) as per the command in Appendix I. Genes that are highly dysregulated are further to the left (downregulated) and right (upregulated) sides, while highly significant changes appear higher on the plot.

All the genes represented by red dots on the volcano plots (Chapter 3, section 3.4.1) were submitted to PANTHER (Thomas et al., 2003) and DAVID (Huang et al., 2005, Huang et al., 2009) for analysis. Contrary to the candidate genes analysed through pyrosequencing and real time PCR, these lists were generated solely on the microarray expression data in attempt to re-examine the expression data independently of the sperm methylation data.

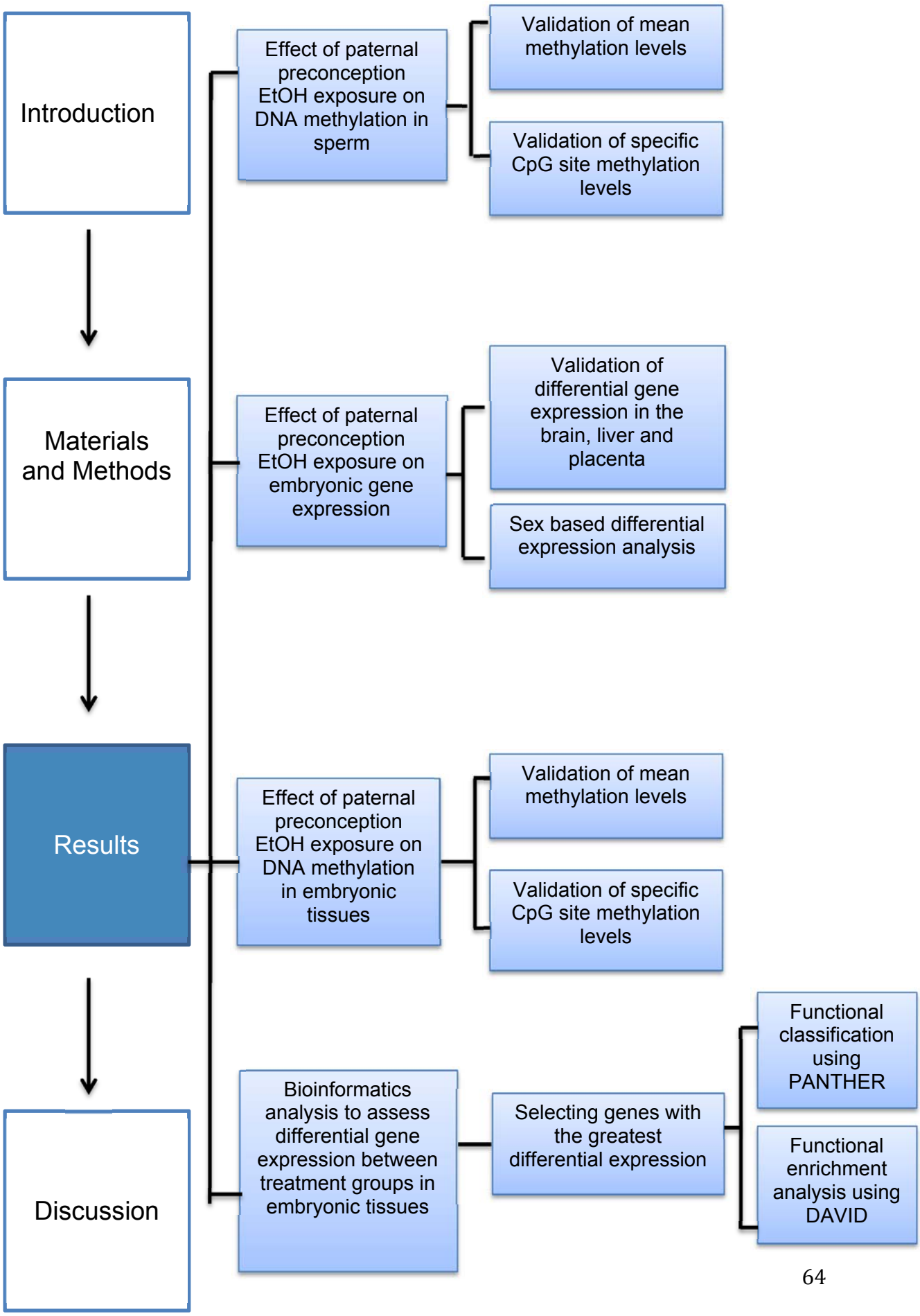
The functional classification analysis performed by PANTHER (Thomas et al., 2003), provides a general overview of the molecular functions and biological processes of the genes submitted for analysis and displays the results as a pie chart (Chapter 3, section 3.4.2). Whereas DAVID (Huang et al., 2005, Huang et al., 2009) offers an integrated

approach to identify functions or pathways significantly enriched in the gene list by comparing it to a background list and then clusters them according to similarity. The principle foundation of an enrichment analysis is that if a biological process or molecular function is altered in in a given study, there is a higher chance that the co-functioning genes would be selected as a relevant group by the high-throughput screening technologies. The Illumina MouseWG-6_V2_0_R2_11278593 was selected as the background list because the whole genome expression array was carried out using this Chip. To ensure that the results provided are not due to chance the gene enrichment analyses corrected for multiple testing and is supported by a modified Fishers exact test. Functions included in the analyses output included: gene ontology terms, Swiss-Prot, InterPro matches, Online Mendelian Inheritance in Man (OMIM) and Kyoto Encyclopedia of Genes and Genomes (KEGG).

Chapter 3

Results

3.0 Results


Introduction
Materials and Methods
Results
Discussion
Effect of paternal preconception EtOH exposure on DNA methylation in sperm
Validation of mean methylation levels
Validation of specific CpG site methylation levels
Effect of paternal preconception EtOH exposure on embryonic gene expression
Validation of differential gene expression in the brain, liver and placenta
Sex based differential expression analysis
Effect of paternal preconception EtOH exposure on DNA methylation in embryonic tissues
Validation of mean methylation levels
Validation of specific CpG site methylation levels
Bioinformatics analysis to assess differential gene expression between treatment groups in embryonic tissues
Selecting genes with the greatest differential expression
Functional classification using PANTHER
Functional enrichment analysis using DAVID

Results

The primary objectives of this study were to: (1) validate significant changes in sperm DNA methylation (as observed by RRBS data analysis) in male mice exposed to alcohol or sucrose for ten weeks; (2) validate significant changes in gene expression (as observed by whole genome expression array data analysis) of their sired embryos for ten genes; (3) quantify DNA methylation in candidate genes in the tissues of embryos sired by ethanol exposed males and compare this to the tissues of embryos sired by sucrose treated male mice. (4) To reanalyse the microarray expression data generated in a previous study to construct a list of genes that displayed a significant ($p<0.001$) increase or decrease in expression with a fold change greater than 1.2 in the embryonic brain, liver or placenta and to identify biological pathways and functions significantly overrepresented among these genes.

DNA methylation in the sperm and embryonic tissues were quantified by pyrosequencing. Gene expression in embryonic tissues was quantified using real time PCR and the SYBR green based method of detection. The functional enrichment analysis of the candidate list of genes obtained from the microarray expression data was performed using PANTHER (Thomas et al., 2003) and DAVID v6.7 (Huang et al., 2005, Huang et al., 2009).

3.1 The effect of preconception EtOH exposure on DNA methylation in sperm of male mice

Sperm DNA methylation was quantified using pyrosequencing to assess the methylation status of specific CpG sites within a list of candidate genes (**Table 3.1**). Quality control measures were applied when generating the methylation data. These measures included amplification of samples in technical triplicates, and amplifying samples in one machine, in order to reduce variation associated with different PCR machines.

Following PCR, primer specificity was assessed using agarose gel electrophoresis to determine amplicon sizes, before pyrosequencing on the PyroMark MD System (Biotage, Uppsala, Sweden). In cases where the incorrectly sized PCR product was amplified, primers were redesigned and the assay was reassessed. The average methylation percentage for each CpG site was calculated by averaging the triplicate values for each CpG site for each sample. When intra-sample variation of greater than 10% was observed, the results were excluded from the study. The average methylation percentage per CpG site was then averaged across all the CpG sites analysed per gene to give the mean methylation percentage per candidate gene.

Batch effects were minimised by using a 96 well plate for the pyrosequencing PCR and randomly selecting samples from the different treatment groups (EtOH or sucrose) in the pyrosequencing run for a specific gene of interest. Each plate also contained a no template control (negative control) and a water sample. As pyrosequencing was conducted in an attempt to validate previous significant ($p<0.05$) RRBS findings, pyrosequencing data are always displayed in parallel with RRBS data. The RRBS and pyrosequencing data for the sperm methylation levels within the candidate genes is graphically represented in **Figures 3.1-3.6**.

3.1.1 Validation of RRBS sperm DNA mean methylation data using pyrosequencing

For sperm samples, DNA methylation data was obtained for five out of the ten candidate genes, due to low DNA concentration and volumes. The mean methylation levels of the candidate genes of each treatment group is shown in **Table 3.1** and graphically displayed in **Figure 3.1**. The non-parametric Mann-Whitney U test was used to assess if there was a significant difference in the DNA methylation in the male mice exposed to EtOH (test) in comparison to the males exposed to sucrose (control). This test was selected due to the small sample size. Each dot on the column scatter graph (**Figure 3.1**) represents the average methylation of a sample that was performed in triplicate. All pyrosequencing data is graphically displayed using navy (sucrose) and dark red (EtOH) contrary to the RRBS data, which is depicted using a light blue (sucrose) and light pink (EtOH).

Table 3.1: Mean methylation levels across the CpG sites of the candiate genes between the treatment groups

Gene	Treatment	N	CpG sites analysed	Mean Methylation Level ± SD (%)	Median (Range)	P-value
Grm7	EtOH	8	1	92.05 ± 1.23	91.68 (90.83,94.73)	0.94
	Sucrose	6		92.11 ± 2.20	91.76 (88.41,94.77)	
Tet1	EtOH	8	2	49.18 ± 4.17	48.53 (42.96,56.07)	0.03 *
	Sucrose	6		44.19 ± 4.09	43.77(36.95,52.43)	
Vwf	EtOH	9	2	33.94 ± 3.75	33.99 (27.82, 40.49)	0.06 ^
	Sucrose	8		36.32 ± 4.79	37.58 (24.58,43.46)	
Odc1	EtOH	7	4	74.26 ± 8.59	74.47 (55.79,90.66)	0.07 ^
	Sucrose	7		68.93 ± 10.67	70.40 (45.57,83.85)	
Depdc1b	EtOH	9	3	53.94 ± 3.90	54.29 (47.44,65.30)	0.01 *
	Sucrose	6		57.45 ± 4.98	57.58 (47.64, 63.87)	

EtOH- Ethanol treated
SD- Standard Deviation
P values generated using the Mann-Whitney U test
Statistically significant ** ($P \leq 0.01$) * ($0.01 < P \leq 0.05$)
Trending significance ^ ($0.05 \leq P < 0.1$)

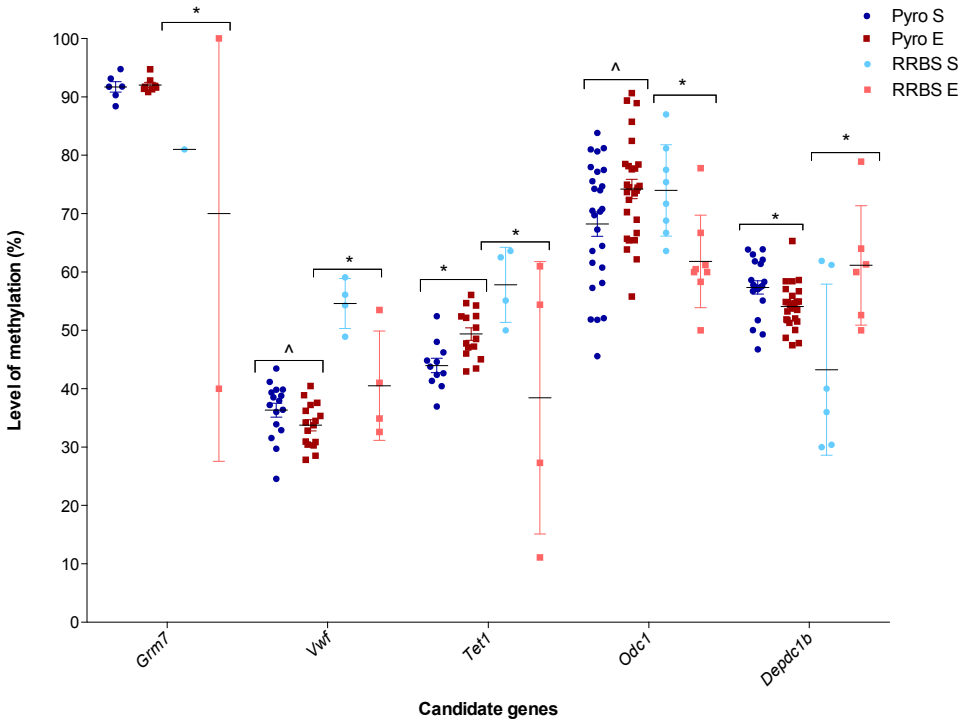


Figure 3.1: Validation of RRBS mean methylation data for five candidate genes using pyrosequencing. Comparison of the mean methylation levels and the techniques used to detect the methylation changes between sires exposed to EtOH or sucrose for ten weeks. Statistically significant ** ($p \leq 0.01$) * ($0.01 < P \leq 0.05$) Trending significance ^ ($0.05 \leq P < 0.1$) Error bars indicate the standard deviation.

According to both the pyrosequencing and RRBS data, *Vwf* in the EtOH exposed male mice (33.94%) showed lower mean methylation levels than their respective sucrose exposed controls (36.32%), validating the observation from the RRBS data. However, *Tet1*, *Odc1* and *Depdc1b*, showed opposite effects in the original experiment and the validation. Quantitative pyrosequencing showed higher mean methylation levels of *Tet1* and *Odc1* in the EtOH group but lower mean methylation by RRBS when compared to the control group. *Depdc1b* exhibited lower mean methylation in the EtOH treated male mice relative to the control group through pyrosequencing but higher mean methylation in the EtOH group according to the RRBS data. Statistical analysis of the pyrosequencing data revealed a significant difference ($p<0.05$) in methylation between the test (EtOH) and control (sucrose) male mice for *Tet1* ($p=0.03$) and *Depdc1b* ($p=0.01$). *Vwf* ($p=0.06$) and *Odc1* ($p=0.07$) appear to be trending toward significance ($p<0.1$). *Grm7* ($p=0.94$) showed no difference in mean methylation between the test and control groups with the

pyrosequencing validation. On a gene-averaged level, only one out of the five genes (*Vwf*) consistently demonstrated a DNA methylation change in the same direction with a trend toward significance (p=0.06).

3.1.2 Validation of RRBS sperm DNA methylation data at specific CpG sites within candidate genes

The methylation levels for the specific CpG sites per candidate gene are shown in **Table 3.2** and graphically displayed in **Figures 3.2 to 3.6**. The non-parametric Mann-Whitney U test was used to assess whether significant differences in methylation between male mice exposed to EtOH in comparison to male mice exposed to sucrose were observed. All CpG sites that are bolded in the x-axis of **Figures 3.2 to 3.6** were analysed by both pyrosequencing (present study) and RRBS (previous study). CpG sites that are not bolded and italicised only have RRBS data available. Each dot on the column scatter graphs represents the average methylation of a sample that was performed in triplicate. All pyrosequencing data points are graphically displayed using navy (sucrose) and dark red (EtOH) against the RRBS data, which is depicted in a light blue (sucrose) and pink (EtOH).

Table 3.2: CpG site methylation analysis within the candidate genes between the treatment groups

Gene	CpG site	Treatment	N	Methylation level ± SD (%)	Median (Range)	P value
Grm7	111543873	EtOH	8	92.04 ±1.23	91.68 (90.83,94.73)	0.94
		Sucrose	6	91.68±2.20	91.76(88.41, 94.77)	
Tet1	62329009	EtOH	8	51.99±3.11	46.22 (42.64, 52.43)	0.03*
		Sucrose	5	46.62±3.87	52.43 (47.09, 56.07)	
	62329050	EtOH	7	46.37±3.11	46.02 (42.96, 52.14)	0.01**
		Sucrose	6	41.77±2.93	41.86 (36.95, 44.84)	
Vwf	125553783	EtOH	9	32.38±3.86	30.86 (27.82, 38.90)	0.54
		Sucrose	8	33.35±4.88	33.41(24.58, 39.35)	
	125553787	EtOH	7	35.49±2.98	35.36 (30.94,40.49)	0.02 *
		Sucrose	8	39.30±2.29	39.16 (26.39, 43.36)	
Odc1	17514120	EtOH	7	73.89±3.65	74.47 (68.95, 78.51)	0.31
		Sucrose	7	69.03±9.33	70.47 (52.07, 80.65)	
	17514135	EtOH	7	80.56±5.98	77.72 (74.76, 90.66)	0.63
		Sucrose	5	78.27±2.99	77.99 (73.99, 81.22)	
	17514144	EtOH	6	75.31±13.22	76.19 (55.79, 89.37)	0.18
		Sucrose	7	63.70±13.40	61.57 (45.57, 83.85)	
	17514154	EtOH	7	67.29 ± 4.54	65.68 (62.18, 73.69)	0.62
		Sucrose	7	64.72±8.85	63.62 (51.72, 61.39)	
Depdc1b	109101087	EtOH	7	52.31±2.96	53.84(47.82, 54.86)	0.001 **
		Sucrose	6	58.25±1.70	57.86 (56.72, 61.39)	
	109101098	EtOH	7	53.61±3.64	54.63(47.44, 58.42)	0.08 ^
		Sucrose	6	58.53±5.63	59.94 (49.31, 63.87)	
	109101330	EtOH	9	55.91±4.36	54.87(51.28, 65.30)	0.83
		Sucrose	7	55.59±6.30	57.35(46.74, 63.84)	

P-values generated using the Mann-Whitney U test.
SD- Standard Deviation
Statistically significant ** ($p \leq 0.01$) * ($0.01 < P \leq 0.05$)
Trending toward significance ^ ($0.05 \leq P < 0.1$)

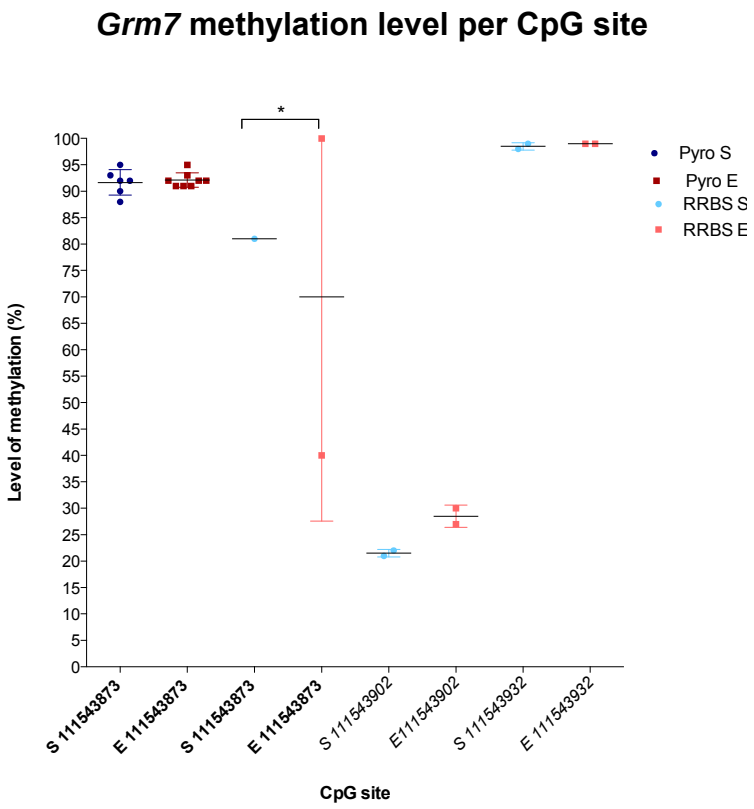


Figure 3.2: Validation of *Grm7* RRBS CpG site-specific sperm DNA methylation data of male mice exposed to EtOH or sucrose using pyrosequencing. CpG site 11154873 was analysed for validation. Statistically significant ** (P ≤ 0.01) * (0.01< P ≤ 0.05) Trending significance ^ (0.05 ≤ P < 0.1) Error bars indicate the standard deviation.

According to the pyrosequencing data in **Figure 3.2**, no difference in methylation was observed between the treatment groups for *Grm7* at CpG site 11154873. This is in contrast to the RRBS data, which demonstrated a significant decrease ($p<0.05$) in methylation in the EtOH treated sires. However, one of the RRBS controls had no coverage for CpG site 11154873. No differences in methylation were observed at CpG sites 111543902 and 111543932 according to the RRBS data.

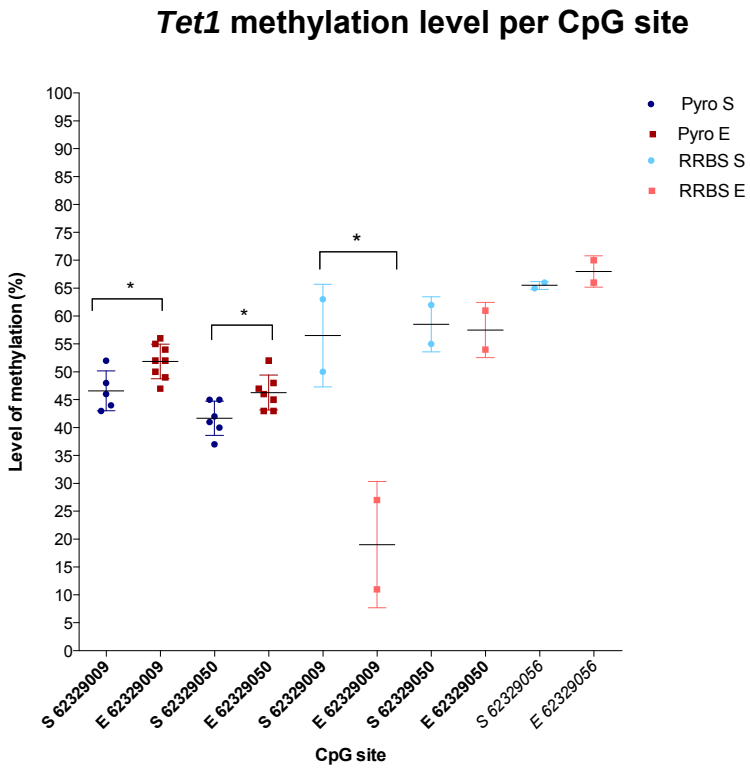


Figure 3.3: Validation of *Tet1* RRBS CpG site-specific sperm DNA methylation data of male mice exposed to EtOH or sucrose using pyrosequencing. CpG site 62329009 and 62329050 were analysed for validation. Statistically significant ** (P ≤ 0.01) * (0.01< P ≤ 0.05) Trending significance ^ (0.05 ≤ P < 0.1) Error bars indicate the standard deviation

Both *Tet1* CpG sites (62329009 and 62329050) (**Figure 3.3**) assessed by pyrosequencing exhibited a statistically significant ($p<0.05$) increase in methylation in the EtOH treated males in comparison to their respective sucrose treated control males. This is contrary to the RRBS data in which the same CpG sites showed decrease in methylation levels in the EtOH treated group in comparison to the sucrose control group. Furthermore, the RRBS data suggests that only CpG site 62329009 displayed a statistically significant ($p<0.05$) decrease in methylation.

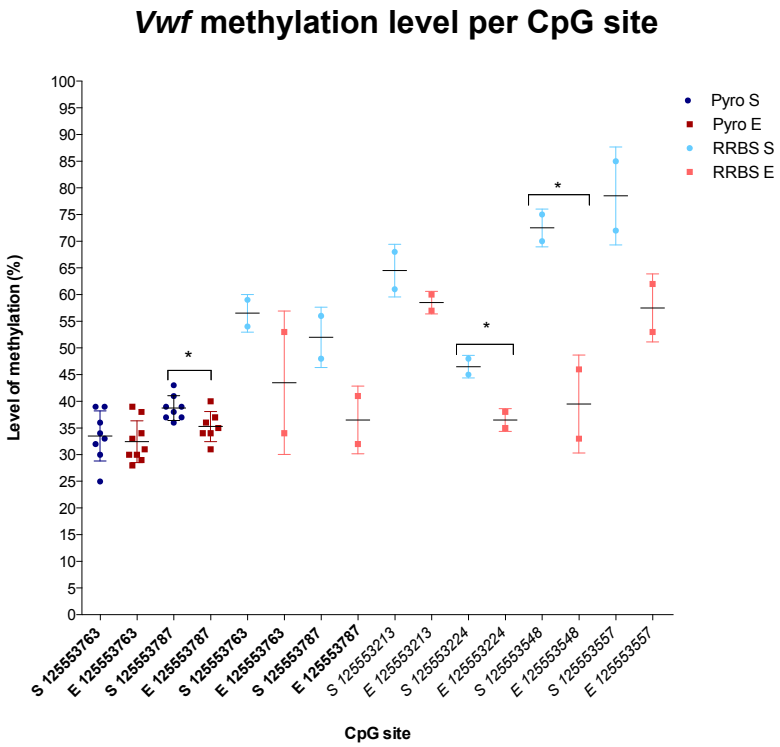


Figure 3.4: Validation of *Vwf* RRBS CpG site-specific sperm DNA methylation data of male mice exposed to EtOH or sucrose using pyrosequencing. CpG sites 125553763 and 125553787 were analysed for validation. Statistically significant ** (P ≤ 0.01) * (0.01< P ≤ 0.05) Trending significance ^ (0.05 ≤ P < 0.1) Error bars indicate the standard deviation.

As shown in **Figure 3.4**, pyrosequencing results revealed that *Vwf* CpG sites 125553763 and 125553787 displayed a decrease in methylation in the male mice exposed to EtOH in comparison to the sucrose-exposed males. Similar results were observed in the RRBS data. Statistical analysis revealed that only CpG site 125553787 showed a significant decrease in methylation (p=0.02) according to the pyrosequencing data, which is in contrast to the statistical analysis of the RRBS data in which the site was not found to be significant. However, the RRBS data found CpG sites 125553224 (p=0.02), and 125553548 (p=0.04) to be statistically significant (sites not analysed by pyrosequencing). Although different CpG sites were statistically significant in the RRBS and pyrosequencing data sets, both suggest that EtOH results in a decrease in methylation within the target region.

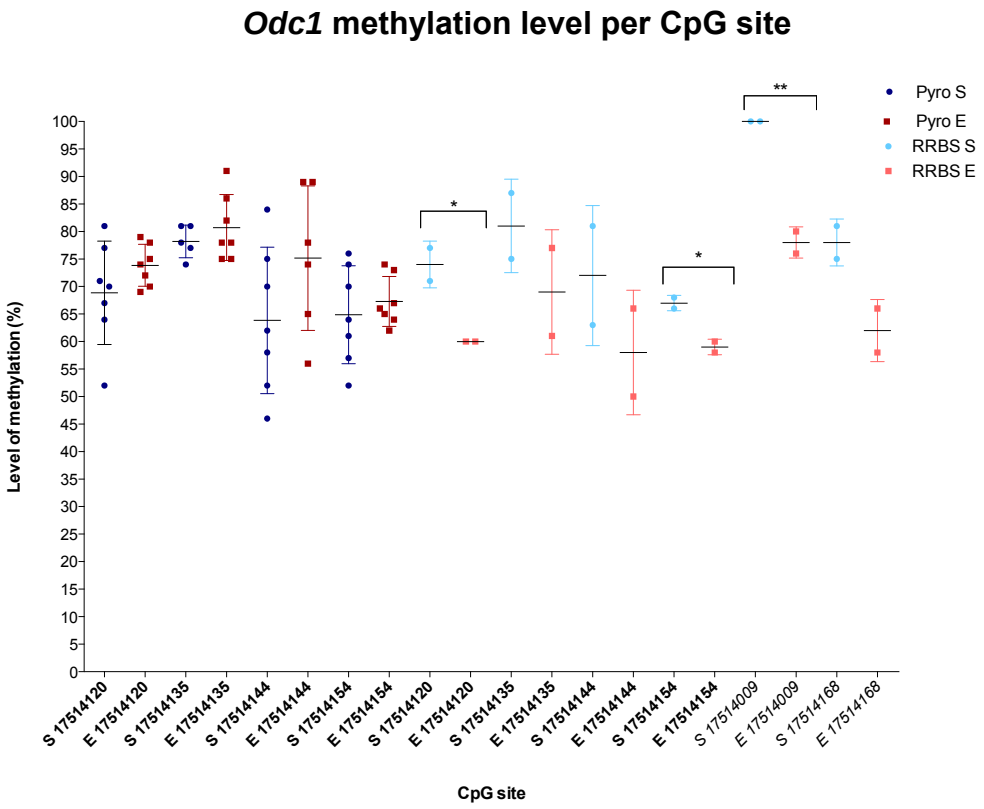


Figure 3.5: Validation of *Odc1* RRBS CpG site-specific sperm DNA methylation data of male mice exposed to EtOH or sucrose using pyrosequencing. CpG sites 17514120, 17514135, 171514144 and 17514154 were analysed for validation. Statistically significant ** (P ≤ 0.01) * (0.01< P ≤ 0.05) Trending significance ^ (0.05 ≤ P < 0.1) Error bars indicate the standard deviation.

The four *Odc1* CpG sites (**Figure 3.5**) analysed by pyrosequencing exhibited an increase in methylation in the sperm of male mice exposed to EtOH in comparison to the sucrose group. This is in contrast to the RRBS data in which the same CpG sites displayed a decrease in methylation in the EtOH relative to the controls. No significance differences in methylation were observed at any of the CpG sites in the pyrosequencing data. The RRBS data found CpG site 17514120, 17514009 and 17514154 to be statistically significant ($p<0.05$).

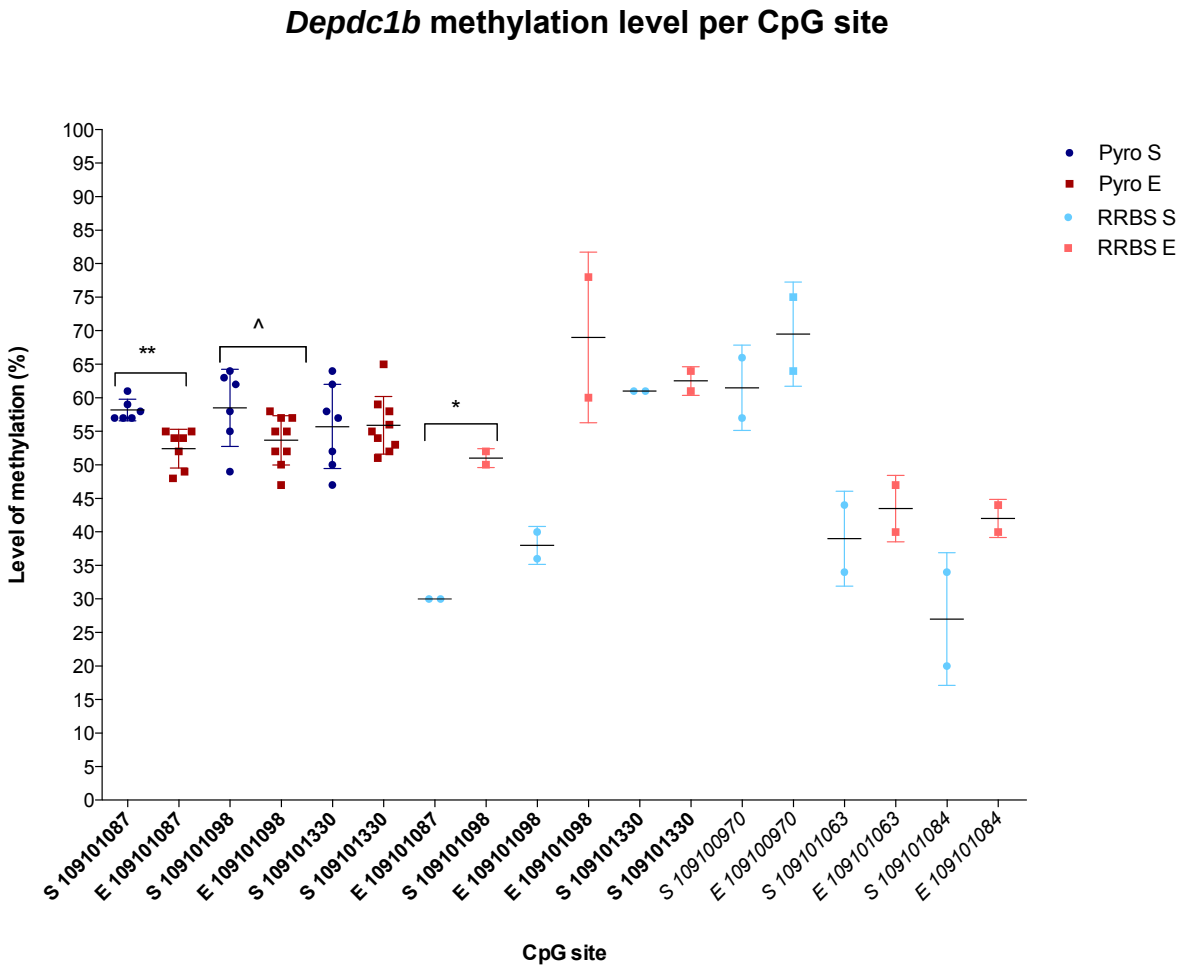


Figure 3.6: Validation of *Depdc1b* RRBS CpG site-specific sperm DNA methylation data of male mice exposed to EtOH or sucrose using pyrosequencing. CpG sites 109101087, 109101098, 109101330 were analysed for validation. Statistically significant ** (P ≤ 0.01) * (0.01< P ≤ 0.05) Trending significance ^ (0.05 ≤ P < 0.1) Error bars indicate the standard deviation.

In line with the RRBS data, the pyrosequencing data for *Depdc1b* showed a significant alteration (p=0.001) in methylation at CpG site 109101087 (**Figure 3.6**). However, the pyrosequencing data suggests a significant decrease in methylation at the specific CpG site whereas the RRBS data displayed a significant increase (p=0.03). According to the pyrosequencing data, CpG site 109101098 displayed a decrease in methylation and is trending toward significance (p=0.08), this site was not found to be statistically significant through RRBS data analysis.

In summary, a CpG site specific analysis confirmed the outcome of the gene averaged analysis. Only *Vwf* showed the same direction of change that was significant. The RRBS

study results were not validated for the other four genes tested (*Tet1*, *Odc1*, *Depdc1b*, *Grm7*).

3.2 The effect of preconception EtOH exposure on gene expression in embryonic tissues

The methodology and quality control measures used in gene expression assessment included RNA samples being quantified, normalised and reverse transcribed to cDNA. This was followed by PCR amplification and agarose gel electrophoresis. Primer specificity was assessed by analysing the melting curves of each primer set. The efficiency of each reaction was evaluated using a standard curve. Each plate included the serially diluted cDNA samples for the standard curve, the test and control samples (assessed in triplicate), a reverse transcriptase negative control, and a water sample. The plate was loaded into the Roche Light Cycler 480 (Roche, Basel, Switzerland). The real-time thermocycler creates an amplification profile for each sample as the amplification occurs. The success of the amplification can be determined from the amplification profile (successful amplifications are displayed as sigmoidal-shaped amplification plots). The C_T values were automatically calculated by the Roche Light Cycler 480 software. Using the C_T values, the differential expression was calculated and normalised by the reference gene (*Actb*).

As the real-time PCR was performed in an attempt to validate previous significant microarray findings, real-time PCR results are always displayed in parallel with the microarray data. The expression array graphs have fewer data points as the samples were pooled and averaged. Twenty-four RNA pools were representative of the three tissue types, (brain, liver and placenta) and the two treatment groups, two biological replicates were used to profile the expression patterns. For validation purposes, all the samples that were run on the expression array were analysed individually in technical triplicates using real-time PCR. The data from both techniques were not compared on the same graphs as different methods were used to analyse the data, and the exact units are not directly

comparable. However, the FC can be compared and is displayed in **Table 3.4**.

3.2.1 Validation of whole genome expression array data from embryonic tissues using quantitative real-time PCR

Expression data were obtained for nine of the ten candidate genes. *Ccnd3* was excluded due to low real-time PCR amplification efficiency. The expression of the candidate genes relative to *Actb* (housekeeping gene) are described in **Table 3.3** and graphically displayed in **Figures 3.7-3.9**. P-values were generated using the non-parametric Mann-Whitney U test. Data are presented as column scatter plots in which individual points represent the average of the samples performed in triplicate.

Table 3.3: Relative gene expression analyis of candidate genes in the tissues of embryos sired by male mice exposed to EtOH or sucrose

Gene	Tissue	Treatment	N	Relative expression± SD	Median (Range)	P value
Grm7	Brain	EtOH	22	3.06 ± 0.93	3.09 (1.35, 5.36)	0.30
		Sucrose	24	3.47 ± 1.15	3.58 (1.71, 5.57)	
Psmg2	Brain	EtOH	22	2.44±0.71	2.54 (1.25, 4.34)	0.54
		Sucrose	24	2.47±0.70	2.48 (1.53, 3.65)	
Pcbp3	Liver	EtOH	23	0.44± 0.10	0.45 (0.23, 0.61)	0.33
		Sucrose	20	0.47±0.17	0.41 (0.26, 0.68)	
Zfp317	Brain	EtOH	22	3.58±1.02	2.06 (0.92, 4.04)	0.45
		Sucrose	24	4.08 ±1.29	1.89 (0.63, 3.28)	
Tet1	Brain	EtOH	22	2.24±0.88	2.06 (0.92, 4.04)	0.56
		Sucrose	23	2.25±1.22	1.89 (0.63, 3.28)	
Vwf	Brain	EtOH	22	1.48±0.39	1.51 (0.87, 2.10)	0.94
		Sucrose	23	1.57±0.37	1.49 (0.60, 2.10)	
	Liver	EtOH	23	2.86±0.65	2.89(1.51, 3.62)	0.16
		Sucrose	23	2.81±1.13	2.70(1.27, 3.91)	
	Placenta	EtOH	23	2.10±1.48	1.19(0.59, 4.64)	0.20
		Sucrose	23	2.50±1.61	2.40 (0.52, 6.25)	
Depdc1b	Brain	EtOH	22	0.77±0.19	0.77 (0.39, 1.15)	0.72
		Sucrose	24	0.78±0.30	0.70 (0.38, 1.35)	
	Liver	EtOH	23	6.26±0.65	6.47 (3.44, 8.32)	0.34
		Sucrose	20	6.59±1.13	5.84 (3.84, 8.96)	
Igf1r	Brain	EtOH	22	1.70±0.62	1.56 (0.88,2.99)	0.65
		Sucrose	21	1.76±0.93	1.39(0.42, 2.99)	
	Liver	EtOH	24	0.82±0.20	0.83 (0.44, 1.24)	0.10
		Sucrose	21	0.76±0.31	0.70 (0.42, 1.16)	
Odc1	Brain	EtOH	22	0.61±0.30	0.65 (0.17, 1.07)	0.62
		Sucrose	24	0.66±0.30	0.63(0.21, 1.17)	
	Placenta	EtOH	24	1.26±0.49	1.08 (0.44, 2.06)	0.86
		Sucrose	23	1.20±0.42	1.10(0.45, 2.03)	

P-values generated using the Mann-Whitney U test
SD- Standard Deviation
N- Number of individual tissue samples tested
Only tissues that showed significant differential expression in the original study were analysed

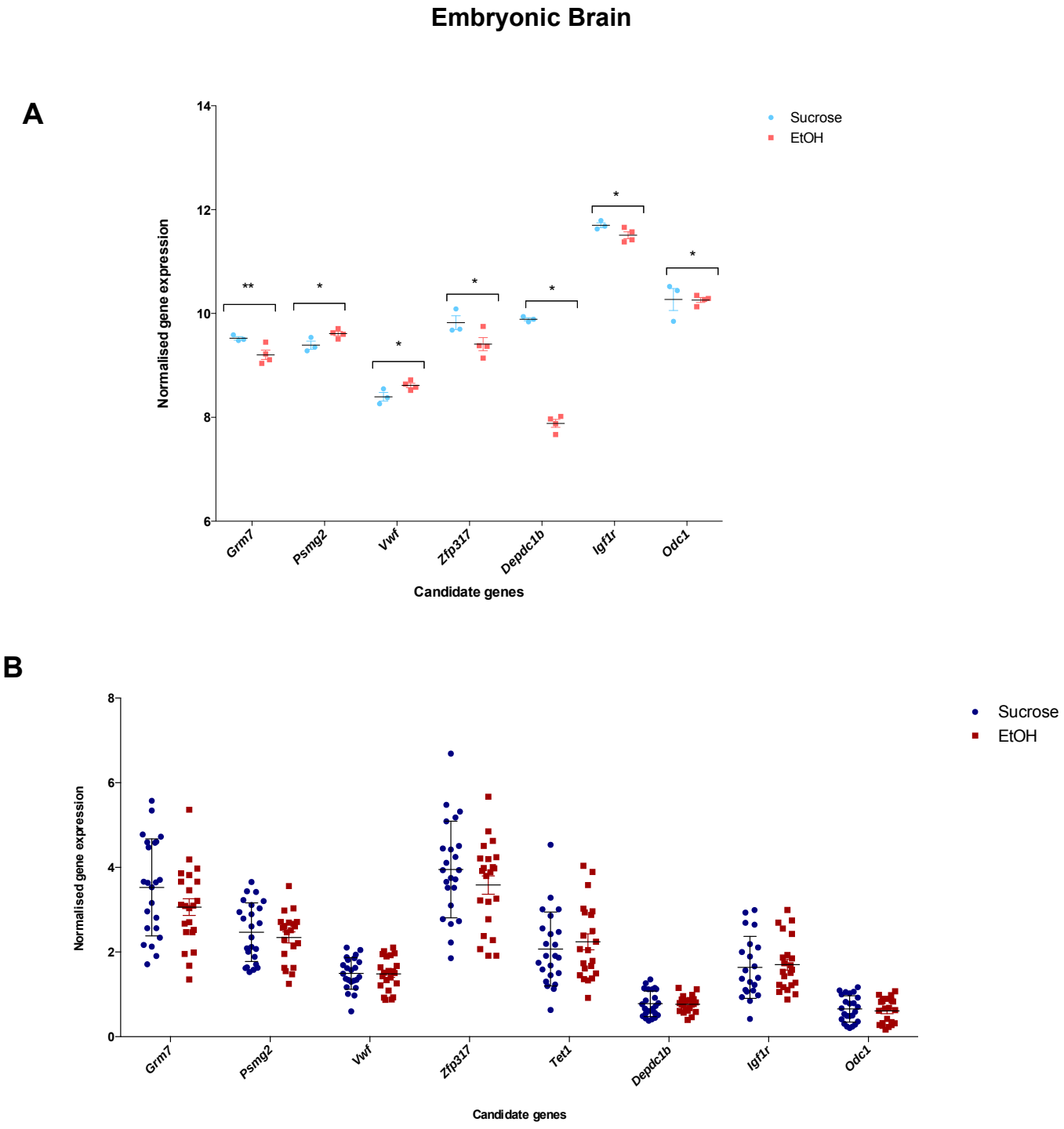


Figure 3.7: Real-time PCR validation of embryonic brain microarray results. A Microarray analysis of the effects of preconception paternal alcohol or sucrose exposure on gene expression in the embryonic brain**. B** Real-time analysis of the effects of preconception paternal alcohol or sucrose exposure on candidate gene expression in the embryonic brain. SYBR green quantitative PCR was used to analyse the gene expression levels relative to the housekeeping gene *Actb*. Statistically significant ** ($P \leq 0.01$) * ($0.01 < P \leq 0.05$) Trending significance ^ ($0.05 \leq P < 0.1$) Error bars indicate the standard deviation.

The expression array revealed a significant ($p<0.05$) decrease in expression in *Grm7*, *Zfp317, Depdc1b, Odc1* and *Igf1r* and a significant increase in expression in *Psmg2* and *Vwf* in the brain of embryos sired by males exposed to EtOH in comparison to sucrose (**Figure 3.7 A**). In order to verify these results, real-time PCR was conducted. Real-time PCR results are graphically represented in **Figure 3.7 B.** The real-time PCR results showed the same direction of change in gene expression for all the genes analysed although none were statistically significant. However, a slight decrease in expression was observed in *Grm7*, *Zfp317* and in the embryos sired by males exposed to EtOH (test) in comparison to sucrose (control). No difference in expression was observed in *Tet1*, *Vwf*, *Psmg2, Depdc1b*, *Igf1r*, and *Odc1* between treatment groups.

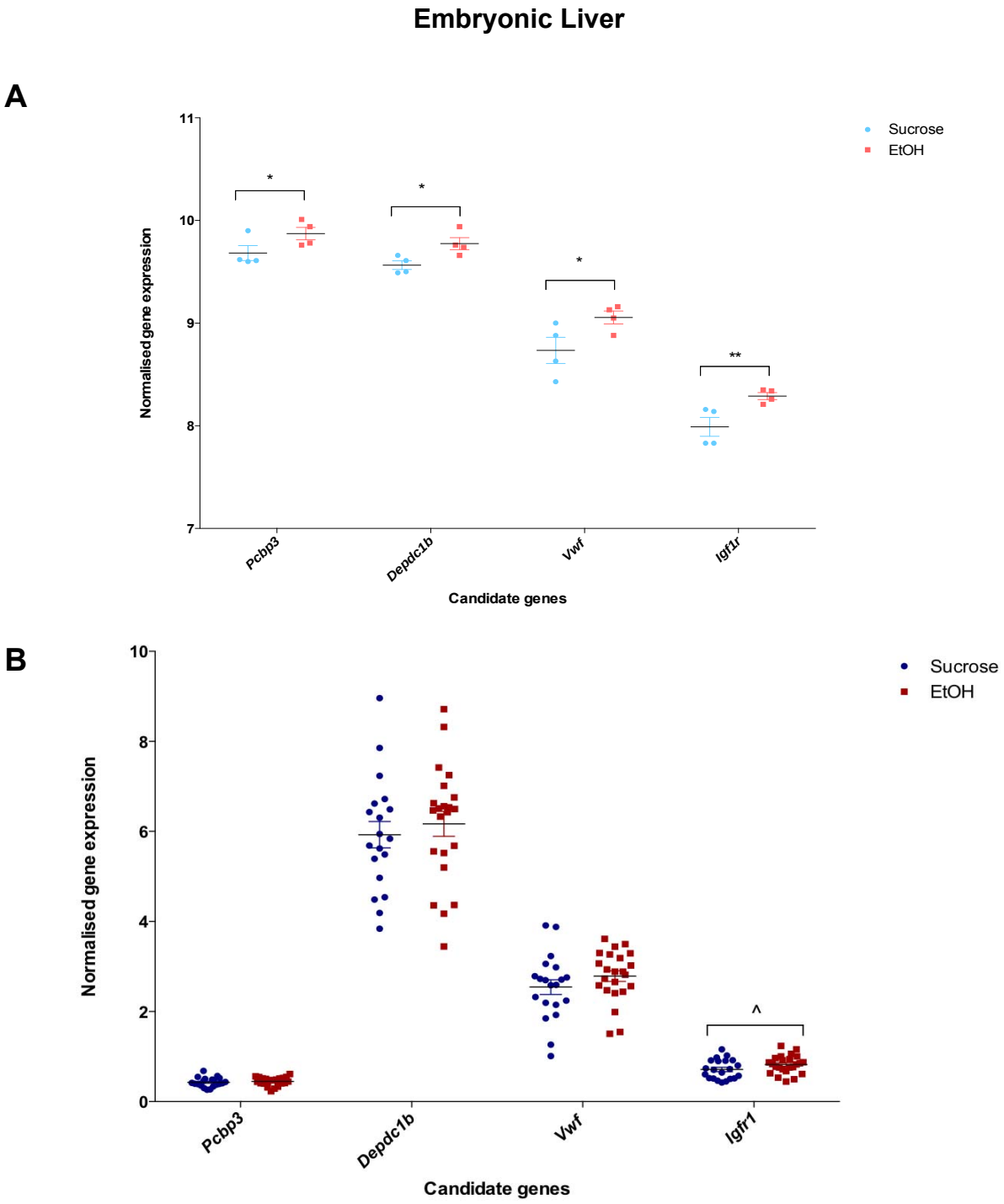


Figure 3.8: Real-time PCR validation of embryonic liver microarray results. A Microarray analysis of the effects of preconception paternal alcohol or sucrose exposure on gene expression in the embryonic liver**. B** Real-time analysis of the effects of preconception paternal alcohol or sucrose exposure on candidate gene expression in the embryonic brain. SYBR green quantitative PCR was used to analyse the gene expression levels relative to the housekeeping gene *Actb*. Statistically significant ** ($P \leq 0.01$) * ($0.01 < P \leq 0.05$) Trending significance ^ ($0.05 \leq P < 0.1$) Error bars indicate the standard deviation.

The expression array results revealed a significant ($p<0.05$) increase in the expression of *Pcbp3*, *Depdc1b*, *Vwf* and *Igf1r* in the liver of embryos sired by males exposed to chronic amounts of EtOH (**Figure 3.8 A**). Real-time PCR was conducted to validate these results. The real-time results (**Figure 3.8 B**) show the same direction of change as the expression array for *Depdc1b*, *Vwf* and *Igf1r*, with *Igf1r* trending towards significance. Collectively, EtOH exposure appears to increase the expression of all of the candidate genes in the liver. Although, when statistically analysed using the Mann-Whitney U test none of the genes assessed by real-time PCR were significant ($p<0.05$).

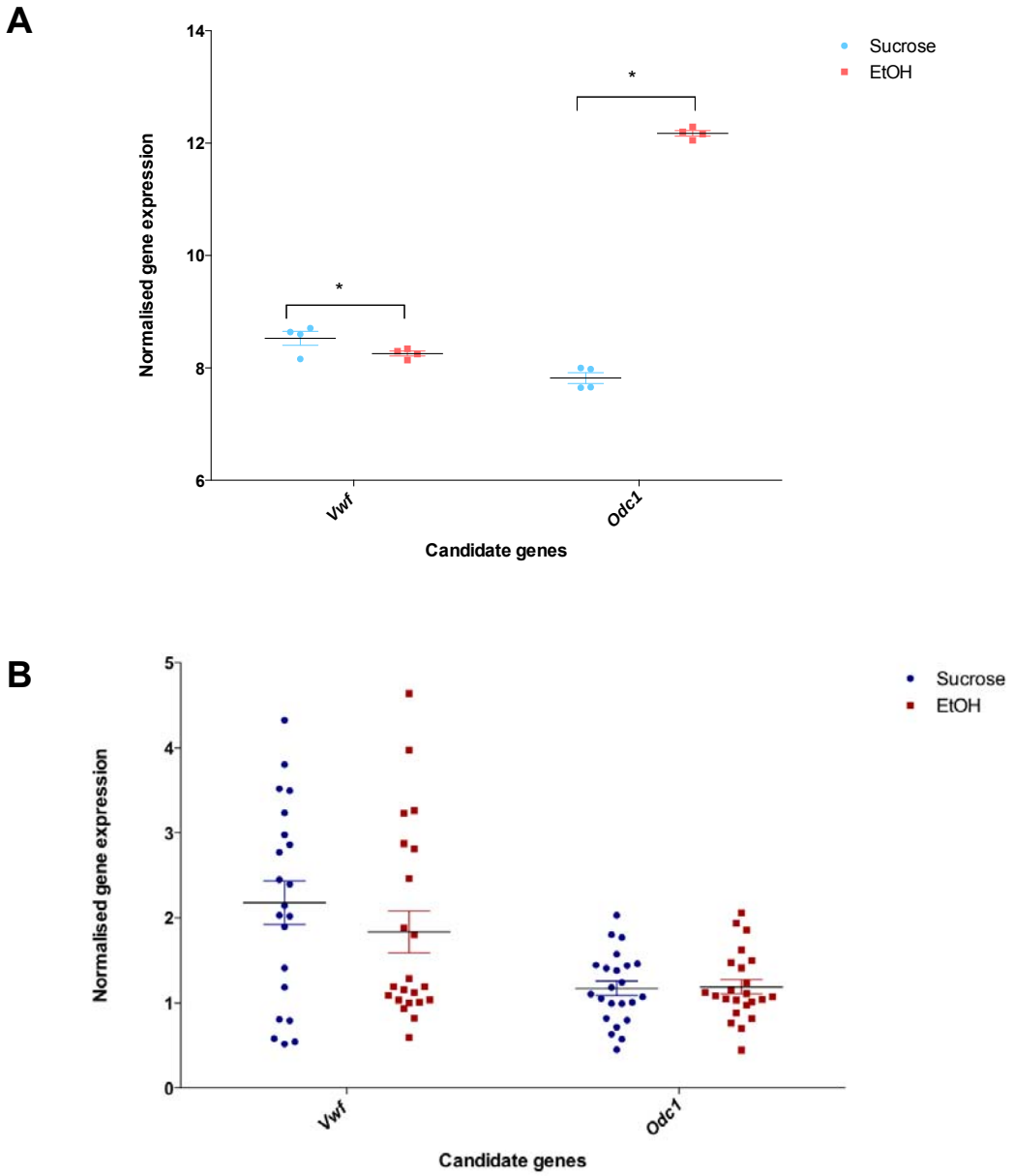


Figure 3.9: Real-time PCR validation of placental microarray results. A Microarray analysis of the effects of preconception paternal alcohol or sucrose exposure on gene expression in the embryonic placenta**. B** Real-time analysis of the effects of preconception paternal alcohol or sucrose exposure on candidate gene expression in the embryonic brain. SYBR green quantitative PCR was used to analyse the gene expression levels relative to the housekeeping gene *Actb*. Statistically significant ** ($P \leq 0.01$) * ($0.01 < P \leq 0.05$) Trending significance ^ ($0.05 \leq P < 0.1$) Error bars indicate the standard deviation.

The microarray analysis (**Figure 3.9A**) revealed that there was a significant decrease in the expression of *Vwf* and a significant increase in the expression of *Odc1* in the placenta of embryos sired by EtOH treated male mice in comparison to embryos sired by the sucrose exposed control group. Real-time PCR was conducted to confirm said results. Similar to the microarray data, the real-time PCR results suggested (**Figure 3.9B**) a decrease in *Vwf*

expression in the embryos sired by EtOH treated males. However, the results produced by real-time PCR did not reach statistical significance according to the Mann-Whitney U test. No change in expression was observed for *Odc1* between the treatment groups.

Comparison of microarray and real-time PCR fold change values

As previously mentioned, the real time PCR technique was carried out to validate the expression array results that showed significant changes between treatment groups. **Table 3.4** compares the P-value, FC and direction of expression in the candidate genes between the different techniques. According to the Mann-Whitney U test, none of the candidate genes selected for validation by real-time PCR showed statistically significant changes. Overall, the FC in the real-time PCR results, were lower than the FC of the whole genome expression array. Although not statistically significant, in six of cases the direction of change for a specific gene was concordant between the array and real time PCR results. In seven of the cases, the validation showed no differences in candidate gene expression between treatment groups (neutral).

Table 3.4: A comparison of the expression array and the real-time PCR data

Gene	Tissue	Array P value	Array FC value	Expression direction	Real-time PCR P value	Real-time PCR FC value	Expression direction
Grm7	Brain	0.008	1.25	Down	0.30	0.88	Down
Psmg2	Brain	0.02	1.16	Up	0.54	0.99	Neutral
Pcbp3	Liver	0.03	1.1	Up	0.33	1.05	Neutral
Zfp317	Brain	0.01	1.33	Down	0.45	0.99	Down
Vwf	Brain	0.02	1.16	Up	0.94	0.99	Neutral
	Liver	0.02	1.25	Up	0.16	1.10	Up
	Placenta	0.05	1.21	Down	0.20	0.81	Down
Tet1	Brain	----	----		0.56	1.08	Neutral
Depdc1b	Brain	0.02	1.18	Up	0.72	0.99	Neutral
	Liver			Up	0.34	1.04	Up
Igf1r	Brain	0.04	1.18	Up	0.66	1.04	Neutral
	Liver	0.04	1.14	Up	0.10	1.15	Up
Odc1	Brain	0.04	1.14	Down	0.62	0.93	Neutral
	Placenta	0.01	1.16	Up	0.86	1.07	Neutral

P-values generated using the Mann-Whitney U test. Statistically significant ** ($p \leq 0.01$) * ($0.01 < P \leq 0.05$).
Neutral- no change in gene expression in embryos whose sires were exposed to EtOH in comparison to the embryos whose sires were exposed to sucrose.
FC- a measure of how much or less a gene is expressed relative to the control.

3.2.2 Sex specific differential expression of candidate genes between treatment groups

To determine whether any of the candidate genes displayed significant changes in expression when separated by sex, data from the male and female embryos were analysed separately. Of the nine genes analysed, only *Grm7* displayed significant sex specific differential expression in the brain of male but not female embryos sired by males exposed to EtOH in comparison to the sucrose exposed group, when applying real-time PCR.

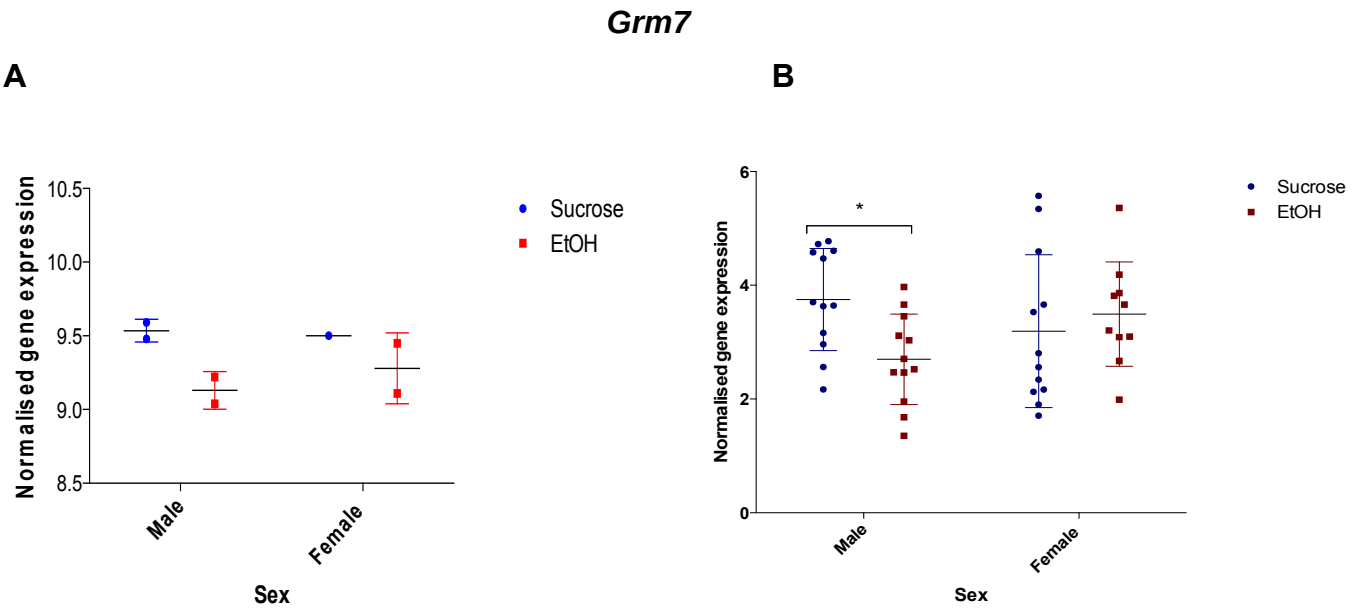


Figure 3.10: Differential *Grm7* expression in male and female embryos sired by male mice exposed to EtOH or sucrose. A. Microarray expression results. B Expression analysed by quantitative real-time PCR. Statistically significant * ($P \leq 0.05$). Error bars indicate the standard deviation.

As shown in **Figure 3.10 A** and **B**, the microarray and real-time PCR expression data for *Grm7* were separated according to sex. The microarray results suggest that male mice that were exposed to EtOH sired male and female embryos that both display a decrease in *Grm7* expression in the brain. Similar results were revealed for the male embryos through real-time PCR results. Furthermore, the real-time PCR results indicate that male embryos sired by males EtOH-treated males show a significant difference (p=0.01) in *Grm7*

expression in the brain compared to the female embryos. However, according to the real-time PCR data female embryos sired by EtOH-treated males had slightly higher expression levels relative to the sucrose controls.

3.3 The effect of preconception EtOH exposure on methylation in embryonic tissues

The DNA methylation levels were quantified in embryonic tissues in genes that showed a significant change in methylation as per the RRBS data and a significant change in microarray data. It must be noted that this section did not serve as a validation as there were no prior analysis of the embryonic tissue DNA methylation levels. The same primers that were used for pyrosequencing of the sperm DNA were used to amplify the embryonic DNA. In addition, primers were designed to amplify and *Igf1r*.

Similar to the methylation analysis of the sperm DNA, the methodology and quality control measures included: samples being amplified in triplicate, as well as samples being amplified within the same PCR machine to reduce variation associated with different PCR machine use. Each PCR sample was assayed on the pyrosequencer. Following PCR and the assessment of products for amplification and contamination using agarose gel electrophoresis, samples underwent pyrosequencing. In cases where the incorrect PCR product was amplified primers were redesigned and the assay was reassessed. The average methylation percentage for each CpG site was calculated by averaging the triplicate values for each CpG site for each sample. Intra-sample variation greater than 10% was excluded from the study. The average methylation percentage per CpG site was then averaged across all the CpG sites analysed per gene to give the mean methylation percentage per candidate gene.

Batch effects were kept to a minimum in both the Pyrosequencing PCR and the run by assembling samples such that 96 well plates contained samples ranging in treatment group

and sex. Each plate also contained a PCR control and a water sample. Pyrosequencing data for candidate genes analysed in the embryonic tissues are displayed in Appendix M. The results below include the mean methylation analysis and the methylation percentage at individual CpG sites.

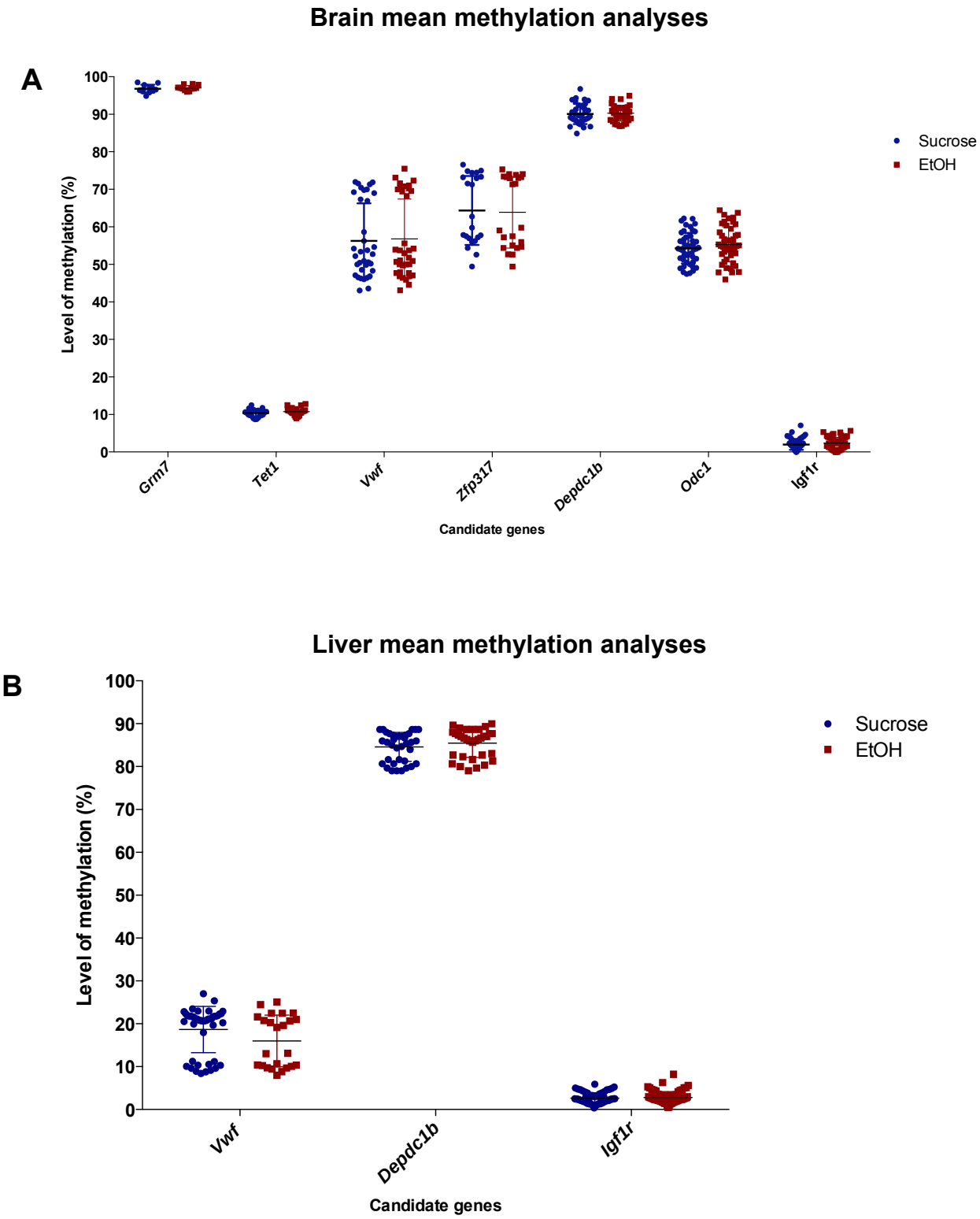
3.3.1 Gene averaged methylation analyses of embryonic DNA

The mean methylation levels of the candidate genes in embryos sired by males exposed to either EtOH or sucrose are depicted in **Table 3.5** and graphically displayed in **Figure 3.11**. The non-parametric Mann-Whitney U test was used to assess if there was a significant difference in the DNA methylation in the male mice exposed to EtOH (test) in comparison to the males exposed to sucrose (control) in all the candidate genes. This statistical test was selected due to the small sample size. Each data point on the column scatter graphs represents the average methylation of a sample that was performed in triplicate.

Table 3.5: Mean methylation data of the candidate genes in the tissues of embryos sired by males that were exposed to EtOH or sucrose

Gene	Tissue	Treatment	N	CpG sites	Mean methylation ± SD (%)	Median (Range)	P value
Grm7	Brain	EtOH	12	1	97.02± 0.69	96.81 (96.01,98.09)	0.59
		Sucrose	12		96.81± 1.06	97 (94.97,98.51)	
Tet1	Brain	EtOH	12	2	10.75± 1.00	10.78 (8.98,12.80)	0.34
		Sucrose	12		10.37± 1.03	10.52 (8.71,12.52)	
Vwf	Brain	EtOH	12	3	56.79±10.61	52.39 (43.10,75.49)	0.78
		Sucrose	12		56.65±9.98	52.72 (43.04,71.95)	
	Liver	EtOH	12	3	17.98±5.82	21.29 (7.98,27.00)	0.76
		Sucrose	12		17.62± 5.70	20.73 (8.43,23.49)	
	Placenta	EtOH	12	3	19.97±7.55	21.35(8.94,32.80)	0.70
		Sucrose	12		20.30±7.63	21.92 (9.35, 31.63)	
Zfp317	Brain	EtOH	12	2	62.26±9.47	59.79 (49.42,75.31)	0.58
		Sucrose	12		65.15±9.17	61.29 (49.42,76.56)	
Depdc 1b	Brain	EtOH	12	2	90.38±2.06	90/50 (86.85,94.96)	0.58
		Sucrose	12		90.00±2.66	89.56 (64.84,96.75)	
	Liver	EtOH	12	2	84.58±3.31	86.67 (79.00,88.67)	0.76
		Sucrose	12		85.46±3.37	85.67(79.00,88.67)	
Odc1	Brain	EtOH	12	4	55.33± 4.65	55.07 (45.97,64.44)	0.21
		Sucrose	12		54.29± 4.06	54.18 (47.45,62.25)	
	Placenta	EtOH	12	4	18.28±1.70	19.96 (15.12,21.34)	0.001 *
		Sucrose	12		19.53±1.76	19.41 (16.03,23.34)	
Igf1r	Brain	EtOH	12	5	2.30±1.45	2.16(0.00,5.31)	0.19
		Sucrose	12		2.02±1.43	1.80 (0.00,7.09)	
	Liver	EtOH	12	5	2.94±1.26	2.68(0.51,8.23)	0.23
		Sucrose	12		2.64±1.44	2.45 (0.33,5.91)	

P values generated using the Mann Whitney U test. Statistically significant ** ($p \leq 0.01$) * ($0.01 < P \leq 0.05$)
CpG sites-the number of CpG sites analysed per target region
SD- Standard Deviation


Brain mean methylation analyses
A
Level of methylation (%)
100
90
80
70
60
50
40
30
20
10
0
Grm7
Tet1
Vwf
Zfp317
Depdc1b
Odc1
Igf1r
Candidate genes
Sucrose
EtOH
Liver mean methylation analyses
B
Level of methylation (%)
100
90
80
70
60
50
40
30
20
10
0
Vwf
Depdc1b
Igf1r
Candidate genes
Sucrose
EtOH

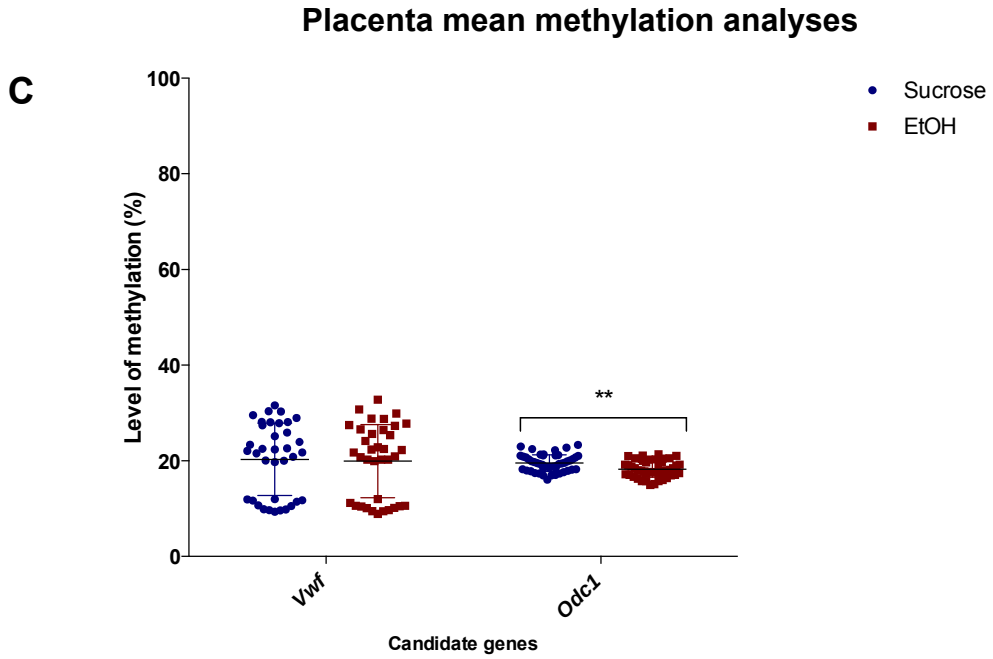


Figure 3.11: Column scatter plots of mean methylation levels of the candidates gene in different embryonic tissues. Embryos were sired by male mice exposed to EtOH or sucrose for ten weeks. Statistically significant ** (P ≤ 0.01) * (0.01< P ≤ 0.05) Trending significance ^ (0.05 ≤ P < 0.1) Error bars indicate the standard deviation.

Figure 3.11 A, **B** and **C** shows the mean methylation of the candidate genes in the brain, liver and placenta, respectively. No significant differences in mean methylation were observed in the candidate genes in the brain and liver of the embryos sired by males exposed to EtOH in comparison to the males exposed to sucrose. However, *Odc1* in the placenta exhibited a significant (p=0.001) decrease in the mean methylation in the embryos sired by males exposed to EtOH, although the effect size was small (mean methylation EtOH group 18.28%; mean methylation in sucrose group 19.53%).

3.3.2 CpG specific methylation analyses of embryonic DNA

In addition to the mean methylation of the target region, methylation levels of specific CpG sites within the region were assessed. The methylation levels for the individual CpG sites per candidate gene are depicted in **Table 3.6** and graphically displayed in **Figures 3.12 to 3.14**. Methylation levels were assessed in either brain, liver or placental DNA. In some

genes, more than one tissue was assessed. The non-parametric Mann-Whitney U test was used to analyse if there was a significant difference in methylation in the candidate's genes of embryos sired by male mice exposed to EtOH in comparison to male mice exposed to sucrose. Each data point on the column scatter graphs represents the average methylation of a sample that was performed in triplicate.

Table 3.6: DNA methylation data of individual CpG sites in the candidate genes of embryos sired by male mice exposed to EtOH or sucrose

Gene	Tissue	CpG site	Treatment	N	Methylation level ± SD (%)	Median (Range)	P value
Grm7	Brain	11543873	EtOH	12	97.02 ± 0.93	96.96 (96.01,98.09)	0.58
			Sucrose	12	96.81± 1.06	96.62 (94.87, 95.51)	
Tet1	Brain	62329009	EtOH	12	10.09 ± 0.76	10.10 (8.98,11.18)	0.29
			Sucrose	11	9.65 ± 0.77	9.60 (8.71,11.09)	
		62329050	EtOH	12	11.40 ± 0.74	11.08 (10.38,12.80)	0.45
			Sucrose	12	11.09 ± 0.72	10.97 (9.97, 12.52)	
Vwf	Brain	125553557	EtOH	10	47.09 ± 2.53	46.96 (43.10, 51,06)	0.77
			Sucrose	12	46.92 ± 2.22	46.67 (43.04,50.45)	
		125553787	EtOH	11	71.03 ± 2.40	70.76 (68.22, 75.49)	0.40
			Sucrose	11	69.95 ± 1.73	69.97 (66.95,71.95)	
		125553763	EtOH	12	52.28 ± 2.02	52.39 (49.71, 55.61)	0.47
			Sucrose	11	53.12 ±2.62	53.04 (49.86,58.62)	
	Liver	125553557	EtOH	9	10.29 ±0.87	10.16 (7.98, 13.11)	0.81
			Sucrose	12	9.85 ± 0.91	9.84 (8.43, 11.22)	
		125553787	EtOH	12	21.99 ± 1.12	21.71 (17.95, 27.00)	0.93
			Sucrose	12	21.66 ± 1.13	21.41 (16.69, 23.49)	
		125553763	EtOH	12	21.66 ± 0.86	21.29 (19.17, 25.05)	0.93
			Sucrose	11	21.36 ± 0.87	21.15(19.93, 22.97)	
	Placenta	125553557	EtOH	12	10.27 ± 0.98	10.30 (8.94, 11.99)	0.32
			Sucrose	12	10.71 ± 1.01	10.62 (9.35, 11.99)	
		125553787	EtOH	12	21.52 ± 1.28	21.35 (19.93, 24.13)	0.80
			Sucrose	12	21.74 ± 1.35	21.92 (19.73, 23.95)	
		125553763	EtOH	12	28.13 ± 1.72	27.63 (25.28,32.80)	0.51
			Sucrose	12	28.45 ± 1.86	28.13(25.13, 31.63)	
Zfp317	Brain	19443712	EtOH	12	55.23 ± 3.21	54.87(49.42, 59.97)	0.26
			Sucrose	10	56.54 ± 3.38	57.07(49.42, 62.78)	
		19443716	EtOH	11	73.28 ± 1.37	73.27(71.32, 75.31)	0.47
			Sucrose	11	73.76 ± 1.62	73.87 (71.27, 73.87)	
Depdc1b	Brain	109101087	EtOH	12	88.16 ± 1.41	88.24(86.85, 89.57)	1.00
			Sucrose	12	88.04 ± 1.47	88.44(84.86, 90.02)	
		109101098	EtOH	12	92.24 ± 2.69	91.98(89.78, 94.96)	0.48
			Sucrose	12	91.65 ± 2.76	91.70 (86.91,96,75)	
		109101330	EtOH	12	90.77 ± 2.34	90.96(88.94,91.93)	0.55
			Sucrose	12	90.32 ± 2.35	90.58(86,70, 93.95)	
	Liver	109101087	EtOH	11	81.21 ± 0.97	81.33(79.00, 83,00)	0.08 ^
			Sucrose	12	80.25 ± 1.01	80.33(79.00,81.67)	
		109101098	EtOH	11	87.85 ± 1.37	87.67 (85.67, 89.67)	0.06 ^
			Sucrose	12	86.75 ± 1.41	86.63 (84.67, 86.83)	
		109101330	EtOH	11	87.33 ± 1.51	87.00 (86.00,90.00)	0.60
			Sucrose	12	86.75 ± 1.57	87.17 (84.00, 88.67)	
Odc1	Brain	17514120	EtOH	12	61.57 ± 2.54	61.37(58.54, 64.44)	0.02*
			Sucrose	12	59.59 ± 2.13	59.55(54.82, 62.25)	
		17514135	EtOH	12	55.19 ± 2.45	54.95(53.33, 57.84)	0.11
			Sucrose	12	53.58 ± 2.40	53.36(49.41, 56.67)	
		17514144	EtOH	12	49.36 ± 2.00	49.30(45.97, 52.93)	0.84

			Sucrose	12	49.74 ± 2.00	49.23(47.45, 53.68)	
		17514154	EtOH	12	55.21 ± 1.78	55.20(52.42, 57,63)	0.12
			Sucrose	12	54.24 ± 1.63	54.31(51.55, 57.26)	
		17514120	EtOH	11	17.45 ± 1.50	17.69 (15.12,20.26)	0.04*
			Sucrose	12	18.22 ± 1.20	18.37(16.03, 20.37)	
		17514135	EtOH	12	17.30 ± 1.58	17.05(14.96, 20.52)	0.04 *
	Placenta		Sucrose	12	18.65 ± 1.38	18.08(16.95, 21.39)	
		17514144	EtOH	12	18.77 ± 1.29	19.13 (16.96, 21.07)	0.05 ^
			Sucrose	12	19.90 ± 1.05	20.09 (17.91, 21.14)	
		17514154	EtOH	12	19.59 ± 1.87	19.79 (17.57, 21.34)	0.005*
			Sucrose	12	21.35 ± 1.51	21.29 (18.55, 23.34)	
		75274289	EtOH	12	2.11 ± 0.73	2.09 (0.98, 3.84)	0.40
			Sucrose	12	1.87 ± 0.38	1.83(1.37,2.69)	
		75274294	EtOH	12	0.51 ± 0.65	0.40 (0.00,2.69)	1.00
			Sucrose	12	0.50 ± 0.50	0.37 (0.00,1.62)	
		75274301	EtOH	12	1.98 ± 1.11	1.61(0.59,4.26)	0.04*
	Brain		Sucrose	12	1.21 ± 0.55	1.17 (0.43,2.37)	
		75274319	EtOH	12	4.34 ±0.57	4.20(3.39, 5.31)	0.29
			Sucrose	12	4.22 ± 1.09	4.01 (3.11,7.09)	
		75274324	EtOH	12	2.55 ±0.67	2.59(1.26,3.62)	0.14
Igf1r			Sucrose	12	2.30 ± 0.69	2.18 (1.67,4.29)	
		75274289	EtOH	12	2.26± 0.43	2.24 (1.65, 2.24)	0.44
			Sucrose	12	2.07 ±0.50	2.01 (1.37,2.97)	
		75274294	EtOH	12	1.40± 0.56	1.38 (0.51,2.28)	0.15
			Sucrose	12	1.17± 0.37	1.27 (0.33,1.84)	
		75274301	EtOH	12	2.79 ±0.67	2.63 (2.06,4.49)	0.29
	Liver		Sucrose	12	2.55± 0.52	2.44 (2.03,3.64)	
		75274319	EtOH	12	5.25 ±1.11	5.08 (4.17, 8.23)	0.15
			Sucrose	12	4.69 ±0.55	4.67 (3.89, 5.92)	
		75274324	EtOH	12	3.03± 0.33	3.11 (2.28,3.60)	0.08 ^
			Sucrose	12	2.74 ±0.52	2.58 (1.82,3.83)	

P values generated using the Mann Whitney U test.
Statistically significant ** (P ≤ 0.01) * (0.01< P ≤ 0.05)
Trending significance ^ (0.05 ≤ P < 0.1)
SD- Standard Deviation

CpG site-specific methylation analyses of Brain DNA

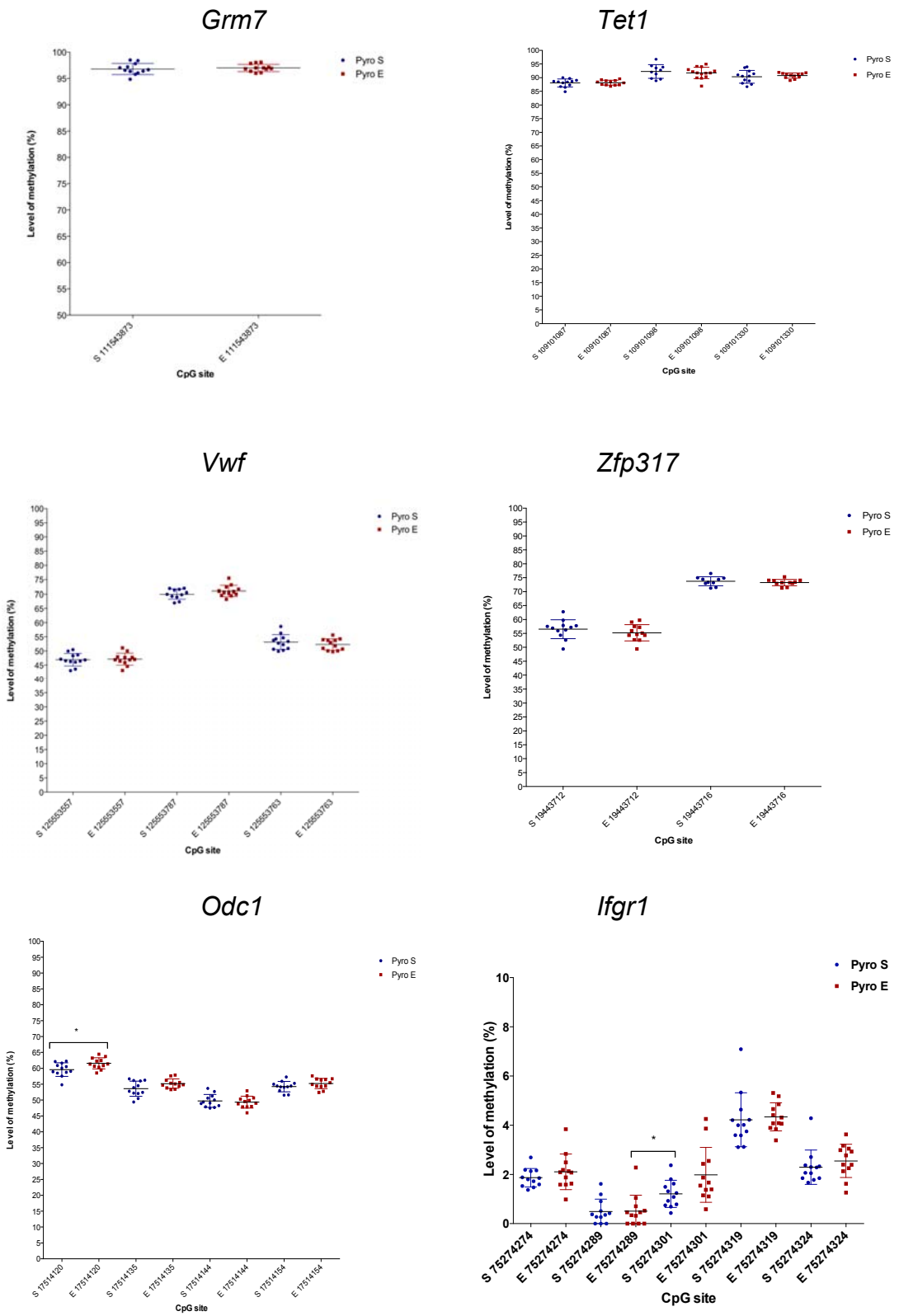


Figure 3.12: Column scatter plots of DNA methylation at individual CpG sites of the embryonic brain. Embryos were sired by male mice exposed to EtOH or sucrose for ten weeks. Statistically significant ** (P ≤ 0.01) * (0.01< P ≤ 0.05) Trending significance ^ (0.05 ≤ P < 0.1) Error bars indicate the standard deviation.

DNA methylation analysis of the individual CpG sites in *Grm7*, *Tet1*, *Vwf*, *Depdc1b* and *Zfp317* in the brain revealed that there is no significant difference in methylation between embryos sired my male mice exposed to EtOH compared to sucrose. A significant difference in methylation was observed at CpG site 17514120 in *Odc1* (p=0.02) and CpG site 75274301 in *Igf1r* (p=0.04). Although significant, the difference was less than 5%.

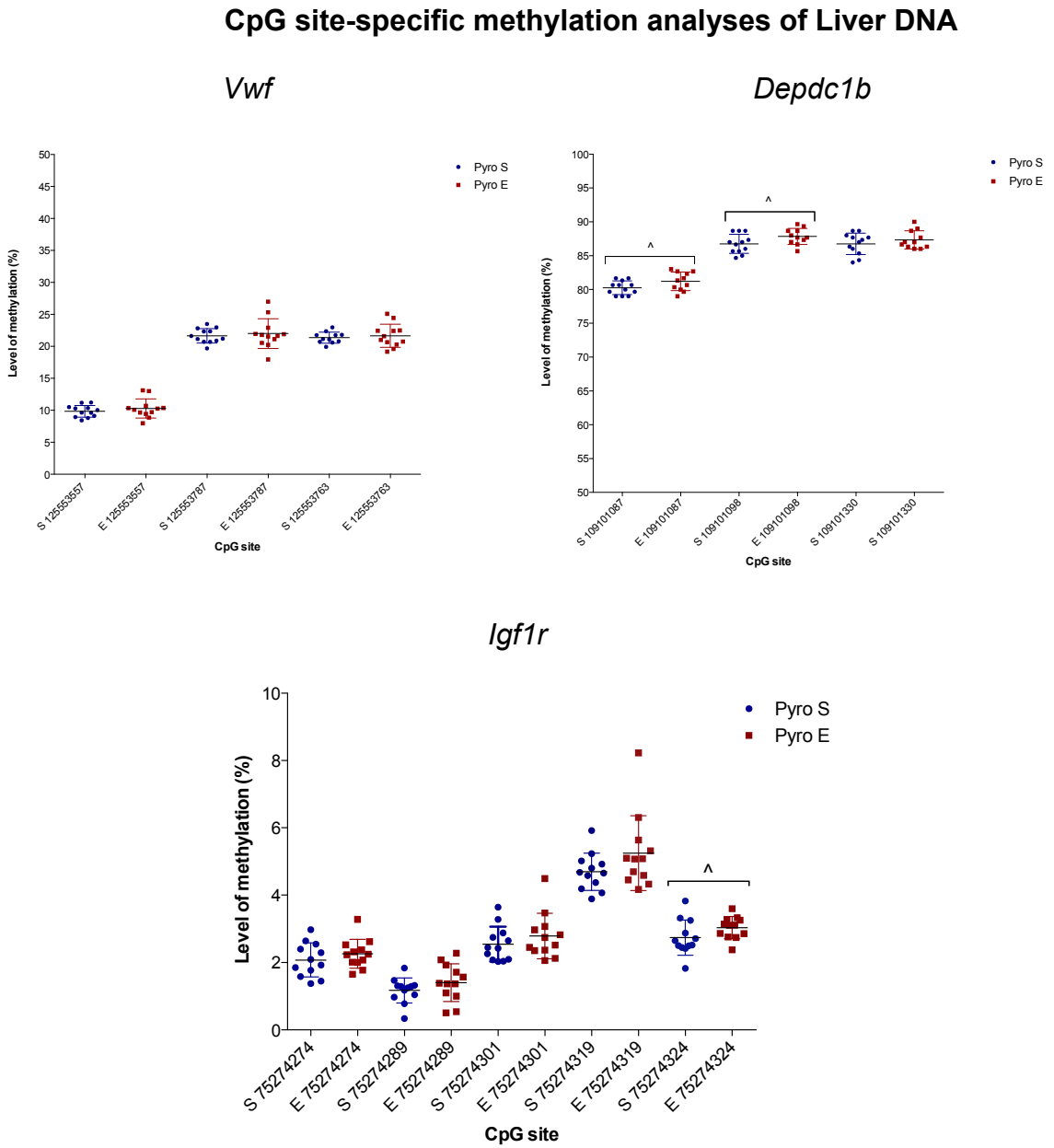


Figure 3.13: Column scatter plots of DNA methylation at individual CpG sites of the embryonic liver. Embryos were sired by male mice exposed to EtOH or sucrose for ten weeks. Statistically significant ** (P ≤ 0.01) * (0.01< P ≤ 0.05) Trending significance ^ (0.05 ≤ P < 0.1) Error bars indicate the standard deviation.

DNA methylation analysis of the individual CpG sites in *Vwf* in the liver revealed that there is no significant difference in methylation between embryos sired by male mice exposed to EtOH compared to sucrose. A trend toward significance was observed at CpG site 109101087 and 109101098 in *Depdc1b* and CpG site 75274349 in *Igf1r*. However, these significant changes are very small (less than 5%), and are therefore unlikely to have a functional impact.

CpG site-specific methylation analyses of placental DNA

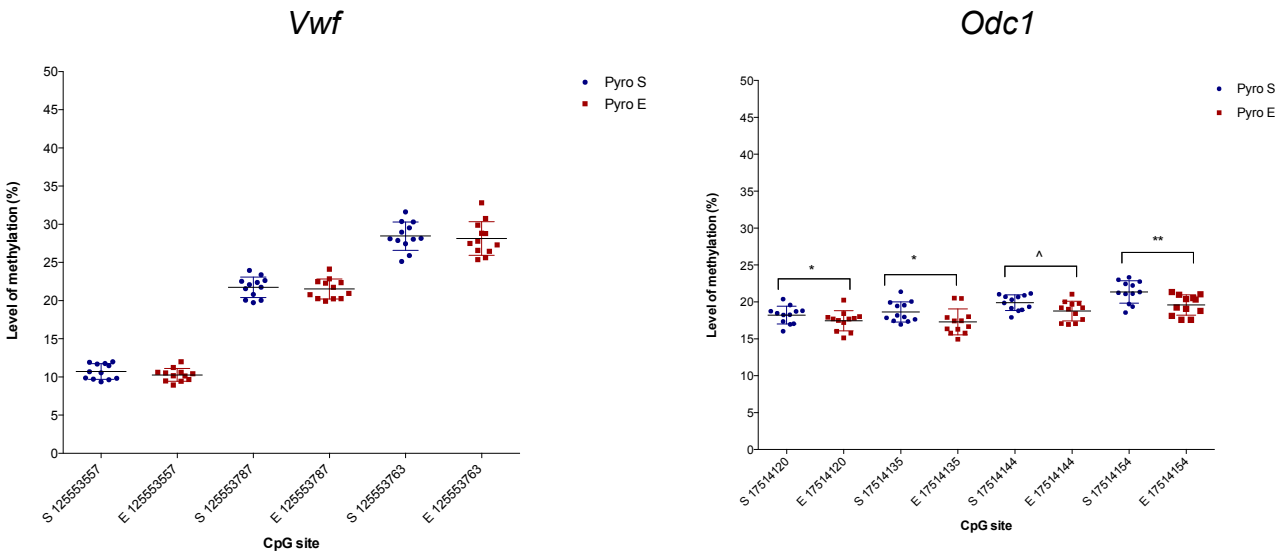


Figure 3.14: Column scatter plots of DNA methylation at individual CpG sites of the placenta. Embryos were sired by male mice exposed to EtOH or sucrose for ten weeks. Statistically significant ** (P ≤ 0.01) * (0.01< P ≤ 0.05) Trending significance ^ (0.05 ≤ P < 0.1) Error bars indicate the standard deviation.

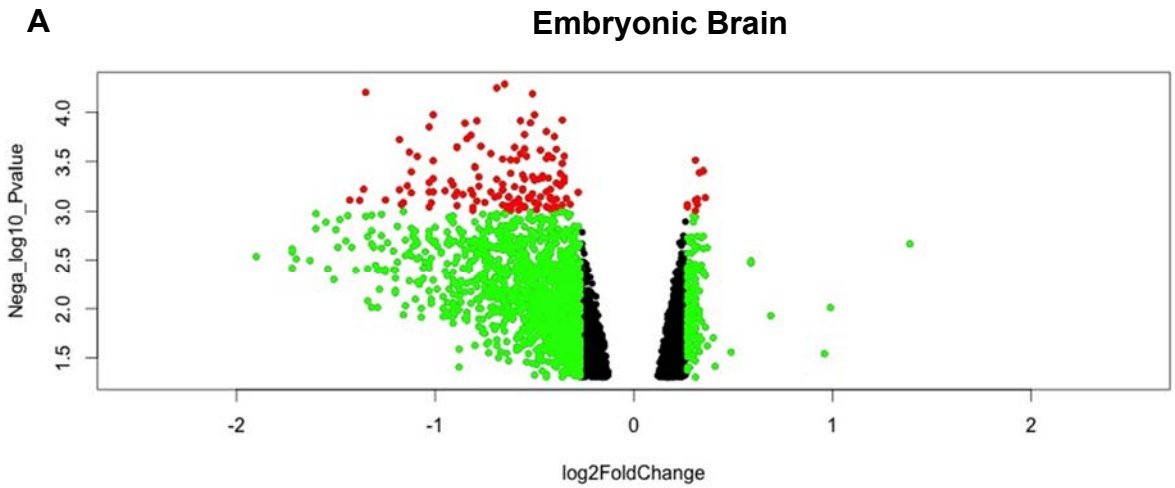
No differences in methylation were observed at any of the CpG sites analysed in *Vwf* in the placenta of embryos sired by the test or control exposed male mice. However, three CpG sites in *Odc1* showed a significant decrease in methylation in the embryos sired by males exposed to EtOH in comparison to the sucrose controls. Furthermore, a trend toward significance was seen at CpG site 17514154 *Odc1*. However, the percentage methylation

changes for the significant sites are very small and are therefore unlikely to be functionally significant.

3.4 Bioinformatics analysis to assess differential gene expression between treatment groups in three embryonic tissues

3.4.1 Volcano plots to identify significantly differentially expressed genes

A volcano plot is a scatter plot of the negative $\log_{10}$ transformed p-values against the $\log_2$ fold change. Genes with statistically significant differential expression will lie above a horizontal threshold. Genes with a large fold change will lie outside a vertical threshold (Cui and Churchill, 2003). Three volcano plots were generated using R (R development core team, 2013); to identify genes that exceeded both thresholds, in the embryonic brain, liver and placenta according to previously generated microarray data by Knezovich et al. (2014). These genes have a significance of $p<0.001$ and a fold change >1.2 and are displayed as red dots in **Figure 3.15 A**, **B** and **C**.



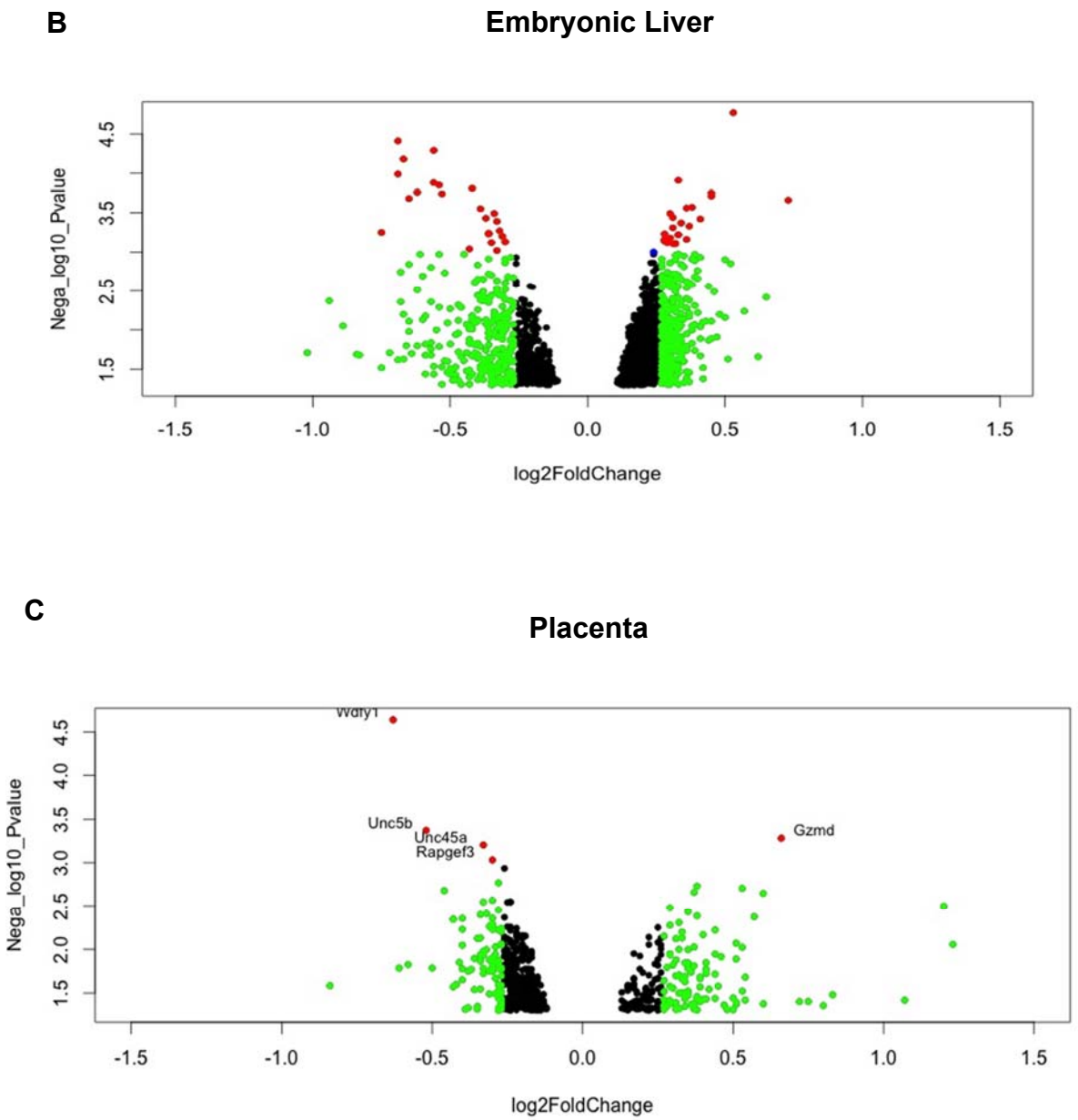


Figure 3.15: Volcano plots of significance against fold change of embryonic brain (A), liver (B) and placental (C) gene expression. Each dot represents a single gene. Highly significant values are toward the top and small expression changes are near the centre. Genes with decreased expression levels will lie on the negative x-axis, while genes that display an increase in expression are located on the positive side. The red dots represent the genes that are highly significant ($p<0.001$) and have a high FC>1.2, the green dots represent genes that have a significance of $p<0.05$ and a FC>1.2 and the black dots represent genes hat have a significance of $p<0.05$ and a FC<1.2. The log fold change is plotted on the x-axis and the negative log10 p-value is plotted on the y-axis.

Three features of the volcano plots stand out. First, the number of genes that are downregulated in the brain (**Figure 3.15 A**) of embryos sired by males treated with alcohol is greater than the genes that are upregulated, as only nine genes were significantly

upregulated genes in the brain exceeded the p<0.001 and FC>1.2 threshold. Secondly, the liver plot (**Figure 3.15 B**) is fairly symmetrical i.e. there are a similar number of genes that are significantly downregulated and upregulated. Thirdly, very few genes (n=5) in the placenta (**Figure 3.15 C**) were significantly dysregulated (p<0.001 and FC>1.2).

Four candidate gene lists were generated using the genes represented as red dots in volcano plots in **Figure 3.15 A and B**; this included the genes that were significantly dysregulated in the brain and liver. Genes are listed in **Table 3.7**. As previously stated, genes that displayed a significant (p<0.001) increase or decrease with a FC>1.2 were regarded as significantly dysregulated. No genes were submitted from the placenta, as there was too few that exceeded the significance (p<0.001) and FC (>1.2) thresholds.

Table 3.7: Genes that showed a significant difference ($p<0.001$) in expression with a fold change of greater than 1.2 in the brain or liver of embryos sired by EtOH-treated males as per the microarray data

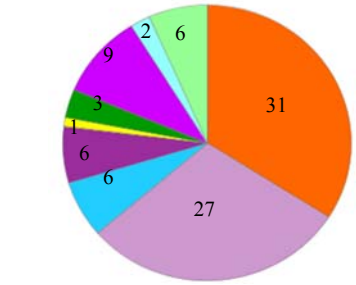
Brain		Liver	
Upregulated (n=10)	Downregulated (n=95)	Upregulated (n=19)	Downregulated (n=23)
Plagl2, Gli3, Mmp2 *Sepn1, BC016423* *Cdca4, Igf2bp2* *Aebp2, Cwf19l2* *Vim*	*Rnasek, Ppa2, Apoa2,* *Grp,6330419J24Rik* *Poldip3, Eno1,2700050C19Rik,* *Ambp, Wdfy1, Serpina1b,* *Hnrpll,Otud6b, Tusc1* *Serpina1b,Tmem55b, Tuba3a,* *Atp6v1a,S100a8, Oxr1, Uck1,* *Terf2ip,Nudcd2, Osbpl2,* *2810453I06Rik, Ndufb9,* *LOC100040413, Faim, Tmed2* *LOC100044170, Kng1, Mrpl20,* *Nef3, Mrpl13, Tcf7l2, Pgk1,* *Abcd3, Apoa2* *5730494M16Rik, Tceal1, Id2,* *Stfa1, LOC674611,,* *9130024F11Rik,5730494N06Rik,* *LOC380665, Rchy1* *Eif4e, Zfand6, Smc5l1, Hsdl1,* *Tmem85, Abce1, Ndufb5,* *Ctnna2, Cd200, Mdh1* *LOC100045776, Ythdf3, Prdx5* *LOC100044087, mt-Nd4l, Ifit2,* *Esd, Dcun1d5, Zhx1, Gad1,* *Vapa, Nkain2, Pdcd5, Tmem66,* *Stfa1,Vstm2a, LOC277856,* *LOC635470, LOC231368* *Cox7a2, Glrx3, LOC381230,* *LOC380707, S100a9,* *LOC668038, Dlx2, Scoc, Ppp2r2c* *Prkcb, Txndc12, Robld3, Rwdd4a* *LOC668837, Cfl2,* *2510003E04Rik* *LOC674706, Nicn1, Tram1,* *2010007L18Rik* *Bhlhb2, Twistnb, Sp1, Rpl18a,* *AI851790* *Selk, Gap43, Efha1, Uhrf1bp1l* *Gjd2, LOC100040671,* *LOC665281* *1700047I17Rik1, LOC668239, Trf* *5730455O13Rik, LOC383125* *Nap1l5, LOC100043906,* *Chmp2a, Efnb2* *LOC100042773, LOC382096,* *Msl2l1* *Tmem176b, 1700108L22Rik,* *Minpp1* *Vps26b, Napa, Vhl, Fam134a,* *Nt5c* *2310036D22Rik, Vps36, Mtrf1l* *2510008P16Rik, Pkig, Serpina1b,* *Cstf3*	*Art4, Mmrn1, Entpd4* *Pip5k1b, Pcyt1b, Emilin2* *Ssx2ip, Mier3, Hemgn* *Pklr, Usp7, Thumpd1* *Cdca8, Kif4, Smarcc1* *Cbfa2t3h, Slc43a1* *Sfrs4, Polh*	*Arg1, Hao1, Pon1* *Cyp2d26, Chi3l4* *LOC667337,* *D4Bwg0951e* *Chi3l3, BC032265* *Tff3, Gsta3, Bhmt2* *Nalp6, Wdfy1, Cml1* *Azgp1, C8g,* *EG665378* *Agmat, Gsta3* *E130112E08Rik,* *Gls2, Atf5*

3.4.2 Functional classification analysis using PANTHER

The four gene lists were first submitted to PANTHER (Thomas et al., 2003) for analyses. The PANTHER classification tool clusters genes that perform similar molecular functions (the function that a protein performs on its direct molecular targets) and biological processes (the biological systems to which a protein contributes) together and displays the results as a pie chart (Thomas et al., 2003), as seen in **Figure 3.16** and **3.17**. Genes were not exclusive to one category.

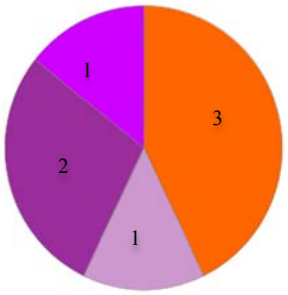
Brain

Downregulated
n=95



binding (GO:0005488)
catalytic activity (GO:0003824)
enzyme regulator activity (GO:0030234)
nucleic acid binding transcription factor activity (GO:0001071)
protein binding transcription factor activity (GO:0000988)
receptor activity (GO:0004872)
structural molecule activity (GO:0005198)
translation regulator activity (GO:0045182)
transporter activity (GO:0005215)

Upregulated
n=10



binding (GO:0005488)
catalytic activity (GO:0003824)
nucleic acid binding transcription factor activity (GO:0001071)
structural molecule activity (GO:0005198)

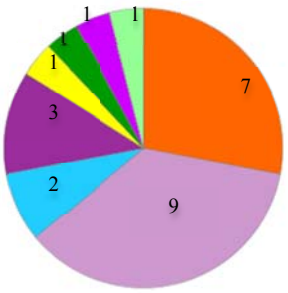
Liver

Downregulated
n=23



binding (GO:0005488)
catalytic activity (GO:0003824)
enzyme regulator activity (GO:0030234)
nucleic acid binding transcription factor activity (GO:0001071)
protein binding transcription factor activity (GO:0000988)
receptor activity (GO:0004872)
structural molecule activity (GO:0005198)
transporter activity (GO:0005215)

Upregulated
n=19



binding (GO:0005488)
catalytic activity (GO:0003824)
enzyme regulator activity (GO:0030234)
nucleic acid binding transcription factor activity (GO:0001071)
protein binding transcription factor activity (GO:0000988)
transporter activity (GO:0005215)

Figure 3.16: Pie charts of the molecular function of genes significantly (p<0.001 & FC>1.2) downregulated and upregulated in the brain and liver of embryos sired by males treated with EtOH. The genes were grouped into these categories by the PANTHER classification system. n indicates the number of genes that were either up or downregulated. The number in each portion of the pie chart indicates the number of candidate genes dysregulated per molecular function.

The PANTHER (Thomas et al., 2003) analysis suggests that the largest molecular function categories of genes dysregulated (increase or decrease in expression) in both the brain and liver of embryos sired by males exposed to EtOH are binding, catalytic and nucleic acid binding transcription factor activity. While genes involved in transporter activity and enzyme regulator activity are dysregulated to a lesser extent. As only ten genes were upregulated in the brain and submitted for analysis, it might explain why only four functions were identified compared to the other groups. Genes were not confined to one category, as they may have numerous molecular functions.

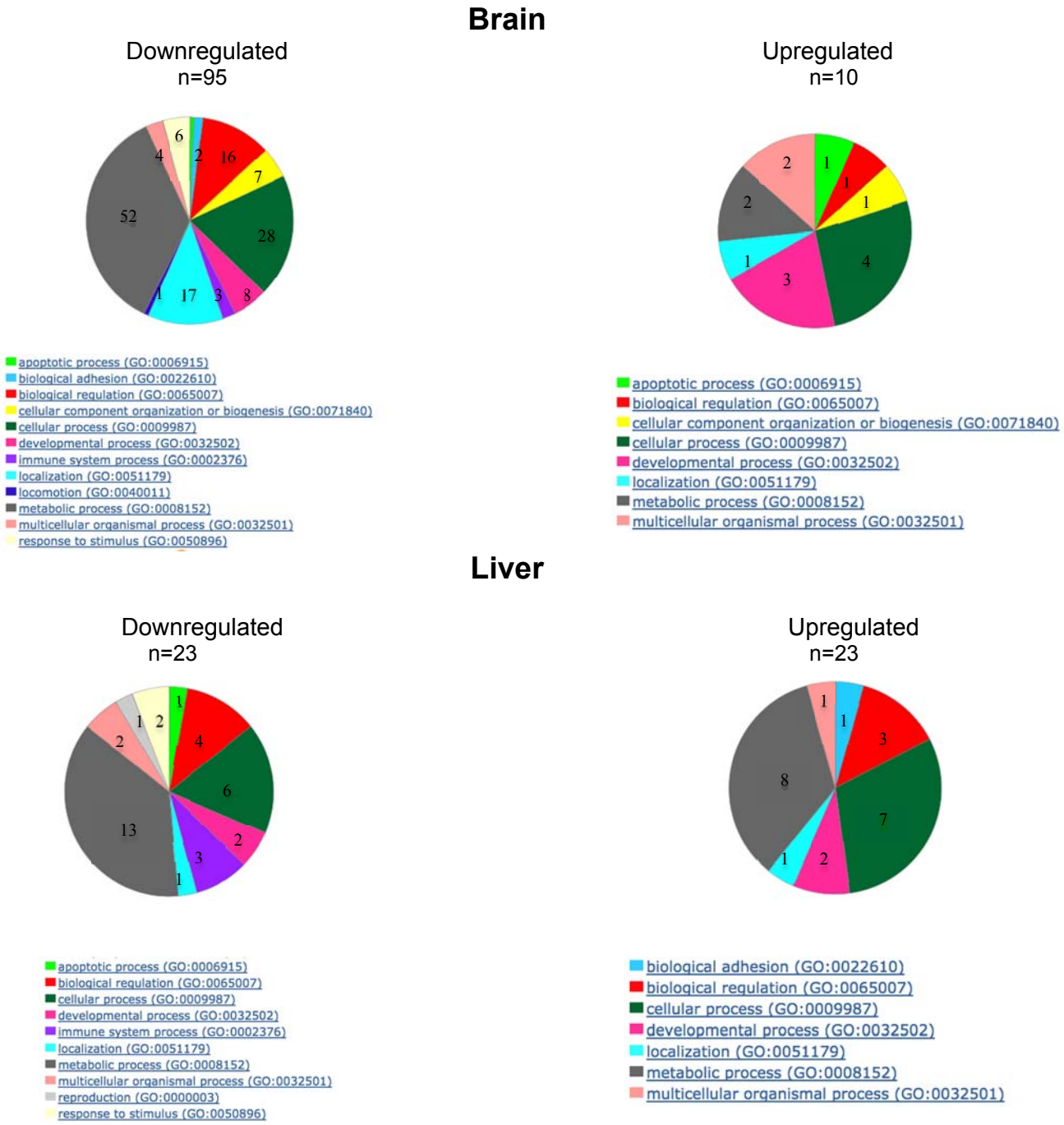


Figure 3.17: Pie charts of the biological processes of genes significantly (p<0.001 & FC>1.2) downregulated and upregulated in the brain and liver of embryos sired by males treated with EtOH. The genes were grouped into these categories by the PANTHER classification system. n indicates the number of genes that were either up or downregulated. The number in each portion of the pie chart indicates the number of candidate genes dysregulated per biological process.

The PANTHER (Thomas et al., 2003) analysis suggests that the largest biological process categories of genes dysregulated (increase or decrease in expression) in the brain and liver of embryos sired by males exposed to EtOH are, metabolic, cellular and developmental processes and biological regulation. Gene's associated with cellular components and biogenesis were exclusive to the brain. Genes involved in response to stimulus and immune system responses appear to be downregulated in response to alcohol in both the brain and liver. Genes were not confined to one category, as they may be involved in various biological processes.

Overall, the PANTHER (Thomas et al., 2003) analyses suggest that preconception paternal alcohol exposure down regulates genes in the embryonic brain and liver, as fewer genes were upregulated. Genes associated with catalytic and binding activity and metabolic and cellular processes appear to be most vulnerable to the effects of alcohol in both the brain and liver.

3.4.3 Functional enrichment analysis using DAVID

The four candidate gene lists were then submitted individually to DAVID v6.7 (Huang et al., 2005, Huang et al., 2009) which identified enriched biological functions or processes relative to the Mouse WG-6_V2_0_R2_11278593_A Chip. The biological functions and pathways that were significantly enriched are displayed in **Table 3.8** and **3.9**. These results were generated using the DAVID (Huang et al., 2005, Huang et al., 2009) functional annotation-clustering tool, which clusters highly similar functional annotations into groups and provides an enrichment score to rank the overall importance of the gene groups. The enrichment score is the geometric mean of the enrichment p-values for each annotation term associated with gene members in the group. An enrichment score of 1.3 is equivalent to non-log scale 0.05 (Huang da et al., 2009). **Table 3.8** and **3.9** summarise the functional annotation clustering data output and the molecular functions and biological processes were based on enrichment scores and enrichment p-values.

The analysis that focused on genes downregulated in the brain of embryos sired by males exposed to EtOH highlighted various molecular functions and biological processes significantly enriched within the gene list, of which majority were associated with energy production such as the mitochondria, oxidative phosphorylation and glycolysis. While ribosomal proteins or processes related to the ribosome appear to be the most significantly enriched within the gene list (highest enrichment score and lowest p value). From the genes significantly upregulated in the embryonic brain only one annotation cluster was identified and this included zinc orientated molecular functions (**Table 3.8**).

Only one cluster was identified in genes downregulated in the liver (**Table 3.9**) and processes within this cluster included two intimately connected processes namely wound healing and inflammatory response and two broad GO terms, extracellular matrix and secretion. No molecular functions or biological processes were statistically enriched in the genes that were upregulated in the embryonic liver. No functions or processes overlapped between the brain and the liver.

Table 3.8: Biological functions or pathways overrepresented among the top ranked candidate genes expressed in the brain. The gene count indicates the amount of genes from the candidate list that were associated with a particular function or pathway. p<0.05 was selected as the significance threshold on DAVID.

Embryonic Brain Gene Expression								
Downregulated					Upregulated			
Annotation Cluster 1	Enrichment Score: 2.2	Genes		P value	Annotation Cluster 1	Enrichment Score: 1.02	Genes	P value
GOTERM_BP_FAT	Translation	8	*Elif4e, Mrp13, Mrp120, Mtrl20, Mtrf1, Rp18a, Rp13A, Kpna2, Rps13*	< 0.01	GOTERM_MF_FAT	Zinc ion binding	4 *Aebp2, Gli3, Mmp2, Plagl2*	0.03
GOTERM_MF_FAT	Structural constituent of ribosome	6	*Mrp13, Mrpl20, Rp18a, Rp113A, Kpna2, Rps13*	< 0.01	INTERPRO	Zinc finger, C2H2-like	3 *Aebp2, Gli3, Plagl2*	0.03
SP_PIR_KEYWORDS	Ribonucleoprotein	8	*Larp1b, Hnrpll, Mrp13, Mrpl20, Rp118a, Rp113A, Kpna2, Rp513*	< 0.01	INTERPRO	Zinc finger, C2H2-type	3 *Aebp2, Gli3, Plagl2*	0.03
GOTERM_MF_FAT	Structural molecule activity	9	*Ctnna2, Mrp13, Mrpl20, Rp18a, Tunc3a, Rp13A, Kpna2, Rp513, Vapa*	< 0.01				
KEGG_PATHWAY	Ribosome	5	*Mrp13, Mrpl20, Rp18A, Rp13A, Kpna2 Rps13*	< 0.01				
GOTERM_CC_FAT	Large ribosomal subunit	3	*Mrp13, Mrpl20, Rp13A*	0.02				
Annotation Cluster 2	Enrichment Score: 1.98	Genes		P value				
SP_PIR_KEYWORDS	Mitochondrion	14	*Abce1, Ndufb5, Ndufb6, Cox7a2, Hnrpll, Mrpll3, Mrpl20, Mtrf1l, Oxr1, Gm9391, Ppa2, Brp441*	< 0.01				
SP_PIR_KEYWORDS	Transit peptide	7	*Ndufb5, Cox7a2, Hnrpll Mrpl20, Mtrf1r, Ppa2, Prdx5*	0.04				
Annotation Cluster 3	Enrichment Score: 1.76	Genes		P value				
KEGG_PATHWAY	Oxidative phosphorylation	6	*Atp6v1a, Ndufb5, Ndufb9, Cox7a2, Gm9391, Ppa2*	< 0.01				
GOTERM_MF_FAT	Inorganic cation transmembrane transporter activity	4	*Atp6v1a, Cox7a2, Gm9391, Trf*	0.02				
Annotation Cluster 4	Enrichment Score: 1.65	Genes		P value				
INTERPRO	Thioredoxin-like	3	*Prdx5, Glrx3, Txndc12*	0.02				
GOTERM_BP_FAT	cellular homeostasis	6	*Gjd2, Prdx5, Glrx3, Prkcb, Txndc12, Trf*	0.02				
Annotation Cluster 5	Enrichment Score: 1.48	Genes		P value				
SP_PIR_KEYWORDS	protease inhibitor	4	*Ambp,Kng1, Serpina1b, Stfa1*	0.02				
GOTERM_MF_FAT	enzyme inhibitor activity	5	*Ambp, Kng1, Serpina1b, Pkig, Stfa1*	0.03				
GOTERM_MF_FAT	endopeptidase inhibitor activity	4	*Ambp, Kng1, Serpina1b, Stfa1*	0.04				
Annotation Cluster 6	Enrichment Score: 1.46	Genes		P value				
GOTERM_BP_FAT	generation of precursor metabolites and energy	7	*Atp6v1a, Ndufb5, Ndufb9, Mdh1, Pgk1, Gm9391, Eno1*	< 0.01				
GOTERM_BP_FAT	glycolysis	3	*Mdh1, Pgk1, Eno1*	0.02				
GOTERM_BP_FAT	glucose catabolic process	3	*Mdh1, Pgk1, Eno1*	0.02				

GOTERM_BP_FAT	hexose catabolic process	3	*Mdh1, Pgk1, Eno1*	0.02
GOTERM_BP_FAT	monosaccharide catabolic process	3	*Mdh1, Pgk1, Eno1*	0.03
UP_SEQ_FEATURE	binding site: Substrate	6	*Nt5c, Gad1, Mdh1, Pgk1, Eno1, Uck1*	0.03
GOTERM_BP_FAT	cellular carbohydrate catabolic process	3	*Mdh1, Pgk1, Eno1*	0.03

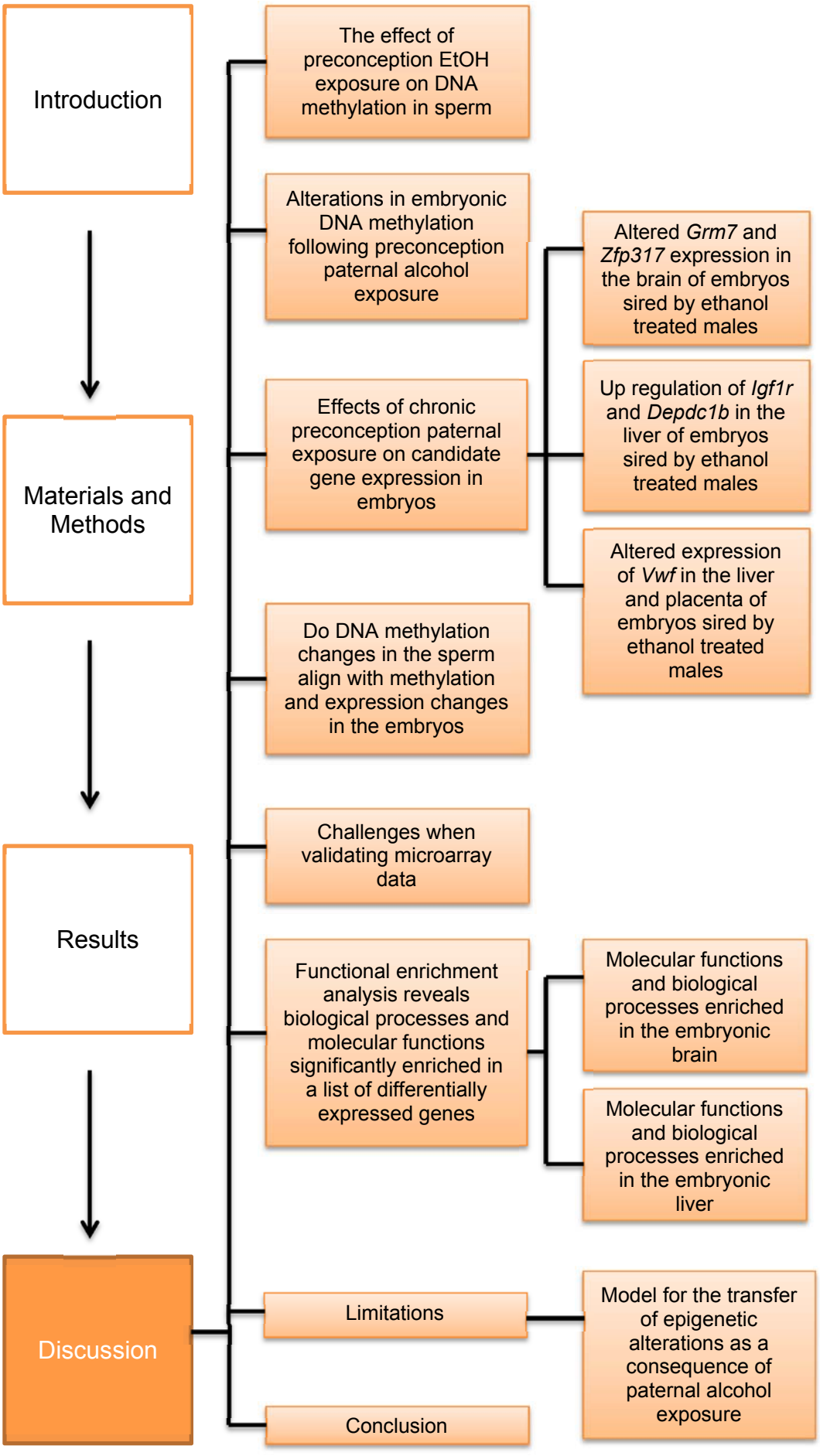
Table 3.9: Biological functions or pathways overrepresented among the top ranked candidate genes expressed in the liver. The gene count indicates the number of genes from the candidate list that were associated with a particular function or pathway. p<0.05 was selected as the significance threshold on DAVID.

Embryonic Liver Gene Expression				
Downregulated				
Annotation Cluster 1	Enrichment Score: 1.28	Genes		P value
GOTERM_CC_FAT	Extracellular region	7	*Azgp1,Arg1, Chi3l4, C8g, Pon1, Chi3l3, Tff3*	< 0.01
GOTERM_BP_FAT	Response to wounding	4	*Arg1, Chi3l4, C8g, Chi3l3*	< 0.01
SP_PIR_KEYWORDS	Secretion	6	Azgp1, Chi3l4, C8g, Pon1, Chi3l3, Tff3	0.01
GOTERM_BP_FAT	Inflammatory response	3	*Chi3l4, C8g, Chi3l3*	0.02

Chapter 4

Discussion

4.0 Discussion



Discussion

Rodent models have been used to study the teratogenic effects of adult and *in utero* alcohol exposure, and have associated these events with growth retardation, central nervous system (CNS) malformations, distinct craniofacial dysmorphology and mental disabilities (as reviewed by Ungerer et al., 2013). The FASD-like phenotype that these rodent models exhibit, have been associated with changes in global gene expression, predominantly in the developing brain (Hard et al., 2005, Kleiber et al., 2012, Hashimoto-Torii et al., 2011). This association, in combination with the role of epigenetic mechanisms in the regulation of gene expression, suggests that ethanol exposure may disrupt the normal epigenetic mechanisms such as DNA methylation, histone modification and non-coding RNAs (ncRNAs) (as reviewed by Ungerer et al., 2013).

The present study was designed to investigate the epigenetic inheritance of aberrant DNA methylation signatures as a consequence of chronic paternal alcohol exposure and the effect on embryonic gene expression in mice. The discussion that follows will examine the outcomes of these findings. Firstly, I will discuss the evidence (or lack thereof) of significant differences in the sperm methylome between ethanol and sucrose exposed males and the possible reasons that the current study (assessed using pyrosequencing data) could not validate the significant findings of the original study (assessed using RRBS data). Secondly, the transmission of the altered DNA methylation marks (assessed using RRBS data) in the sperm to the somatic tissues of the embryos (assessed using pyrosequencing data) will be addressed. Thirdly, alterations in embryonic gene expression (assessed using real-time PCR data) as a consequence of paternal alcohol exposure will be discussed. Fourthly, the correlation between methylation changes in the sperm and embryos to the gene expression changes in embryos will be addressed. Lastly, the reanalysis of the microarray data and the findings of the functional enrichment analysis of the differentially expressed genes will be explored.

4.1 Effect of alcohol exposure on DNA methylation in the sperm of male mice

Over the past decade there has been a rapid accumulation of evidence suggesting that DNA methylation marks that were once thought to be remnant patterns of spermatogenesis and contributors to the transcriptionally silent state of mature sperm, have an important role in guiding stages in embryonic development. They are likely to be influenced by the environment, thereby conferring phenotypic effects to offspring (Jenkins et al., 2015).

To address the question as to whether alcohol exposure has the ability to perturb global methylation levels in sperm of male mice exposed to alcohol, Knezovich et al. (2014) used RRBS to examine the sperm methylome of male mice exposed to chronic amounts of ethanol (n=10) or sucrose (n=10) for ten weeks. RRBS enriches for CGIs across the genome. Methylation levels at captured CGIs in the ethanol-exposed sires were compared to the sucrose controls to identify significantly differentially methylated regions (Chapter 1, section 1.7). This generated a list of differentially methylated genes. This list was then examined for genes that also showed differential expression in the embryonic tissues, to produce the candidate gene list (Chapter 1, section 1.7) used in the present study. According to the RRBS data, eight of the ten candidate genes (*Grm7*, *Vwf*, *Tet1*, *Ccnd3*, *Odc1*, *Pcbp3*, *Igf1r*, *Zfp317*) were significantly hypomethylated ($p<0.05$) and two significantly hypermethylated ($p<0.05$) (*Psmg2* and *Depdc1b*). The present study used pyrosequencing to validate the significant RRBS findings by Knezovich et al. (2014).

However, due to limited amounts of sperm DNA, only six test and six control samples could be used and only five (*Grm7*, *Vwf*, *Odc1*, *Depdc1b*, *Tet1*) of the ten candidate genes were analysed by pyrosequencing. Two approaches were used when analysing the RRBS and pyrosequencing sperm data in parallel: (1) quantification of the candidate gene mean methylation levels and (2) quantification of candidate genes CpG site-specific methylation levels. For both approaches, only one (*Vwf*) out of the five genes appeared to consistently

show the same direction of change with a statistically significant difference (Mean methylation p value: 0.06, CpG site 125553737 p value: 0.02).

While statistically ($p<0.05$) altered mean and site-specific methylation profiles were observed in the sperm DNA of ethanol exposed sires for *Tet1*, *Odc1*, *Depdc1b* using pyrosequencing, this was not considered a validation of the original experiment because in all three cases there was discordance in the direction of altered methylation levels in comparison to the RRBS data. The reason for the discordance in direction is unclear, however, it is likely that the concentration of the sperm DNA was too low to replicate the RRBS DNA methylation findings. Similar inconsistent pyrosequencing results were observed in a study conducted by White (2006) in which the *Small Nuclear Ribonucleoprotein Polypeptide N* (*SNRPN)* gene methylation levels were quantitatively analysed as a diagnostic test for Prader-Willi and Angelman Syndromes. Part of the study involved analysing the effect of DNA concentration on the reproducibility of methylation data. The results revealed that low amounts of DNA template have the potential to introduce bias into the PCR reaction because of allele specific preferential amplification. As seen in **Figure 4.1**, when 2.5ng [5] and 5ng [4] of template DNA was added to the pre pyrosequencing PCR reaction, the methylation results were highly variable within replicates and contradicted the results generated when 50ng [1], 20ng [2] and 10ng [3] of sample DNA was used. This study indicates that a minimum of 10ng of template DNA should be added to the PCR reaction if reproducible results are to be obtained. The present study used sperm DNA ranging from 2ng to 5ng, which might account for the discrepancies observed between the techniques.

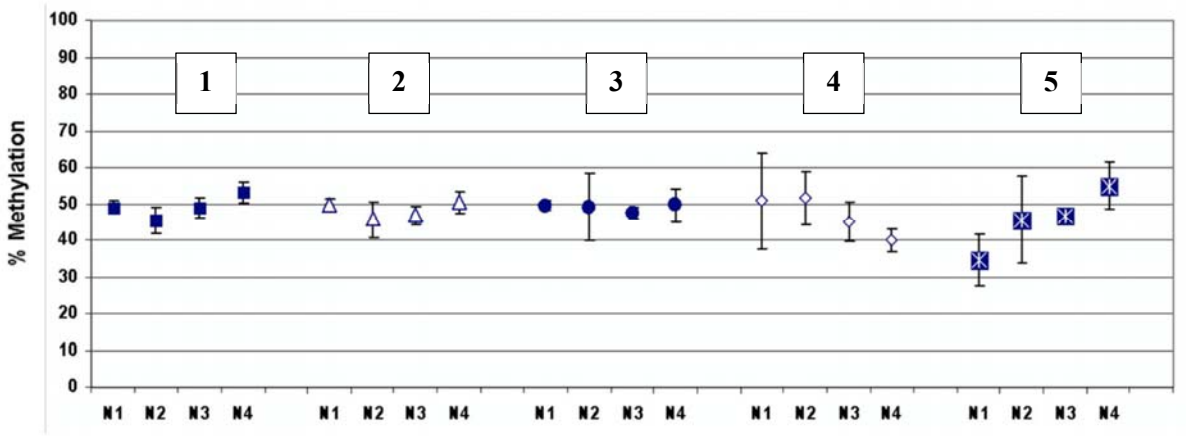


Figure 4.1: The effect of bisulfite treatment and DNA concentration on the quantification of DNA methylation, as determined by pyrosequencing. Analysis of four control samples (N1 to N4) in which 50ng [1], 20ng [2], 10ng [3], 5ng [4], and 2.5ng [5] of template DNA was added to the PCR reaction. The values for percentage of methylation were corrected for methylation bias. The error bars indicate the standard deviations for triplicate analyses (White, 2006).

Pyrosequencing data from the current study indicated no difference in DNA methylation levels at *Grm7* in the sperm of ethanol-exposed male mice relative to the sucrose-exposed controls. Although *Vwf* appeared to be validated, since it consistently showed the same direction of change as RRBS at a significant level, the low concentration of DNA, the small sample size and the discordance of the other genes cast reasonable doubt on the reliability of this finding.

Since the pyrosequencing data generated may not accurately represent the methylation differences in the sperm between the ethanol and sucrose exposed sires, as a result of the limited quantity and concentration of the DNA available, the pyrosequencing data does not validate nor does it discredit the RRBS data. Therefore, for the purpose of evaluating DNA methylation and gene expression changes in the embryonic tissues, and determining whether these findings correlated to methylation changes in the sperm (as the per the RRBS data), the RRBS data were considered to be the more reliable results and the possible biological consequences were explored.

As previously described, the RRBS data suggested that eight of the ten candidate genes were hypomethylated in the ethanol-exposed males. These findings are consistent with previous studies in which alcohol exposure was associated with hypomethylation of sperm DNA. For instance, in 2009 Ouko et al. aimed to investigate whether an association existed between alcohol use in men and hypomethylation of paternally imprinted loci in their sperm DNA, regions considered critical for embryonic development. This was done in an attempt to provide a mechanism for paternal effects in the etiology of FASD. Their study demonstrated a correlation between chronic paternal alcohol consumption and hypomethylation (lower than normal methylation) within imprinting regions (*H19* and *IG-DMR*) of the sperm DNA, suggesting that these epigenetic changes have the potential to be transmitted through fertilisation, which may alter the gene expression patterns in their offspring, which may then manifest as a FASD-like phenotype (Ouko et al., 2009). In a similar study, Knezovich and Ramsay (2012) investigated the DNA methylation patterns at two paternally methylated Imprinting Control Regions (ICRs) (*H19* and *Rasgrf1*) in the sperm of exposed male mice and somatic DNA of sired offspring following preconception paternal alcohol intake. They observed significant reductions at the *H19* CCCTC-binding factor (CTCF) 1 ($p = 0.0027$) and CTCF 2 ($p = 0.0009$) binding sites in the offspring of ethanol-treated males, which was correlated with a reduction in weight at postnatal days 35-42 ($p < 0.05$).

Although the majority of the candidate genes were significantly hypomethylated as a consequence of ethanol exposure, the RRBS revealed two candidate genes (*Psmg2* and *Depdc1b*) that were significantly ($p<0.05$) hypermethylated in the sperm DNA of the ethanol-exposed males. These findings are similar to two studies conducted by Liang and colleagues in which both hyper- and hypomethylated regions were observed in the sperm of male mice exposed to chronic amounts of ethanol (Liang et al., 2014, Liang et al., 2015).

In the first study Liang et al. (2014) investigated whether paternal alcohol exposure could alter the sperm methylome of male mice and whether these changes could be transmitted

to offspring giving rise to developmental disorders, specifically FAS. Male mice were treated with ethanol (intragastric administration: 0, 1.1, 3.3g/kg) and behavioral changes of their offspring were assessed. One paternally (*H19*) and three maternally (*Peg3*, *Ndn* and *Snrpn*) imprinted genes were evaluated, and the DMR methylation status of the four imprinted genes were assessed in the sperm and cerebral cortices of their offspring. A significant increase in methylation was observed at the *Peg3* DMR in the 3.3g/kg ethanol-exposed male mice compared to the control group. Differences in embryonic methylation were detected at the CpG7 and CpG11 in *Peg3* and *Snrpn* in the cerebral cortices of ethanol-exposed offspring that exhibited circling behavioural patterns (behavioural abnormality) relative to the control group. Similarly, Knezovich and Ramsay (2012) observed a significant decrease in methylation at *H19* in the ethanol-exposed male mice, suggesting that this locus is particularly sensitive to ethanol exposure.

Following this, Liang et al. (2015) used a similar mouse model to evaluate the effects of chronic paternal ethanol exposure with hearing loss in offspring. The methylation of *H19*, *Peg3*, *Ndn* and *Snprn* were assessed in the sperm of male mice and the cerebral cortices of deaf mice. Similar to their previous study they identified a significant decrease in methylation levels at the *H19* DMR and significant increase at the *Peg3* DMR in the 3.3g/kg ethanol-exposed male mice compared to the control. Their data suggested alterations in methylation of *Peg3* (CpG 3, 7 and 9) in the male sperm and in the cerebral cortices of deaf mice. Collectively, these studies suggest that alcohol exposure acts in a locus specific manner.

Majority of studies have associated alcohol exposure with hypomethylation (Knezovich and Ramsay, 2012, Ouko et al., 2009, Choi et al., 1999), and have begun to explore the mechanisms that may underlie this effect, while alcohol induced hypermethylation remains poorly understood. Alcohol is known to disrupt the folate-dependent pathway and subsequent activity of methionine synthase. The altered activity of methionine synthase results in an accumulation of DNMT inhibitors causing disruption to the establishment of

maintenance of normal methylation patterns (Halsted et al., 2002). It has also been demonstrated that *Dnmt1* knockouts and mutations of *Dnmt1* result in severe global hypomethylation in mice (Spada et al., 2007, Trinh et al., 2002, Gaudet et al., 2003).

As previous studies indicate that all three epigenetic mechanisms are sensitive to alcohol exposure, and since their functions are interconnected, it is likely that alcohol could have altered all three mechanisms and together they may explain the methylation differences in sires exposed to ethanol. While the nature of what makes these regions of the sperm methylome vulnerable to the environment is difficult to elucidate, studies have speculated that it may be one or a combination of factors related to chromatin architecture, developmentally important genes or potentially a consequence of timing and dosage of alcohol exposure (Radford et al., 2014, Pacchierotti and Spano, 2015).

In summary, the present study could not serve to replicate the previously observed significant RRBS findings due to the low concentration and absolute quantity of the DNA available. Therefore, it remains unclear whether the methylation differences observed through RRBS were reliable. However, if one considers the RRBS data to be reliable, the aberrant DNA methylation observed in the mature sperm DNA would indicate that the DNA methylation errors occurred as a consequence of alcohol exposure, which was not corrected during spermatogenesis and sperm maturation. Thus, the possibility exists for these methylation alterations to be transmitted to sired offspring.

4.2 Alterations in embryonic DNA methylation following preconception paternal alcohol exposure

Although still in its infancy, experiments using rodent models have begun to determine whether exogenous exposures to some measurable factor have the potential to alter epigenetic mechanisms in the germline. These studies have investigated changes in DNA methylation at a global, gene-specific and genome-wide level (as reviewed by Pacchierotti

and Spano, 2015). The vast majority of studies hypothesise that methylation changes in gametes could resist zygotic reprogramming and have functional consequences in the offspring sired by exposed animals (as reviewed by Pacchierotti and Spano, 2015).

Therefore, in an attempt to determine whether the alcohol-induced differentially methylated CpG sites (according to the RRBS data) in the sperm of the sires were inherited by their offspring, the methylation status of the CpG sites were analysed in embryonic brain, liver and placenta samples using pyrosequencing. The differentially methylated sites identified by RRBS were not always located within a gene, and were therefore allocated to the closest gene (Chapter 1, section 1.7). The number of differentially methylated CpG sites differed between genes, ranging between three and six. Using the pyrosequencing software (Biotage, Uppsala, Sweden) primers were designed to capture the CpG sites of interest. However, due to primer binding constraints not all the CpG sites that were captured by RRBS were analysed by pyrosequencing. While some advocate that the methylation status at one or two CpG sites could be used as a proxy for the rest of the region, this may not always be true (Barrera and Peinado, 2012). The present study in combination with other studies, has shown that using a few CpG sites as a proxy for the region may be misleading as CpG sites next to one another or in close proximity can differ significantly in their methylation levels (Marjonen et al., 2015, Radford et al., 2014), and using a few CpG sites as a proxy of the region can often mask opposing effects. That is, regions that show increased levels of DNA methylation as a result of alcohol exposure, are averaged out by regions that show decreased levels when they are added together.

While it is true that the RRBS approach provides a comprehensive view of the methylation status at a genome-wide and site-specific level, the pyrosequencing primers were unable to capture the complete region, as primer binding efficiency decreases with larger amplicons. Therefore, it is preferable to design more than one primer per candidate gene, which would allow for the analysis of all the significantly differentially methylated CpG sites captured by RRBS.

According to the embryonic pyrosequencing data, exposure of male mice to ethanol does not appear to cause dramatic effects in the tissue (brain, liver and placenta) methylation profiles of their embryos. However, a few significant differential DNA methylation findings were observed. For instance, *Odc1* exhibited a significant increase (p=0.001) in mean and site-specific methylation levels (CpG 17514120 p=0.04, CpG 17514135 p=0.04, CpG 17514154 p =0.005) in the placenta of embryos sired by ethanol-treated males. Furthermore, a significant decrease (p=0.04) in methylation was observed at *Odc1* CpG site 17514120 in the brain of the test group relative to the controls. *Igf1r* displayed a significant decrease in methylation at one CpG site (CpG 75274301 p=0.04) in the embryonic brain and a trend towards significance at one CpG site in the liver (CpG 75274324 p=0.08) of the embryos sired by ethanol treated males. Two CpG sites in *Depdc1b* exhibited a slight increase in methylation in the liver of embryos sired by males treated with ethanol, which were trending toward significance (CpG 109101087 p=0.08, CpG 109101098 p=0.06). While the changes were statistically significant, they all had a small effect size (less than 5% DNA methylation difference). Although, no current consensus exists on how large a methylation change needs to be to have a functional impact at the gene expression level, the level required may be context dependant, rather than absolute.

As significant differences in DNA methylation were observed at *Odc1*, *Igf1r* and *Depdc1b* in the sperm (according to the RRBS data) and the embryonic tissues (according to the pyrosequencing data), the correlation suggests that the alterations induced by ethanol exposure in the sperm may have been inherited as epimutations (altered DNA methylation profiles) by sired offspring. However, the RRBS revealed methylation differences of greater than 10%, whereas the methylation differences in the embryos were less than 5%.

It is currently unclear why the methylation status of CpG sites in the same region may vary substantially from each other. Given that differential methylation at a single CpG site is sufficient to alter the expression of a gene (Zhang et al., 2010, Nile et al., 2008), it is

believed that these regions could have a role in gene regulation. For instance, they may be transcription factor binding sites. Hypermethylation as a consequence of alcohol exposure may disrupt the binding of transcription factors and thus suppress gene expression. Whereas hypomethylated regions allow for transcription factors to easily access the DNA thereby increasing gene expression.

To date, there have been a limited number of studies that have provided an association between paternal alcohol exposure and DNA methylation aberrations in sired offspring, and in cases where an association was detected; it is often within an imprinted gene. In 2011, Stouder et al. evaluated the effects of prenatal alcohol exposure on five imprinted genes (*H19*, *Gtl2*, *Peg1*, *Snrpn* and *Peg3*) in somatic and sperm DNA in male offspring. Alterations in *H19* methylation were observed in the sperm of the F1 males as well the brain cells of their offspring, with a good correlation between the demethylated CpG sites between the generations (Stouder et al., 2011). Similarly, previously mentioned studies conducted by Liang and colleagues identified significant differences ($p<0.05$) in methylation in *Peg*3 in the sperm of ethanol-exposed male mice and in the cerebral cortices of their sired offspring in comparison to the sucrose controls (Liang et al., 2015, Liang et al., 2014).

Interestingly, in 2013 Wei et al. used a paternal high-fat induced mouse model to show the transmission of glucose intolerance and insulin resistance to their offspring. Using Methylated DNA immunoprecipitation sequencing (MeDIP-Seq), several insulin signalling genes exhibited differential methylation levels in the test offspring relative to the controls, and these changes correlated with changes in gene expression. The analysis of cytosine methylation profiles in the sperm of the prediabetic fathers showed several alterations and a large proportion of differentially methylated genes overlapped with that of the offspring's pancreatic islets. Bisulfite sequencing of some of these genes in blastocysts showed that they resisted global post fertilisation demethylation and largely inherited cytosine methylation from the sperm, suggesting that there might be an intergenerational transmission of methylation profiles (Wei et al., 2014). In a study investigating the

transgenerational effect of tamoxifen in rats, the offspring sired by tamoxifen-treated males displayed methylation differences in *H19* ICR similar to those observed in the sperm of exposed sires (Pathak et al., 2009).

Due to the complexity of the epigenome it is difficult to deduce why only three of the candidate genes in the embryonic tissues appear to have sites of significant methylation differences; why methylation differences are not consistent through different embryonic tissues; and why the effect size in the embryos is so small relative to the changes observed in the sperm of the sires. However, there are plausible explanations. Firstly, it may be partially attributed to the complexity and limitation of studying sperm epigenetics. Sperm studies rely on an estimation of the sperm population rather than individual spermatozoa and the epigenetic variability between different sperm cells from the same population remains unclear (Flanagan et al., 2006, Jenkins et al., 2015). Therefore, when methylation levels were quantified in a sperm sample, the entire sample would be regarded as hyper- or hypo-methylated based on the majority of the sperm within a population. However, only one sperm cell within the population would fertilise the egg. This explanation could account for discrepancies between methylation in sperm and embryo samples. Secondly, sperm samples are haploid and are only representative of the paternal line whereas embryos have received both maternal and paternal epigenetic profiles, which may balance out or mask any significance differences that may have been induced by paternal ethanol exposure. Thirdly, the lack of significant methylation differences in other candidate genes may be due to the erasure and reestablishment of methylation marks during meiosis and offspring development. This germline epigenetic reprogramming allows not only for the resetting of parental imprints in successive generations, but also prevents the propagation of epimutations (Reik et al., 2001).

In summary, no large-scale differences in methylation were observed in any of the candidate genes in any of the embryonic tissues. However, significant alterations of small effect size (less than 5%) were observed at individual CpG sites in the brain (*Odc1*, *Igf1r*)

liver (*Depdc1b*, *Igf1r*) and placenta (*Odc1*). Since methylation regulates gene expression, the functional impact of the methylation differences on the levels of mRNA expression of the genes under investigation was determined using real-time PCR.

4.3 Effects of chronic preconception paternal alcohol intake on candidate gene expression in embryos

The teratogenic actions of alcohol have been studied in mice models using both acute and long-term gestational exposure. Several studies have shown that alcohol has the potential to disrupt gene expression (direct exposure). However, the effect of preconception alcohol exposure on embryonic gene expression is poorly understood. Therefore, Knezovich et al. (2014) used a microarray approach to provide evidence that excessive alcohol intake by males prior to conception may affect gene expression in their offspring. This generated a large number of significantly differentially expressed genes in the test embryos relative to the controls.

Therefore, to validate the significance of the microarray results generated by Knezovich et al. (2014), the expression levels of nine candidate genes in specific tissues (brain, liver or placenta) were quantified using real-time PCR and analysed using the relative standard curve method. Real-time PCR was chosen for this study, as it has become the gold standard, because it is used in the majority of published microarray studies, presumably because it is a rapid approach with a sensitive throughput that requires small amounts of input RNA.

Of the nine candidate genes chosen, seven were analysed in the brain, four in the liver and two in the placenta. Differences in embryonic gene expression between the test and control groups were observed in two genes (*Grm7* and *Zfp317*) in the brain, three genes (*Igf1r*, *Depdc1b* and *Vwf*) in the liver and one gene (*Vwf*) in the placenta. Importantly, the direction of change was consistent with the microarray results. However, according to the non-

parametric two-tailed Mann-Whitney U test, none of these genes showed statistically significant changes. The failure to reach statistical significance could be a consequence of the small sample size (n=48) or the modest microarray gene expression fold change (discussed in section 4.5). The remainder of the genes (*Pcbp3*, *Psmg2*, *Tet1*, *Odc1*) showed no alteration in expression between the cases and controls.

4.3.1 Altered *Grm7* and *Zfp317* expression in the brain of embryos sired by ethanol-treated male mice

Many birth abnormalities occur during organogenesis, a time of great vulnerability to the effects of alcohol exposure. The brain is the first organ to develop, the last to mature and is sensitive to the effects of alcohol (Hard et al., 2005). Ethanol has the ability to alter neurodevelopmental expression (Green et al., 2007, Zhou et al., 2011).

According to the whole genome expression array conducted prior to this study, *Grm7* displayed a significant decrease in expression in the embryos sired by males treated with alcohol in comparison to the sucrose treated controls. This is consistent with the real-time PCR data generated in this study, which demonstrates a trend for reduced expression in the EtOH-treated group (**Figure 3.7B)**. However, when statistically analysed, the alterations in expression did not reach statistical significance (p=0.30).

Grm7 spans 880, 291bp, includes 11 exons and is located on the distal portion of the short arm of chromosome 3 (Gene Cards, 2015b). *Grm7* encodes the 915 amino acid glutamate receptor protein, a major regulator of glutamate transmission in the central nervous system as the glutamate pathway has been considered to play an important role in neural plasticity, neural development and neurodegeneration (Scott, 2014). Interestingly, *Grm7* has been associated with brain developmental defects such as ADHD, a phenotype associated with FASD, autism, and has recently been linked to alcoholism (Vadasz et al., 2007, Elia et al., 2012). Metabotropic glutamate receptors are also potential targets for

neuropsychiatric disorders and regulate neuronal excitability by the release of neurotransmitters (Pilc et al., 2008).

ADHD is a common heritable neuropsychiatric disorder of unknown aetiology. Elia et al. (2012) performed a whole-genome copy number variation (CNV) study in a cohort of ADHD cases and healthy children. CNVs impacting metabotropic glutamate receptors were significantly enriched across ADHD cohorts ($P=2.1x10^{-9}$). Among them, deletions occurred in *Grm5*, *Grm7* and *Grm8* at a higher frequency than the controls. Given that they identified variants in the glutamate receptor gene and associated them with ADHD, and that ADHD is part of the FAS phenotype that has been implicated in a mouse model following chronic alcohol exposure (Kleiber et al., 2012), this suggests that altered *Grm7* expression within the brain of ethanol-treated embryos may contribute to phenotypes that resemble ADHD.

Interestingly, down regulation of *Grm7* was shown to reduce the number of progenitor cells that differentiate into neurons. This was demonstrated by Xia et al. (2015) who investigated the role of *Grm7* in early cortical development by down regulating its expression in neuronal progenitor cells of the cerebral ventricle of embryos via in utero electroporation (IUE). The group determined that *Grm7* knockdown ultimately leads to a reduction in the number of progenitor cells that differentiate into neurons. Other studies showed that *Grm7* knock out mice have deficits in fear learning and aversive behaviour (Gee et al., 2014) and exhibit reduced anxiety and depression-like behaviours (Cryan et al., 2003).

Not only has *Grm7* been associated with neurodevelopmental disorders but it has recently been linked to alcoholism in a study that found that mice that carried *Grm7* rs30557150_C allele had lower expression levels in the brain and increased alcohol consumption (Vadasz et al., 2007). It is interesting to note that the regions of the brain vulnerable to alcohol related abnormalities (hippocampal, cerebral cortex and cerebellum) have been associated with the highest levels of *Grm7* expression (Oscar-Berman and Marinkovic, 2003, Makoff et al., 1996). Although *Grm7* has not been associated with alcoholism in humans, a gene

sequence comparison between rats and humans showed a 92% sequence identity and 99% protein identity (Makoff et al., 1996).

While more studies need to be conducted to determine how *Grm7* expression levels influence alcohol dependence, the current literature suggests that glutamate expression levels in the brain, including *Grm7* levels play an important role in addiction (Simonyi et al., 2004). Considering that alcohol abuse has been associated with decreased levels of *Grm7* expression it seems plausible that individuals who carry the *Grm7* variant might be born with an imbalance in glutamate expression levels making them more susceptible to alcohol dependence (Valenzuela, 1997).

Furthermore, when the embryonic gene expression data of the current study was stratified by sex and analysed using the non-parametric Mann-Whitney U test to determine whether there were any significant sex differences, *Grm7* was the only gene that showed sex specific differences. A significant ($p=0.01$) decrease in expression was observed in the male embryos sired by males that were exposed to alcohol. This finding gains support by Rinn et al. (2004) who examined post-mortem samples from the hypothalamus of humans with an average age of 70 years and found that men had significantly lower levels of *Grm7* expression than women. Interestingly, this finding may contribute toward understanding the significantly decreased brain ($p=0.0213$) and liver weight ($p=0.0233$) of the male embryos observed by Knezovich et al. (2014).

Considering that alcohol abuse has been associated with decreased expression of *Grm7*, in males particularly, and that men have been shown to have higher rates of alcohol dependence than women, these findings suggest that *Grm7* expression could contribute to differences that have been observed between the sexes in the prevalence of alcoholism.

Taken together, the data suggest that *Grm7* is not only associated with alcoholism but is associated with neurodevelopment deficits that are associated with FASD. Our data

suggest that male embryos sired by ethanol-treated males display significant decreases in *Grm7* expression in the brain. However, no DNA methylation differences were observed at *Grm7* in the brains of the male mice, making it difficult to identify the mechanisms perturbed by alcohol exposure that ultimately contributed to the differential expression.

Based on our data and previous studies it may be hypothesised that chronic paternal alcohol exposure altered epigenetic mechanisms in the sperm (unlikely to be DNA methylation), and these alterations were then inherited by the male embryos, which consequently resulted in significantly reduced gene expression and brain weight.

The quantitative analysis of *Zfp317* expression revealed slightly decreased expression levels in the brains of embryos sired by ethanol treated males in comparison to embryos sired by sucrose treated males. This trend is consistent with the microarray data, which showed a significant ($p=0.01$) reduction in gene expression. However, when applying the two-tailed Mann-Whitney U test the difference was not statistically significant ($p=0.45$).

The *Zfp317* gene located on chromosome 9, spanning base pairs 19641224- 19648048, encodes a zinc finger protein (Cunningham et al., 2015). Currently, the function of this protein remains unknown. However, in general, zinc finger proteins constitute the largest family often transcriptional repressors encoded in the genome of higher organisms (Lupo et al., 2013). The zinc proteins contain a Kruppel associated box positioned at the NH2 terminus, which interacts with the KAP-1 scaffold protein to recruit transcriptional factors that lead to the repression of genes that bind to the zinc finger proteins (Lupo et al., 2013). The relevance of zinc finger mediated repression is reflected in the large number of KRAB zinc proteins however, despite their abundance, little is known of their gene targets. Recent studies suggest that transcriptional repression mediated by KRAB-ZFPS functions in cell proliferation, differentiation, and apoptosis (Lupo et al., 2013).

Given that the function of *Zfp317* remains unknown, and that it's expression was only slightly altered in the embryonic brain, with no changes in methylation in the embryo, and no evidence of altered expression levels have previously been implicated in any disease phenotype, it is difficult to explain what the possible functional impact (if any) would be. However, as the general family of proteins function as transcriptional repressors, alterations in expression, although unlikely may lead to the transcriptional repression or activation of the targets of *Zfp317*, which may then have a functional impact on cell proliferation or differentiation.

4.3.2 Upregulation of *Igf1r* and *Depdc1b* in the liver of embryos sired by ethanol treated males

The liver is the primary site of alcohol metabolism and is highly vulnerable to injuries due to chronic alcohol abuse. A number of studies have explored the effects of alcohol on gene expression in liver tissue (Deaciuc et al., 2004, Clugston et al., 2011). In addition, alterations in genes expressed in the liver were observed in embryos exposed to alcohol *in utero* (Kaminen-Ahola et al., 2010, Yao and Nyomba, 2008).

According to the microarray data generated by Knezovich et al. (2014) a significant increase in gene expression was observed in *Igf1r*, *Vwf*, *Depdc1b*, *Pcbp3* in the liver of ethanol sired offspring. The real time-PCR analysis data did not show any of these changes to be statistically significant. Although, consistent with the direction of change observed in the microarray, *Igf1r*, *Vwf* and *Depdc1b* exhibited a trend for increased expression in the ethanol-sired group.

Igf1r belongs to the receptor tyrosine kinase family and mediates its action by binding to insulin-like growth factor 1 and 2 (De Meyts et al., 2004). The activated receptor is involved in cell growth and survival, plays an important role in childhood growth and continues to have anabolic effects in adults (Froesch et al., 1986). Mice lacking the *Igf1r* receptor die

late in development and show a dramatic reduction in body mass, attesting to the strong growth-promoting effect of this receptor (Liu et al., 1993, Okubo et al., 2003). Mice carrying only one functional copy of *Igf1r* are normal, but exhibit a 15% decrease in body mass (as reviewed by Cloud Clone Corp, 2015).

Reduced body weight is consistent with a FAS phenotype. Klug et al. (2003) reported that children with FAS have significantly lower weight and length at birth or in childhood. Day et al. (2002) found that 14 year olds whose mothers consumed alcohol during their first and second trimesters of pregnancy had lower body weight whereas mothers that drank during their first trimester had reduced body length. Furthermore, in a study that used a computational approach to identify candidate FASD susceptibility genes, *Igf1r* was selected as a likely candidate (Lombard et al., 2011)

Alterations in *Igf1* (which produces a protein that binds to the Igf1r receptor) expression as a consequence of alcohol exposure have previously been described. Dobson et al. (2014) hypothesized that chronic prenatal ethanol exposure alters expression of insulin and IGF signaling molecules in the prefrontal cortex and liver of the adult guinea pig offspring. Pregnant Dunlin-Hartley strain guinea pigs received ethanol or sucrose through gestation. In the liver, chronic prenatal ethanol exposure decreased *Igf1*, *Igf1r* and *Igf12* and increased *IRS* mRNA expression in male offspring compared to controls. They hypothesized that alterations in the liver may contribute to metabolic dysregulation in adult offspring.

Igf1r is particularly interesting as the RRBS data suggested a significant decrease in methylation in the sperm of ethanol exposed males, in addition trends toward significance were observed in *Igf1r* in the embryonic liver (CpG site 75274324 p=0.08). Trends for altered expression were observed in the liver (p=0.10) of exposed embryos and male embryos sired by ethanol-treated males exhibited significantly decreased brain and liver weights (Knezovich et al., 2014). Altogether, these findings suggest that alcohol may have

induced methylation changes in the sperm that were inherited by the embryos, altering gene expression, which in turn affected *Igf1r* growth-related processes. Disturbance of these processes may then contribute to the significantly reduced brain and liver sizes of the male embryos.

Similar to *Igf1r*, *Depdc1b* exhibited slightly increased levels of expression in the embryonic liver of embryos sired by males exposed to alcohol relative to their control, which was not statistically significant (p=0.34). The directional change was consistent with the microarray data in which significantly increased levels of expression were detected.

Depdc1b located on chromosome 13 has been proposed to participate in cell migration, intracellular signal transduction and positive regulation of the Wnt signal transduction pathway (Mouse Genome Informatics, 2015). Several studies have associated differential expression of *Depdc1b* in various cancers including papillary thyroid cancer, breast cancer, synovial cancer and prostate cancer (Carolina, 2013, Kanehira et al., 2007, Katagiri et al., 2009).

Considering that metabolites of alcohol are likely carcinogens, and that leukemia, the most common childhood cancer and can arise *in utero*, numerous studies have hypothesized and have provided evidence that chronic preconception and prenatal maternal alcohol consumption is associated with several childhood cancers including acute myeloid leukemia (Severson et al., 1993, Shu et al., 1996), and acute lymphocytic leukemia (van Duijn et al., 1994). Furthermore, in 2013, Murugan et al. used a rat model to determine whether any functional abnormalities and increased cancer susceptibility exist in the prostate of fetal exposed male rats during the adult period. Their study showed for the first time that ethanol exposure induces histophysiological changes in the prostate and increases the susceptibility to develop neoplasia during adulthood.

Since *Depdc1b* is critical to cell proliferation and has been associated with several cancers and that ethanol is known to dysregulate gene expression and is a potential risk factor for childhood cancers, it may be plausible to suggest that dysregulation of *Depdc1b* in conjunction with *Zfp317* (potential mediator of cell proliferation) could contribute to dysregulation in cell proliferation in subsequent generations.

4.3.3 Altered expression of *Vwf* in the liver and placenta of embryos sired by ethanol treated males

The real-time PCR quantitative gene expression analysis of *Vwf* expression revealed slightly reduced expression in the liver and slightly increased expression in the placenta of embryos sired by ethanol treated males in comparison to the sucrose treated males. This trend is consistent with the microarray data. However, according to the two-tailed Mann-Whitney U test the difference in either tissue was not significant (liver $p=0.16$; placenta $p=0.20$). The microarray results suggested a significant decrease of expression in the brain of ethanol treated embryos, however no change in expression was observed between the cases and controls using real-time PCR.

Vwf is approximately 178kb long, contains 52 exons and is responsible for encoding the von Willebrand factor, a large multimeric glycoprotein that plays a vital role in the blood coagulation system, as it serves as a mediator of platelet vessel interaction and adhesion (a critical function in response to vascular injury), and as a carrier for coagulation factor VIII (Kniffin, 2015). Reduced or abnormal VWF activity results in von Willebrand disease, a common and complex hereditary bleeding disorder (Ginsburg et al., 1985).

Although limited, there are a few studies that have investigated the effect of ethanol on coagulation factors and platelet adhesion (Salem and Laposata, 2005, Renaud and Ruf, 1996). These studies suggest that alcohol consumption may inhibit or enhance platelet

activation. Alcohol may also reduce the levels of Vwf and factor VII (Mukamal et al., 2001), although the mechanism remains unknown.

Besides its role in hemostasis, *Vwf* has been associated with cell growth and proliferation and is suggested to function in the spread and dissemination of cancer. Given that coagulation cascade factors and platelets actively interact with tumor cells to escape from the vasculature and enter into different tissues thereby enhancing the capacity of the malignant cells to spread, makes this suggestion plausible.

Although, collectively these studies suggest that ethanol exposure has the potential to disrupt the function of *Vwf*, which may alter hemostasis and potentially promote the spread of cancer, no methylation changes were observed in any of the embryonic tissues and only a small difference in gene expression was observed in the embryonic liver and placenta. It is therefore unlikely that *Vwf* would be responsible for any of the associated phenotypes on its own, but could potentially disrupt hemostasis or cell proliferation in conjunction with other genes that regulate these processes that may have been perturbed by alcohol exposure.

4.4 Do DNA methylation changes in the sperm align with methylation and expression alterations in the embryos?

As shown in the previous sections there is growing evidence demonstrating the ability of environmental factors to dysregulate methylation in the germline and that these methylation changes are correlated with methylation changes in the embryo. It is also becoming increasingly clear that environmental factors can alter gene expression during *in utero* development. However, what is unclear and what this study aimed to do was to provide evidence that paternal preconception environmental exposure (ethanol) could alter DNA methylation in the sperm, that these alterations could be transmitted to the embryos and that they may be responsible for dysregulating gene expression.

However, the present study did not identify a clear correlation between the sperm and embryos as none of the candidate genes consistently showed statistically significant methylation changes in sperm and embryonic tissue, nor did this result in the anticipated changes in gene expression. Although they were not statistically significant, correlations between DNA methylation status and gene expression were observed in *Igf1r* and *Depdc1b*. Both showed significantly reduced methylation in the sperm (according to RRBS data), trends for reduced significance in the embryo (according to the pyrosequencing data) and trends for increased expression in the embryo (real-time PCR data). These trends follow what was expected, as alcohol exposure is generally associated with hypomethylation, and hypomethylation is traditionally associated with increased gene expression.

Furthermore, *Grm7*, *Zfp317*, and *Vwf* showed trends for altered expression in the embryonic tissues but there were no correlating methylation changes in the embryonic tissues. Similar observations were noted in previous studies. For example, Radford et al (2013) evaluated the effects of undernourished dams on the epigenetic profiles of their offspring. Several genes in the sperm of the F1 generation were hypomethylated, as well as differentially expressed metabolic genes in the brain and liver of the exposed foetuses. Similar to the present study, differentially expressed genes mapped closely to differentially methylated regions in the sperm, however differential methylation was not transgenerationally retained.

The authors concluded that it was unlikely that the expression changes were directly mediated by methylation but it was rather a cumulative effect of epigenetic patterns early in development that may have resulted in sustained alterations in chromatin architecture, transcriptional regulatory networks, cell type or tissue structure (Radford et al., 2014). Similar conclusions were drawn by Lambrot et al. (2013) who used a genome-wide methylation approach to investigate methylation changes in the sperm of mice exposed to a folate-deficient diet and these changes were correlated with expression and methylation

profiles within their sired offspring. They observed significant methylation changes in the sperm and significant expression changes in the placenta, however the affected differentially methylated regions in the sperm and embryos did not correlate. Similar to Radford et al. (2014) the authors attribute these inconsistencies to the disturbance of alternative epigenetic mechanisms.

Although significant differences were observed in the expression of *Grm7* in the male brain sired by ethanol treated males in the current study, as previously mentioned, it is unlikely that this is a result of methylation changes as no *Grm7* methylation changes were observed in the embryonic brain. Based on the studies above, it is plausible that the significant differences in expression in the male brains of embryos sired by ethanol treated males were mediated by an additive effect of dysregulated epigenetic mechanisms in early development.

It is also clear that the direction of methylation change in the sperm samples does not always correlate with the direction of methylation changes in the embryos. The most likely explanation for this is that genes have tissue specific methylation patterns and whilst they might have been hyper- or hypomethylated in the haploid sperm samples, the embryo is diploid and is therefore a representative of the methylation profiles inherited from both parents.

The present study, together with previous studies suggest that epigenetic inheritance via the gametes is likely to be more complex than direct transmission of DNA methylation, and likely involves the cross-talk between different epigenetic mechanisms.

4.5 Challenges when validating gene expression microarray data with real-time PCR

Although microarray technology has revolutionized gene expression studies and allowed for the generation of high volumes of data on thousands of genes in a single experiment, it has come with significant challenges with regard to analysis and interpretation.

The lack of statistically significant real-time PCR results casts reasonable doubt on both the expression array results and the criteria for selecting genes for validation. The reliability of microarray data has also been questioned by other studies that have failed to replicate microarray results and have found discordance between microarray platforms or in some studies the same platforms (Filion, 2012). In addition, although it is accepted that microarray results need to be validated, the chosen validation platform to use remains a contentious issue.

Numerous studies have attempted to determine what factors contribute to the variation in microarray data compared to the real-time PCR data (Morey et al., 2006, Filion, 2012). Similar to the present study fewer correlations have consistently been reported for genes exhibiting small degrees of fold change, generally less that 2-fold, compared to those showing greater than 2-fold change (Rajeevan et al., 2001, Dallas et al., 2005). To the contrary, numerous studies have yielded significant results with a fold change of at least 1.4 (Morey et al., 2006). Wurmbach et al. (2003) reported 100% validation of their array results exhibiting a fold change of at least 1.6. However, they defined validation as directional confirmation only and large discrepancies in the level of the change were not addressed in their study. On the other hand, some studies consider a result valid if the FC is greater than 2 for both techniques (Rajeevan et al., 2001). The majority of studies reason that together with FC, differentially expressed microarray genes should have a significant p value ($p<0.001$) and be of biological relevance.

Other explanations for the discrepancies between techniques include the differences in experimental design and the notion that they may not be comparable. For example, the distance between the location of the PCR primers used for real-time PCR and the location of the microarray probes for a given gene may be different. However, this would not serve as an explanation of the lack of validation in the present study as the PCR primers were designed in the same region as the microarray probes (Chapter 2, section 2.7). Furthermore, Beckman et al. (2004) found that array spot intensity could potentially contribute to discrepancies between microarray and real-time data, as they found that low intensity spots had considerably lower correlation with real-time data in comparison to the high intensity spots.

In summary, this study could not validate previously generated significant microarray expression data using real-time PCR. The current study in combination with other studies has attributed the discrepancies in the two approaches to fold change and p values of the candidate genes, microarray platforms, and the data analysis. Due to a failure to validate the significant embryonic gene expression data, the next section will discuss the reanalysis of the microarray data using bioinformatics tools that generated candidate gene lists (Chapter 3, **Table 3.7**). The candidate gene lists were then submitted to PANTHER and DAVID to identify and examine biological functions, processes and pathways significantly overrepresented.

4.6 Bioinformatics approach reveals molecular functions and biological processes significantly enriched in a list of differentially expressed genes in embryonic tissues

The gene expression array data were reanalysed to focus on genes with significant differential expression on the microarray array, not taking into account the DNA methylation data. Volcano plots (Chapter 3, section 3.4.1) were generated to identify genes that were significantly dysregulated ($p<0.001$) with a fold change greater than 1.2 in the embryonic

brain and liver tissues. As only four genes in the placenta passed the significance threshold (Chapter 3, **Figure 3.15**), these genes were not further analysed.

The volcano plots revealed that the total number of genes significantly dysregulated ($p < 0.001$ and $FC > 1.2$) in the brain of embryos sired by ethanol-treated males exceeded those dysregulated in the embryonic liver and placenta. However, the number of genes that were downregulated in the brain was greater than those upregulated in the brain, whereas the number of genes up and downregulated in the liver was almost equal. Furthermore, only five genes were significantly dysregulated ($p < 0.001$ and $FC > 1.2$) in the placenta of ethanol-sired embryos. These findings collectively, suggest that genes in the brain may be vulnerable to the effects of alcohol whereas the placenta is resilient to the effects of alcohol. Contrary to this, Haycock and Ramsay (2009) observed a significant decrease in placental weight and methylation in the paternal allele (*H19*) of the ethanol-treated embryos in comparison to the saline-exposed controls. However, this was observed when female mice were prenatally exposed to alcohol.

Significantly upregulated and downregulated gene lists for the brain and liver tissues were submitted to PANTHER (Thomas et al., 2003). Using PANTHER, the bioinformatics tool clustered genes that perform similar molecular functions and biological processes together and displayed the results as a pie chart, without doing any statistical analyses. Lastly, the gene lists were submitted to DAVID (Huang et al., 2005, Huang et al., 2009) for a functional enrichment analysis. The functional enrichment analysis used statistical methods to find functional annotations (e.g. metabolic pathways, biological processes and cellular components) that when tested against a background set of genes were significantly enriched. This analysis relies on biological databases (e.g. Gene Ontology, KEGG, etc.) and is a promising strategy that increases the likelihood for investigators to identify biological processes most pertinent to the biological phenomena under investigation.

The broad range and large number of functions and processes identified by PANTHER (Thomas et al., 2003) and DAVID (Huang et al., 2005, Huang et al., 2009) may be explained by the fact that unlike other teratogens, ethanol has a wide host of effects on multiple targets through its effect on gene expression alteration. It does not have a single receptor but instead may disturb cellular activity through its interaction with various proteins including G-protein-coupled receptors, ligand-gated ion channels and intracellular signalling proteins (Smith et al., 2014). The diversity of targets enables alcohol to alter several signalling pathways and biological processes crucial for normal development. Disruption of these processes may contribute to the dysregulation of cellular events that are central to morphogenesis such as proliferative expansion, migration and cell survival (Smith et al., 2014), especially in the developing brain. While the enrichment of genes associated with GO terms and biological processes is significant and intriguing, the biology behind the enrichment is difficult to elucidate.

4.6.1 Molecular functions and biological processes identified by gene ontology analysis of differential gene expression in embryo brains

Significantly enriched biological processes and molecular functions in the list of genes dysregulated in the brains of embryos sired by males treated with alcohol have been broadly placed into six categories including: cellular homeostasis, mitochondrial and ribosomal functions, oxidative phosphorylation, glycolysis, zinc orientated functions and the generation of precursor metabolites and energy. Although it is becoming increasingly apparent that ethanol exposure acts in a tissue and locus specific manner, the majority of studies have associated ethanol exposure with hypomethylation. Given that hypomethylation is generally associated with increased gene expression, we would expect that the majority of genes would be upregulated in the embryos. Contrary to this, the volcano plots based on the microarray data revealed that the number of genes

downregulated in the brain of embryos sired by ethanol treated males exceed the number of genes upregulated in the brain.

Control and regulation of cellular homeostasis are essential for normal embryo development and maintenance of viability (Lane and Gardner, 2000). Intracellular levels of calcium, protons, magnesium and phosphate regulate a multitude of cellular functions such as cell division, protein synthesis, differentiation, cell-cell communication, cytoskeletal dynamics, metabolism, and energy production (Lane and Gardner, 2000). Disruption of these processes has been shown to reduce ATP production, causing a cascade of events that lead to disrupted cellular function and perhaps ultimately disturbed epigenetic mechanisms leading to aberrant placental and foetal development (Lane and Gardner, 2000). Maternal alcohol exposure has been shown to disrupt several forms of homeostasis including, cellular homeostasis, foetal iron homeostasis (Smith and Huebner, 2015), cholesterol homeostasis (Guizzetti and Costa, 2007), retinoic homeostasis (Kim et al., 2015b), and glutamine homeostasis (Rathinam et al., 2006) and these changes have been associated with FASD-like phenotypes (Kim et al., 2015b, Guizzetti and Costa, 2005). Given that cellular homeostasis functions to regulate mitochondrial function, this may potentially explain why the mitochondrion was identified as a significantly enriched molecular function through the DAVID (Huang et al., 2005, Huang et al., 2009) analysis (**Table 3.8**).

The mitochondria are membrane-enclosed organelles found in most eukaryotic cells, as they are responsible for generating most of the cells ATP, which is used as a source of chemical energy. Its importance lies in that it contains Kreb cycle enzymes, β-oxidation, oxidative phosphorylation and components of the electron transport chain (as reviewed by Manzo-Avalos and Saavedra-Molina, 2010). There is evidence that ethanol alters the structure and function of mitochondria in various organs including, the liver, heart and brain, both in laboratory animals and humans (Klein and Harmjanz, 1975, Regan, 1990, Pachinger et al., 1975). These changes affect the mitochondrial function by decreasing

respiratory rates and ATP levels and increase the production of the reactive oxygen species (ROS) (as reviewed by Manzo-Avalos and Saavedra-Molina, 2010). ROS can damage membranes, DNA, and mitochondrial oxidative phosphorylation enzymes, resulting in declined mitochondrial function (as reviewed by Fosslien, 2001). Individuals who consume chronic amounts of alcohol commonly suffer from neuropathy or neurocognitive deficits that are believed to be associated with mitochondrial dysfunction in the brain (Haorah et al., 2013). Furthermore, prenatal alcohol exposure has been associated with mitochondrial dysfunction in embryos that may contribute to a FAS phenotype (Goodlett and Horn, 2001, Marin-Garcia et al., 1996). By way of example, Sanchis and Guerri (1986) and Sari et al. (2010) used rodent models to evaluate the mitochondrial structure, enzyme activity and mRNA levels in embryos exposed to ethanol prenatally. Together, their data showed a significant reduction in mitochondrial ATP synthase activity in both the brain and liver of ethanol exposed embryos and the numbers of mitochondria were lower and elongated. Both authors postulate that their findings could contribute to the abnormal neurological developments and growth phenotypes observed in FAS cases.

Alterations in the structure and function of mitochondria may lead to abnormal mitochondrial biosynthesis, which in turn disrupts oxidative phosphorylation, an important process in a variety of organs, particularly in the brain. Since oxidative phosphorylation is dependent on the mitochondria, this may explain why it was identified as a significantly enriched process (**Table 3.8**). As oxidative phosphorylation provides most of the ATP, the chemical energy required for a cell's metabolism, each step of the mitochondrial energy conversion must function efficiently for cell survival (as reviewed by Fosslien, 2001). Using a mouse model Mansouri et al. (2001) evaluated the effects of acute intragastric ethanol administration (5g/kg) on the brain, heart, skeletal muscle and liver mitochondrial DNA (mtDNA) in mice. Ethanol administration led to mtDNA depletion in all organs and was

associated with oxidative degradation. Their findings suggest that these disturbances could contribute to the development of brain injury and cardiomyopathy in alcoholic patients.

Interestingly, glycolysis, another metabolic reaction that generates ATP, was identified as a significantly enriched biological process. During glycolysis, glucose is broken down into two molecules of a compound referred to as pyruvate. Through a series of reactions that occur in the mitochondria, pyruvate ultimately results in the generation of ATP molecules (Cunningham and Van Horn, 2003). To determine whether chronic alcohol intake impairs glycolysis, researchers have measured the levels of pyruvate in the livers of animals exposed to alcohol. These analyses revealed that the concentration of pyruvate in the hepatocyte of alcohol-treated animals was reduced in comparison to the control animals, which suggests that alcohol consumption may reduce glycolysis (Baio et al., 1998, Van Horn and Cunningham, 1999). Reductions in glycolysis may lead to reductions in ATP, which could have a functional impact on several cellular processes that require a large amount of energy during embryonic development.

As described in the sections above, alcohol has the potential to disturb oxidative phosphorylation and glycolysis, two processes central to the generation of the cells energy. As both processes were identified to be significantly enriched biological processes through the DAVID (Huang et al., 2005, Huang et al., 2009) analysis, this may explain why generation of precursor metabolites and energy was revealed as a significantly enriched biological process in the embryonic brain (**Table 3.8**).

It may seem counterintuitive that energy metabolism biological processes were enriched in the embryonic the brain and not the liver. However, there is complex cross talk between the organs. The central nervous system plays a key role in controlling both energy and glucose homeostasis, and metabolic processes in the liver are tightly regulated via the sympathetic and parasympathetic nervous system, which directly innervate the liver (Rui, 2014). Therefore, although functions enriched in the brain and liver were discussed

separately, it is important to remember that epigenetic alterations in one organ are likely to have a systemic impact, influencing other organs and their functions.

Interestingly, the ribosome and processes reliant on its function were ranked as the most significant cluster in the genes that were downregulated in the embryonic brains, generating the highest enrichment score (2.2) through the DAVID (Huang et al., 2005, Huang et al., 2009) analyses (**Table 3.8**). The ribosome is a large and complex molecular machine, found within all cells that serve as the site of biological protein synthesis (translation).

Smith et al (2014) studied genetic factors that could contribute to the risk of developing craniofacial outcomes following prenatal alcohol exposure. As part of the study RNA-Seq data from ethanol vulnerable and ethanol resistant neuroprogenitors of chick embryos were compared. They identified significant enrichment of 38 genes that encode ribosome proteins. Interestingly, two independent mice model studies also identified altered ribosomal gene clusters in the neuroprogenitors in response to ethanol exposure at comparable developmental stages to the chick embryos (Downing et al., 2012, Green et al., 2007). The repeated emergence of ribosomal protein clusters from multiple comparisons of ethanol-treated neuroprogenitors suggests that ribosome activity may be an integral component of cellular ethanol processes, however it remains unclear what makes ribosomal genes vulnerable to ethanol exposure. Human ribosomopathies (human disorders of ribosome dysfunction) are not embryo lethal but are associated with anaemia, short stature, limb and heart defects and significant craniofacial deficits that are synonymous to those seen in FASD cases.

Our data in conjunction with previous studies suggests that ethanol exposure may have disturbed processes involved in ATP production, which in turn altered ribosomal functions, potentially perturbing protein synthesis.

Zinc ion binding/zinc fingers proteins were the only significant clusters identified in genes upregulated in the embryonic brain. These are very broad terms, making it difficult to interpret the biological significance of their enrichment. Generally, zinc is essential for growth, development and differentiation of all types of life, including microorganisms, plants and animals, as it is required for the activity of more than 300 enzymes (McCall et al., 2000). According to the enrichment analysis (**Table 3.8**), genes that were associated with both zinc ion binding and zinc finger proteins include *Adipocyte Enhancer-Binding Protein 2* (*Aebp2*) and *GLI family Zinc Finger 3* (*Gli3*). *Aebp2* is a zinc finger gene that is known to interact with the Polycomb Repressive Complex 2 (PRC2) (Kim et al., 2009). As a PRC2 binding protein, *Aebp2* function includes recruiting the histone-modifying complex to various developmental genes for their repression via H3K27 (Kim et al., 2015a). *Aebp2* may regulate the migration and development of the neural crest cells through the PRC-mediated epigenetic mechanisms. Diseases associated with *Aebp2* include Waardenburg syndrome and Hirschsprung syndrome (Kim et al., 2009).

Gli3 is a zinger finger gene that encodes a protein that belongs to the C2H2-type zinc finger proteins subclass that are characterised as DNA-binding transcription factors and mediators of Sonic hedgehog (Shh) signalling (Gene Cards, 2015a). *Gli3* may function as a transcriptional activator or repressor. As a repressor, it down regulates dHand and Gremlin, which are central to the development of digits (te Welscher et al., 2002). Studies have shown that mutant mice *Gli3* knockout mice have several abnormalities including lung and CNS defects and polydactyly (Franz, 1994, Schimmang et al., 1992, Rash and Grove, 2007).

Overall, if these enriched interconnected biological processes and molecular functions that serve to regulate processes crucial to embryonic development are sensitive to environmental exposures, particularly ethanol, then it is plausible that their dysregulation in the embryonic brain may cumulatively contribute to resemble FASD related phenotypes in the embryos sired by ethanol-treated males. If several of the genes have sex specific

expression differences, this may explain the significant decrease in brain weight of the male embryos sired by ethanol-treated males as observed by Knezovich et al. (2014).

4.6.2 Molecular functions and biological processes identified by functional enrichment analysis of differential gene expression in embryonic liver

Chronic alcohol consumption is one of the most common causes of chronic liver failure. However, the molecular mechanism underlying the progression of liver injury by ethanol consumption is not well understood. Epigenetic alteration is emerging as one of the key mechanism in the development and progression. That being said, it is understandable that a large number of genes were dysregulated in the liver of embryos sired by ethanol-treated males. The functional enrichment analysis performed using DAVID (Huang et al., 2005, Huang et al., 2009) identified two biological processes and two GO terms enriched for in genes that displayed a significant change in expression in the livers of embryos sired by males exposed to ethanol. The biological processes included wound healing and inflammatory response, and two very broad GO terms, extracellular region and secretion.

The inflammatory response is regarded as the first of four processes that constitute wound healing. Throughout the process there is an elaborate coordination of events that requires communication amongst multiple cell types and soluble factors, including cytokines and chemokines, keratinocyte proliferation, angiogenesis and the extracellular matrix (Ti et al., 2014). Given the inherent dependency of these processes, it explains why wound healing, inflammatory response and the extracellular matrix were all significantly enriched in the functional enrichment analysis. This gains support by the fact that the same three genes were significantly enriched in all three processes (**Table 3.9**).

Wound healing begins with haemostasis, which is evident by the appearance of platelets that activate innate immune mechanisms that ultimately initiate an inflammatory response.

Inflammation is the first defensive response against harmful stimuli and is necessary for proper wound healing (Ti et al., 2014). Studies have shown that epigenetic modifications contribute to inflammatory gene expression and signal transduction. Specifically, alcohol induced epigenetic alterations may disturb the function of developmental pathways of several types of immune cells (e.g. granulocytes, macrophages and T-lymphocytes) and through these and other mechanisms promote exaggerated inflammatory responses (Ti et al., 2014). For instance, Kendrick et al. (2010) used a human monoblastic cell line to assess alcohols effect on the inflammatory response of macrophages. Their study revealed that prolonged exposure of ethanol to the cells results in dramatically enhanced pro-inflammatory cytokine responses that were associated with increased acetylation of histone H3 and H4, as well as increased acetylation of specific cytokine gene promoters, including those encoding *IL-6* and *TNF*.

Aroor and colleagues have conducted various studies directed toward identifying the molecular mechanisms underlying liver injury following chronic ethanol consumption (Aroor et al., 2011, Aroor et al., 2010, Aroor et al., 2012). Their studies have shown that chronic ethanol followed by repeated binge alcohol exposure causes augmentation of liver injury as reflected by increased alanine aminotransferase levels, macrosteatosis, inflammation and neutrophil infiltration. They have also found that ethanol binge after chronic ethanol exposure alters the profile of site-specific histone modifications (Aroor et al., 2010).

The studies above, in conjunction with several other studies have implicated epigenetic mechanisms as one of the molecular mechanisms that contribute to the development and progression of liver disease following chronic ethanol consumption (French, 2013, Mandrekar, 2011). Given that inflammation is a hallmark of the disease and inflammatory responses are central to wound healing, it is likely that alcohol may have dysregulated inflammatory response genes in male sires, which were then inherited by their sired offspring, which may explain the enrichment of these processes in the liver of their offspring. It is thus possible that dysregulation of these interconnected processes may have

resulted in altered wound healing or phenotypes that resemble chronic liver disease in the embryos sired by ethanol-treated males.

In summary, the functional enrichment analysis of the gene expression profiles of embryos sired by ethanol-treated males identified energy metabolism and generation processes to be significantly enriched in the brain, whereas wound healing and inflammatory responses were significantly enriched in the liver. Due to the fact that these processes have been shown to be sensitive to alcohol, are important for development and have previously been associated with FAS like phenotypes, it may be hypothesised that chronic paternal preconception alcohol exposure could have dysregulated all three epigenetic mechanisms, which together were responsible for the significantly enriched processes and may have ultimately led to the development of a FAS like phenotype in the embryos had they lived to term.

4.7 Limitations and future studies

This study used effective statistical methods and well-established platforms for analysing alterations in DNA methylation and expression in candidate genes. Regardless, there are several limitations to this study.

One limitation is that the candidate genes were chosen because they showed both significant methylation and gene expression changes in the previous study (Knezovich et al., 2014). By placing both selection conditions on the choice of candidate genes meant that the most significantly altered genes for the two independent studies were excluded from the list of candidates. This caused highly significant ($p<0.001$) genes demonstrating either differential DNA methylation or differential gene expression to be excluded from validation. Instead, genes with more moderate effects were selected. This may have contributed to the failure to validate the differential expression data. Another approach for the selection of genes to validate for the expression studies, would be to select genes with

a fold change greater than 2, as previous studies suggest that these are more likely to validate.

Sufficient resources for effective validation were also a limitation. The concentration of the sperm DNA and the limited number of genes that could be pyrosequenced potentially contributed to the conflicting results observed between the RRBS and pyrosequencing data.

4.7.1 Future model for the transfer of epigenetic alterations as a consequence of chronic preconception paternal alcohol intake

As this study failed to confirm a significant correlation between alcohol-induced differences in sperm methylation that were inherited by offspring which ultimately dysregulated gene expression, certain recommendations could be made if the study were to be redesigned and repeated (**Figure 4.3**). These changes could potentially address the limitations and discrepancies of the previous and present study.

Firstly, the sample size of the sires could be reassessed. A larger sample would increase the power to detect small effects. Secondly, as it is clear that all three epigenetic mechanisms are connected, it would be critical to analyse all three to get a holistic understanding of the effect of alcohol on the epigenome.

The methodology of the future study is described in **Figure 4.2**. C57BL/6 male mice would be exposed to alcohol (3g/kg) (test) or sucrose (control) by gavage for ten weeks. Male mice would then mate with unexposed C57BL/6 females to generate embryos. Once female mice are inseminated, male mice would be sacrificed and mature sperm would be extracted from the epididymis. DNA would be extracted and the methylation levels in the sperm would be analysed using RRBS. RNA would be extracted and miRNAs would be analysed using RNA-Seq. The histones will not be analysed in the sperm as the majority have been replaced by protamines.

Embryos would then be harvested and the placenta, liver and brain would be extracted from each embryo. Using the extracted DNA and RNA from the tissues of each embryo the following techniques would be conducted:

- RRBS would be conducted to analyse the methylation levels in embryos sired from ethanol-treated males in comparison to embryos sired by sucrose-treated males.
- Chromatin immunoprecipitation would be used to identify modified histones in test embryos in comparison to the sucrose-treated controls.
- A whole genome expression array would be performed to identify differentially expresed genes between the treatment groups.
- RNA seq could be used to anlayse differentially expressed miRNAs between cases and controls.

Following this, bioinformatics analyses would be conducted to identify significantly ($p<0.05$) differentially methylated regions and differentially expressed miRNAs in the sperm of ethanol-treated males compared to sucrose-treated males. Similarly, the offspring data would be analysed to identify differentially methylated and expressed genes, differentially modified histones and differentially expressed miRNAs. From the genes, miRNAs and modified histones that pass the significance thresholds in the bioinformatics analysis, a candidate gene list could be constructed for validation or further analyses such as functional studies.

The candidate gene list would include:

- Genes that were significantly ($p<0.05$) differentially methylated in the sperm of ethanol exposed males.
- miRNAs that were significantly differentially expressed in ethanol exposed males.
- Genes that were significantly differentially methylated in the embryos of ethanol-treated sires compared to embryos sired by sucrose-treated males.

- Genes that were differentially expressed in the embryos of ethanol-treated sires compared to embryos sired by sucros-treated males.
- miRNAS that were differentially expressed in the embryos of ethanol-treated sires compared to embryos sired by sucrose-treated males.
- Histones that were modfified in the embryos of ethanol-treated sires compared to embryos sired by sucrose-treated males.
- Genes/miRNAs that significantly altered in both the sperm and embryos.
- Genes that are regulated by more than one epigenetic mechanism.

The candidate list would then be filtered by biological relevance, fold change and the most significant changes. The list of genes could then be validated using site-specific techniques such as pyrosequencing for DNA methylation analysis, PCR based methods for histone modifications and real-time PCR for miRNAs and gene expression changes.

Alternatively, to address the phenotypic effect that result as a consequence of alterations in gene expression future researchers may opt to evaluate additional phenotypes related to growth, behaviour and development of the offspring.

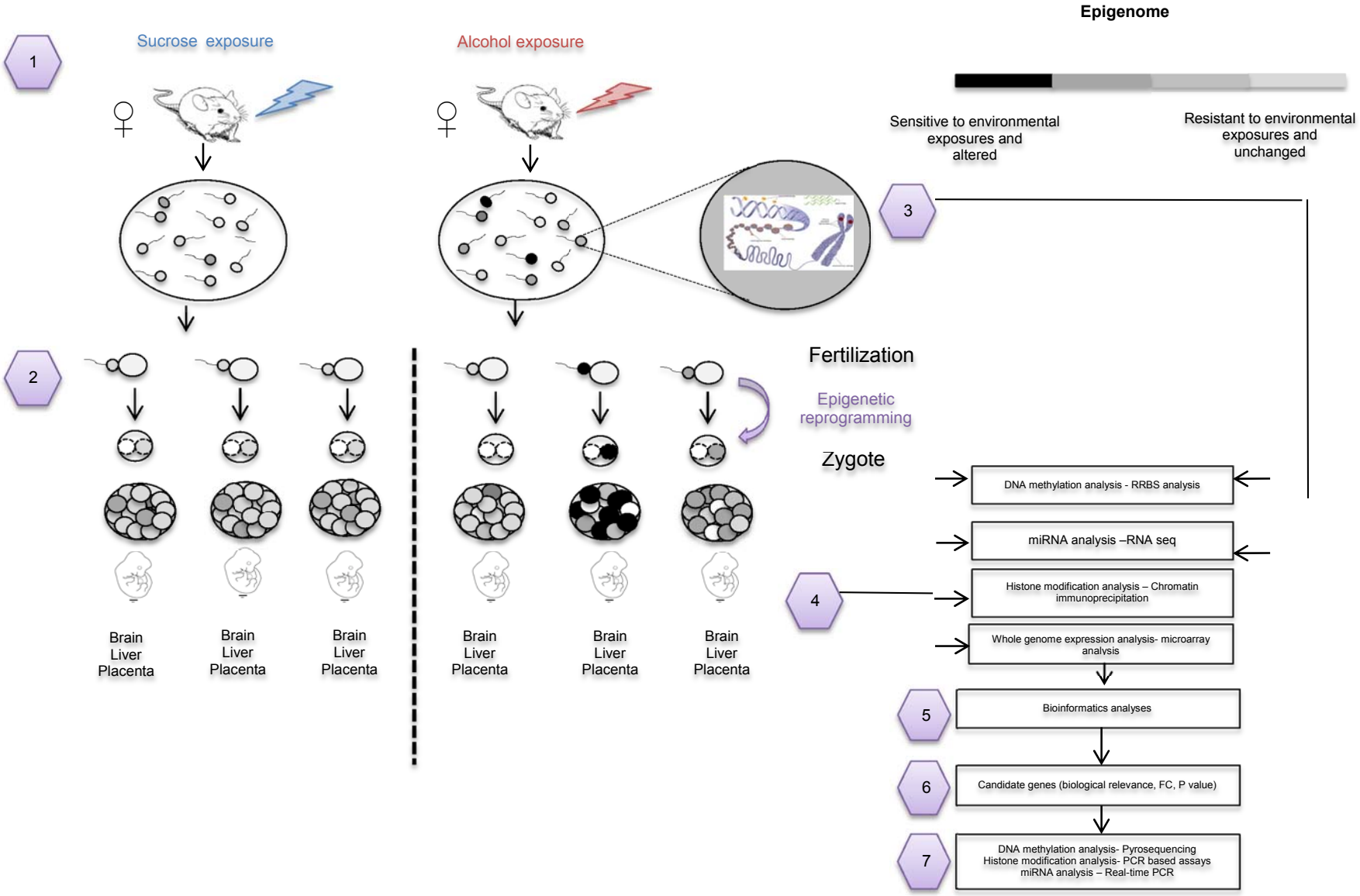


Figure 4.2: A simplistic future model to investigate the effect of preconception paternal alcohol exposure on epigenetic mechanisms in mouse offspring. (1) Expose C57BL/6 male mice to alcohol (3g/kg) (test) by gavage for ten weeks and to sucrose (control). (2) Mate male mice with unexposed