

Investigating the Molecular Mechanisms of the Metabolic Syndrome

Tiffany Jaye Morris MSc

Hughes Hall



Department of Pathology

University of Cambridge



This dissertation is submitted to the University of Cambridge in accordance
with the requirements of the degree of Doctor of Philosophy in the
School of the Biological Sciences.

This research programme was carried out in the Department of Pathology,
University of Cambridge, under the supervision of Professor Nabeel Affara
and in collaboration with the Liggins Institute, University of Auckland,
in Auckland, New Zealand.

Submitted: 7th of May 2010

Abstract

Investigating the Molecular Mechanisms of the Metabolic Syndrome

by Tiffany Jaye Morris

This thesis aims to highlight molecular mechanisms that have been altered by prenatal undernutrition and may be involved in the metabolic syndrome. Two separate studies were conducted both using a rat model developed through manipulation of the maternal diet to provoke the key features of the metabolic syndrome in adult offspring. Microarray technology was used to detect changes in gene expression in target tissues between offspring of control (normally fed, AD) and undernourished (UN) mothers to obtain a broader picture of the cellular functions and genetic pathways that may be implicated in the metabolic syndrome.

The first study compared gene expression differences in liver, skeletal muscle, and white adipose tissue between 55 day old male offspring of AD and UN mothers. No significant changes were found in muscle or adipose tissue; however, the differences in the liver suggested the UN animals had been metabolically programmed to favour fat as an energy source.

To investigate whether DNA methylation might be responsible for the observed transcriptional changes, pooled liver samples from the first study were used with the MspI restriction enzyme assay to determine full, partial, incomplete, or no methylation between AD and UN. Two differentially expressed genes (*Zfand2a* and *Mapk4*) showed methylation changes.

The same liver samples were hybridised to a miRNA array. Two miRNAs showed a nearly 2-fold upregulation in the UN livers. Both were found to be either directly or indirectly associated with the metabolic syndrome. MiR-335 has been shown to be upregulated in the livers of obese/diabetic mice. By association with miR-27a, miR-451 might be involved in aspects of lipid metabolism in adipose tissue.

A second study used microarray to analyse the liver tissues of day 170 female offspring of the same rat model with additional insults (neonatal leptin treatment and post-weaning high-fat (HF) diet). Leptin has been shown to reverse the programming effects of the restricted maternal diet and this study aimed to highlight mechanisms that could be involved in this reversal. The results revealed the importance of the interaction between treatments. Significant gene expression changes were only present when two or more treatments were combined. This study revealed significantly, differentially expressed genes involved in immune function, regulation of the circadian rhythm, and metabolism.

These findings provide a number of interesting genes and pathways for further studies and also highlight the need to conduct a thorough study in multiple tissues at different time-points to pinpoint the window of developmental plasticity.

‘...Two roads diverged in a wood, and I I took the one less traveled by, And that
has made all the difference.’ Robert Frost, 1920.

This thesis is dedicated to my family: Scott, Deborah, Crystal and Heather
(and Shelby too!)

Acknowledgements

I would like to begin by acknowledging my supervisor, Professor Nabeel Affara for his guidance throughout my PhD. My life would be very different right now if it was not for him. When I arrived in England with my 6-month work visa, he gave me the opportunity to work in his lab, offered me a PhD studentship, and helped me to find funding. This has been the experience of a lifetime and it would not have been possible without his support.

It was nothing less than serendipitous that I ended up in Cambridge, England working with Prof. Affara. I owe that to Dr. Robert Erickson, from the University of Arizona. I worked in his lab as an undergraduate. He introduced me to Bioinformatics and encouraged me to pursue it in my graduate studies. I arrived in Cambridge hoping to work at the Sanger Institute, but naive to UK work laws, I searched in vain for a lab that would give me a temporary position on such short notice. Suddenly, I remembered the box of chocolates that said, 'To: Tiffany From: England' and immediately emailed Dr. Erickson to find out if he knew of a lab in Cambridge...the rest is history!!

An important thank you goes to my PhD mentor, Professor Stewart Gilmour, for his help in getting me through my crash course in biochemistry. His contribution to my work with his ideas and discussion was immensely important.

I would like to thank the many members of the lab that have helped me throughout my studies: Dr. Carole Sargent and Dr. Anthony Brown for their advice and training, particularly when I first joined the lab; Julien Bauer for his help and advice on the analyses in this thesis; Dr. Oriane Chausiaux for her constant support, advice, and willingness to travel; Dr. Emily Clemente, Dr. Gina Oliver, Claire Wyllie, Sharon Morris, Kerry Cline, Dr. Alexandra Lopes, Chris Reitter, Dr. Claire Quilter, Dr.

Lydia Ferguson, Dr. Peter Ellis and Meena Bagga are all past or present members of the lab that impacted my experience. It is with pleasure that I will always remember the tea breaks in the canteen.

Anytime not spent in the lab was usually spent on the river. Thank you to the girls that took my mind off work and rowed with me in the cold early winter mornings: Ailsa Benton, Catherine Blake, Ciara Metcalfe, Emma Smith, and many more through the years!

My PhD was funded by the University of Auckland Liggins Institute because of Dr. Peter Gluckman whom I am indebted to for giving me this opportunity. I spent a total of six months in New Zealand on two separate occasions, which was an amazing experience. I loved working at the Liggins Institute and am especially thankful to: Dr. Mark Vickers for his help in my understanding of the physiological side of my project; Dr. Deborah Sloboda for her advice and support; and Dr. Alice Coveny for the coffee breaks, the chats, the mountain biking and the fun we had at the Perth Zoo!

I am also appreciative to Dr. Jonathan Hayes for his latex troubleshooting skills that proved to be very useful as I was writing this thesis. Also his support, advice, friendship and spaghetti carbonara!!

I am grateful to have completed this writing process in the Vicarage; a special place full of special people (Nelsy, Ultan, Charm, Nick, Scott, Sandra, and Kaushik). I have known them only a short time yet they have made an impact. It has been the perfect place to end this journey and look forward to a new one.

Finally, a big thank you to my supportive and loving family: my sisters Crystal (especially for her thorough proofreading) and Heather; my parents Scott and Debbie. Their constant belief in me carried me through the most difficult times in this process. I have talked about getting a PhD since I was eleven and they are as relieved as I am that it is finally complete.

This thesis was created using the L^AT_EX typesetting system (MiKTeX 2.8 package for Microsoft Windows). The figures were prepared with CorelDraw 10.0 software.

Author's Declaration This dissertation is the result of my own work that was performed between October 2006 and October 2009. It contains nothing which is the outcome of work done in collaboration with others, except as specified in the text. The results presented in this thesis have not been submitted for any other degree or diploma. The thesis does not exceed the prescribed word limit of 60,000 words as set by the School of the Biological Sciences.

Tiffany Jaye Morris MSc

Contents

Table of Contents	i
List of Tables	vii
List of Figures	xi
Abbreviations	xii
1 Introduction	1
1.1 The Metabolic Syndrome	1
1.1.1 Obesity	2
1.1.2 Diabetes	2
1.2 Developmental Plasticity	3
1.3 DoHaD and Maternal Nutrition	5
1.4 An Evolutionary Perspective	5
1.4.1 Thrifty Genotype/Thrifty Phenotype	6
1.4.2 Predation Release Hypothesis and Genetic Drift	7
1.4.3 Summary	8
1.5 Epidemiological Data	9
1.5.1 Dutch Hunger Winter	9
1.5.2 Pima Indians	10
1.5.3 Cohort Studies	10
1.6 Animal Models of the Metabolic Syndrome	12
1.6.1 Maternal Diets	13

1.6.2	Postnatal Diets	16
1.6.3	Hormonal Insults	16
1.6.4	Surgical Interventions (Reduced Placental Transfer by Uterine Artery Ligation)	18
1.6.5	Genetic Models	18
1.6.6	Summary	19
1.7	Mechanisms of Nutritional Programming	19
1.8	Microarray	20
1.8.1	microRNA Arrays	22
1.8.2	Microarray Applied to Developmental Plasticity	22
1.9	Structure and Aims of the Thesis	23
2	General Methods	25
2.1	Animal Care/Sample Collection	25
2.2	RNA Extraction	25
2.2.1	For Oligonucleotide Expression Array (Undernutrition Study and Leptin Study)	25
2.2.2	For microRNA Expression Array	26
2.2.3	Quality Assessment	26
2.3	Microarray	27
2.3.1	Experimental Design	27
2.3.2	Quality Control	28
2.3.3	Reducing Error	29
2.3.4	Data Analysis	30
2.3.5	Validation of Microarray Data: Quantitative Real Time PCR .	30
2.3.6	Mouse Exonic Evidence Based Oligonucleotide Expression Array	30
2.3.7	Illumina BeadArray	34
2.3.8	microRNA Expression Array	37
2.3.9	Filtering, Data Visualisation and Normalisation	38
2.3.10	Selection of Differentially Expressed Genes	40

2.3.11	Gene Ontology Analysis	42
2.4	Polymerase Chain Reaction	43
2.4.1	Genomic PCR	43
2.4.2	Comparative Quantitative Real Time RT-PCR	43
2.5	DNA Extraction	44
2.5.1	From Tissue (Undernutrition Study)	44
2.5.2	Following RNA extraction	44
2.6	Methylation PCR	45
2.7	Bisulphite Sequencing	46
3	Expression Study of rats undernourished <i>in utero</i>	47
3.1	Introduction	47
3.1.1	Energy Balance and Starvation	48
3.1.2	Genes/Pathways Highlighted in Other Studies	56
3.1.3	Initial Study	57
3.2	Results	58
3.2.1	Previous Findings: Phenotypic Assessment	58
3.2.2	Microarray	60
3.2.3	QRT-PCR for Microarray Verification	87
3.3	Discussion	89
3.3.1	Glucose Metabolism	92
3.3.2	Insulin and IGF-1 Signalling	93
3.3.3	Mitochondrial Activity	93
3.3.4	Fat	94
3.3.5	Ribosomal Proteins and Protein Turnover	95
3.3.6	Effects on Gene Transcription	96
3.4	Conclusions	97
4	Leptin Reversal Study	108
4.1	Introduction	108
4.1.1	Leptin	109

4.1.2	Leptin Receptor	111
4.1.3	Genetics of Obesity	111
4.1.4	Energy Balance: Interactions of Leptin and the Brain	112
4.1.5	Leptin and the Circadian Rhythm	115
4.1.6	Leptin During Pregnancy, Lactation and Puberty	119
4.1.7	Leptin Differences in the Sexes	119
4.1.8	Overview	120
4.1.9	Initial Study	121
4.2	Results	126
4.2.1	Illumina Microarray	126
4.2.2	qRT-PCR for Microarray Validation	158
4.2.3	Comparison of Treatment Groups	158
4.3	Discussion	171
4.3.1	Cluster 10	173
4.3.2	Genes Involved in the Immune Response	175
4.3.3	Relevant Genes in Other Clusters	176
4.3.4	Candidate Genes	176
4.3.5	Adipose Tissue	177
4.4	Conclusions	177
5	Methylation Study	179
5.1	Introduction	179
5.1.1	Epigenetics	180
5.1.2	Epigenetic Change, Developmental Programming, and the Metabolic Syndrome	183
5.1.3	Available Techniques for Methylation Analysis	185
5.1.4	Summary	187
5.1.5	Study Design	187
5.2	Results	188
5.3	Discussion	194

5.3.1	<i>Zfand2a</i>	196
5.3.2	<i>Mapk4</i>	196
5.4	Summary	197
6	microRNA Expression Study	200
6.1	Introduction	200
6.1.1	microRNA	200
6.1.2	Study Design	206
6.2	Results	207
6.2.1	Data Quality Control	207
6.2.2	Normalisation	208
6.2.3	Differential Expression	218
6.3	Discussion	225
6.3.1	miR-335 Implicated in Lipid Metabolism	226
6.3.2	miR-451 Expression in Human Cancer Cells and Possible Association with Adipocyte Differentiation	227
6.4	Summary	229
7	General Discussion and Future Work	230
7.1	Conclusions	230
7.1.1	Gene Expression Profiling Male 55 day old Rats Exposed to Maternal Undernutrition <i>in utero</i>	231
7.1.2	Gene Expression Profiling Female 170 day old Rats Exposed to Maternal Undernutrition <i>in utero</i> , Postnatal Leptin and post-weaning High Fat Diet	231
7.1.3	Methylation Assay on Male 55 day old Rats Exposed to Maternal Undernutrition <i>in utero</i>	232
7.1.4	MiRNA Expression Profiling of Male 55 day old Rats Exposed to Maternal Undernutrition <i>in utero</i>	233
7.2	Future Directions	234

A Primer Lists	235
B Publications and Presentations	241
C Supplementary Data	242
References	243

List of Tables

3.1	Day 55 Phenotype Data	59
3.2	Day 110 Phenotype Data	60
3.3	Illumina Gene List	86
3.4	Fold Change Values Comparing Microarray and QRT-PCR	89
3.5	Functional Analysis	107
4.1	Day 170 Phenotype Data	125
4.2	Biological Processes of Clusters	149
4.3	Leptin Gene List	157
4.4	Leptin Gene Biofunctions	170
5.1	Methylation	193
6.1	miRNA Results	225
A.1	Expression Primers1	235
A.2	Expression Primers2	236
A.3	Lillycrop Expression Primers	236
A.4	Methylation Primers2	237
A.5	Methylation Primers3	238
A.6	Methylation Primers5	239
A.7	Lillycrop Methylation Primers	239
A.8	Bisulphite Primers	240

List of Figures

2.1	Rat Samples on Mouse MEEBO	31
2.2	Bead Array Image	35
3.1	Major Sites of Insulin Action	49
3.2	Major Sites of Glucagon Action	49
3.3	Basal State	51
3.4	Gluconeogenesis	52
3.5	Tissue Interrelationships During Fasting	54
3.6	Starved State	55
3.7	Effects of Maternal Nutrition on Lipid Metabolism	56
3.8	Study Design of Rats Undernourished in Utero	58
3.9	MEEBO Dye Swap	61
3.10	Pre-Normalisation MEEBO Density Plot	62
3.11	Amplitude in Red Channel	62
3.12	Amplitude in Green Channel	63
3.13	MEEBO MA plot	64
3.14	Illumina Microarray Study Design	66
3.15	Illumina Internal Controls.	68
3.16	Hierarchical Clustering of All Tissues Before Normalisation.	69
3.17	Pairwise Comparison of AD Liver Samples Before Normalisation.	70
3.18	Pairwise Comparison of UN Liver Samples Before Normalisation.	71
3.19	Pairwise Comparison of AD Muscle Samples Before Normalisation.	72
3.20	Pairwise Comparison of UN Muscle Samples Before Normalisation.	73

3.21	Pairwise Comparison of AD WAT Samples Before Normalisation. . .	74
3.22	Pairwise Comparison of UN WAT Samples Before Normalisation. . .	75
3.23	Boxplot Before Normalisation	77
3.24	Density Plot Before Normalisation	77
3.25	Boxplot After Quantile Normalisation.	78
3.26	Density Plot After Quantile Normalisation.	78
3.27	Microarray Validation by qRT-PCR	88
3.28	Microarray Validation by qRT-PCR	88
3.29	Genes that Influence Glucose, Fat and Energy Metabolism Pathways and IGF-1 and PI3 Signalling Pathways.	91
4.1	Leptin and the Starvation Response	115
4.2	Leptin and the Homeostatic Pathway	116
4.3	Leptin Receptor Expressing Neurons	117
4.4	Insulin Leptin Reward	118
4.5	Design of Leptin Reversal Study	122
4.6	Results From Vickers 2005 <i>et al.</i>	124
4.7	Microarray Design for Leptin Reversal Study	126
4.8	Hierarchical Clustering of Raw Leptin Array Data	128
4.9	Summary of Leptin Internal Controls	129
4.10	Pairwise Comparison of ADSALCHOW Samples Before Normalisation.	130
4.11	Pairwise Comparison of UNSALCHOW Samples Before Normalisation.	131
4.12	Pairwise Comparison of ADLEPCHOW Samples Before Normalisation.	132
4.13	Pairwise Comparison of UNLEPCHOW Samples Before Normalisation.	133
4.14	Pairwise Comparison of ADSALHF Samples Before Normalisation. . .	134
4.15	Pairwise Comparison of UNSALHF Samples Before Normalisation. . .	135
4.16	Pairwise Comparison of ADLEPHF Samples Before Normalisation. . .	136
4.17	Pairwise Comparison of UNLEPHF Samples Before Normalisation. . .	137
4.18	Boxplot before Normalisation	138
4.19	Density Plot Before Normalisation	138

4.20	Boxplot After Quantile Normalisation	139
4.21	Density Plot After Quantile Normalisation	139
4.22	Venn Diagram of Overlap Between FSPMA Gene Lists	140
4.23	Genes Present in Pairwise Comparisons	141
4.24	Hierarchical Clustering of Data Analysed by One-Way ANOVA	144
4.25	Biological Processes Found by FSPMA	145
4.26	Molecular Functions Found by FSPMA	146
4.27	Cellular Components Found by FSPMA	147
4.28	qRT-PCR Data for Cyclophilin Control	159
4.29	Microarray Expression Data 1	160
4.30	Microarray Validation by qRT-PCR 1	161
4.31	Microarray Expression Data 2	162
4.32	Microarray Validation by qRT-PCR 2	163
4.33	Comparison of Treatment Effects for <i>Gck</i> and <i>Per2</i>	164
4.34	Comparison of Treatment Effects for <i>Igfbp2</i> and <i>Igfbp4</i>	164
4.35	Comparison of Treatment Effects for <i>Rt1-ba</i> and <i>Rt1-149</i>	165
4.36	Comparison of Treatment Effects for <i>Rt1-m6-2</i> and <i>Ca3</i>	165
4.37	Gluckman <i>et al.</i> qRT-PCR Comparison	166
4.38	Effect of Treatment on Array Signals for Candidate Genes	167
4.39	Differentially Expressed Genes for Two Studies.	168
4.40	Comparison of Treatment Effects for <i>Klf15</i> and <i>Enpp2</i>	175
5.1	Methylation	181
5.2	Diet and Methylation	185
5.3	Methylation Techniques	188
5.4	Methylation Analysis Pipeline	189
5.5	PCR of McrBC Methylation Assay 1	194
5.6	PCR of McrBC Methylation Assay 2	195
6.1	miRNA Biogenesis	203
6.2	miRNA Processes	205

6.3	miRNA Internal Controls	209
6.4	Hierarchical Clustering of Raw miRNA Data	210
6.5	MA Plot of Raw miRNA Data	211
6.6	MA Plot 2 of Raw miRNA Data	212
6.7	MA Plot 3 of Raw miRNA Data	213
6.8	Boxplot of Raw miRNA Data	214
6.9	Density Plot of Raw miRNA Data	215
6.10	Heatmap of Raw miRNA Data	216
6.11	Mean SD of Raw miRNA Data	217
6.12	Pairwise Comparison of AD Liver miRNA Samples After Normalisation	219
6.13	Pairwise Comparison of UN Liver miRNA Samples After Normalisation	220
6.14	Boxplot miRNA Samples After Quantile Normalisation.	221
6.15	Density Plot of AD miRNA Samples After Quantile Normalisation.	222
6.16	Density Plot of UN miRNA Samples After Quantile Normalisation.	223
6.17	miR335 Expression in WT Tissues	228
6.18	miR335 Liver Expression in Diabetes/Obesity Mouse Models	228

List of Abbreviations

AD	ad-libitum fed
ALSPAC	Avon Longitudinal Study of Parents and Children
ANOVA	analysis of variance
ATP	adenosine triphosphate
BLAST	Basic Local Alignment Search Tool
cDNA	complementary deoxyribonucleic acid
CHOW	post-weaning normal chow diet
CpG	areas of sequence rich in C's and G's
Ct	threshold cycle value
Cy3/Cy5	cyanine fluorescent dyes (cyanine 3/5)
DAVID	Database for Analysis, Visualization, Integrated Discovery
DNA	deoxyribosenucleic acid
dNTP	deoxyribosenucleic triphosphate
DoHaD	developmental origins of health and disease
EDTA	ethylenediamine tetra acetic acid
EST	expressed sequence tag
FDR	false discovery rate
FSPMA	Friendly Statistical Package for Microarray Analysis
g	grams
HF	post-weaning high fat
KDE	Knowledge Discovery Environment Software
LEP	leptin treatment
log ₂	base-2 logarithmic
MEEBO	Mouse Exonic Evidence Based Oligonucleotide
mg/kg	milligrams/kilograms
Milli Q	Milli Q grade water
miRNA	micro RNA
ml	milli litre
mM	milli molar
mRNA	messenger ribonucleic acid
μg	micro gram
μl	micro litre
MgCl ₂	magnesium chloride
Na ₂	sodium
NaCl	sodium chloride
NCBI	National Centre for Biotechnology Information
ng	nanograms

nm	wavelength
PAR	predictive adaptive response
PCR	polymerase chain reaction
PolyA	poly deoxyadenosine
p-value	probability value
qRT-PCR	quantitative reverse transcription polymerase chain reaction
RNA	ribonucleic acid
rpm	revolutions per minute
SAL	saline treatment
SD	standard deviation
SSC	saline sodium citrate
ssDNA	single stranded DNA
TAE	tris-acetate albumin
TCA cycle	The Citric Acid Cycle
T _m	melting temperature
Tris	tris(hydroxymethyl)aminomethane
tRNA	transfer ribonucleic acid
UN	undernourished
UTR	untranslated region
VST	variance-stabilising transformation

Chapter 1

Introduction

This thesis aims to investigate the molecular mechanisms behind the rising incidence of metabolic syndrome related disorders. The work included here investigates the metabolic effects of the nutritional environment by comparing the genetic makeup of different phenotypes and considering this in light of the evolution of diet. The intended outcome of this research is to enable us to gain a better understanding of the molecular mechanisms that signal the body to utilise nutrients, store fat, and feel hunger. This introduction provides the background on the theories behind the process of development, the selection for our current biological makeup, and the evolution of diet that has driven humans to develop this way. The background on microarray will give details of the technique that is most important to the data presented in this thesis. Each chapter in this thesis will begin with an introduction of its own to present more detailed background on relevant topics.

1.1 The Metabolic Syndrome

In the 1920s, observations revealed that risk factors including obesity, hyperlipidaemia, and hypertension tended to occur together. This was termed ‘syndrome X’ and it was proposed that insulin resistance might be the underlying mechanism [1]. Today, this clustering of metabolic abnormalities is referred to as the metabolic syndrome. It occurs when several risk factors occur together in a single person. These

risk factors include: obesity, insulin resistance or intolerance to glucose, proinflammatory state, prothrombotic state, atherogenic dyslipidemia, and higher blood pressure. Metabolic syndrome was defined in 2006 by the International Diabetes Federation (IDF) as central obesity and two of the following: raised triglycerides, reduced HDL cholesterol, elevated blood pressure, and elevated plasma glucose.

Understanding the metabolic syndrome has become important as studies have shown that individuals with the metabolic syndrome are six times more likely to develop diabetes, 2-3 times more likely to develop cardiovascular disease, and are likely to suffer from obesity, particularly visceral obesity [2].

1.1.1 Obesity

The rate of obesity is quickly increasing worldwide. Before the 20th century obesity was rare, but in 2005 it was estimated that 9.8% of adults are obese with rates in the United States, Australia, and Canada increasing rapidly compared to the overall rate (WHO organisation fact sheet). The growing obesity problem is impacting human morbidity, mortality and quality of life, and results in costly healthcare. With the increasing number of young obese women, obesity during pregnancy is also a problem. This can have negative maternal health effects, can cause risks during pregnancy, and can lead to persistent problems in the developing child. Type-2 diabetes is also becoming more common and affects particular groups of people with a higher frequency.

1.1.2 Diabetes

Type-2 diabetes mellitus is characterised by high blood glucose associated with insulin resistance and relative insulin deficiency. It is related to the inability to make or respond to insulin. Fat and muscle cells require insulin to absorb glucose. If these cells fail to respond to insulin, blood glucose levels will rise. The liver helps regulate glucose levels by reducing the secretion of glucose in the presence of insulin. However, if an individual is insulin resistant the liver's production of glucose will not decrease.

Insulin resistance normally refers to the reduced ability of insulin to lower glucose. Other functions of insulin can also be affected: insulin resistance in adipocytes reduces uptake of circulating lipids and increased mobilisation of stored lipids in these cells elevates free fatty acids in the blood plasma. The elevated blood fatty-acid concentrations, reduced muscle glucose uptake, and increased liver glucose production all contribute to elevated blood glucose levels. A major component of the metabolic syndrome is the high plasma levels of insulin and glucose due to insulin resistance. If an individual is insulin resistant they can develop hyperglycemia (excessive blood glucose) after a meal. The inability of the pancreatic β -cells to produce sufficient insulin to maintain normal blood sugar levels is what causes the transition from insulin resistance to type-2 diabetes [3].

The occurrence of diabetes often coincides with obesity and other symptoms of the metabolic syndrome. Diabetes has been correlated with low birth weight and has also been tied to poor maternal nutrition [4]. A better understanding of the cause of metabolic syndrome and the metabolic pathways that have been altered could lead to better treatment and hopefully prevention.

1.2 Developmental Plasticity

Developmental plasticity describes the time during development when organs and body systems are sensitive to the environment and can change structure and function [5]. Intrauterine life is a critical point in development. Environmental factors can have a large impact on fetal growth. Gestation is considered a time of developmental plasticity because it is a period when organs and body systems are plastic and particularly sensitive to developmental cues. The changes that occur during these sensitive developmental windows are often irreversible. This permits a range of phenotypes to develop from a single genotype depending on environmental conditions [6, 7]. Different environments can directly induce changes in an individual's behaviour, morphology and physiology. This plasticity is adaptive, in that individuals that show a plastic response have higher fitness (potential to reach reproduction)

than those that do not [8]. The responses can be immediate and short term or they can be expressed in the offspring rather than the parent. Many organisms are known to express specific predictive adaptive responses (PAR) to their environment. A PAR is specifically defined as a type of developmental plasticity that evolved as an adaptive response to an environmental cue acting early in the life cycle, but where the advantage of the induced phenotype is primarily manifest in a later phase of the life cycle [9].

There are a plethora of examples of this concept in nature. One example is the freshwater crustacean *Daphnia*. If the mother is exposed to chemical traces left behind by a predator then the offspring will be born with a defensive ‘helmet’ that offers protection against the predator. However, the helmet takes considerable resources during development and therefore, it can be detrimental to the success of the organism in a predator-free environment where it is unnecessary [10]. A second example is present in the plague locust *Schistocerca gregaria*. This locust can develop into two very different phenotypes: the migratory form or the solitary form. These phenotypes are so different that they could be mistaken for two different species. The migratory form has bigger wings, different mouthparts, a different camouflage colour, and a different metabolism. They also display different behaviours, as the migratory form does not hide but congregates in large swarms. The selection for one phenotype over the other depends on the population density. The signal comes from the mother who secretes a chemical on the eggs that she lays [11]. In this example, the migratory form takes more resources to develop. Therefore, it only develops when there are not enough resources to support the solitary form. A higher initial investment must be made to raise the chances of survival to reproduction. These examples, like others, occur because of a cue during development that predicts the environment that the offspring will encounter. The primary purpose of this type of adaptive response is to increase the likelihood that an individual will survive to reproduction. These examples are of major changes in morphology, but differences could also be much more subtle with the effects magnified in later life. These are examples of trade-offs. One phenotype takes more resources but allows for a chance of survival in a

harsher environment, whereas the other takes less resources and is likely to result in a longer life span. In humans, longevity runs in a family which indicates a genetic component. It is also possible to select animals artificially for longevity. In mice, prenatal undernutrition leads to reduced longevity [12], and postnatal undernutrition leads to a marked prolongation of the lifespan [13].

1.3 DoHaD and Maternal Nutrition

DoHaD (developmental origins of health and disease) is an emerging field of study investigating the concept that insults during development and early life can lead to certain adult diseases (type-2 diabetes, obesity, hypertension). Studies in this field have focussed on nutritional changes (low protein, caloric restriction) at different windows during gestation. The developing foetus is completely dependent on its mother, and so it follows that maternal nutrition during pregnancy in particular has a strong influence on the intrauterine environment. The developing foetus responds to undernutrition by changing the trajectory of development and slowing growth. It is proposed that development is being induced by cues in the environment that are preparing the developing organism for the type of environment into which it will likely be born [10]. Thus the cue is acting as a predictor of the nature of this environment. The DOHaD hypothesis proposes that the metabolic syndrome originates through developmental plasticity in response to undernutrition during fetal life and infancy [6]. Two common explanations for how this occurs have been developed: the thrifty phenotype hypothesis proposed by Hales and Barker [14] and the predation release hypothesis proposed by Speakman [15].

1.4 An Evolutionary Perspective

The thrifty phenotype hypothesis and the predation release hypothesis are two attempts to explain this relationship between poor fetal and infant growth and the increased risk of metabolic syndrome in later life. The thrifty phenotype hypothesis

proposes that undernutrition during development leads to physiological and metabolic changes that promote survival at the expense of health later in life [14]. Human diseases are increasingly correlated with growth patterns early in life implicating early-life nutrition as the underlying mechanism [16]. The predation-release hypothesis suggests that the genes responsible for increased susceptibility to obesity were not selected for but have been the result of random drift. Now that western lifestyles have changed drastically in a short period of time the potential of these genes is being realised [17].

1.4.1 Thrifty Genotype/Thrifty Phenotype

In 1962, James Neel [18] proposed the thrifty genotype hypothesis. This theory proposed that ‘thrifty genes’ were selected during times of famine that enhanced an individual’s ability to store fat. The theory hypothesises that in the current environment these genes now predispose those individuals to the metabolic syndrome.

In 1992, Hales and Barker [14] proposed another hypothesis, the thrifty phenotype hypothesis. The theory is sometimes called ‘Barker’s hypothesis’ after David J.P. Barker at the University of Southampton who published the theory in 1997 [19]. The theory proposed that early-life metabolic adaptations promote survival, with the developing organism responding to cues of environmental quality by selecting an appropriate trajectory of growth. If the fetal environment is poor (i.e. undernutrition), there may be an adaptive response, which optimises the growth of important body organs sacrificing the health of others [20]. This enhances postnatal survival under similar poor conditions but becomes detrimental when nutrition is more abundant in the postnatal environment than it had been in the prenatal environment [21]. This theory is also used to explain insulin resistance. If individuals that are undernourished prenatally continue to be poorly nourished and remain thin and insulin sensitive, the poor functional capacity for insulin secretion would not be detrimental; however, increased food intake and decreased energy expenditure trigger glucose intolerance and lead to obesity [20]. The offspring can be said to have been ‘programmed’

for the environment that the mother was living in at the time of pregnancy. The term programming was chosen to describe the link between patterns of infant feeding and later changes in physiology and disease risk. It now includes broader early life events that contribute to altered growth patterns and altered risk of metabolic disease later in life.

1.4.2 Predation Release Hypothesis and Genetic Drift

Speakman *et al.* [15, 17] states that the thrifty genotype hypothesis does not account for the large number of thin people in western societies. The thrifty gene hypothesis suggests that obesity was selected for as being beneficial for fecundity after times of famine. Speakman argues that the frequency of alleles does not support that interpretation. He presents an alternative hypothesis which he calls the ‘predation release hypothesis’ or when combining it with similar theories involving genetic drift, he calls them the ‘drifty gene’ hypotheses.

Speakman’s predation release hypothesis considers the evolutionary history of humans to understand and put into context the way the body utilises and stores nutrients. For many centuries, there was evolutionary pressure to select against the very lean as those that could store fat most effectively were more likely to survive until the next meal could be found. In this hunter gatherer society, physical work and exercise was an unavoidable part of daily life. This would put an additional selective pressure against excessively lean people that would be too weak, but would also prevent people from putting on excess weight. In early societies, although excessive fat and sugar intake were not part of the diet, there was still a selective pressure against an ability to store too much fat which would have led to decreased ability to move quickly and escape predators. This balance was maintained by establishing upper and lower set-limits. The lower limit was set by the risk of starvation and the upper limit by the risk of predation. These set-limits evolved over many thousands of years until the skill of fire-starting was acquired. This ‘release from predation’ relaxed the selection against a fatter phenotype. This relaxation in the set-point allowed the upper weight

limit to be affected by genetic drift (the change in allele frequencies in a population). Even though this upper weight limit was not under selective pressure the lifestyle required a significant amount of physical activity and so those individuals with a higher set-point would be unlikely to reach that potential and become overweight or obese. Now very recently (the last 50-100 years) the abundance, palatability, and cheapness of food have changed the way food is consumed. Westerners consume more calories than are required to maintain their health and have reduced physical activity at work. So those people with a higher set point are able to reach that potential and those are the people more prone to obesity. Today, lifestyle is very different to the one that our bodies adapted to over the millennia [17, 22]. This hypothesis gives one potential explanation for the variation in susceptibility to obesity seen in the human population. This variation in set-point (described in more detail in Chapter 4 in relation to leptin) may provide an explanation for the difficulty particular individuals have in avoiding obesity and for the tendency for obesity to run in families.

1.4.3 Summary

It is likely that the rapid rise in obesity, diabetes, and metabolic syndrome involves both of these theories. As discussed in more detail in Chapter 4, humans do have varying set-points at which the feedback mechanism controlling obesity will function properly. Speakman has shown that the rate of obesity in human society is not high enough to indicate it has been selected for; however, the predicative adaptive response has been observed in many species. The metabolic syndrome is a complex syndrome that involves many disorders. It is likely that some aspects have been selected for and others are the result of genetic drift. Several epidemiological studies have been done showing the effect of maternal nutrition on the future health of the offspring and the knock-down effects this has on generations to come.

1.5 Epidemiological Data

The idea that maternal health may influence the future health of the baby has been around since Victorian times [23]. Epidemiological data supports the hypothesis that poor growth in fetal life and infancy is associated with type-2 diabetes, coronary heart disease, stroke, hypertension and obesity [24]. Epidemiological studies by Forsdahl *et al.* in Norway, first showed a causative link between early life environmental factors and subsequent disease [25]. In 1986, David Barker and colleagues suggested that poverty, poor nutrition, and general health of mothers produced high rates of infant mortality, and also a lifelong risk of coronary heart disease. Epidemiological data from several human populations have provided human data on the importance of maternal nutrition during pregnancy and its influence on disease in the adult offspring.

1.5.1 Dutch Hunger Winter

The Dutch Hunger Winter was a short-term famine that occurred from 1944-1945. Toward the end of the Second World War, the German occupation forces in the Netherlands cut off the food supplies coming into the northern part of the country. The population of much of the Netherlands suffered severe food shortages for six months. The epidemiological data from women pregnant during this famine shows that babies born to mothers who suffered severe starvation during the last three months of their pregnancies grew up to have a dramatically reduced ability to deal with high levels of sugar. These offspring often grew up nutritionally richer in the post-war environment and experienced an increased risk of developing diabetes [26]. This indicates that undernutrition in the last trimester of pregnancy may programme later obesity in men and women [27, 28]. These studies suggest that perturbations of central endocrine regulatory systems established in early gestation may contribute to the development of obesity in later life.

1.5.2 Pima Indians

The Pima Indians, of the Gila River Indian community in Arizona, lived successfully in the desert environment by maintaining the 2,000 year old tradition of irrigation and agriculture. This was disrupted in the late 19th century when their water supply was diverted. This disruption caused poverty, malnutrition and even starvation. The community was forced to survive on the lard, sugar and white flour provided by the U.S government for survival [29]. By 1929, irrigation was gradually reintroduced and now agriculture is once again a major source of food, but the type of farming has changed. The Pima Indians are now the subject of numerous studies into the population's high prevalence of obesity and particular chronic diseases. The high risk of diabetes can largely be explained by the presence of maternal diabetes during pregnancy [30]. In the Pima Indians, higher instances of diabetes have been correlated with both high and low birth weight. This is in contrast to studies in the USA, Sweden, France, Norway and Finland that demonstrate a significant correlation between low birth weight and the later development of adult diseases [31]. McCance *et al.* proposed that the population may have developed a genetic predisposition for insulin resistance as it provided a survival advantage in their previous environment.

1.5.3 Cohort Studies

In addition to these retrospective studies, prospective studies are being conducted worldwide. Here is a summary of one that was conducted in the UK.

Southampton Women's Survey

The Southampton Women's Survey (SWS) is a large study in Southampton, England. The study began at the University of Southampton, where David Barker developed the thrifty phenotype hypothesis, to investigate the diet and lifestyle factors that affect women's health and the health of their children. It was established to measure characteristics of women living in the city that are aged 20-34 before pregnancy. The survey recruited 12,500 women and followed them through their sub-

sequent pregnancies. The study includes 3,000 live-born infants that will be followed throughout childhood. The objectives of the project are twofold. First, to characterise the influence of a mother's own fetal growth on her dietary balance, body composition, and endocrine profile before and during pregnancy on: (i) the early trajectory of fetal growth; (ii) the maintenance of this growth trajectory in late gestation; and (iii) placental and fetal adaptive responses, including altered regional blood flow and body composition of the foetus. Second, to examine how maternal and intrauterine influences interact with the offspring's genes and postnatal environment to determine: (i) weight gain, head growth and linear catch-up growth in infancy; (ii) the pathways of growth and development during childhood that lead to poor adult health; and (iii) cardio-respiratory function and asthma in childhood, as well as levels of risk factors for adult coronary heart disease, type 2 diabetes mellitus and osteoporosis.

The studies in Southampton have demonstrated that the balance between carbohydrate intake in the first trimester and animal protein intake in the last trimester influences placental growth and neonatal body composition. Furthermore, studies of early pregnancy, assisted reproductive technology and animal experiments, indicate that maternal body composition, endocrine profile, diet and physical activity around the time of conception, are important in establishing both the fetal growth trajectory and the fetal supply line.

ALSPAC-Avon Longitudinal Study of Parents and Children

The ALSPAC study is a longitudinal population-based birth cohort study that recruited 14,541 pregnant women in 1991-92, with 14,062 live-born children. These women were residents of one of three former (Avon) Health Districts in the UK. The women represented 85% of the eligible population. The goal was to understand the ways in which the physical and social environments interact with the genotype to affect, health, behaviour and development [32]. The study recorded information for growth, onset of obesity, respiratory function, traits relevant to adult-onset diseases, infections, motor and mental ability, educational achievements, sexual development, accidents and injuries, and atopic diseases (i.e. asthma, eczema, allergies, mood

behaviour, and temperament). Also, from late adolescence into adult life the study will monitor the development of type-2 diabetes, markers of coronary heart disease, schizophrenia and other psychotic disorders, criminal behaviour, ability to hold down a job, onset of drug and alcohol abuse, and reproductive success and failure.

1.6 Animal Models of the Metabolic Syndrome

Animal models are very important in the study of the metabolic syndrome. The metabolic syndrome involves the interaction of many factors (genetics, fetal programming, eating habits, energy expenditure, age, and gender) that would be impossible to manipulate and control in a human population. Several models have been developed to investigate the metabolic syndrome. These models have been done in a variety of different organisms including mice, rats, guinea pigs, rabbits, hamsters, pigs, sheep, primates, horses, dogs, and frogs. The most common models utilise a maternal diet manipulation during pregnancy (low protein and general calorie restriction, or high fat/junk food diet). In addition, models given genetic or surgical manipulations, hormonal insults, or a particular postnatal diet can also be informative. Several studies on the impact of severe undernutrition have been done by our collaborators using small animal species (rats) as well as larger mammals (sheep). Cost, size, time, and ethical issues are major reasons for choosing particular animal models over others. Also, popularity can be a good reason for choosing a model. If many other studies have been done on a particular model then animals will be easier to obtain and there will be more data for comparison. The difference in gestation time and variation in timing of the insult makes it possible to gain more information regarding the crucial ‘windows’ for fetal programming. A recent review by Warner and Ozanne [33] details the importance of animal models in studying fetal programming. Many of the different models and the major findings associated with each are summarised here.

1.6.1 Maternal Diets

Rodents are a useful model in maternal diet manipulation studies as the shorter gestation time and lifespan. The development of the brain involved in regulating feeding and energy expenditure develops in the early postnatal period in rodents, whereas in humans it develops in the third trimester of pregnancy [34]. A guinea pig would be a slightly better model as they are born with a well-developed endocrine, cardiovascular, and central nervous systems [35]. Sheep are also a good model as their pregnancies are usually singleton or twin so nutrient allocation between offspring during pregnancy is more analogous to human. The examples listed below will focus on rodent animal models, particularly rats, as that is the model used in this thesis.

Protein Restriction

The maternal low protein model is extensively studied due to the similarities found between studies using this model and studies of individuals with type-2 diabetes and/or the metabolic syndrome. Studies have shown that amino acids found in a protein diet are key to normal fetal growth, but can also be detrimental at high concentrations [33]. The commonly used low protein rat model involves feeding the rats a low (5-8%) protein diet during pregnancy [36]. This mimics the nutritional situation in many underdeveloped countries and results in growth restriction of the offspring. If these offspring are suckled by the same protein-restricted mothers they remain permanently growth restricted even when weaned to a normal chow diet (20% protein). Cross-fostering to mothers being fed a normal chow diet during lactation until weaning causes rapid catch up growth. Catch up growth is the rapid growth after birth that brings a low-birth weight offspring up to the weight of control offspring. In rats, catch up growth affects longevity by causing accelerated loss of kidney telomeric DNA. Maternal protein restriction has long-term effects on insulin-sensitive tissues [37] and causes reductions in pancreatic β -cell mass, skeletal muscle mass and central adipose tissue weights. Offspring subjected to a low-protein diet during gestation demonstrate improved glucose tolerance at a young age (6 weeks to 3 months) [38, 39].

However, males undergo a gradual loss in glucose tolerance and by 17-months develop Type 2 diabetes and insulin resistance [40]. The female offspring experience a more gradual loss in glucose tolerance and do not develop Type 2 diabetes and insulin resistance until 21 months [41].

Caloric Restriction

A calorie restricted rat model is utilised in this thesis (see Chapter 3 and Chapter 4). This involves feeding rats 30-50% of their normal diet during pregnancy. This maternal dietary restriction during pregnancy in rats results in offspring that suffer severe growth restriction and show endocrine and metabolic abnormalities. Studies have shown that the timing of the restricted diet during development is critical in the programming of metabolic disorders [33]. Work done by Garofano *et al.* has shown that a 50% restriction of a normal chow diet in the last week of pregnancy results in low birthweight offspring with decreased β -cell mass. The rat offspring demonstrate this reduced β -cell mass and insulin content in adulthood [42] even after returning to a normal diet and normal body and pancreatic weight. In addition, extended nutrient restriction into the nursing period results in more permanent metabolic abnormalities (permanent β -cell mass and loss of glucose tolerance [43, 44]) in offspring. The model used in this thesis was more severely restricted (30% *ad libitum*) during pregnancy. Vickers *et al.* has shown that rats exposed to this level of maternal restriction and then switched to a normal postnatal diet gain weight quickly and develop the metabolic syndrome (obesity, hyperphagia, hypertension etc.) [45]. This effect is exacerbated by a postnatal high fat diet. Even a restricted diet after birth (3-6 weeks) will result in impaired insulin secretory response at 12 weeks of age [37]. Many studies have been done using caloric restriction. Diets differ between laboratories using different proportions of macronutrients, which can help elucidate the importance of different nutrients in fetal programming.

Iron Restriction

Iron deficiency is a very common nutritional disorder and pregnant women are among the most susceptible. In the study by Gambling *et al.*, rats were given 7.5 mg/kg compared to the control rats that were given 50 mg/kg. This mild iron deficiency *in utero* caused a significant increase in the blood pressures of both male and female offspring [46]. In addition, the offspring have decreased iron concentrations in brain tissue and behavioural differences have been observed [47].

Overnutrition

High fat diets aim to imitate westernised diets. Relatively few studies have focused on the long-term consequences of a maternal high-fat diet and maternal obesity during pregnancy or lactation compared to other nutritional studies. Studies have revealed effects on adiposity, cannibalism (presumably related to failure of lactation), and an increased body weight and visceral fat depot [2]. In a study by Gorski *et al.* [48], rat offspring of normally fed mothers were cross-fostered onto obese dams, which resulted in increased obesity later in life suggesting the suckling period is still an important developmental window in rats.

Maternal High-Fat Diet Offspring that experience maternal overnutrition during pregnancy have been shown to have increased risk of developing Type 2 diabetes and cardiovascular disease later in life [49]. These offspring develop abnormal cholesterol and lipid metabolism, hyperinsulinaemia, insulin resistance and have an increased risk of hypertension and cardiovascular disease [49, 50, 51]. Bayol *et al.* [52] developed a rat model using a 'junk food' diet with processed, palatable food items developed for human consumption. The diet contained high levels of fat, sugar, and salt with low protein levels and thus resembled a low protein model. The study found that the diet influenced muscle development, feeding behaviour, and adiposity. This preliminary study suggested that the observed changes in the muscles of the offspring whose mothers had been on the junk food diet may have a decreased ability to move effectively (i.e. exercise) and may therefore be more prone to obesity.

Maternal Obesity Increase rates of obesity in western countries have led to an increased prevalence of maternal obesity, which can lead to high-risk pregnancies and complications at birth and in later life [2]. Animal models of maternal obesity can help us understand the mechanisms that may transfer this susceptibility to obesity and features of the metabolic syndrome from mother to offspring and into later life. Offspring of obese mouse mothers have been shown to have an increased fat-to-lean-mass ratio and hyperphagic behaviour. In addition, they were insulin resistant at 3 months and males had developed impaired glucose tolerance by 6 months. An increase susceptibility to cardiovascular disease was demonstrated by hypertension and signs of endothelial cell dysfunction [53]. Another study showed that both male and female offspring of obese mothers developed metabolic syndrome phenotypes and increase adiposity. Males also demonstrated insulin resistance and poor glucose tolerance [54].

1.6.2 Postnatal Diets

High fat postnatal diets are often used in conjunction with undernutrition during pregnancy (as done in the study described in Chapter 4. This is done to create a mismatch between pre- and postnatal life that is often associated with the metabolic syndrome as described by the thrifty phenotype hypothesis. Vickers *et al.* [45] has shown that the postnatal high fat diet exacerbates the effects caused by the maternal undernutrition diet.

1.6.3 Hormonal Insults

Glucocorticoid Treatment

It is established that glucocorticoid treatment during pregnancy in animals and humans leads to reduced birth weight, and it is more recently being realised that this has long-term effects on the offspring [37]. Exposing pregnant females to glucocorticoids is another treatment used to induce the metabolic syndrome. A recent study using the nonhuman primate marmoset, revealed that glucocorticoid exposure

resulted in increased expression of *11 β -hsd1* in liver, pancreas, and subcutaneous fat. The increase in *11 β -hsd1* occurred before the onset of obesity [55].

Adrenalectomy

Gardner *et al.* [56] conducted a study combining adrenalectomy (the removal of one or both adrenal glands) in the rat with a glucocorticoid treatment and a low protein diet to confirm that maternal glucocorticoids are involved in the programming of hypertension in the offspring. The maternal diet-induced hypertension is dependent on an intact adrenal gland postnatally and glucocorticoids are key nutrients responsible for maintaining high blood pressure.

Leptin Treatment

Leptin functions to sense and regulate body energy stores in humans and in animals. Leptin administration after birth has been used to investigate it as a possible treatment to prevent obesity or to reverse the programming effects of maternal undernutrition during pregnancy. Vickers *et al.* conducted a study utilising a leptin treatment that reversed the effects of prenatal undernutrition in female rats [57] (these samples are used in Chapter 4. However, the same protocol showed no significant effect in male rats [58]. Yura *et al.* [59] conducted a similar study in mice that did not result in a reversal of the programming effects, but rather caused a growth surge in the mice that received normal maternal nutrition during gestation.

Growth Hormone

Growth hormone treatment was tested as a possible treatment for fetal programming. Vickers *et al.* found that GH treatment after weaning in male rats that were exposed to maternal undernourishment during gestation resulted in increased body weight, but significantly reduced fat pad weight. Systolic blood pressure was markedly decreased and the rats had an increased heart-to-body weight ratio indicating GH treatment reduces hypertension and improves cardiovascular function.

1.6.4 Surgical Interventions (Reduced Placental Transfer by Uterine Artery Ligation)

A common cause of intrauterine growth restriction (IUGR) is due to placental disease particularly in developing countries. The sheep is an ideal placental insufficiency model as the foetus can be monitored throughout gestation. However, these studies usually only focus on fetal development as cost prohibits taking the offspring to adult life. There is also a problem of the confounding of the factors of reduced maternal and fetal blood flow, nutrient restriction and hypoxaemia that vary in severity between species (i.e. sheep, rodents and humans) [35].

1.6.5 Genetic Models

The ability to introduce or eliminate genes from the genome of rodents has made rats and mice excellent complex, genetic models. Rodents are especially useful because of their large family size and easy measurement of phenotypic parameters. There are several mouse models with knock-out genes that are useful for investigating the metabolic syndrome. Two genetic models that will be described in detail in Chapter 4 are the *ob/ob* leptin deficient mouse and the *db/db* leptin-receptor deficient mouse. Also mentioned briefly in Chapter 6 is the *KKAy* lethal agouti gene mouse model. A recent publication by Artinano *et al.* [60] reviewed nine genetic rat models of the metabolic syndrome that develop a number of disorders including obesity, hypertension, hyperphagia, fatty liver, etc. The most common rat mentioned is the Zucker rat that has a mutation in the leptin-receptor similar to that of the *db/db* mouse. A few of the other genetic models available for investigating the metabolic syndrome include: IGF-I knockout mice, IGF-II transgenic mice and IRS-I disrupted mice. The IGF-II transgenic mouse highly over-expresses IGF-II resulting in increased fetal size, organ overgrowth, and has been used to investigate the development of the endocrine pancreas [61].

1.6.6 Summary

This is a broad overview of the many animal models available for the investigation of the metabolic syndrome. Of course none of these models manifest the disease in the same way humans do, the combination of models helps to understand the role of fetal programming *in utero* as a cause of adult disease [62].

1.7 Mechanisms of Nutritional Programming

Either in the introduction or in chapter 5 you need to include much more detail about the proposed mechanisms of nutritional programming. This should include a general introduction to epigenetics, how methylation blocks gene transcription, histone modification, the reciprocal relationship between methylation and histone modification and the evidence that early life environment can alter epigenetics processes (in vitro culture, IUGR, maternal behaviour and the now wealth of evidence that nutrition can alter DNA methylation and the fact that methylation changes have been seen in your model).

The molecular mechanisms behind the developmental programming of disease in response to insults during gestation are unclear. Waterland *et al.* [63] came up with three criteria for considering possible mechanisms: 1) Nutritional programming is an adaptive process, so only mechanisms consistent with an adaptive response are considered; 2) Programming occurs during critical windows of development, so only mechanisms that act in these windows can be considered and 3) The potential mechanism must explain a programming process and not be a reflection of the outcome of the process.

The mechanisms that have been proposed and generally meet this criteria are: organ structure, altered cell number, clonal selection, cell type, metabolic differentiation, autoregulatory pattern of DNA binding proteins, precocious activation of the HPA axis, increased local glucocorticoid axis and endocrine sensitivity, epigenetic regulation of gene expression via altered chromatic structure or DNA methylation, hepatocyte polyploidisation, impaired mitochondrial function, and reduced oxidative

capacity [64, 63]. More than one mechanism is likely to be involved. These mechanisms may vary between tissues and depending on the duration and timing of the insult. One mechanism that has gained significant support is the epigenetic regulation of gene expression via DNA methylation. The understanding of how, when, and what changes are made to the epigenome could have significant clinical implications. The identification of methylation patterns that increase disease susceptibility could lead to identification of individuals with these patterns and allow advanced preventative treatment or lifestyle adjustments [65].

This thesis aims to take an exploratory look into whether DNA methylation may be involved in any of the gene expression changes highlighted by the microarray studies.

1.8 Microarray

Since the sequencing of the human genome the field of molecular biology has been transformed. This has led to a new field of biology referred to as genome biology or genomics. New technology combined with our growing knowledge enables scientists to probe the genome investigating the structure and function of a very large number of genes simultaneously. The types of investigations performed can be grouped into subfields including structural genomics, comparative genomics and functional genomics. Structural genomics includes the genetic mapping, physical mapping and sequencing of genes in the genomes of various organisms. Comparative genomics aims to make inferences based on the comparison of the genomes of different organisms. Finally, functional genomics involves investigating the role that individual genes or subsets of genes play in the development and life of an organism [66]. The work in this thesis would fall in the subfield of functional genomics.

Many high-throughput techniques have been developed to investigate properties of the genome. Microarray is a technique that enables the user to probe thousands of genes simultaneously to look for changes in expression levels, methylation or SNPs (single nucleotide polymorphisms).

This thesis utilises microarray technology to probe the rat genome expression profile for novel genes that are affected by the nutritional insult during pregnancy. Microarray enables this large scale screening of gene expression by detecting the presence and abundance of labelled nucleic acids. It makes it possible to measure the expression of thousands of genes simultaneously. It also makes it possible to compare the expression in different cell types (i.e. tissues) and to compare the effects of different treatments.

An array is a glass slide on which single stranded DNAs (ssDNA) with various sequences, referred to as the probe, have been placed or synthesised in a grid-like pattern. This can be done using a variety of techniques including spotting, photolithographic synthesis, inkjet synthesis, or most recently, BeadArray technology. The array is washed in a solution containing ssDNA that is generated from a particular biological sample that is being studied and is referred to as the target. The DNA in the solution contains sequences complementary to the sequences of DNA on the surface of the array and will hybridise to those complementary sequences. The solution is labelled with a fluorescent dye so that the hybridisation spot can be detected and quantified easily.

The DNA target hybridised on the array is obtained by a reverse transcriptase reaction from the mRNA extracted from a tissue sample. This DNA is fluorescently labelled with a dye and a subsequent illumination with an appropriate source of light will provide an image of the array of features (set of spots). The intensity of each spot or the average difference between matches and mismatches can be related to the amount of mRNA present in the tissue and in turn, with the amount of protein produced by the gene corresponding to the given feature. The array is then used to answer a specific question regarding the printed DNAs.

A number of microarray platforms have been developed using the different techniques mentioned above; with different formats (one-channel/ two-channel, in-house/ commercial, cDNA/ oligonucleotide); by different companies (Affymetrix, Illumina, Agilent and Nimblegen); with a different number of probes; for different species (most common human, rat and mouse). To choose a particular platform considerations have

to be made concerning the organism of study, the number of samples to hybridise, the cost and the amount of extracted RNA available. One channel arrays are becoming more common and are largely preferable because there is less variation introduced due to dye bias. Homemade arrays allow for total flexibility and are low cost if enough of them will be produced; however, they do not offer standardised protocols and tested analysis methods that the commercial arrays offer[66].

In this thesis, two different types of mRNA arrays were used. The MEEBO oligonucleotide array is a printed array that was made in-house by the Department of Pathology Centre for Microarray Resources. The Illumina array is a BeadChip that uses a bead coated in ssDNA inserted into the glass slide. Both are described in more detail in Chapter 2.

1.8.1 microRNA Arrays

In addition to the mRNA arrays, this thesis also utilised microRNA (miRNA) arrays. miRNAs are small noncoding RNAs that regulate gene expression by targeting mRNAs for cleavage or translational repression. Recent developments in microarray technology have included the manufacture of arrays with miRNA targets. Illumina briefly released a miRNA expression array that had 380 sequences from the mouse genome (described in more detail in Chapter 2). The miRNA array was used for this thesis in Chapter 6 to investigate the effects of maternal undernutrition on the miRNA expression levels in the livers of 55 day old male offspring.

1.8.2 Microarray Applied to Developmental Plasticity

By using microarray to find gene expression changes, it is possible to take a global gene approach to investigating which genes/mechanism might be involved in programming. Previous studies have used a candidate gene approach. By approaching the problem with less bias, it may be possible to identify novel genes and pathways important in programming. In this study, genes will initially be identified by microarray and then the methylation status of the promoter regions of these genes of interest

will be analysed.

1.9 Structure and Aims of the Thesis

This thesis aims to pinpoint potential molecular mechanisms that underpin the development of the metabolic syndrome. The thesis utilises microarray technology and epigenetic assays to probe the genome for information. All the techniques and materials used are introduced and described in Chapter 2.

Chapter 3 utilises samples from a previous study (Vickers unpublished study) to compare the gene expression profiles of offspring of rats that were undernourished during pregnancy (UN) to those whose mothers ate a normal chow diet during pregnancy (AD). Liver, white adipose, and skeletal muscle tissue samples are examined from male animals that were 55 days old, before the development of the metabolic syndrome phenotype. The chapter gives the background on metabolism and relevant metabolic pathways. The chapter also summarises the results of the initial phenotypic study before detailing the results of the expression profiling and the associated pathway analysis. The aim of the study presented in this chapter was to highlight genes differentially expressed between AD and UN rats that may be involved in predisposing the UN animals to the development of the metabolic syndrome in later life.

Chapter 4 utilises the samples from a previous study [57] to compare the gene expression profiles of rats subjected to three different treatments: undernourishment during pregnancy (AD/UN); leptin treatment (SAL/LEP); and postnatal high-fat diet (CHOW/HF). Liver samples were from female animals that were 170 days old and had already developed metabolic syndrome in particular treatment groups. The chapter introduces leptin and its role in appetite signalling and energy homeostasis. The results of the initial phenotypic study are described to put the results of the expression profiling study in context. The results of the phenotypic study had suggested that leptin treatment in UN animals resulted in a reversal of programming that made these animals susceptible to the metabolic syndrome. This chapter aims to highlight

genes differentially expressed between treatment groups and potentially involved in the metabolic programming and the reversal of the programmed phenotype.

Chapter 5 introduces the hypothesis that epigenetic mechanisms are responsible for the fetal programming. The preliminary investigation presented here determined the methylation status of particular genes by conducting a restriction enzyme methylation assay (McrBC) on samples that were highlighted in a preliminary analysis of the results presented in Chapter 3. This chapter aims to investigate whether epigenetic effects through methylation at CpG dinucleotides in promoter sequences are responsible for gene expression changes and could in turn be responsible for the resulting phenotypic changes.

Chapter 6 utilises a newly developed miRNA microarray to investigate the effect of maternal undernutrition on the microRNA (miRNA) expression levels using the liver samples of male rats that were 55 days old (the same samples as utilised in the study in Chapter 3). This chapter aims to highlight miRNAs involved in differentially regulating genes as a consequence of maternal undernutrition during pregnancy.

Chapter 7 summarises the results of the work presented in this thesis and discusses the implications of these results. This chapter aims to outline the general conclusions and to state plans for future work.

Chapter 2

General Methods

2.1 Animal Care/Sample Collection

Tissues (liver, retroperitoneal white adipose fat, and biceps femoris skeletal muscle) for the data in Chapter 3 were collected and snap-frozen in liquid nitrogen by collaborators at the Liggins Institute, University of Auckland. They were stored at -80 °C until being shipped to the University of Cambridge on dry ice. Liver tissue for the data in Chapter 4 was collected and snap-frozen in liquid nitrogen by collaborators at the Liggins Institute, University of Auckland. All tissue was stored at -80 °C until RNA extraction was done at the Liggins Institute. RNA was shipped to the University of Cambridge on dry ice. All animal manipulations were performed under the ethics and approval protocol number CR328 from the Animal Ethics Committee of the University of Auckland.

2.2 RNA Extraction

2.2.1 For Oligonucleotide Expression Array (Undernutrition Study and Leptin Study)

Total RNA was isolated from tissue using TRIzol reagent (Invitrogen, Paisley, Scotland, UK). For each sample, 25 mg of frozen tissue was dropped in 1 ml of

TRIzol solution. The tissue was homogenised at maximum speed for 30 seconds. The homogeniser head was thoroughly cleaned between each sample. Muscle and fat tissue was then centrifuged at 11,500 g (12,000 rpm) for 10 minutes at 4 °C. This caused the extracellular membranes, polysaccharides, and high-molecular weight DNA to form a pellet at the bottom of the tube with the RNA in the supernatant. In fat samples, an excess of fat will collect in the top layer which can be removed. The cleared homogenate solution can then be moved to a fresh tube. These samples along with liver samples, were then left at room temperature for 5 minutes to promote the dissociation of the nucleoprotein complexes. In the fume hood, 200 μ l of chloroform was added and the sample was mixed with a vortex on high speed for 15 seconds, then incubated at room temperature for 3 minutes. Tubes were then spun at 11,500 g (12,000 rpm) for 15 minutes at 4 °C. The aqueous phase was carefully transferred to a new tube. Equal volume of 70% ethanol (50% for liver samples) was added to the aqueous phase for total RNA precipitation. The subsequent steps were all preformed at room temperature using the RNeasy Mini Column purification kit (Qiagen, UK) following the manufactures instructions.

2.2.2 For microRNA Expression Array

Purification of total RNA and a separate miRNA-Enriched Fraction from 25 mg of frozen liver tissue was performed using the miRNeasy kit (Qiagen, UK) following the manufacturer's protocol. This protocol also utilises Qiagen's RNeasy MinElute Cleanup Kit.

2.2.3 Quality Assessment

Total RNA quality was assessed by running 100–500 ng of the RNA sample on the Nano Lab-on-a-chip system (Agilent, Palo Alto, CA.) to check for RNA degradation. MicroRNA quality was assessed using the Small Lab-on-chip system. The Agilent Bioanalyser 2100 separates the total RNA sample such that 18s and 28s ribosomal RNA will dominate the electropherogram. When good quality RNA is analysed the

ratio between the two peaks ranges from 1.5–2.0. The ND-1000 spectrophotometer (Nanodrop) was used to measure the total RNA concentration and purity. A 280/260 nm (ratio of absorbance) of 1.98 was used as the quality threshold, as values less than this would indicate protein contamination. A 260/230 nm of 1.8 threshold was used to avoid samples with organic contamination.

2.3 Microarray

2.3.1 Experimental Design

An important aspect of a microarray experiment is the experimental design. A designed experiment involves a series of tests in which purposeful changes can be made to the input variables of a process or a system, so that the results can be observed and the reasons for the changes in the output can be identified. Factors that may contribute to noise in an experiment must be identified and if possible, controlled. The appropriate statistical analyses should be chosen in advance. The three main principles of experimental design are replication, randomisation, and blocking. Replication is the repetition of an experiment which allows an estimate of the experimental error. In microarray, there are two types of replication: biological and technical. Biological replication involves the repetition of an experiment using different biological samples that have gone through the same experimental conditions. This will reveal biological variation between individuals in the population being studied. Technical replication is an exact repetition of an experiment using the same biological sample. This will reveal differences noise introduced in the experiment based on differences in equipment or experimental conditions [66]. Randomisation requires the random choice for every factor (nuisance factors) that is not of interest but might influence the outcome of the experiment. Examples of this involve the placement of spots or beads on an array, and the use of control and treatment samples on the same slide of batch of slides. Randomisation can control for the confounding of factors. If all treatment samples were run on one batch of slides and all the control samples on a

different batch than the variation introduced by the experiment and that introduced by the slides would be impossible to separate [66].

Blocking is a technique for experimental design that aims to increase the accuracy with which the influences of the nuisance factors are assessed in a given experiment. A block is a subset of experimental conditions that are expected to be more similar than other conditions. Blocking is used to eliminate the variability due to the difference between blocks. An example of a block is the chip itself. All spots in an array or on a slide are subject to the same factors during slide processing; it is expected that the spots on a single chip will have less variance than the measurements across the entire experiment. The two channel array process deals with this by hybridising the treatment and control samples on the same array; however, it introduces the nuisance factor of the dyes, which in this experiment is dealt with by running replicates that are dye swapped. This means that the same experiment is run twice with the control and treatment hybridised with each dye. The BeadArray deals with blocking by having more than one array on a single slide [66].

2.3.2 Quality Control

As the image analysis is fully automated, it is important to have good quality assessment methods. With the spotted arrays, this is very crucial as spotting errors can cause problems with spot shape regularity, spot area to perimeter ratio, displacement, spot uniformity, and spot signal area to spot area ratio. With spotted arrays, it is useful to visually inspect slides and images to check for any obvious problems with the printing or hybridisation.

Many of these problems are overcome with BeadArray Technology. However, with both methods there are still a variety of other potential problems that can be detected using visualisation tools. BlueFuse® performs a quality control assessment and uses a flagging system to highlight potentially unreliable data, while Illumina uses a set of internal controls for the same purpose. Visualisation tools are then used to look for differences in slides or outlying samples, or to assess the information provided by

BlueFuse® or Illumina.

2.3.3 Reducing Error

Choosing the right statistical methods is an essential part of microarray analysis. It is important to identify the potential sources of variability during the quality control step and then choose normalisation methods that will correct for this without covering the biological variability of interest. There is a large potential for error when using microarray as cost restraints reduce the number of replicates that can be done while the number of variables remains large. Data normalisation aims to remove the external noise to reveal biologically relevant differences in gene expression. A lack of statistical significance may indicate low experimental sensitivity rather than an absence of biological effect. Low sensitivity may be caused by an inadequate number of replicates and/or a failure to control noise that contributes to random error [67].

Data Filtering

Following the quality control analysis, it is important to remove samples that are extreme outliers. If the experimental design has incorporated enough biological and technical replicates, it should be possible to detect outliers that may negatively affect the results.

Between Array Normalisation

The characteristics of each array should be similar. Proper experimental design should have randomised samples in such a way that any sources of error will be diffused when replicates are combined. If it is assumed that these precautions have been successful and that filtering has removed samples that are extreme outliers, then the within-array-normalisation step should centre all arrays on the same mean/median.

2.3.4 Data Analysis

Once data has been filtered and normalised it is analysed to determine which genes differ significantly in expression level between the treatment and the control. In this thesis, a linear model is used to compare samples. This method requires two matrices: one is the *design matrix*, which indicates what RNA samples have been applied to the array and the second, is the *contrast matrix*, which specifies what comparisons should be made between the samples. The differential expression is then determined by calculating a p-value using an analysis of variance method (ANOVA) or a pairwise comparison.

2.3.5 Validation of Microarray Data: Quantitative Real Time PCR

The goal of microarray is to explore a large subset of the genome to find genes that might be involved in a particular biological process related to the test sample utilised. A list of genes with a chosen level of statistical significance will be revealed, but it is absolutely essential to test the validity of these results by testing the expression of that gene individually with an alternate method. RT-PCR and northern blotting are two methods often used at this step. Quantitative RT-PCR is the method of choice for validation of microarray experiments, as it has been shown to be the most sensitive method for quantifying changes in expression. Comparative quantitative RT-PCR has been used for validation in this thesis.

2.3.6 Mouse Exonic Evidence Based Oligonucleotide Expression Array

Initially, the samples were hybridised to a spotted array, specifically the Mouse Exonic Evidence Based Oligonucleotide (MEEBO) array. This array is comprised of DNA oligonucleotides representing the mouse genome that are pre-spotted to the array using robotic spotting. This chip probes approximately 25,000 genes in the

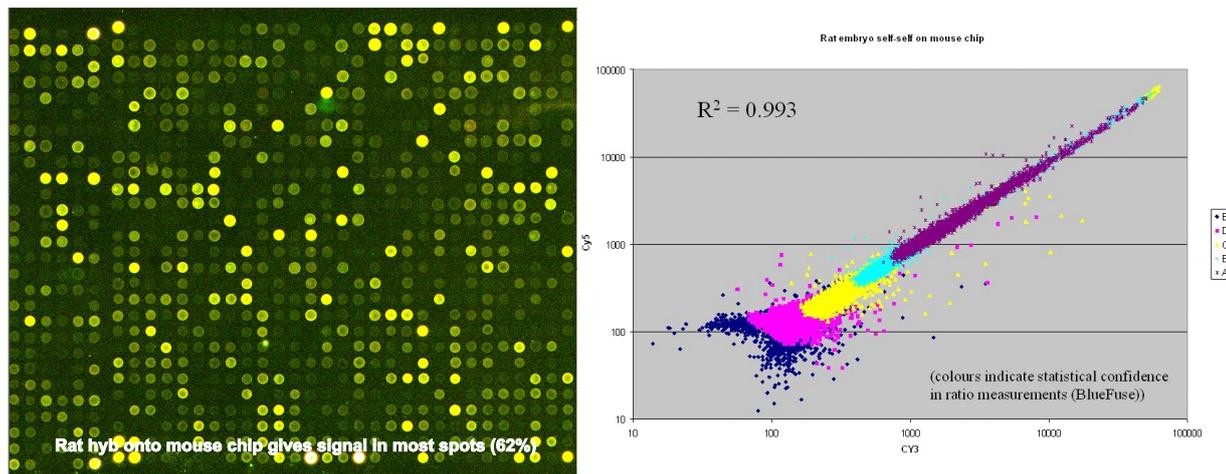


Figure 2.1: Rat embryo DNA hybridised to the mouse MEEBO chip. The rat DNA gave signal in approximately 62% of the features and scatter plots of data revealed a R^2 value of 0.993. This work was done by Peter Ellis.

mouse genome and uses a dual colour hybridisation protocol. Dual-colour hybridisation is used in a two-channel experiment by labelling the target and the probe with Cy3 and Cy5 dyes. By using this chip, we were cross hybridising rat samples to a chip with mouse oligonucleotides. To test the accuracy of hybridising rat samples to a mouse chip, we performed several test hybridisations. Rat embryo DNA was amplified, labelled and self-self hybridised to the MEEBO chips. The rat embryo DNA gave signal in approximately 62% of the features and scatter plots of data revealed a R^2 value of 0.993 as shown in figure 2.1 (work done by Peter Ellis). These are very good results, but not as ideal as having an actual rat chip. For this reason, once the technology was available, samples were also hybridised to the Illumina Bead Array described below.

Smart PCR amplification, Klenow Labelling and Array Hybridisation

Extracted RNA samples were run on the Agilent to ensure RNA quality. Reverse transcription was performed using Powerscript (Clontech), from 1 μg of total mRNA mixed with 1 μl of cDNA synthesis primer (10 μM), and 1 μl of template switching primer and RNase-free water up to 5 μl . This mix was incubated for 2 minutes at 72 $^\circ\text{C}$. To this mixture 2 μl of 5x First Strand Buffer, 1 μl DTT (20 μM), 1 μl

dNTPs (10 μ M), and 1 μ l Powerscript RT were added and this mix was incubated for 60 minutes at 42 °C . The cDNA amplification used the SMART PCR amplification protocol. A mix containing 75 μ l of RNase-free water, 10 μ l 10x PCR Buffer II, 2 μ l 10 mM dNTPs, 4 μ l IIA Primer, 5 μ l 25 mM MgCl₂, and 2 μ l AmpliTaq (5 U/ μ l) was made and added to 2 μ l of product from the 1st Strand synthesis. Amplification was performed at 95 °C for 1 minute in order to activate the enzyme, then for 14 cycles, 95 °C for 5 seconds, 65 °C for 5 seconds, and 68 °C for 6 minutes. Labelling was performed using the Klenow Labelling kit (BioPrime) following the manufacturer's instructions. 20 μ l of 2.5x Random Primer Buffer and 1 μ l Klenow (40 U/ μ l) was added to 22 μ l of amplified DNA. The mix was incubated at 95 °C for 5 minutes. This mix was placed on ice while 5 μ l 10x Low-C dNTPs and 1 μ l cyanine dye (Cy3 or Cy5) was added and the mix was incubated for 2 hours at 37 °C. 5 μ l of STOP Buffer was added to stop the reaction. Labelled cDNA was purified using G50 Columns. Labelling and synthesis was checked using the Nanodrop and purified Cy3 and Cy5 labelled products were combined. The sample was then pooled with 1 μ l Cot1DNA, 1 μ l yeast tRNA (Invitrogen), and 1 μ l of poly-dA (Sigma) as blocking reagents. Ethanol precipitation was performed by adding 250 μ l of 100% Ethanol and 10 μ l 3M NaAcetate (pH 5.2), inverting the mix, placing on dry ice for a 10 minutes, and spinning at 11,500 g (12,000 rpm) for 15 minutes at 4 °C . The pellet was washed with 500 μ l of 75% Ethanol and spun for 5 minutes at 11,500 g (12,000 rpm). The supernatant was discarded and the excess Ethanol was tapped away. The pellet was immediately resuspended in 50 μ l of hybridisation buffer (40% formamide, 5x Denhardt's solution, 5x SSC, 1.65 mM sodium pyrophosphate, 50 mM Tris-Cl pH 7.4, and 0.1% SDS) at 50 °C for 10 minutes. The samples were denatured at 95 °C for 5 minutes, quickly pipetted onto slides, and covered with lifter-slips. The slides were placed in hybridisation boxes containing tissues soaked with 50 μ l of water to ensure constant humidity, but not so much moisture that cover slips slide off. The boxes were placed in a water bath at 48 °C and hybridised for 17 hours. After incubation, the cover slip was separated from the array by immersing it into a solution of 1x SSC. The labelled array was washed with agitation for 5 minutes each in baths of 1x SSC,

followed by 0.5x SSC, and finally 0.1x SSC. The array was rinsed quickly in water (no more than a few seconds to avoid denaturation) and in ethanol, and spun at 3000 x g to dry. Slides were then stored desiccated in a dark box until scanning and were stable for at least a week. When possible, the arrays are scanned on the same day as the washing step to avoid differential degradation of the dyes and accumulation of dust.

Image Capture, Quantification and Quality Control

The MEEBO microarray image was captured using an optical scanner, the ArrayWorx® biochip reader (Applied Precision). The preliminary scan was performed at the minimum exposure time in each channel (0.06 seconds), and the optimum exposure calculated to bring the brightest features on the array to the saturation limit of the scanner, thus, increasing the sensitivity of the image analysis for the less bright pixels. The final scan was performed with the new exposure times. The array was scanned to produce a digital record of the red and green fluorescence emissions at each point on the array. This digital record typically takes the form of a pair of 16-bit TIFF images. Images must be analysed to extract numerical foreground and background intensity for the red and green channels for each spot on the array. The localisation of each spot on the array was done automatically by the software Bluefuse® using the basic array layout and the number of spots on the array. Bluefuse® averages intra-slide replicate spots to assess the reproducibility of the duplicates and provide a confidence score. The setting of time-exposures for Cy3 and Cy5 is a way of adjusting, therefore, correcting for intensity variances between both channels. Images were analysed with the Bluefuse® software. Bluefuse® performs a quality control assessment of the data and allocates flags accordingly (A: highly reliable, E:highly unreliable). This information can be used to filter bad data. For some analyses, a balanced design is required and therefore flagged data was not filtered. However, the flag was considered once the list of significant genes was obtained and genes with E flags were considered unreliable. Data files were then imported into excel, where relevant pieces of information could be cut out and transferred to an analysis software

program.

2.3.7 Illumina BeadArray

The Illumina Bead Array technology is based on the random self-assembly of an array of beads onto a patterned substrate. The 3-micron silica beads self assemble in microwells on planar silica slides (see Figure 2.2). Each bead is covered with hundreds of thousands of copies of a specific oligonucleotide that act as probe sequences. The random loading of beads and subsequent decoding allows extreme miniaturisation and high-density packing which increases the redundancy of each element on the array. The DNA-decoding method utilises sequential hybridisations of dye-labelled oligonucleotides to create a combinatorial decoding method. By using an error checking scheme, the median random error rate for the decoding algorithm was estimated to be $<1 \times 10^{-4}$. This has negligible impact on the results of the hybridisations which give the >20 -fold redundancy and the five-fold minimum redundancy of each bead type [68]. This system has a specific rat chip. The specific Illumina chip we use in our hybridisation is the Rat-Ref12 chip, as shown in Figure 2.2. This chip was designed from the NCBI RefSeq database and contains 12 arrays, each comprised of 22,500 oligonucleotide 50-mers and probes 21,910 genes covering almost the entire rat transcriptome.

Each array has up to 1,536 different bead types. Each bead type is represented by an average of ≈ 30 copies in any array. This means that each array is unique. Due to the ≈ 30 -fold oversampling (50,000 beads/1,536 bead types), it is insured that decoded arrays have greater than or equal to five beads of each type in the array, all sequences are represented. Each bead type has $\approx 700,000$ copies of a particular oligonucleotide probe covalently attached to the bead by the sequence's 5' end. These chimeric oligonucleotides are approximately 75-nucleotides in length, comprising a ≈ 25 -nucleotide identifier sequence and a 50-nucleotide gene-specific probe. Because the population of beads in an array is a random sampling of a starting bead pool containing 1,536 bead types, the representation of the bead types in the array follows

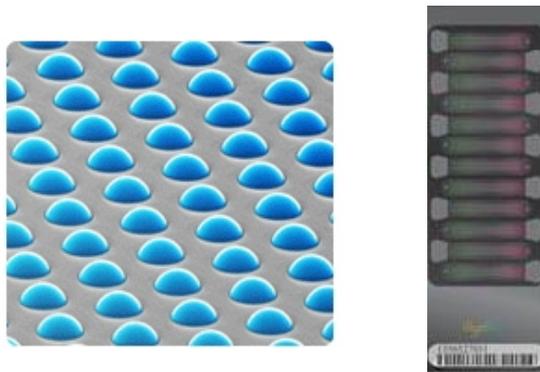


Figure 2.2: Bead Array Image

the Poisson distribution. That is, there is a variable number of each of the 1,536 bead types both within and between arrays. This randomness and redundancy provides two important advantages. The randomness minimises the effects of spatially localised artefacts and the redundancy increases measurement precision and robustness. These factors combine to increase measurement accuracy [69].

It is also important that the identifier sequence of the oligonucleotide does not interfere with the hybridisation. This is done by screening the sequences to avoid similarity that would lead to cross-hybridisation. In addition, the identifier sequences are half the length of gene-specific probes and have lower T_m 's (melting temperatures). It is estimated that the signal provided by the identifier sequences is not enough to affect the analysis [69].

RNA Amplification, Labelling and Array Hybridisation

The RNA samples were amplified following the Illumina TotalPrep RNA amplification protocol. This process begins with reverse transcription using the T7 Oligo(dT) Primer to synthesise first strand cDNA containing a T7 promoter sequence. Second strand cDNA synthesis converts the single-stranded cDNA into a double-stranded DNA (dsDNA) template for transcription. The reaction uses DNA polymerase and RNase H to simultaneously degrade the RNA and synthesize the second strand cDNA. cDNA purification removes RNA, primers, enzymes, and salts that would inhibit in vitro transcription. In vitro transcription generates multiple copies

of biotinylated cRNA from the double-stranded cDNA templates; this is the amplification and labelling step. cRNA purification removes unincorporated NTPs, salts, enzymes, and inorganic phosphate. After purification, cRNA is ready for use with Illumina's direct hybridisation array kits. The samples were hybridised overnight at 55°C to the RatRef-12 expression chip available from Illumina (sampling 21,290 genes from the rat transcriptome).

Image Capture, Quantification and Quality Control

The BeadArray Reader and associated BeadStudio software was used for the image analysis of the Illumina BeadArrays. The scanner uses a confocal-type imaging system with $-0.8\text{-}\mu\text{m}$ resolution and 532 and 635 nm laser illumination. Scans were performed in the 532-nm channel [69]. The scanner completes the following steps during the scanning of the BeadArrays: (i) All pixel intensities are altered using a sharpening transformation. The intensity of a particular pixel is made higher/lower if its intensity is high/low in comparison to the intensities of the pixels surrounding it. (ii) Foreground intensities are calculated as a weighted average of signals obtained using the four pixels nearest to each bead centre as a virtual bead centre. Sharpened pixel intensities are used in the calculation. (iii) The local background, an average of the five dimmest pixels (unsharpened intensities) within the 17x17 pixel area around each bead centre, is subtracted.

It is essential to upload the slide specific decode file before scanning the slide, as this gives the software the details of the location of the bead types on the slide. Once the scan is complete, the files are transported to BeadStudio for the quality assessment.

The Illumina Bead Studio software was used for the initial quality control assessment. Bead Studio includes a spectrum of internal controls for determining data quality. These check for consistency in signal intensity, expected background and noise levels, mismatches versus perfect matches, and signal in housekeeping genes.

2.3.8 microRNA Expression Array

RNA Amplification, Labelling and Array Hybridisation

DNA samples were run with the Illumina MicroRNA Profiling Assay using the Mouse microRNA Panel which contains 380 sequences. The manufacturer provides a detailed protocol for the assay which is an extension of the proven DASL (cDNA-mediated Annealing, Selection, Extension, and Ligation) Assay. First, a stretch of polyA tail was added to the 3' end of each sequence in the RNA sample. Then sufficient RNA from each sample was reverse-transcribed. The biotinylated cDNAs were combined with microRNA-specific oligos (MSOs), hybridisation reagents, and paramagnetic particles in an Assay Specific Extension (ASE) plate. The plate was heated to allow the MSOs for each sequence target of interest to anneal to the biotinylated cDNA samples. The cDNA was simultaneously captured by paramagnetic particles. After the oligos were hybridised to the cDNA, mis-hybridised and excess oligos are washed away. Next, an extension and ligation master mix (consisting of extension and ligation enzymes) was added to each cDNA sample. The extension and ligation reaction occurs at 45 °C . The DNA polymerase and the Uracil DNA Glycosylase were added to the master mix for PCR. The PCR reaction used three universal primers. Two were labelled with fluorescent dyes and the third was biotinylated. The biotinylated primer captures the PCR product and allows the strand containing the fluorescent signal to be eluted. The PCR plate was thermal cycled to fluorescently label and amplify the templates generated in the Pre-PCR process. The double-stranded PCR products were immobilised by binding the biotinylated strand to paramagnetic particles. The solution was transferred to a filter plate and incubated at room temperature so that the PCR product could bind to the paramagnetic particles. The single-stranded fluor-labelled PCR product from the filter plate was washed and then eluted into an intermediate (INT) plate. The product from this plate was hybridised to the BeadChip. The BeadChips were hybridised overnight using the Illumina Hyb Chamber with a temperature ramp from 60 °C to 45 °C . The BeadChips were then removed from the Hyb Chamber and washed three times

with UB2 and XC4 reagents.

Image Capture and Quantification

The Illumina BeadArray scanner was used to capture and quantify the hybridisation signals as described for the expression arrays. The Illumina Bead Studio software was used for the initial quality control assessment. Subsequent analysis was done using the R package Limma as previously described for the expression arrays.

2.3.9 Filtering, Data Visualisation and Normalisation

Once the data has been captured and quantified, the raw data was imported into an analysis software program. Most of the analysis done in this thesis utilised R. R is an open source language and environment for statistical computing and graphics. R provides a variety of statistical (linear and nonlinear modelling, classical statistical tests, time-series analysis, classification, clustering, etc.) and graphical techniques, and is highly extensible. Bioconductor is an add-on package for R that was developed for Computational Biology and Bioinformatics [70]. As Bioconductor is commonly used for microarray analysis, additional add-on packages are frequently being developed by users. These packages can be platform or analysis specific. The two packages utilised most for analysis in this thesis are Limma and Lumi, which are described later.

Inforsense Knowledge Discovery Environment 4.2 (KDE; www.inforsense.com) was also occasionally used for data manipulation and analysis. KDE is a statistical package with a unique interface that appears as a pipeline of tasks. Although less flexible than R, it is useful for visually manipulating datasets. The software also includes a wide variety of normalisation methods, classifications, clustering, and graphical options.

Filtering

Illumina data was filtered in R for those showing expression on at least one array. Filtering out only those that were not detected at all is the least stringent filtering available. A detect call was used to include a count of how many slides showed an expression value for each gene. This information was useful in analysing the data to see the robustness of the results for a particular gene.

Data Visualisation

Data visualisation is important to assess the success of the experiment, the quality of the data, and to choose an appropriate normalisation method. Plots comparing the arrays are useful to determine that all arrays performed in a similar way. A selection of graphs including MA plots, pairwise comparisons, box plots, density plots and hierarchical clustering were produced for all datasets.

Normalisation Methods

Normalisation was performed in R (Bioconductor) using the Limma [71] and Lumi [72] Bioconductor packages for microarray analysis. Limma is a package that was designed for the analysis of gene expression microarray data using linear models to assess differential expression. Lumi is a package that was designed especially for Illumina BeadArray Data. It includes algorithms specific to the unique aspects of the Illumina BeadArray system. In particular, the variance-stabilising transformation (VST) algorithm takes advantage of multiple technical replicates (on average 20–30 per array) unique to the Illumina BeadArray system [73]. It is therefore, better suited to the analysis of Illumina data than a base-2 logarithmic (\log_2) transformation, which is traditionally used in microarray analysis. Variance stabilisation is one of the primary reasons for \log_2 transformation of microarray data. Larger intensities tend to have larger variances when repeatedly measured and can cause a problem when data is subsequently analysed using ANOVA models. The VST algorithm utilises the within-array technical replicates to model the mean-variance relationship directly.

This allows the direct calculation of parameters for optimal data transformation from each array.

This was followed by quantile normalisation which was chosen to remove any remaining technical variability between arrays. Quantile normalisation [74] method transforms the distribution of intensities from one distribution to another.

Multiple Testing Correction

Due to the large number of genes being tested simultaneously it is important to perform a multiple testing correction to correct for the chance of obtaining false positives. Here, we have chosen the false discovery rate (FDR) to correct for false positives. The FDR method orders genes in increasing order of the p-values provided from the independent tests. The p-values are then compared to a threshold value related to the genes position in the list. The null hypothesis is rejected for genes that have a p-value less than their corresponding threshold. The FDR correction is not one of the most stringent and is a good choice because it does not assume all genes are independent.

2.3.10 Selection of Differentially Expressed Genes

Pairwise Comparison-Linear Modelling

Limma was used for pairwise comparisons between treatment groups using a linear model to compute p-values, which were adjusted using the Benjamini and Hochberg multiple testing correction. The derivation of the p-value reflects the degree of variance between biological replicates and thus, is a measure of the confidence in assigning significance of small fold changes in gene expression. Genes with a p-value < 0.05 were considered to be significantly differentially expressed.

Mixed Model ANOVA

The Leptin Reversal Study (Chapter 4) involved three separate treatments. To analyse the data effectively, it was necessary to find a method that could separate

the effects of each treatment and provide a separate gene list for each. ANOVA based methods are particularly suited to estimating variance in studies with several experimental steps. The appropriate model for this experiment is a Mixed Model ANOVA, which allows random and fixed effects and estimates the variability by variance components. Peter Sykacek, previously at University of Cambridge, wrote a software package for R called FSPMA: Friendly Statistical Package for Microarray Analysis [75] that relies on another mixed model ANOVA package for R, YASMA: Yet Another Statistical Microarray Analysis [76], which allows the precise specification of nested effects. FSPMA requires a balanced study design and is run by entering the specifics of the experimental design into a definition file specifically designed for one channel microarray experiments. In this particular experiment, the fixed effects were the three treatments (*preg*, the AD (normal) or UN (undernourished) diet during pregnancy; *lepsal*, the leptin (LEP) or saline (SAL) injection at days 3-13 of life; and *diet*, the normal (CHOW) or high-fat (HF) postweaning diet) and the random effect was the biological replicates for each treatment group. The technical replicates were taken into account during the normalisation. Once the definition file has been accurately completed, including the list of genes after filtering and the list of data files, the definition file is then loaded into R and produces output files that rank the genes that were found to be significant (p-value < 0.05). The p-value is then adjusted in R using the FDR multiple testing correction. (FSPMA lecture) The model used is:

$$y_{ijkl} = \mu + \alpha_i + \beta_{ij} + A_{ijk} + \epsilon_{ijkl}$$

where μ models the gene-specific global mean, α_i models the main gene effect, β_{ij} models the gene-time interaction, A_{ijk} models the gene-time-sample interaction and ϵ_{ijkl} models the residual error. The effects that correspond to gene and time are fixed. The effects corresponding to sample and replicate are random.

$$i = 1, \dots, n; (n = \text{number of genes})$$

$$j = 1, \dots, m; (m = \text{number of timepoints})$$

$$k = 1, \dots, a; (a = \text{number of samples})$$

$$l = 1, \dots, b; (b = \text{number of technical replicates})$$

2-way ANOVA with an Interaction Term

The 2-ways ANOVA considers the interaction between two of the effects. This model was also used for the Leptin Reversal Study (Chapter 4) to investigate the interaction between the different treatments.

$$X_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \epsilon_{ijk}$$

This equation is similar as to that of the Mixed Model ANOVA. Here the term $(\alpha\beta)_{ij}$ represents the interaction between the i -th factor α and the j -th factor β . This model was implemented using Limma.

Principle Component Analysis

Principle Component Analysis (PCA) is a multivariate analysis method used for exploratory analysis. It transforms a number of possibly correlated variables into a smaller number of uncorrelated variables called principle components. This was implemented using KDE.

2.3.11 Gene Ontology Analysis

Gene ontology analysis was done using DAVID: Database for Analysis, Visualization, Integrated Discovery [77], Ingenuity Pathway Analysis software (Ingenuity, Stanford USA) and Onto-Express [78].

2.4 Polymerase Chain Reaction

2.4.1 Genomic PCR

Genomic DNA PCR was carried out in a total volume of 20 μl containing 2 μl 10x buffer, 3.2 μl dNTPs (2.5 mM each), 2 μl DNA (100 ng/ μl), 2.4 μl primer mix (5mM each forward and reverse), 0.2 μl Hot Star Taq, and water up to 20 μl . PCR conditions were adjusted based on the melting temperatures of the primers used in a particular reaction.

2.4.2 Comparative Quantitative Real Time RT-PCR

The comparative expression levels of genes identified as differentially expressed from our microarray data were verified by qRT-PCR, together with the expression levels of several candidate genes not present on the array. Cyclophilin was used for the generation of a standard curve and normalisation of the RNA concentration. Each qRT-PCR was performed in triplicate. Real-time RT-PCR was performed in 96-well white plates (Abgene) using the Verso SYBR Green 2-Step qRT-PCR Fluorescein Kit (Thermo Scientific, UK) according to the manufacturer's protocols; the resulting fluorescence was quantified using an iCycler system (Bio-Rad). The Ct was obtained for each well as the cycle number at which the measured fluorescence crossed the arbitrary threshold of 150 units (all values are in log phase). The average Ct was calculated for each gene in each sample. Data were normalised to cyclophilin, with ΔCt calculated as follows: $\Delta\text{Ct} = \text{Ct}_{(test)} - \text{Ct}_{(cyclophilin)}$. $\Delta\Delta\text{Ct}$ values were then calculated as the change in ΔCt for the UN sample relative to the ΔCt value for the AD sample using the equation: $\Delta\Delta\text{Ct} = 2^{\text{Ct}_{(test)}} - 2^{\text{Ct}_{(cyclophilin)}}$. A melting curve was also performed to verify the specificity and identity of the RT-PCR product.

2.5 DNA Extraction

2.5.1 From Tissue (Undernutrition Study)

To extract DNA from tissue, 25 mg of frozen tissue was placed in separate white capped 10 ml tubes each containing 1 ml of homogenising buffer for each sample. The tissue was homogenised for 30 seconds. The homogeniser head was thoroughly cleaned between each sample. To pellet the nuclei, the homogenate was slowly centrifuged for 5 minutes at 500 g (2500 rpm). The pellet was then resuspended in 500 μ l homogenising buffer. If the pellet was not dissolving immediately a pipet-tip was used to break it up. To yield a final concentration of 10 mM EDTA, 10 μ l 0.5 M EDTA (pH 7.0) was added to the solution. The solution was mixed thoroughly and 20 mg/ml of proteinase K (5 μ l) was added to yield a final concentration of 100 μ g/ml. The solution was mixed thoroughly, and 10% SDS (51.5 μ l) was added to give a final concentration of 1% SDS. The solution was again mixed thoroughly but gently by inversion. The samples were incubated overnight at 55 °C. The next day 1 volume of 5 M ammonium acetate was added and the solution was mixed by inversion. The solution was chilled on ice for 30 minutes and then centrifuged at 12,000 x g for 3 minutes. The supernatant was transferred to a clean tube and 0.6a volume of isopropanol was added to precipitate the DNA. The solution was mixed by inversion. The samples were then placed in the freezer to aid in precipitation of DNA, if necessary. The precipitated DNA was carefully removed and washed by soaking in 70% EtOH for 10 minutes. The DNA was allowed to air dry briefly, but not completely, sometimes this took a long time. The DNA was dissolved in TE buffer (approximately 100–300 μ l). Finally, DNA samples were purified using the DNeasy Column purification kit (Qiagen, UK) following the manufacturer's instructions.

2.5.2 Following RNA extraction

For some samples, DNA was extracted from the frozen remnants (phenol and interphase mix) of the TRIzol RNA extraction protocol. This frozen phenol and

interphase mix was centrifuged for 15 minutes at 4 °C. It was important to make sure the aqueous phase had been thoroughly removed before adding 300 ml of 100% ethanol per 1 mL TRIzol Reagent originally used. This was mixed by inversion and stored at room temperature for 2–3 minutes. The solution was then centrifuged at no more than 2000 x g for 5 minutes at 4 °C. The phenol-ethanol supernatant was removed and saved for protein isolation (approx 800 mL per 1 mL TRIzol). The DNA pellet was washed twice in 0.1 M sodium citrate in 10% ethanol. For large pellets an additional wash was required. For each 1 mL of TRIzol Reagent originally used, 1 mL of the sodium citrate solution was used. At each wash, the DNA pellet was stored in the washing solution for 30 minutes at room temperature with periodic mixing. The solution was centrifuged at 2000 x g for 5 minutes at 4 °C. After two washes, the DNA was suspended in 75% ethanol (1.5–2 mL per 1 mL TRIzol Reagent). The solution was then left for 10–20 minutes at room temperature with periodic mixing before centrifuging at 2000 x g for 5 minutes at 4 °C. The DNA was then air-dried for 5 minutes in an open tube and then was dissolved in 8 mM NaOH such that the concentration of DNA was 0.2–0.3 $\mu\text{g}/\mu\text{l}$. For DNA isolated from 50–70 mg of tissue, 300–600 μl of 8 mM NaOH was added. The insoluble (gel-like) material (fragments of membranes, etc.) was separated by centrifugation at >12,000 x g for 10 minutes. The supernatant containing the DNA was transferred to a new tube. The A260 value of an aliquot of the DNA prep was measured in NaOH mixed with water. The expected yield of DNA per mg of liver tissue was 3–4 μg .

2.6 Methylation PCR

The HpaII-McrBC PCR [79] was used to discriminate patterns of methylation. In this method, HpaII is used to cut unmethylated alleles

(cutting at sites: 5'...C[^]CGG...3'/3'...GGC[^]G...5'),

while McrBC is used to cut methylated alleles

(cutting at sites: 5'...Pu^mC(N_{40–3000})Pu^mC...3').

MspI was used as a negative control as it cuts independent of methylation at HpaII

sites. PCR was performed using the HotStarTaq PCR kit and optimised using ‘touch-down’ techniques for individual primers. Products were separated by gel electrophoresis using 1–1.5 agarose with 0.5 mg/ml EtBr. The methylation changes were inferred from the intensity of the bands from different digests. The subsequent digestions of HpaII and McrBC gave four possible results. These results indicate full, null, incomplete, or composite methylation. This provides a general idea of where methylation sites are located. The exact location of the methylation sites can then be examined in more detail.

2.7 Bisulphite Sequencing

DNA samples were subjected to sodium bisulphite modification using EpiTect Bisulphite conversion kit (Qiagen, UK) and used as template in PCR.

Chapter 3

Expression Study of rats undernourished *in utero*

3.1 Introduction

The data presented in this chapter has resulted in a manuscript that was published in PLoS ONE ([80]). The results have all been discussed in this chapter and certain points have been examined in more detail for the purpose of this thesis.

In this chapter, an established rat model of balanced maternal undernutrition has been used for investigation of gene expression differences in target tissues (liver, retroperitoneal white adipose fat, and biceps femoris skeletal muscle) between offspring of control and undernourished mothers. Male rats are able to breed as early as 35 days and can live to nearly 550 days. We studied young male adult rats at day 55 when they have not yet developed the abnormal phenotype in order to identify any gene expression changes that may play a role in predisposing these animals to the development of the metabolic syndrome. The genes that showed significant gene expression differences in the liver suggest that the rats have suffered a substantial impairment of their ability to utilise carbohydrate. They may have lost the flexibility to switch between carbohydrate and fat as an energy source. The results suggest that maternal undernutrition leads to offspring that favour fat as an energy source thus, resulting in mitochondrial dysfunction. This chapter begins with a review of fuel

metabolism as a background for the discussion of the results. This is followed by a description of experimental design and data quality. Finally, the analysis is explained and discussed.

3.1.1 Energy Balance and Starvation

In this study, pregnant rats are subjected to severe undernutrition (30% of their normal diet). This is approaching a level of starvation. To understand the effects this might have on their offspring, it is important to understand the basic biochemistry of fuel metabolism (specifically lipid and carbohydrate metabolism) and the changes that occur during starvation.

Metabolic Homeostasis

Cells require a constant supply of fuel in order to create adenosine triphosphate (ATP) that is required for normal cell function and growth. A balance must be maintained between carbohydrate, fat, and protein intake; their storage when present in excess, and their mobilisation; and synthesis when in demand. This is metabolic homeostasis: the two major hormonal regulators are insulin and glucagon. Insulin is a major anabolic hormone that promotes the storage of fuels and the utilisation of fuels for growth. Glucagon is the major fuel mobilisation hormone. Other hormones, such as adrenaline respond to stress and can increase the availability of fuels. These hormones contribute to homeostasis by responding to changes in the levels of circulating fuels that are determined by diet. Figure 3.1 shows the major sites of insulin action on fuel metabolism in liver, adipocytes, and skeletal muscle and figure 3.2 shows the major sites of glucagon action in the same tissues.

Insulin and glucagon are very important in regulating blood glucose levels. After a high-carbohydrate meal, blood glucose rises quickly. After the meal is digested and absorbed, blood glucose levels decline because cells continue to metabolise glucose. If blood glucose continued to rise after a meal, the high level of blood glucose would cause the release of water from tissues including the brain. However, if blood glucose

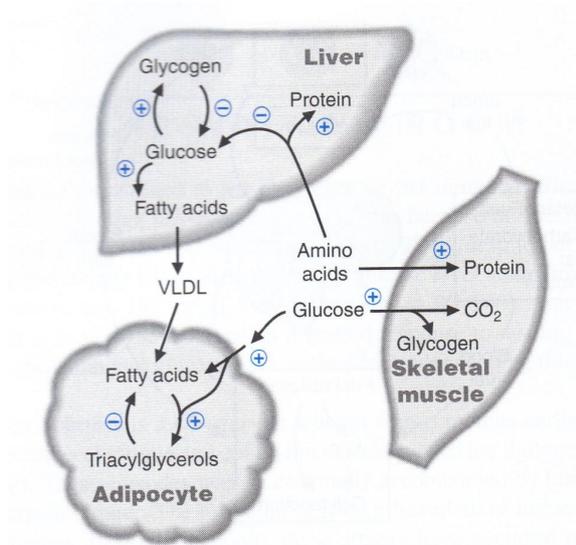


Figure 3.1: This figure shows the major sites of insulin action in fuel metabolism. The plus sign represents pathways that are stimulated by insulin; the minus sign represents pathways that are inhibited by insulin. This figure was adapted from Lieberman *et al.* [81]

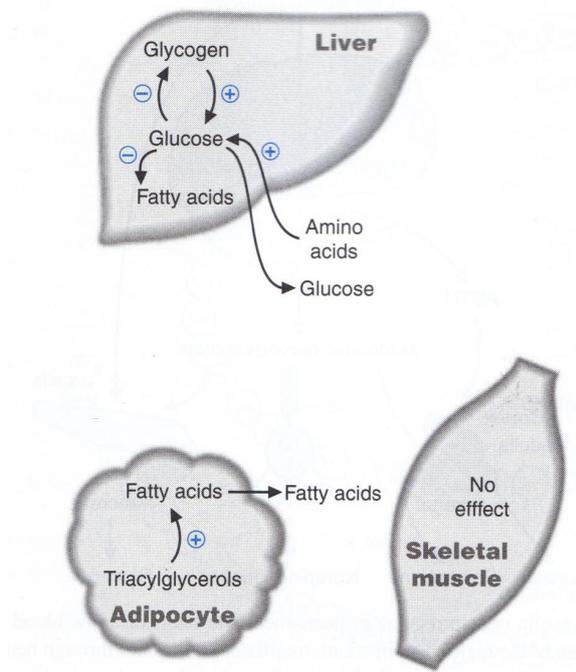


Figure 3.2: This figure shows the major sites of glucagon action in fuel metabolism. The plus sign represents pathways that are stimulated by glucagon; the minus sign represents pathways that are inhibited by glucagon. This figure was adapted from Lieberman *et al.* [81]

levels drop after a meal, glucose-dependent tissues would suffer from a lack of energy. Insulin helps to regulate these levels so that even after 5-6 weeks of starvation blood glucose levels do not decrease dramatically [81].

Metabolic Fasting

In the initial stages of fasting, stored fuels are used for energy. This state is called the basal state and is illustrated in figure 3.3. During fasting, fatty acids released from adipose tissue serve as the body's major fuel. The liver oxidizes most of its fatty acids only partially, converting them to ketone bodies, which are released into the blood. Muscle and many other tissues are able to use fatty acids or ketone bodies; however, red blood cells, the brain, and other neural tissues use mainly glucose. The liver is the organ that maintains blood glucose levels during fasting, and therefore, its role in survival is crucial. When blood glucose levels drop, the liver replenishes blood glucose via gluconeogenesis. Figure 3.4 illustrates the key reactions of gluconeogenesis. During this process, lactate, glycerol, and amino acids (particularly alanine) can be used as carbon sources to synthesize glucose. Most of the amino acid is supplied by degraded muscle protein. The nitrogen of the amino acids can form ammonia, because this is toxic, the liver converts it to urea [81].

Metabolic Starvation

The hormonal changes that occur during fasting stimulate the breakdown of adipose triacylglycerols. This results in the release of fatty acids and glycerol into the blood. Figure 3.5 shows the tissue interrelationships during fasting. Glycerol is the major source of carbon for gluconeogenesis. Fatty acids become the major fuel for the body and are oxidized by various tissues, which enable these tissues to reduce their consumption of glucose. Fatty acids are also oxidized to acetyl CoA in the liver to provide energy for gluconeogenesis. Acetyl CoA can be converted to ketone bodies. If this pattern of fuel utilisation was prolonged, the body's protein would be rapidly consumed. Metabolic changes occur during prolonged fasting (starvation) to conserve muscle protein. The starvation state is illustrated in figure 3.6. After 3-5

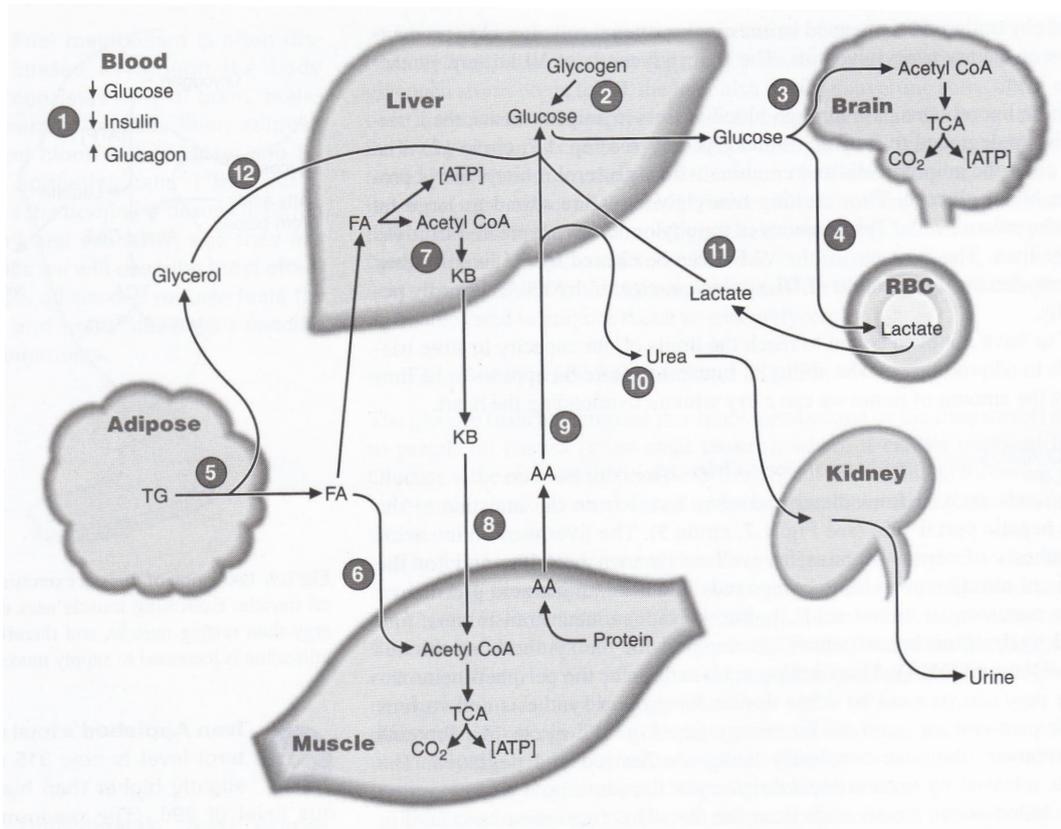


Figure 3.3: This figure represents the basal state. This state occurs after an overnight (12-hour) fast. The circled numbers serve as a guide indicating the approximate order in which the processes begin to occur. KB=ketone bodies; TG=triacylglycerols;FA=fatty acid; AA=amino acid; RBC=red blood cell. This figure was adapted from Lieberman *et al.* [81]

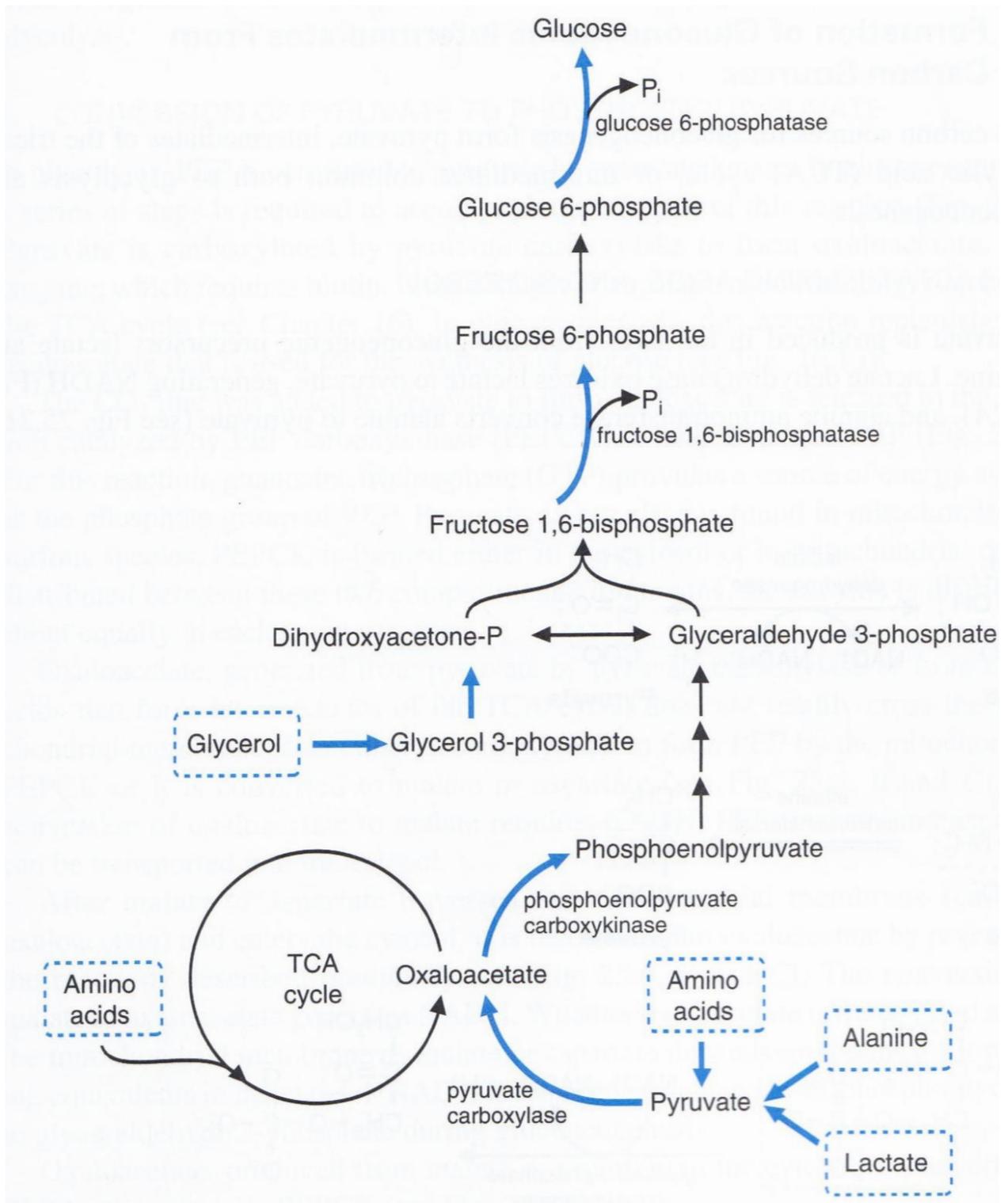


Figure 3.4: This figure depicts the key reactions that occur during gluconeogenesis. The precursors are amino acids (particularly alanine), lactate, and glycerol. Heavy arrows indicate steps that differ from those of glycolysis. The conversion of phosphoenolpyruvate to glyceraldehyde 3-phosphate requires the reversible enzymes of glycolysis. This figure was adapted from Lieberman *et al.* [81]

days of fasting the muscle decreases its use of ketone bodies and depends mainly on fatty acids for fuel, while the liver continues to convert fatty acids to ketone bodies. The result is that the concentration of ketone bodies in the blood rises. The brain begins to take the ketone bodies from the blood and oxidise them for energy. Gluconeogenesis in the liver is now the only process by which the liver can supply glucose to the blood if fasting continues. However, due to the ketone body utilisation the rate of gluconeogenesis can decrease, in turn decreasing the amount of protein degraded to supply amino acids to gluconeogenesis. As the use of ketone bodies increases, the body needs roughly one-third as much glucose. As a result of reduced glucose utilisation, the rate of gluconeogenesis in the liver decreases, and the production of urea go down. For this reason, urea can be used as a measure of the extent of protein degradation. Proteins function as enzymes, as structural proteins, and in muscle contraction. If tissue protein is severely degraded then body function can be compromised. If starvation continues and no other problems occur the individual can die of severe protein loss that causes organs to malfunction. For this reason, the increase in ketone body levels that result in the conservation of body proteins allows individuals to survive starvation for extended periods.

As fatty acids are released from the adipose tissue during fasting, they travel in the blood complexed with albumin. These fatty acids are oxidized by various tissues particularly muscle. In the liver, fatty acids are transported into mitochondria because acetyl CoA carboxylase is inactive, malonyl CoA levels are low, and carnitine pamlitoyltransferase I is active. Acetyl CoA, produced by β -oxidation, is converted to ketone bodies.

Our energy expenditure is equivalent to our oxygen consumption. The majority of our oxygen consumption (90-95%) is used for the synthesis of ATPases in the electron transport chain by oxidative phosphorylation. Most of the enzymes for the Citric Acid (TCA) cycle, electron transport, and other pathways for oxidation are located in the mitochondrial matrix. Fatty acid oxidation is the process in which very long chain fatty acids are oxidized by peroxisomal β and α -oxidation pathways, which are essentially chain-shortening pathways. Ketone bodies are synthesised in the mitochondrial

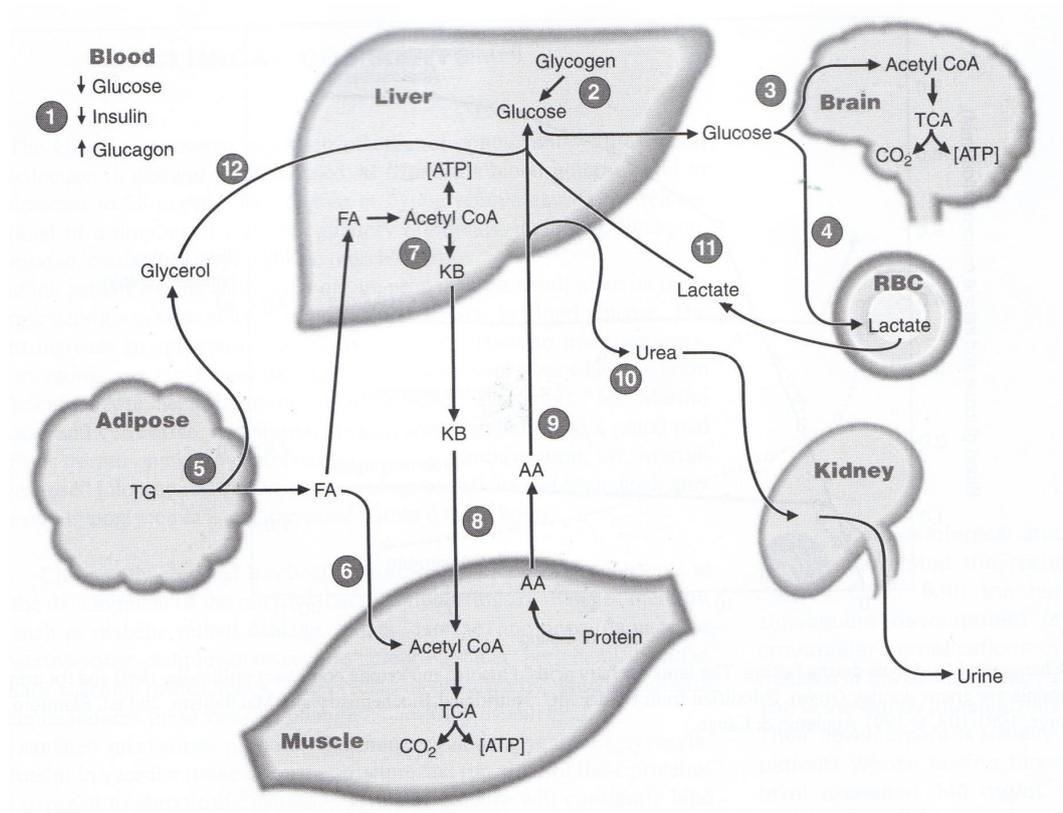


Figure 3.5: This figure illustrates the tissue interrelationships during fasting. (1) Blood glucose levels decrease, decreasing insulin, and raising blood sugar levels. (2) Glycogenolysis is induced in the liver to raise blood glucose levels. (3,4) The brain and red blood cells use the glucose released by the liver. (5) Adipose tissues are signalled to release free fatty acids and glycerol from stored triglycerides. (6) The muscle and liver use fatty acids for energy. (7,8) The liver converts fatty acid derived acetyl CoA to ketone bodies for export, which the muscles and brain can use for energy. (9) Protein turnover is induced in muscle, and amino acids leave the muscle and travel to the liver for use as gluconeogenic precursors. (10) The high rate of amino acid metabolism in the liver generates urea, which travels to the kidney for excretion. (11) Red blood cells produce lactate, which returns to the liver as a substrate for gluconeogenesis. (12) The glycerol released from adipose tissue is used by the liver set gluconeogenesis. KB=ketone bodies; FA=fatty acid; AA=amino acid; TG=triacylglycerols. This figure was adapted from Lieberman *et al.* [81]

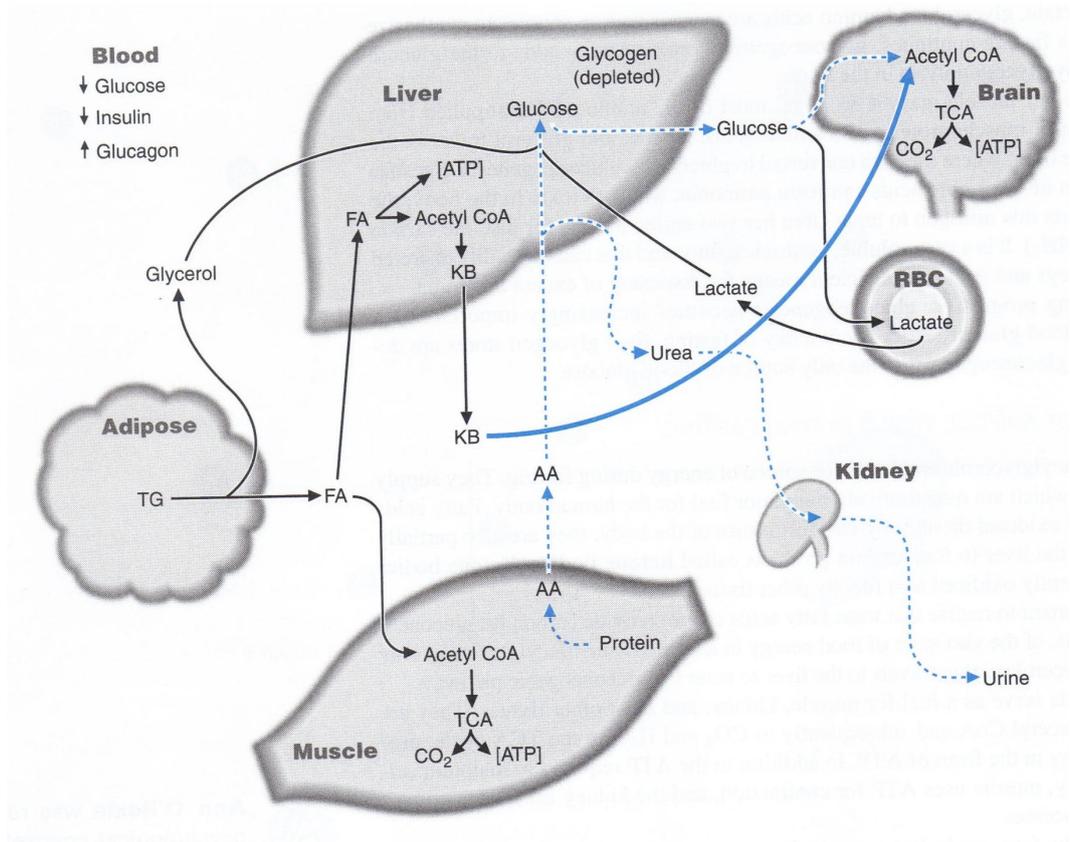


Figure 3.6: This figure represents the starved state. Broken lines indicate processes that have decreased, and the heavy solid line indicates a process that has increased relative to the fasting state. KB=ketone bodies; FA=fatty acid; AA=amino acid; TG=triacylglycerols; RBC=red blood cell. This figure was adapted from Lieberman *et al.* [81]

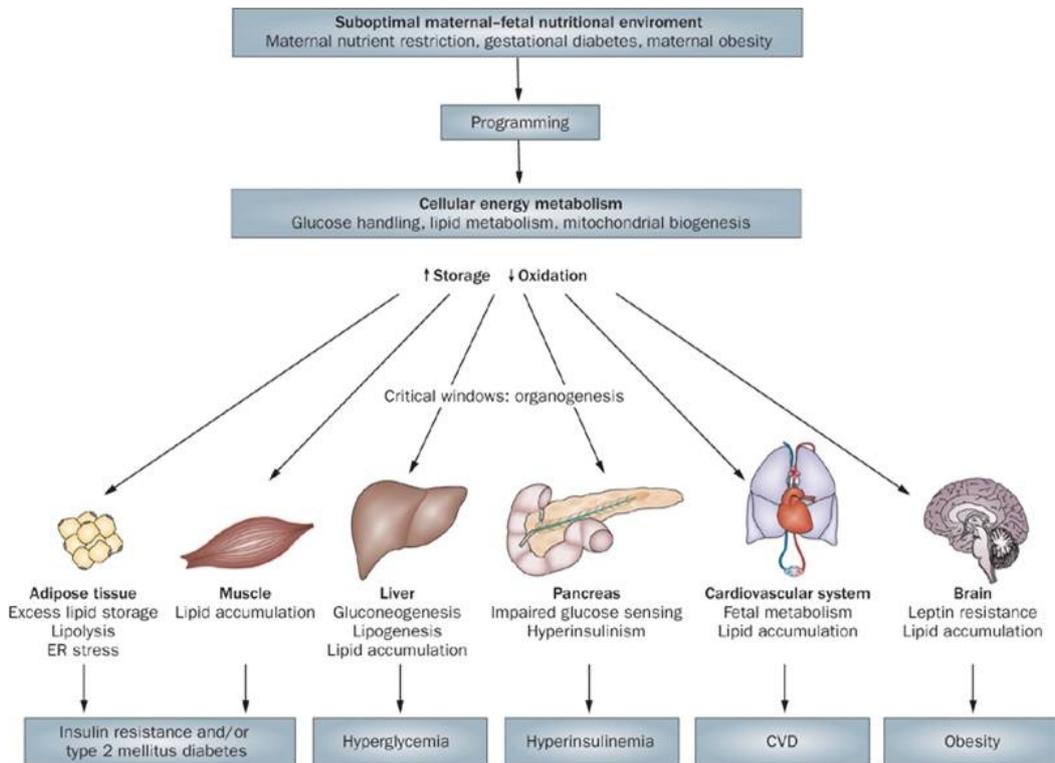


Figure 3.7: This figure from Symonds *et al.* [64] illustrates the importance of the maternal-fetal nutritional environment in contributing to the developmental programming of cellular energy metabolism in favour of lipid storage.

matrix from acetyl CoA generated from fatty acid oxidation. These examples just briefly highlight the role of mitochondria in metabolism. Mitochondrial dysfunction could significantly impact many different aspects of metabolism and energy balance [81].

In the study presented in this chapter liver, muscle, and fat were used for molecular analysis based on their importance in metabolism as highlighted in the preceding biochemical background.

3.1.2 Genes/Pathways Highlighted in Other Studies

Several candidate genes have been highlighted as showing altered expression levels in animal studies using restricted diets. Lillycrop 2005 showed that maternal protein restriction during pregnancy in rats increased glucocorticoid receptor (*GR*) expression and decreased *11β-hsd2* expression and altered the expression of *Ppara*

and acetyl-CoA oxidase (*Aox*) [82]. These genes are of particular interest because changes in their expression levels have been linked with disturbances in cardiovascular and metabolic control in humans and animals. *GR* activity is important for the regulation of blood pressure, *Ppara* is central to lipid and carbohydrate homeostasis. *Aox* is the rate-limiting factor in the peroxisomal β -oxidation pathway and is directly regulated by *Ppara*. *Pepck* is the rate-limiting factor in gluconeogenesis, whereby glucose is synthesised. Studies with mice have shown that over expression of *Pepck* can lead to diabetes mellitus type 2.

3.1.3 Initial Study

This study exploited a well defined model of developmental programming via maternal undernutrition [45, 83]. Virgin Wistar rats (age, 100 ± 5 days) were time mated using a rat oestrous cycle monitor to assess the stage of oestrous of the animals before introducing the male. After confirmation of mating, rats were housed individually in standard rat cages with free access to water. All rats were kept in the same room with constant temperature maintained at 25°C and a 12 h light/12 h darkness cycle. Animals were assigned to one of two nutritional groups: a) undernutrition (30% of ad libitum) of a standard chow diet throughout gestation (UN group), b) standard chow diet ad libitum throughout gestation (AD group). The study design is illustrated in Figure 3.8. Food intake and maternal weights were recorded daily until the end of pregnancy. After birth, pups were weighed, and litter size was adjusted to 8 pups per litter to assure adequate and standardised nutrition until weaning. A minimum of four litters per group were used. Pups from undernourished mothers were cross-fostered onto dams that had received AD feeding throughout pregnancy. The animals were culled at day 55 following an overnight fast to represent an age known to precede the development of increased adiposity and altered insulin sensitivity in UN offspring. Tissues (liver, retroperitoneal white adipose, and biceps femoris skeletal muscle) were collected as described in Chapter 2. A subcohort of animals ($n=8$ per group) were maintained until postnatal day 110 to represent an age where



Figure 3.8: Study Design of Rats Undernourished in Utero. This figure summarises the study design. During pregnancy rats were fed either ad libitum a standard chow diet (AD) or 30% of a standard chow diet (UN). Pups from UN mothers were cross-fostered onto dams that had received AD feeding throughout pregnancy.

phenotypic changes in growth and metabolism between AD and UN offspring have previously been reported [45, 58].

3.2 Results

3.2.1 Previous Findings: Phenotypic Assessment

Physiological measurements were taken by the collaborators at the Liggins Institute as part of the initial unpublished study. These findings are summarised here. Serial dual-energy x-ray absorptiometry (DEXA) analysis was performed on the animals at day 55 to establish body composition and assessment of bone parameters. In addition to DEXA, body composition was analysed via standard techniques (fat depot dissection and weighing) following post-mortem. Body weight and food and water intake were monitored. Plasma analysis incorporated these potential indicators of metabolic syndrome: IGF-1, insulin, C-peptide (a measure of insulin secretion), and lipid profiles. For comparison to an animal that had developed the phenotype, we have included the phenotypic assessment of male rats (AD and UN) at day 110. These animals were from a different cohort, and were given a saline injection ($2.5 \mu\text{g/g}\cdot\text{d}$) from day 3-13, but received the same nutritional treatment. Statistical analyses were performed using KDE software. Differences between groups were determined by a T-test for difference in variance with an FDR correction, and data are shown as an adjusted p-value \pm standard error (tables 3.1 and 3.2).

Measurement	AD Mean	UN mean	Adjusted p-value
Body weight (grams)	282.4 ± 5.6	234.5 ± 5.5	0.001
Body length (mm)	203.6 ± 1.7	192.63 ± 1.9	0.005
Liver weight (% body weight)	3.7 ± 0.1	3.0 ± 0.1	0.014
Spleen weight (% body weight)	0.32 ± 0.01	0.33 ± 0.02	NS
Heart weight (% body weight)	0.36 ± 0.01	0.38 ± 0.01	NS
Fat pad weight (% body weight)	0.80 ± 0.06	0.59 ± 0.04	0.039
Plasma glucose (mmol/l)	6.24 ± 0.28	6.67 ± 0.48	NS
Urea (mmol/l)	4.76 ± 0.24	6.18 ± 0.29	0.010
Free Fatty Acids (mmol/l)	0.79 ± 0.08	0.68 ± 0.06	NS
Glycerol (mmol/l)	0.21 ± 0.03	0.19 ± 0.01	NS
Total Protein (g/DL)	5.45 ± 0.12	5.39 ± 0.07	NS
Lipase (U/l)	9.10 ± 0.97	9.06 ± 1.19	NS
C-Peptide (pg/ml)	132.18 ± 15.63	162.85 ± 24.72	NS
Triglycerides (mmol/l)	0.78 ± 0.1	0.77 ± 0.06	NS
IGF-1 (ng/ml)	1221.50 ± 55.06	1277.63 ± 63.72	NS
Creatinine (mmol/l)	20.15 ± 0.49	23.64 ± 1.36	0.014
Insulin (ng/ml)	0.33 ± 0.06	0.53 ± 0.07	0.017
Total Fat (%)	24.09 ± 0.95	23.66 ± 0.47	NS
LDL (mmol/l)	0.27 ± 0.03	0.46 ± 0.04	0.006
HDL (mmol/l)	1.27 ± 0.06	1.34 ± 0.07	NS
LDL:HDL ratio	0.21 ± 0.02	0.35 ± 0.02	0.003

Table 3.1: Phenotypic measurements relevant to metabolic syndrome measured for each of the eight animals in each treatment group at day 55 and shown as mean ± SEM. The p-value was calculated with a t-test and an FDR correction. NS = not significant.

Measurement	AD Mean	UN mean	Adjusted p-value
Body weight (grams)	511 \pm 23.30	489 \pm 24.45	NS
Body length (mm)	238 \pm 2.52	218 \pm 2.61	0.05
Total Fat (%)	27.81 \pm 1.64	35.89 \pm 2.51	0.05
Fat pad weight (% body weight)	1.45 \pm 0.14	2.14 \pm 0.20	0.05
Leptin (ng/ml)	9.63 \pm 1.56	22.53 \pm 5.49	0.005
Insulin (ng/ml)	0.26 \pm 0.08	0.50 \pm 0.08	0.001
C-Peptide (pg/ml)	466 \pm 108.6	729 \pm 96.8	0.05
Plasma glucose (mmol/l)	7.0 \pm 0.44	7.2 \pm 0.24	NS
Free Fatty Acids (mmol/l)	0.86 \pm \pm 0.10	0.99 \pm 0.08	NS
Liver weight (% body weight)	2.92 \pm 0.07	2.78 \pm 0.07	NS
IGF-1 (ng/ml)	1236 \pm 70.87	1145 \pm 72.69	NS

Table 3.2: Phenotypic measurements relevant to metabolic syndrome measured for each of the eight animals in each treatment group at day 110 and shown as the mean \pm SEM. The p-value was calculated with a t-test and an FDR correction. NS = not significant.

3.2.2 Microarray

Tissue samples were shipped to the University of Cambridge on dry ice for RNA extraction. Eight biological replicates were used for each of the two treatment groups (AD and UN) and for each of three tissues (liver, skeletal muscle, and white adipose tissue).

MEEBO

The liver samples were hybridised to the MEEBO chip as described in Chapter 2. The experimental design used a reference design with a dye swap as shown in Figure 3.9. Each MEEBO slide contained one UN individual hybridised against the reference. The reference was a control pool of all eight AD samples. Only four of the UN samples were hybridised. A technical replicate was done for each of the four biological replicates, in addition to two dye swaps. So a total of 16 MEEBO slides were hybridised, four for each of the four UN samples. RNA extraction, amplifications, and microarray hybridisations were done as described in Chapter 2 and were randomised to avoid sample bias.

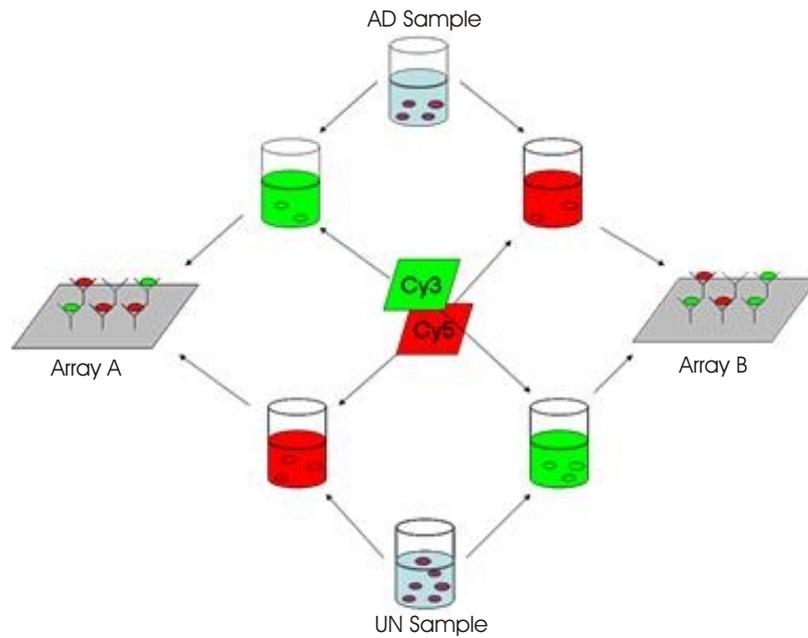


Figure 3.9: Design of the Dye Swap Experiment for the Samples Hybridised on the MEEBO Array. Control (AD) and treatment (UN) samples were labelled with cy3 or cy5 dye and hybridised to MEEBO arrays. Samples were then labelled with alternate dyes and again hybridised to arrays. These technical replicates allowed experimental control over dye bias.

Data Quality Control

It is important to first check the quality of the hybridisations (described in Chapter 1). This can be done using a variety of visualisation tools. Figure 3.10 is a density plot of the signal intensity against the density for each sample in each channel. This is useful for visualising huge differences in intensity between samples or even between dyes. Figures 3.11 and 3.12 show box plots of the amplitudes of intensities for each dye separately across all samples. Both of these methods have revealed that although there is variation in intensity across the samples, there does not seem to be a dye bias.

Further quality control assessments can be made by comparing replicates within treatment groups. This was determined using M-A plots which are shown for all the samples in Figure 3.13. For most of the samples, the plot shows a funnel shape indicating that there is higher variance at the lower intensities.

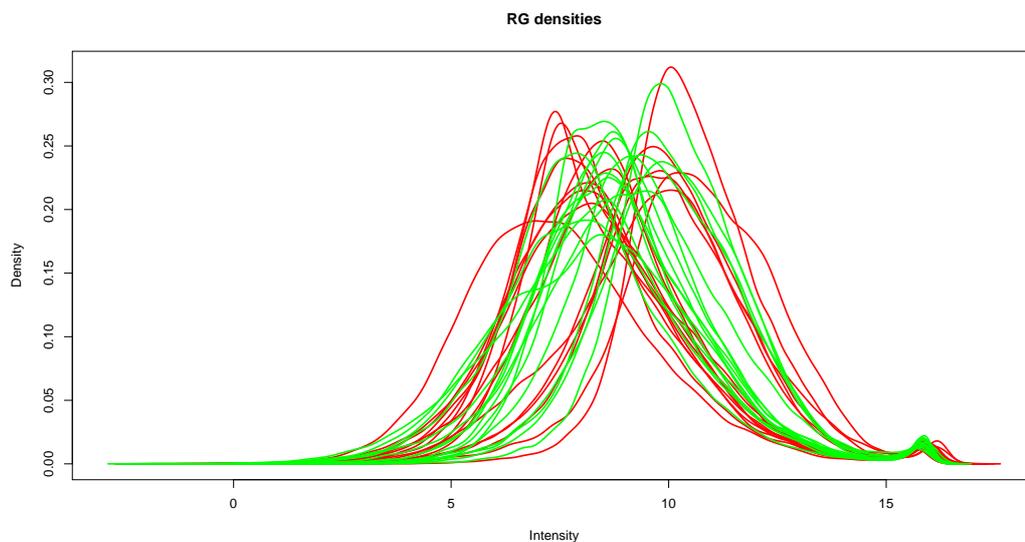


Figure 3.10: This is a density plot of all MEEBO samples in each channel before normalisation. This figure shows the density of intensities for the red and green channels on each array. There is clearly variation in intensity, which can be corrected with normalisation; however, there is no dye bias.

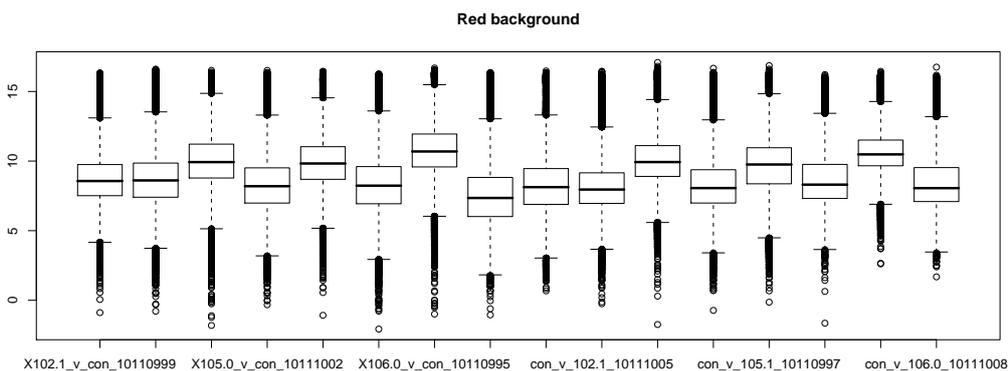


Figure 3.11: This is a boxplot showing the amplitude of the intensities for each sample in the red channel.

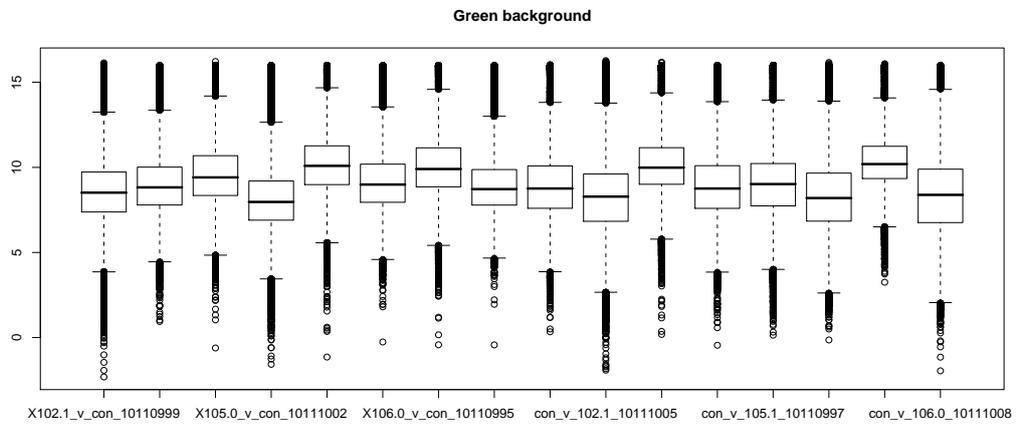


Figure 3.12: This is a boxplot showing the amplitude of intensities for each sample in the green channel.

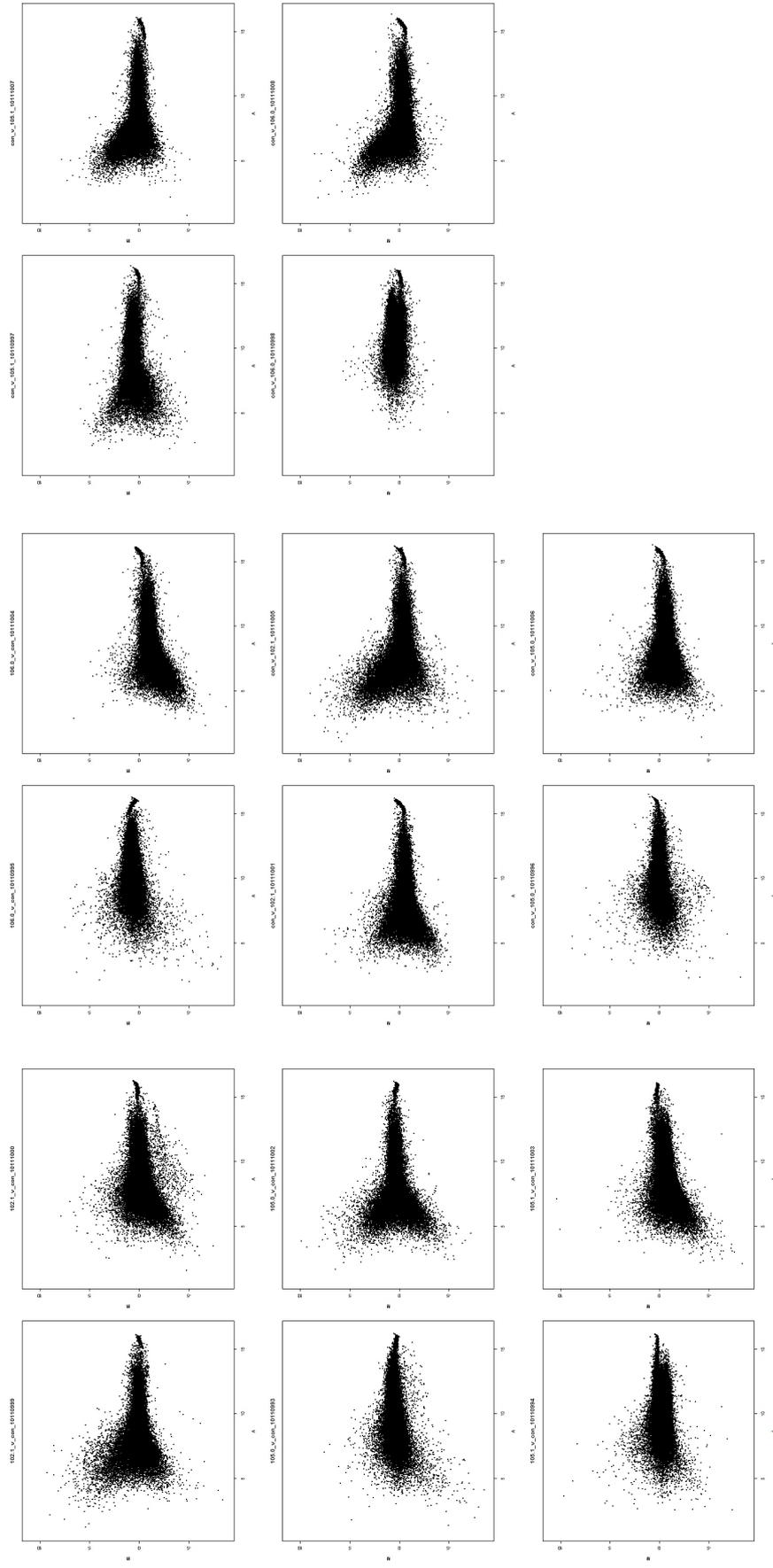


Figure 3.13: This shows an MA plot before normalisation for each of the 16 MEEBO arrays. The shape of the plot shows that the majority of the genes have low expression levels. This is normal, as you would expect only a small number to be highly expressed.

Normalisation/Differential Expression

Two different analyses of the data were performed. The initial analysis was done using the R library FSPMA [75] for normalisation and to calculate p-values; FSPMA is based on the mixed model ANOVA library YASMA [76] and is described in more detail in Chapter 2. FSPMA averaged technical and biological replicates and normalised the data using a loess normalisation. The analysis revealed 75 upregulated genes (p-value < 0.001) and 100 significantly downregulated genes (p-value < 0.001) (see Appendix C, Table 1).

The second analysis utilised the R library Limma. The data was normalised using a loess within-array background correction and a quantile between-array normalisation. The differentially expressed genes were determined by pairwise comparison with a FDR multiple testing correction to calculate p-values. This analysis revealed 956/38976 significant genes (see Appendix C, Table 2).

For both of these analyses the significant genes largely had A-flags with a few B-flags which confirm that the data is reliable. This chip included several candidate genes that are of interest in the literature including (*Ppargc1b*, *Ppard*, *Pparc1a*, *Pparg*, *Pparbp*, *Ppara*, *Acox1*, *Nr3c1*, and *Hsd11b2*). Although all of these genes had A-flags indicating successful hybridisation, none of them had significant p-values for differential expression.

Illumina

After the initial hybridisation on the MEEBO chip, the Illumina platform was acquired so the study was repeated and expanded using this technology. Subsequent and more detailed analysis have focused on this dataset. Each Rat-Ref12 Illumina chip has twelve arrays. As it is a single colour system there is no need for a dye swap. In addition, as samples are hybridised independently AD samples were individually hybridised rather than combined into a pool. In total, 4 Illumina slides (48 arrays) were hybridised for the three tissue samples and two treatment groups. Figure 3.14 illustrates the experimental design. RNA was extracted from liver, muscle, and fat

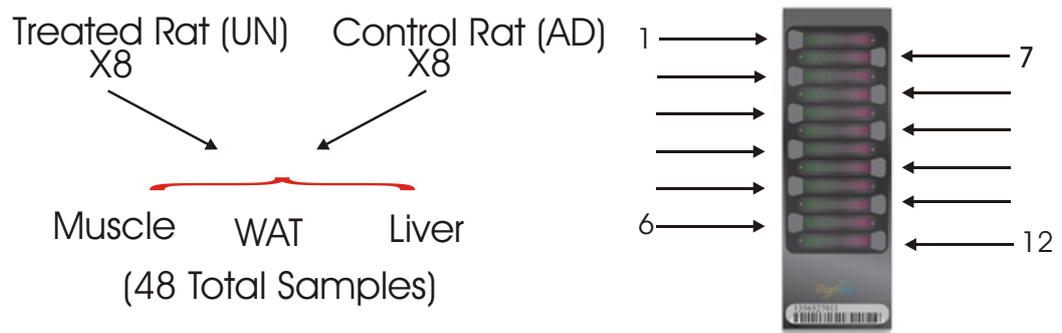


Figure 3.14: Microarray Design. This study included two treatment groups and three different tissues. There were eight biological replicates for each treatment group. RNA was extracted from the liver, muscle, and fat of each animal. The samples were hybridised independently on Illumina RatRef12 chips. Each chip has 12 arrays; the 48 samples filled 4 Illumina slides.

tissue and hybridised to Illumina microarrays as described in Chapter 2.

Data Quality Control

The Illumina system includes several internal controls for estimating hybridisation quality. Figure 3.15A-F shows six graphs of the quality control data for the liver, muscle, and fat samples. Figure 3.15A compares the low, medium, and high intensities across arrays. Ideally, arrays should have signal intensities in the same range. It is expected that there be a linear increase of signal between the three intensities. In figure 3.15A, there is a slight dip at medium intensity. To investigate this, it is necessary to look at the individual data points for the medium intensity signals. Figure 3.15B shows the detailed data for each array for the medium intensity signals. This shows that all samples are not completely consistent. The medium intensity signals, for the liver tissue samples, all fall in the tight range of 14,000-16,000 while the muscle (12,000-22,000) and WAT (6,000-18,000) are much more spread out. Figure 3.15C shows the high stringency of the data. For high quality data, it is expected that the biotin signal would be 3–4 fold lower than the high stringency signal, as shown for this data. Figure 3.15D shows the low stringency of the data, by comparing the signal intensity of two mismatches to the signal intensity of a perfect match. The perfect matches should have a 3–4 fold higher signal, as shown for this data. No

signal intensity indicates the hybridisation failed. Figure 3.15E shows the data for the negative control. The background signal should be approximately 70. This value is accounted for in the normalisation algorithms. The noise signal should be very low for high quality data. Figure 3.15F compares the signal from the housekeeping genes to the signal from all genes. As housekeeping genes are always expressed they should have a much higher signal compared with all genes. This shows that this data was generally high quality although a few of the muscle and WAT arrays had intensities slightly higher or lower than most of the arrays.

To compare the samples in more detail the data was exported from BeadStudio and imported into R Bioconductor. Here the raw data was plotted in several graphs (pairwise, MA, density plot, boxplot), then normalised and re-plotted to compare pre and post normalisation. In addition, the arrays were clustered using hierarchical clustering. The hierarchical clustering in figure 3.16 shows that with few exceptions the data clusters based on tissue type. This indicates there are more differences between tissues than between treatment groups. In addition, it is a good quality control check to confirm that samples do not cluster based on day of experiment or other factors. Figures 3.17 to 3.22 show pairwise comparisons for each of the two treatment groups for each of the three tissues. A tight diagonal line is expected for replicates, as this indicates low variability between samples. Figure 3.20 and 3.22 reveal some data that do not look good. The UN-muscle samples 101.1 and 105.0 and the UN-fat sample 105.0 show much more scatter than the others. Taking another look at figure 3.16 reveals that two of these samples are the ones that did not cluster in the same way as the other samples.

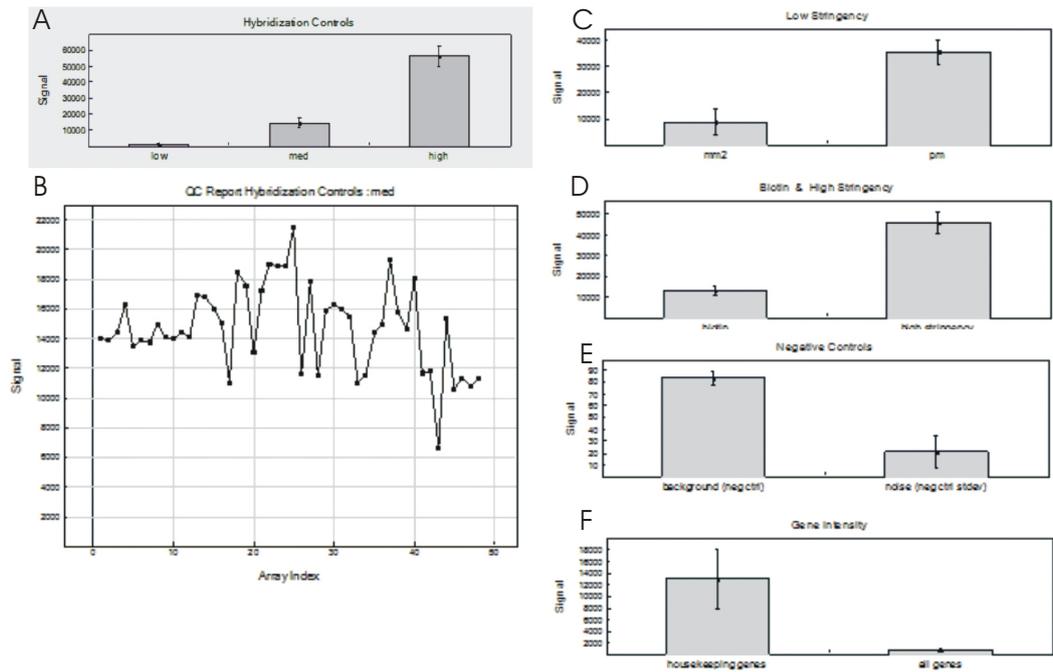


Figure 3.15: Illumina Internal Controls. This figure shows the details of the internal quality control tools for the Illumina platform. (A) Low, medium and high intensities across arrays. Ideally, arrays should have signal intensities in the same range. It is expected that there be a linear increase of signal between the three intensities. In this data, there is a slight dip at medium intensity. (B) Detailed data for each array, for the medium intensity signals. Individual array data is available for all internal control measures, this one is being shown as an example. It shows that all samples are not completely consistent. The medium intensity signals for the liver tissue samples all fall in the tight range of 14,000-16,000 while the muscle (12,000-22,000) and WAT (6,000-18,000) are much more spread out. (C) High stringency of the data. For high quality data, it is expected that the biotin signal would be 3–4 fold lower than the high stringency signal as shown for this data. (D) Low stringency of the data by comparing the signal intensity of two mismatches to the signal intensity of a perfect match. The perfect matches should have a 3–4 fold higher signal as shown for this data. No signal intensity indicates the hybridisation failed. (E) Data for the negative control. The background signal should be approximately 70. This value is accounted for in the normalisation algorithms. The noise signal should be very low for high quality data. (F) Comparison of the signal from the housekeeping genes to the signal from all genes. As housekeeping genes are always expressed, they should have a much higher signal compared with all genes. This shows that this data was generally high quality although a few of the muscle and WAT arrays had intensities slightly higher or lower than most of the arrays.

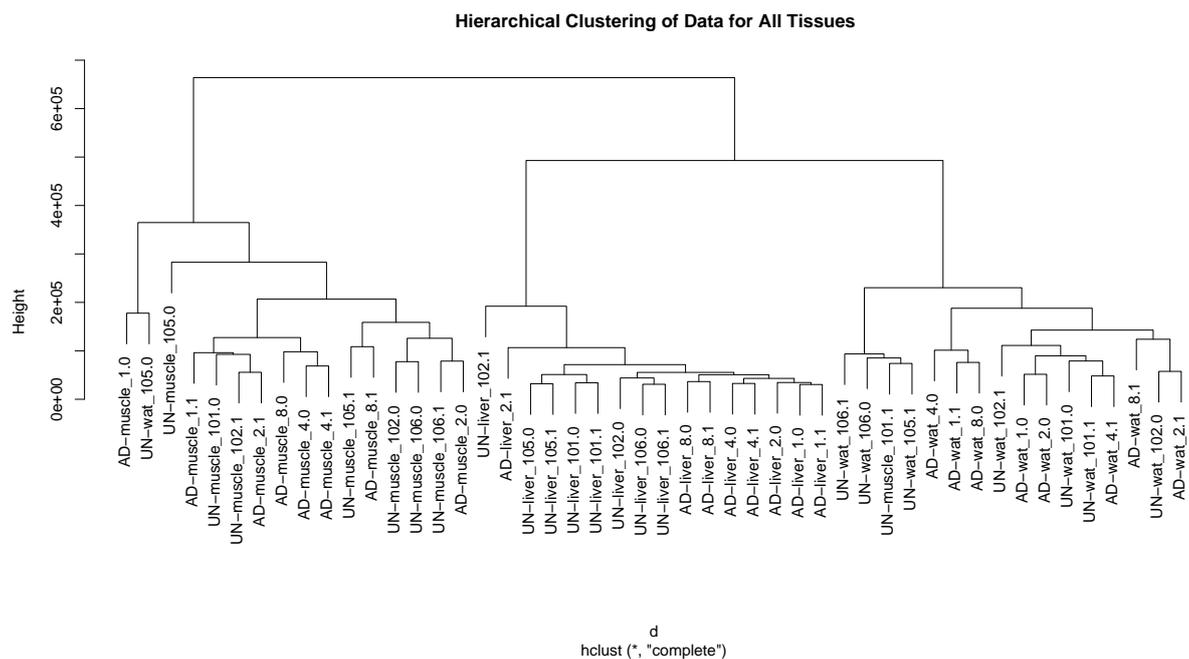


Figure 3.16: Hierarchical Clustering of All Tissues Before Normalisation. We would expect the biological replicates to cluster together. Here samples have clustered based on tissue type. The samples that have clustered out of place (UN-muscle sample 101.1 and the UN-WAT sample 105.0) are also two of the samples that show a larger variability in the pairwise plots in figure 3.20 and 3.22

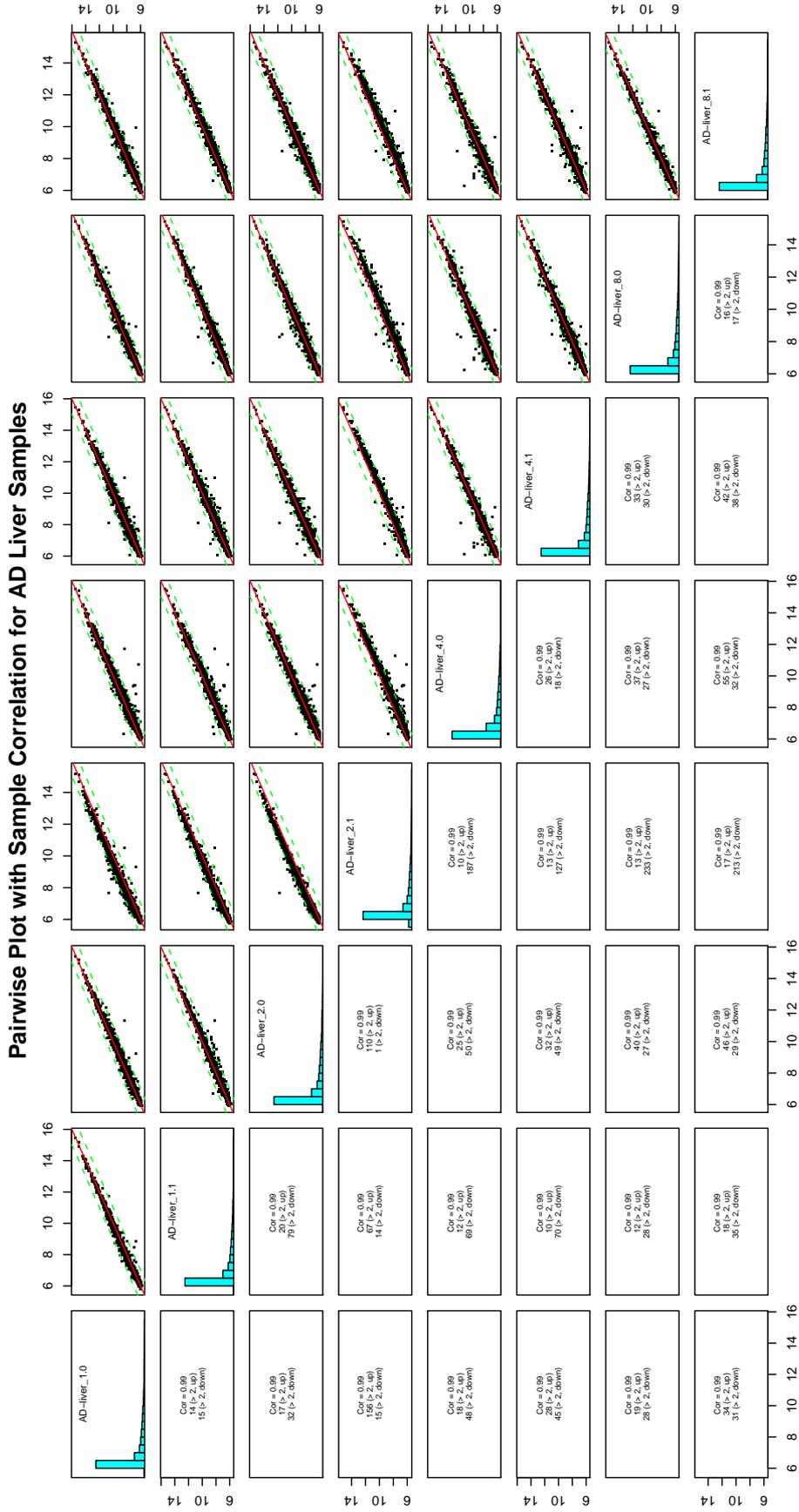


Figure 3.17: Pairwise Comparison of AD Liver Samples Before Normalisation. This plot compares all eight AD liver biological replicates against each other. These replicates should be very similar and should be tightly distributed on the x=y axis, as these samples are. The correlation for all comparisons ≥ 0.99 .

Pairwise Plot with Sample Correlation for UN Liver Samples

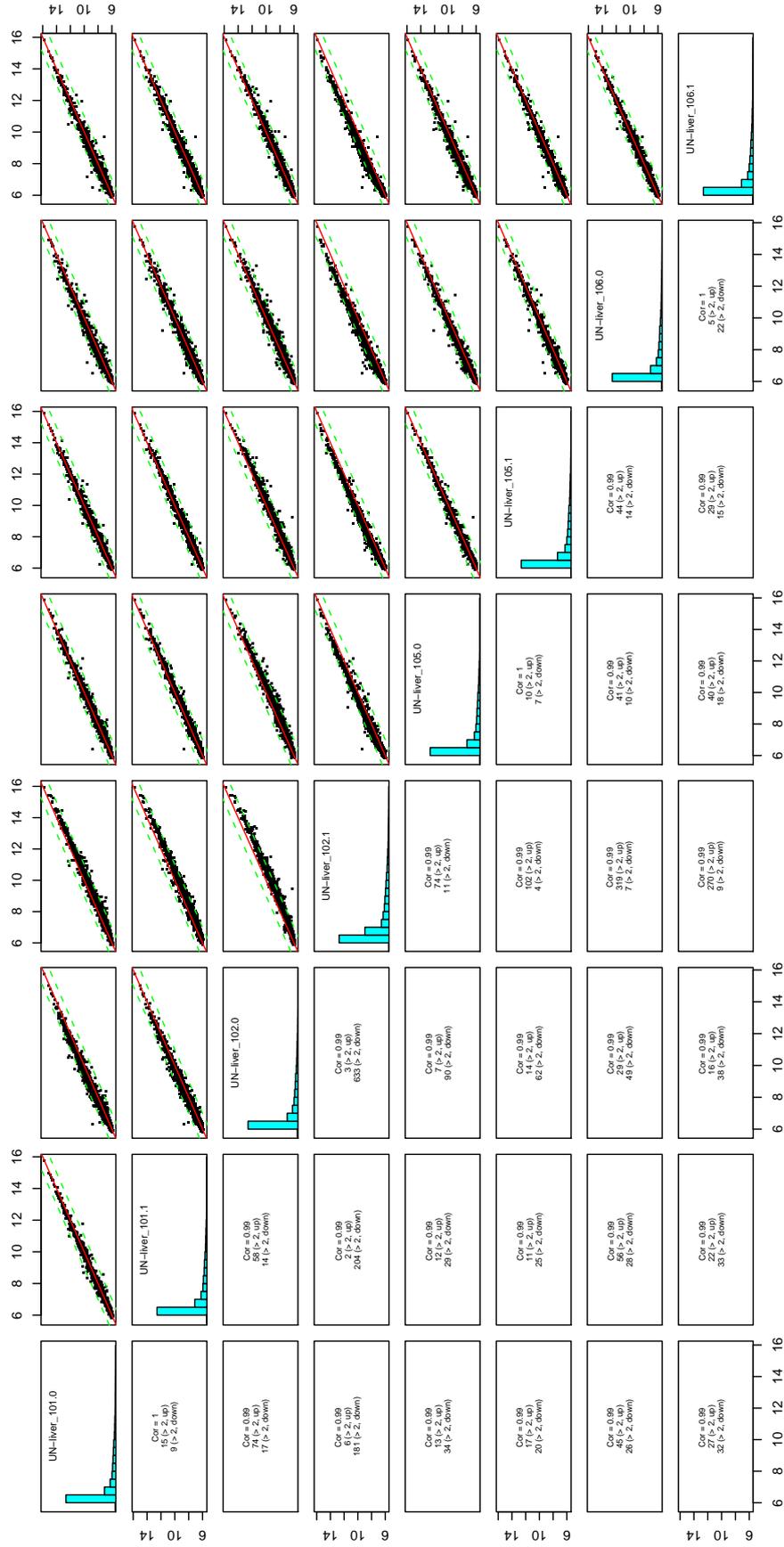


Figure 3.18: Pairwise Comparison of UN Liver Samples Before Normalisation. This plot compares all eight UN liver biological replicates against each other. These replicates should be very similar and should be tightly distributed on the x=y axis, as these samples are. The correlation for all comparisons ≥ 0.99

Pairwise Plot with Sample Correlation for AD Muscle Samples

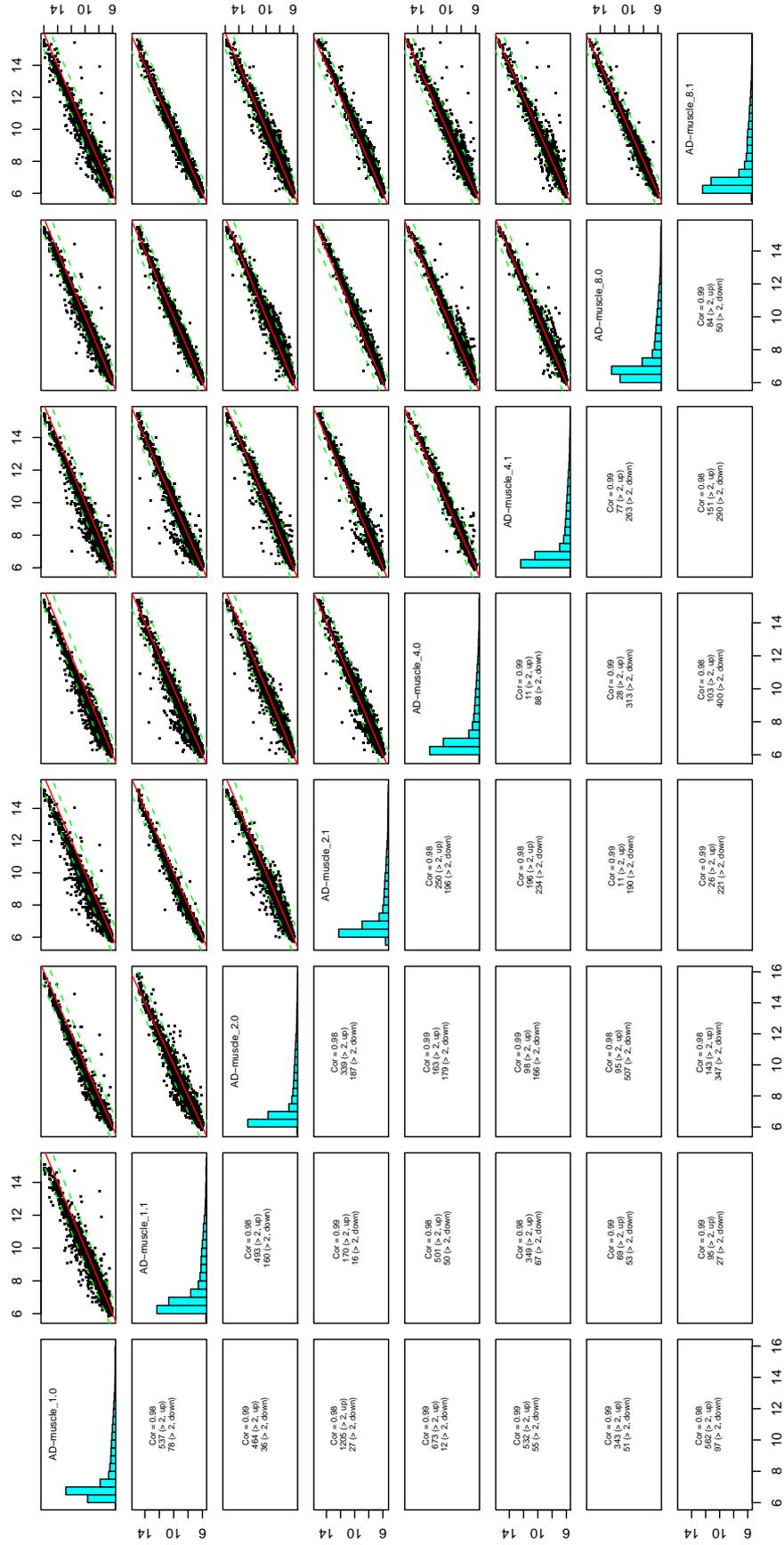


Figure 3.19: Pairwise Comparison of AD Muscle Samples Before Normalisation. This plot compares all eight AD muscle biological replicates against each other. These replicates should be very similar and should be tightly distributed on the x=y axis, as these samples are. The correlation for all comparisons ≥ 0.98 .

Pairwise Plot with Sample Correlation for UN Muscle Samples

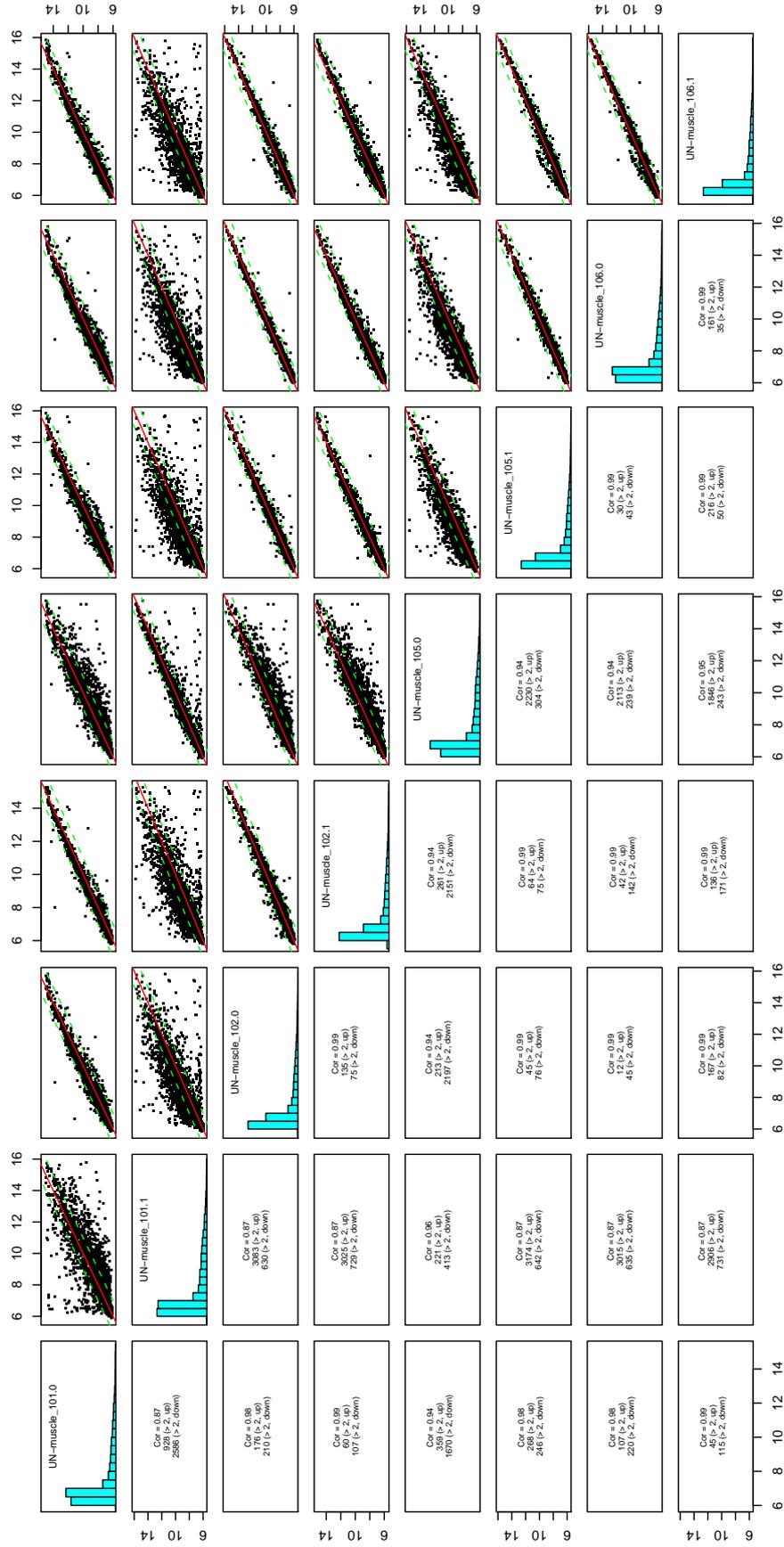


Figure 3.20: Pairwise Comparison of UN Muscle Samples Before Normalisation. This plot compares all eight UN muscle biological replicates against each other. These replicates should be very similar and should be tightly distributed on the x=y axis. These comparisons show higher variability of samples 101.1 and 105.0. All other comparisons have a correlation ≥ 0.98 .

Pairwise Plot with Sample Correlation for AD WAT Samples

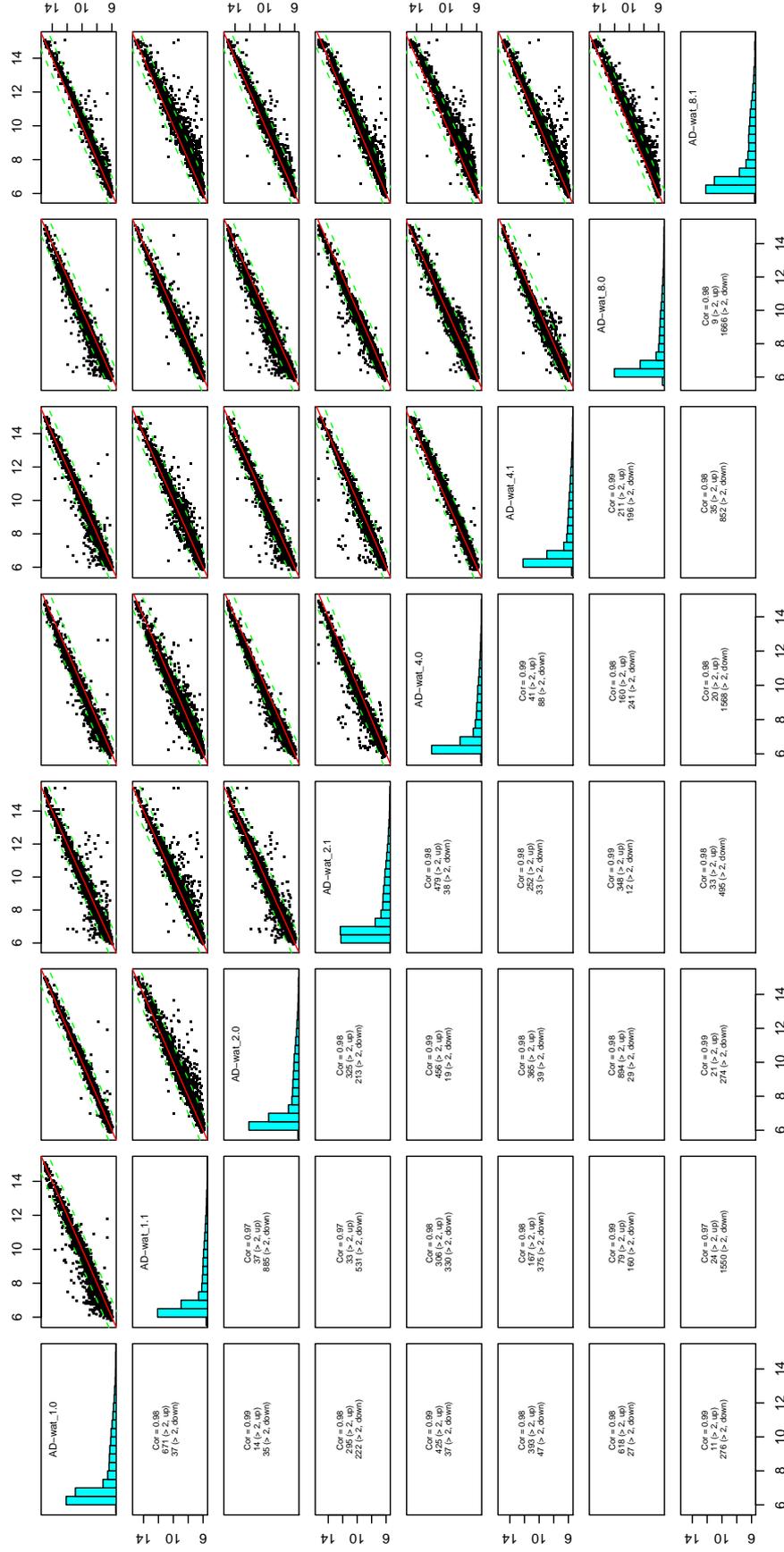


Figure 3.21: Pairwise Comparison of AD WAT Samples Before Normalisation. This plot compares all eight AD white adipose tissue biological replicates against each other. These replicates should be very similar and should be tightly distributed on the x=y axis, as these samples are. The correlation for all comparisons ≥ 0.97 .

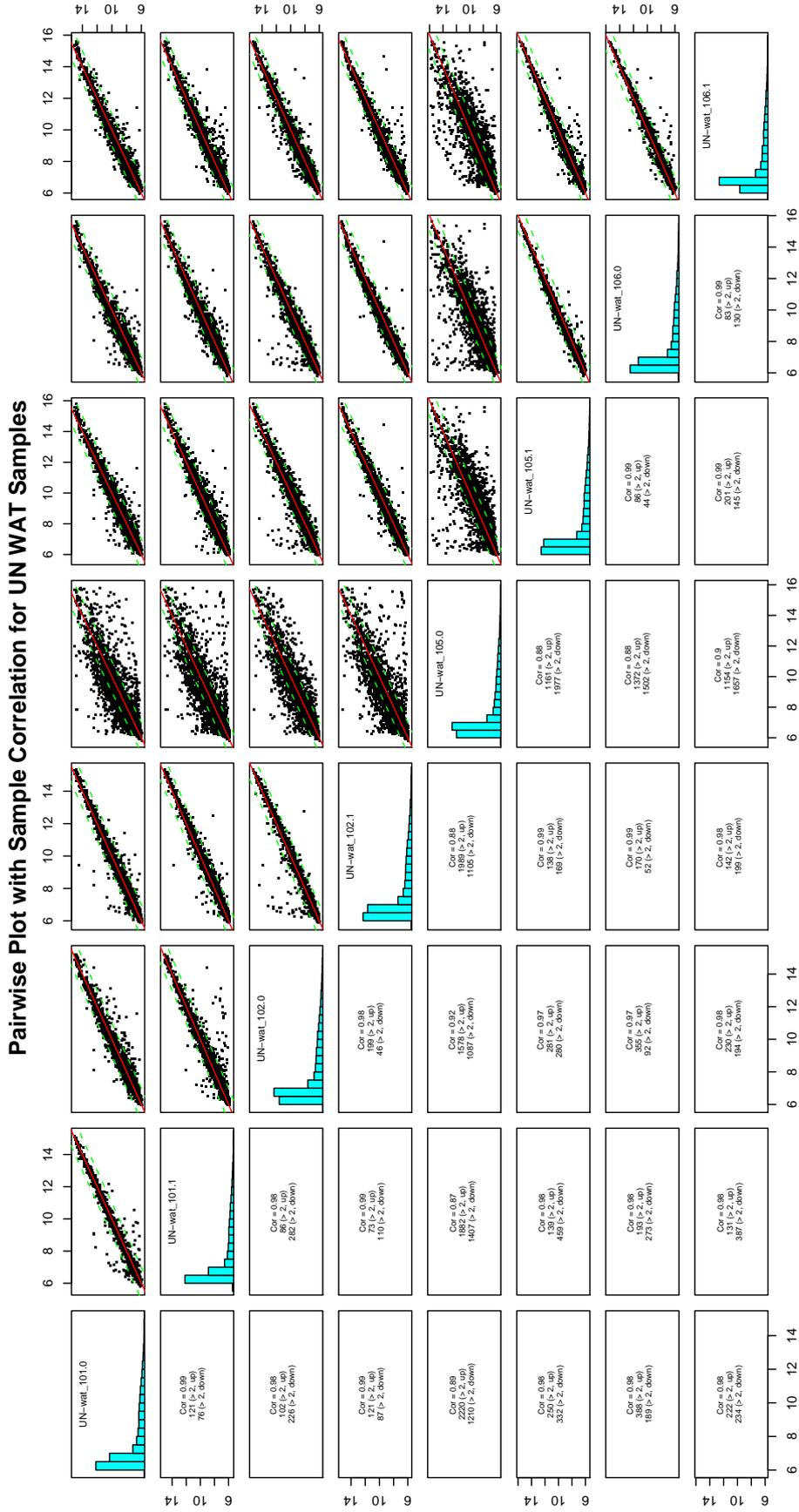


Figure 3.22: Pairwise Comparison of UN WAT Samples Before Normalisation. This plot compares all eight UN white adipose tissue biological replicates against each other. These replicates should be very similar and should be tightly distributed on the x=y axis. These comparisons show higher variability of sample 105.0. All other comparisons have a correlation ≥ 0.97 .

Figures 3.23 and 3.24 are plots of the arrays before normalisation. Figure 3.23 is a boxplot of the amplitude across and arrays and figure 3.24 is a density plot of the intensity for each array. These plots show that there is variation between arrays, therefore, normalisation is necessary.

Normalisation/Differential Expression

The three samples that had been shown as being poor data in the quality control assessment were removed from subsequent analysis. In total, there were 22,226 genes represented on the Illumina RatRef Chip. The data was filtered for genes expressed on at least one array. This left 12,951 genes. Data was then transformed with a variance stabilising transformation and normalised with a quantile normalisation (described in more detail in Chapter 2). Figure 3.25 and 3.26 show a boxplot and density plot of all samples after normalisation.

A pairwise comparison produced p-values and a list of differentially expressed genes. This analysis revealed no significantly, differentially expressed genes amongst the muscle and fat samples. This suggested that the larger variation among these samples might be affecting the normalisation. Filtering and normalisation was repeated separately for samples from each tissue. Muscle was filtered to 10,644, fat was filtered to 9,941, and liver was filtered to 10,887 genes. When the same transformation and normalisation was done to each data set, still no genes were revealed as being differentially expressed in muscle and fat. For the liver samples, genes with a p-value < 0.05 were considered to show significant differential expression between the two groups. The comparison for liver revealed a list of 249 differentially expressed genes, which is shown in table 3.3 with associated p-values and fold changes between the two treatment groups. Appendix C Tables 1, 2, and 3 summarise the expression values for all genes represented on the array for liver, muscle, and WAT.

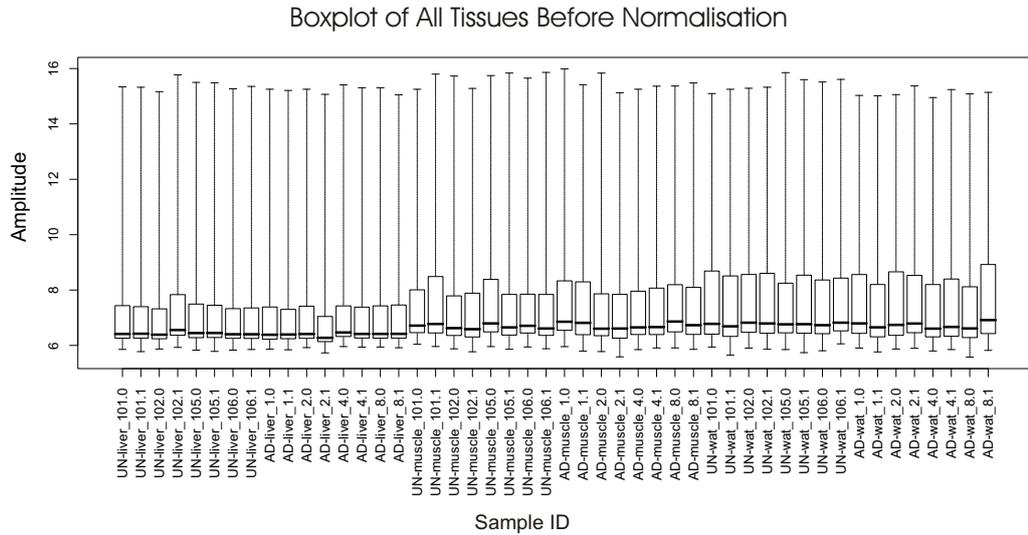


Figure 3.23: Boxplot Before Normalisation. The box itself contains the middle 50% of the data points. The line in the box indicates the median value of the data. If the median line within the box is not equidistant from the hinges, then the data is skewed. The ends of the vertical lines or ‘whiskers’ indicate the minimum and maximum data values of the data from each array. The plot compares the amplitude of all expression values on each array. This shows that the majority of the signal is at lower expression levels. It also shows that there is some variability between each array. This analysis was done in Lumi.

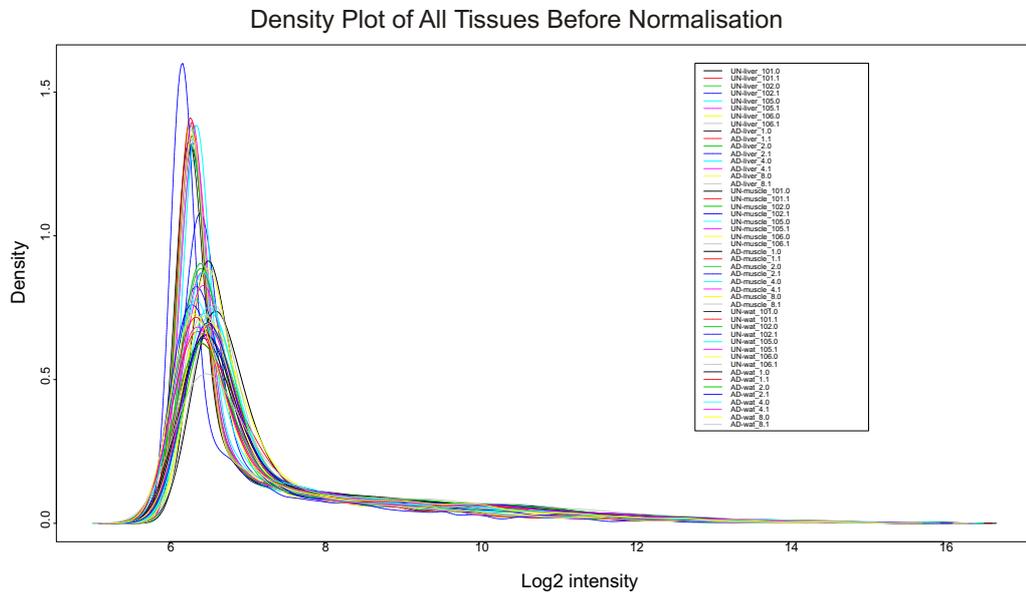


Figure 3.24: Histogram of Density of Log2 Intensities For All Arrays. This is a histogram of each array showing the density of intensities (log2). This plot shows that the majority of intensity values are low and that there is significant variation between samples. This analysis was done in Lumi.

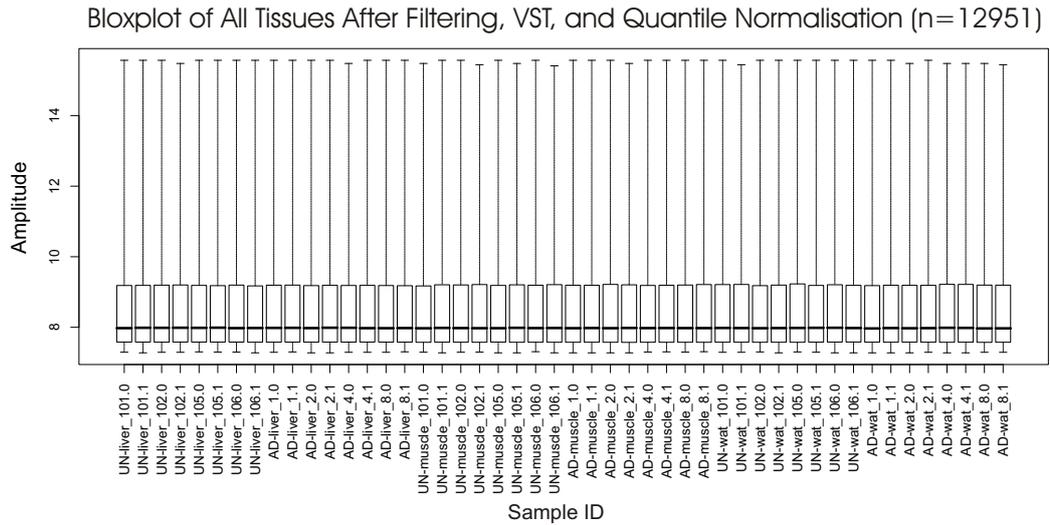


Figure 3.25: Boxplot After Quantile Normalisation. This plot compares the amplitude of all expression values on each array. This shows that the majority of the signal is at lower expression levels and that there is no variability between the arrays. The normalisation has effectively adjusted the expression values to remove the variability. This analysis was done in Lumi.

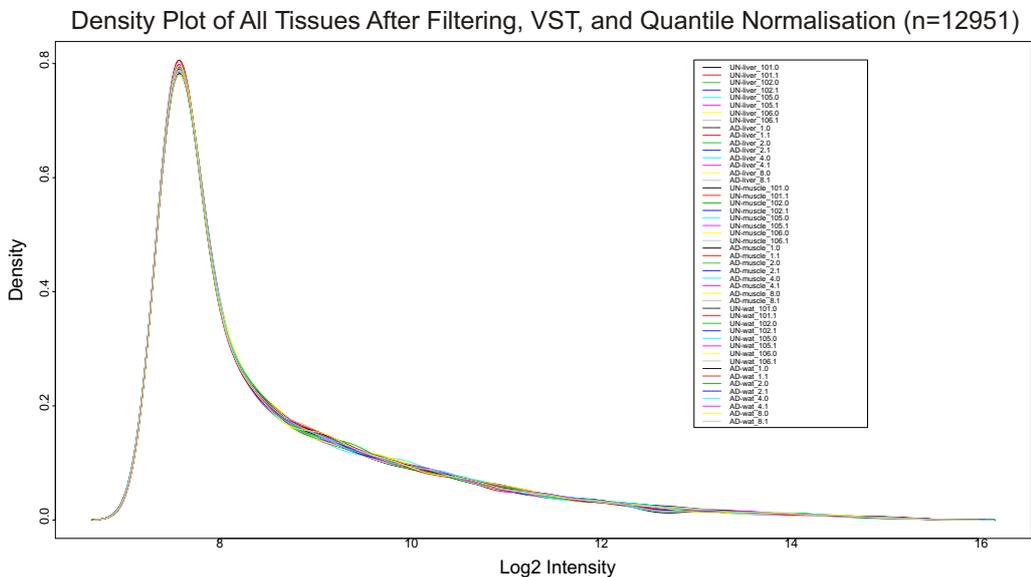


Figure 3.26: Histogram of Density of Log2 Intensities For All Arrays After Quantile Normalisation. This is a histogram of each array showing the density of intensities (\log_2). This plot shows that the majority of intensity values are low and that there is no variation between samples. The normalisation has effectively adjusted the expression values to remove the variability. This analysis was done in Lumi.

Log Fold Change	Adjusted p-value	Gene Accession	Gene Symbol
1.1003	0.0315	NM_013122.1	Igfbp2
0.7059	0.0114	NM_012733.3	Rbp1
0.6239	0.0483	NM_013089.1	Gys2
0.577	0.038	NM_001008363.1	RGD1310991 - Zfand2a
0.4877	0.0278	NM_053907.1	Dnase1l3
0.4662	0.024	NM_030832.1	Fabp7
0.4596	0.0331	NM_001007732.1	MGC94010 - Serpinb9
0.4265	0.0252	NM_175756.1	Fcgr2b
0.3684	0.0319	NM_024400.1	Adamts1
0.3412	0.038	NM_017334.1	Crem
0.3406	0.038	XM_573570.1	Crlz1_predicted
0.3322	0.0492	NM_001001507.1	Oit3
0.3299	0.0259	XM_235156.3	Ptprb_predicted
0.3254	0.0278	XM_579533.1	Ugcg
0.3248	0.0319	NM_199085.1	Serpinb6
0.3181	0.0252	XM_576256.1	LOC500859 - similar to 60S ribosomal protein L7a
0.3151	0.0278	NM_001007657.1	RGD1359127 - similar to RIKEN
0.3125	0.0252	NM_022396.1	Gng11
0.3065	0.0278	XM_579502.1	LOC497936- Tmem88
0.3027	0.0319	XM_221702.3	LOC304138 - Cyyr1
0.2884	0.0286	XM_235041.3	LOC314733 - similar to ribosomal protein S19
0.2863	0.0395	XM_575053.1	Sh3glb1_predicted
0.2854	0.038	XM_223786.3	Zfp503_predicted
0.2798	0.0407	XM_235308.3	Col14a1_predicted
0.2789	0.0309	NM_138974.1	Gstp2
0.2774	0.0315	NM_022510.1	Rpl4
0.2716	0.0328	NM_057137.1	Ebp
0.2702	0.0112	XM_216482.3	LOC298370 - Txndc12
0.2696	0.0326	XM_578896.1	LOC367566 - similar to Y-linked testis-specific protein
0.2684	0.0317	NM_199372.1	Eif4a1
0.2602	0.0326	NM_001004279.1	Ppid
0.2535	0.0418	XM_213058.2	LOC298785 - similar to ribosomal protein S26

Log Fold Change	Adjusted p-value	Gene Accession	Gene Symbol
0.2529	0.0319	NM_198765.1	BicD2
0.2526	0.0252	XM_217279.3	Scap_predicted
0.2521	0.0395	XM_577114.1	LOC501709 - similar to 60S ribosomal protein L23a
0.2515	0.0252	NM_031552.1	Add3
0.2494	0.0492	XM_235687.3	LOC315324 - Kb40
0.2494	0.0252	NM_001009674.1	Itm2c
0.2483	0.033	NM_053525.1	Ddx52
0.2481	0.0326	XM_579342.1	Tacr1
0.2479	0.0423	XM_233462.3	Tie1
0.2433	0.0252	XM_343771.2	LOC363450 - Ftsj1
0.2425	0.0492	NM_031646.1	Ramp2
0.2417	0.0492	NM_173116.1	Sgpl1
0.2406	0.0417	NM_001004085.2	Crat
0.2382	0.0326	NM_001009620.1	MGC105601 - Tmem204
0.2355	0.0278	XM_215041.3	Rpl27a_predicted
0.2348	0.0417	NM_031093.2	Rala
0.2325	0.0006	XM_212922.3	LOC294700 - similar to ribosomal protein L21
0.2313	0.0405	XM_343002.2	Cebpz_predicted
0.2313	0.033	XM_235689.3	LOC315329 - similar to expressed sequence AW556797
0.2288	0.0259	XM_341612.2	Sema6a_predicted
0.2277	0.0395	XM_224350.3	Fndc3_predicted
0.2257	0.0423	XM_217105.3	Ei24_predicted
0.2252	0.0282	NM_001009702.1	RGD1306106 - Rrp15
0.2251	0.0259	XM_213960.3	LOC289324 - Cnih4
0.2248	0.0328	NM_017178.1	Bmp2
0.2219	0.0315	XM_578657.1	LOC366887 - similar to ribosomal protein L31
0.2214	0.0497	NM_017233.1	Hpd
0.2205	0.0259	XM_345933.2	LOC367077 - similar to 40S ribosomal protein S26
0.2205	0.0278	XM_232937.3	LOC313196 - similar to KIAA0368
0.2205	0.0457	XM_214553.3	Atp8b1_predicted
0.2155	0.038	XM_342911.2	LOC362593 - Gnl2

Log Fold Change	Adjusted p-value	Gene Accession	Gene Symbol
0.2138	0.0233	NM_022605.1	Hpse
0.2124	0.0211	NM_001004245.1	Esam
0.2121	0.0457	XM_227556.3	LOC310760 - Ctnbp2nl
0.208	0.0315	NM_031745.2	Rsn
0.2067	0.0211	NM_017199.1	Ssr4
0.2062	0.0417	NM_013135.1	Rasa1
0.2057	0.0252	XM_221473.3	LOC498078 - similar to 60S ribosomal protein L7a
0.204	0.0469	XM_343850.1	LOC363531 - similar to 40S ribosomal protein S19
0.202	0.0424	XM_579528.1	B4galt6
0.2015	0.0315	XM_214191.3	Xpo4_predicted
0.2013	0.0319	NM_054004.1	Cand1
0.1992	0.0407	XM_580138.1	LOC500939
0.1982	0.0259	NM_012875.1	Rpl39
0.1952	0.0252	XM_216499.3	LOC298425 - RIKEN
0.1945	0.0211	XM_576096.1	LOC500714 - similar to ribosomal protein L6
0.1915	0.0282	XM_213403.3	LOC287541 - Ift20
0.1907	0.0226	XM_215053.3	LOC293454 - Ubfd1
0.1893	0.0395	XM_343490.2	LOC363151 - Ccdc12
0.1891	0.0252	XM_577309.1	LOC364108 - similar to ribosomal protein S17
0.1887	0.0492	XM_573428.1	LOC498211-MAT-II
0.1883	0.0315	NM_053302.1	Admr
0.1882	0.0423	XM_575861.1	LOC366411 - similar to ribosomal protein S24
0.1866	0.0259	NM_031085.2	Prkch
0.1859	0.0492	NM_133528.1	Prei3
0.1815	0.0278	NM_198787.1	Rutbc3
0.181	0.0312	NM_053720.1	Aatf
0.181	0.0315	XM_573168.1	LOC497974 - similar to novel protein
0.1802	0.0081	NM_001009696.1	RGD1307008 - Dcun1d5
0.18	0.029	NM_022005.1	Fxyd6
0.1739	0.0487	XM_228760.3	LOC317376 - Tcfe3
0.1734	0.0315	XM_574016.1	Serpib6b_predicted
0.1711	0.0314	NM_053345.1	Gtf2a2

Log Fold Change	Adjusted p-value	Gene Accession	Gene Symbol
0.1706	0.0407	XM_576398.1	Tmem24_predicted
0.1683	0.0252	NM_021681.1	Epb4.1l1
0.1674	0.0416	NM_133297.1	Sep-15
0.1637	0.0417	XM_576285.1	LOC500885 - similar to 40S ribosomal protein S19
0.1586	0.0492	XM_213394.2	LOC287477 - Tmem93
0.1585	0.028	XM_233982.3	LOC313974 - Trib2
0.1581	0.034	NM_053927.1	Epb4.1l3
0.158	0.0417	XM_219373.3	LOC309009 - similar to cDNA sequence BC025641
0.1577	0.0407	XM_341863.2	Gtf2h1_predicted
0.1552	0.0395	XM_343037.2	Ddx1
0.1539	0.0315	XM_227527.3	Wdr3_predicted
0.1524	0.0417	NM_012500.1	Apeh
0.1514	0.009	XM_215404.2	Bxdc1_predicted
0.1501	0.0315	XM_232531.3	Surb7_predicted
0.1501	0.0278	NM_012745.2	Klrd1
0.1501	0.0395	NM_138532.1	Nme7
0.1488	0.0435	NM_001008290.1	RGD1310861 - similar to RIKEN cDNA 1500011H22
0.1484	0.0492	XM_573294.1	LOC498088 - similar to hypothetical protein A
0.1452	0.0492	NM_031735.1	Stk3
0.1451	0.0418	NM_021682.1	Negr1
0.1446	0.0278	NM_001005536.1	p49/STRAP
0.1429	0.0335	XM_579788.1	LOC498118 - RGD1564468
0.1426	0.0435	XM_232995.3	Rnf20_predicted
0.1419	0.0278	XM_573749.1	LOC498489 - RGD1559923
0.1416	0.0315	NM_022700.1	Arl3
0.1338	0.0287	XM_214983.3	RGD1310022_predicted
0.1294	0.0309	XM_216759.3	Med6_predicted
0.1279	0.0444	XM_225053.2	LOC290916 - similar to 60S ribosomal protein L23a
0.1251	0.0252	NM_001009689.1	Cdc42ep2
0.1228	0.0492	NM_001008357.1	Hcfc2
0.1223	0.0418	XM_341918.2	LOC361639 - RGD1307507

Log Fold Change	Adjusted p-value	Gene Accession	Gene Symbol
0.1217	0.0259	XM.578086.1	LOC502599 - similar to 40S ribosomal protein S26
0.1215	0.0331	XM.218415.2	Gpr4_predicted
0.1209	0.0411	XM.573303.1	LOC498097 - Wdr53
0.1205	0.0328	XM.236438.3	LOC315843 - similar to WD repeat domain 11 protein
0.1202	0.0278	NM.145680.2	Gimap5
0.1195	0.0326	XM.215017.3	LOC293181 - Galnt14
0.1159	0.0252	XM.214874.3	LOC292724 - Ccdc97
0.1154	0.038	XM.222868.3	Olfml2b_predicted
0.1146	0.0278	XM.574766.1	LOC499443 - Lims1
0.1126	0.0252	XM.342168.2	Cetn3
0.1034	0.0497	XM.219309.2	LOC293472 - similar to 60S ribosomal protein L13
0.1024	0.0278	XM.579982.1	LOC499580 - RGD1561695
0.097	0.0395	XM.574926.1	LOC499600 - RGD1560944
0.0923	0.0492	XM.236644.3	Smarcc1_predicted
0.0923	0.0492	XM.215902.3	RGD1311678_predicted
0.0921	0.0407	XM.573169.1	LOC497975 - Znhit3
0.083	0.0481	XM.218452.2	Lgtn_predicted
-0.087	0.0331	XM.579952.1	LOC499365 - RGD1561823
-0.092	0.0315	XM.237146.2	LOC301388 - RGD1562317
-0.101	0.0417	XM.218346.1	LOC308444 - Axl
-0.102	0.0305	NM.145789.1	Il13ra1
-0.105	0.0416	XM.573373.1	LOC498160 - Zkscan1
-0.112	0.038	NM.138976.1	Mfn1
-0.116	0.0417	XM.230291.3	Fnbp4_predicted
-0.116	0.038	XM.221043.3	Tex2_predicted
-0.119	0.0319	NM.198757.2	Srr
-0.13	0.033	XM.341579.2	Npc1_predicted
-0.134	0.038	XM.212982.3	LOC296724 - RGD1565758
-0.137	0.0328	XM.580139.1	LOC500955 - RGD1561356
-0.137	0.0328	XM.226417.2	Thap11_predicted
-0.137	0.0494	XM.237326.2	LOC316550 - similar to Rab18
-0.138	0.0402	XM.215626.3	LOC295245 - Rag1ap1
-0.138	0.0379	XM.232647.3	LOC312946 - Tmem68

Log Fold Change	Adjusted p-value	Gene Accession	Gene Symbol
-0.139	0.0278	XM_342238.2	RGD1310132_predicted
-0.141	0.0278	NM_139332.1	Tpcn1
-0.147	0.0379	NM_031030.1	Gak
-0.148	0.0278	NM_013177.1	Got2
-0.148	0.0492	XM_216255.3	Crbn_predicted
-0.15	0.0278	XM_574939.1	LOC499612 - similar to NADH dehydrogenase (ubiquinone) 1, subcomplex unknown, 1
-0.153	0.0259	XM_342838.2	Mapk9
-0.155	0.038	XM_342838.2	Fcmd_predicted
-0.156	0.0407	XM_575280.1	LOC499935 - RGD1562582
-0.158	0.0315	XM_576379.1	Atpi
-0.159	0.0395	XM_233196.3	Hook1_predicted
-0.161	0.0326	XM_218848.3	Stard5_predicted
-0.161	0.038	NM_133387.1	Tmlhe
-0.163	0.0252	NM_013173.1	Slc11a2
-0.166	0.0335	XM_220076.3	Atad1_predicted
-0.166	0.0226	NM_080787.1	Dgka
-0.167	0.0315	NM_001009290.1	Ndufc2
-0.171	0.0429	XM_220047.3	Tnks2_predicted
-0.174	0.0497	XM_343513.2	Aplp2
-0.174	0.0382	NM_020076.1	Hao
-0.179	0.0252	NM_031353.1	Vdac1
-0.18	0.0497	XM_222022.3	Trfr2_predicted
-0.181	0.0287	XM_575959.1	Tardbp_predicted
-0.181	0.0481	NM_053357.2	Ctnnb1
-0.184	0.0423	NM_138883.1	Atp5o
-0.184	0.0417	XM_579951.1	LOC499357 - RGD1562275
-0.187	0.0492	XM_214276.2	LOC290555 - Spcs1
-0.188	0.0305	XM_216928.3	LOC299949 - Wdr67
-0.188	0.0417	XM_342268.2	Trim2_predicted
-0.189	0.0315	NM_001008304.1	Ptov1
-0.191	0.038	XM_573163.1	LOC497971 - withdrawn
-0.194	0.0392	NM_017033.1	Pgm1
-0.194	0.0326	XM_579608.1	LOC497802 - Manea
-0.196	0.0407	XM_578010.1	LOC502525 - similar to homolog of yeast TIM14 isoform c

Log Fold Change	Adjusted p-value	Gene Accession	Gene Symbol
-0.197	0.0483	NM_139081.1	Oaz1
-0.198	0.0497	XM_573972.1	LOC498687 - RGD1564910
-0.198	0.033	NM_017022.1	Itgb1
-0.2	0.0487	XM_224617.3	Capn7_predicted
-0.205	0.0372	XM_223745.3	LOC305633 - similar to Antxr2 protein
-0.205	0.0328	XM_235768.3	Tmprss6_predicted
-0.207	0.0252	NM_001009646.1	Qprt
-0.211	0.0319	XM_233808.3	Prkcn_predicted
-0.212	0.0287	NM_172091.1	Gcgr
-0.214	0.0492	NM_172008.1	Canx
-0.214	0.038	XM_575138.1	Axot_predicted
-0.215	0.0417	XM_216378.3	Ndufb6_predicted
-0.216	0.0417	XM_341854.2	LOC361571 - Aldh16a1
-0.216	0.0309	XM_344301.1	LOC364258
-0.217	0.0252	XM_342695.2	LOC362370
-0.22	0.0492	XM_579362.1	Nfia
-0.221	0.0319	XM_579711.1	LOC497803
-0.222	0.0326	XM_219895.3	RGD1305246_predicted; Fam108b1
-0.228	0.038	XM_214409.3	LOC290964
-0.23	0.0278	XM_228158.3	LOC309788
-0.235	0.0405	NM_199404.1	Man2b1
-0.24	0.033	XM_576497.1	LOC501085 - rSULT1C2
-0.249	0.0476	NM_173123.1	Cyp4f4
-0.253	0.0252	NM_053600.1	Fez2
-0.255	0.0407	XM_216044.3	Ndufa8_predicted
-0.257	0.0315	XM_213943.3	Mgst3_predicted
-0.263	0.0435	NM_178105.2	Gpm6a
-0.266	0.0114	XM_343935.2	RGD1307010_predicted; Glod4
-0.272	0.0497	XM_341824.2	Capns1
-0.278	0.0278	NM_031013.1	Abcc6
-0.29	0.0407	XM_345167.2	LOC365699 - similar to hypothetical protein FLJ30596
-0.298	0.0114	XM_576040.1	LOC500662 - GST 8-8
-0.302	0.0278	XM_346854.2	Besh3
-0.32	0.0372	NM_013146.2	Cald1
-0.335	0.0315	XM_579586.1	LOC497846 - Magt1

Log Fold Change	Adjusted p-value	Gene Accession	Gene Symbol
-0.337	0.0319	XM_575256.1	LOC499912
-0.35	0.0315	XM_343776.2	Pls3
-0.358	0.038	XM_576370.1	LOC500959 - triosephosphate isomerase
-0.359	0.0226	NM_001009636.1	RGD1308082 - Prelid1
-0.36	0.0328	XM_214551.3	Cidea_predicted
-0.363	0.0435	NM_153318.1	Cyp4f6
-0.367	0.0417	NM_031144.2	Actb
-0.371	0.0031	BC088177	BC088177 - Qprt
-0.371	0.0392	NM_022272.1	Fbxl20
-0.377	0.0326	XM_341557.2	Dhtkd1_predicted
-0.38	0.0226	NM_130403.1	Ppp1r14a
-0.389	0.0278	XM_574498.1	Dgat2
-0.394	0.0326	XM_219925.3	Ankrd15_predicted
-0.413	0.0326	XM_231739.3	RGD1306512_predicted
-0.427	0.029	NM_012992.2	Npm1
-0.444	0.0252	NM_016991.2	Adra1b
-0.446	0.022	XM_230613.3	Hao1
-0.471	0.0417	XM_221747.3	Gbe1_predicted
-0.553	0.0418	XM_342965.2	Arhgef19_predicted
-0.662	0.038	XM_221358.3	LOC303861 - Cpn2
-0.875	0.0252	XM_575821.1	LOC500457

Table 3.3: Differentially Expressed Genes from Illumina Microarray. This table shows the 249 genes that were differentially expressed on the Illumina Rat-Ref12 chip in the livers of the 55 day old male offspring of mothers that had been undernourished during pregnancy. The table shows the log fold change, the adjusted p-value, the genbank accession ID, and the gene symbol.

3.2.3 QRT-PCR for Microarray Verification

Quantitative RT-PCR analysis of selected genes was performed in order to confirm the expression data derived from array analysis. Seven genes (*Igfbp2*, *Fabp7*, *Zfand2a*, *Dgat2*, *Gys2*, *Adra1b*, and *Rbp1*) found to be differentially expressed were selected for qRT-PCR on the pooled RNA samples for the biological replicates in each treatment group. qRT-PCR was performed as described in Chapter 2. Cyclophilin was used as an invariant control for normalisation.

Figure 3.28 summarises the data and illustrates that the differences in $\Delta\Delta C_t$ values between the AD and UN RNA populations for each gene are consistent with the expression changes determined by array analysis.

Previous analysis has provided evidence that the expression of the *Ppara* α , *GR*, and *Aox* genes is altered in the liver of offspring from normally nourished mothers compared to mothers who had either a low protein diet or balanced undernutrition during pregnancy [82, 84]. It has also been suggested, that *Pepck* (phosphoenolpyruvate carboxykinase) and *11 β -hsd2* (11 β -hydroxysteroid dehydrogenase type 2) may also be genes influenced by diet [85, 86]. As these genes were not represented by any of the oligonucleotides on the RatRef12 array, they were assessed by qRT-PCR in the AD and UN RNA populations using cyclophilin as the normalisation control. The qRT-PCR expression results of the five candidate genes are shown in figure 3.27 for the 8 biological replicates in each treatment group. For *Ppara* α and *Pepck*, there is no evidence for any differences in expression levels between the two treatment groups. However, *GR*, *Aox*, and *11 β -hsd2* show 1.5 to 2-fold increases in expression in the UN treatment group. As previously mentioned, some of these genes were represented on the MEEBO chip, but without significant p-values. It is interesting to note that the study done by Sedova *et al.* [87] used a diet high in sucrose and therefore also low in protein, making it similar to the diet used by Lillycrop *et al.*. Sedova found early indicators of the metabolic syndrome consistent with the findings of Lillycrop *et al.*, which suggests that sucrose may be involved in dietary programming.

Gene Expression Validation by qRT-PCR

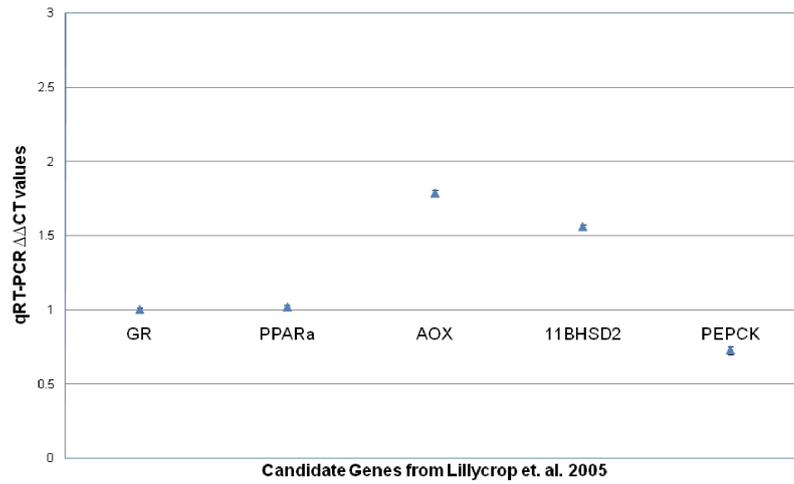


Figure 3.27: Five previously published candidate genes from Lillycrop *et al.* [82] were analysed by qRT-PCR. Each of the eight individual samples was run in triplicate (24 data points for each treatment group). The $\Delta\Delta$ CT values are shown for each gene along with error bars for standard error. Due to the large number of data points the error bars are very small and difficult to see.

Gene Expression Validation by qRT-PCR

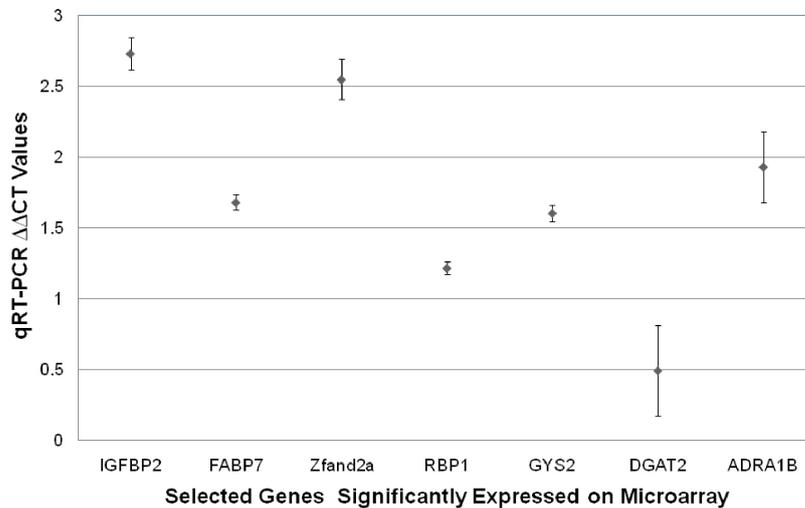


Figure 3.28: Seven genes were chosen from the 249 significantly differentially expressed genes on the Illumina microarray. The RNA from the 8 biological replicates was pooled for each of the two treatment groups. qRT-PCR for the two samples was run in triplicate with cyclophilin as a control. The $\Delta\Delta$ CT values are shown for each gene along with error bars for standard error. The values have been converted into a fold change for comparison to the array results shown in table 3.4.

Gene Name	Illumina Microarray Expression (Fold Change)	QRT-PCR Expression (Fold Change)	Meebo Microarray Expression (Fold Change)
ADRA1B	0.7	1.9	NS
DGAT2	0.8	0.5	NS
FABP7	1.4	1.7	0.887
GYS2	1.5	1.6	NS
IGFBP2	2.1	2.7	NS
Zfand2a	1.5	2.5	NS
RBP1	1.6	1.2	NS

Table 3.4: Microarray and QRT-PCR Data. Fold change for the QRT-PCR data and the Illumina and MEEBO expression data is compared for each of the seven genes from the Illumina array that were analysed by QRT-PCR. NS = not significant

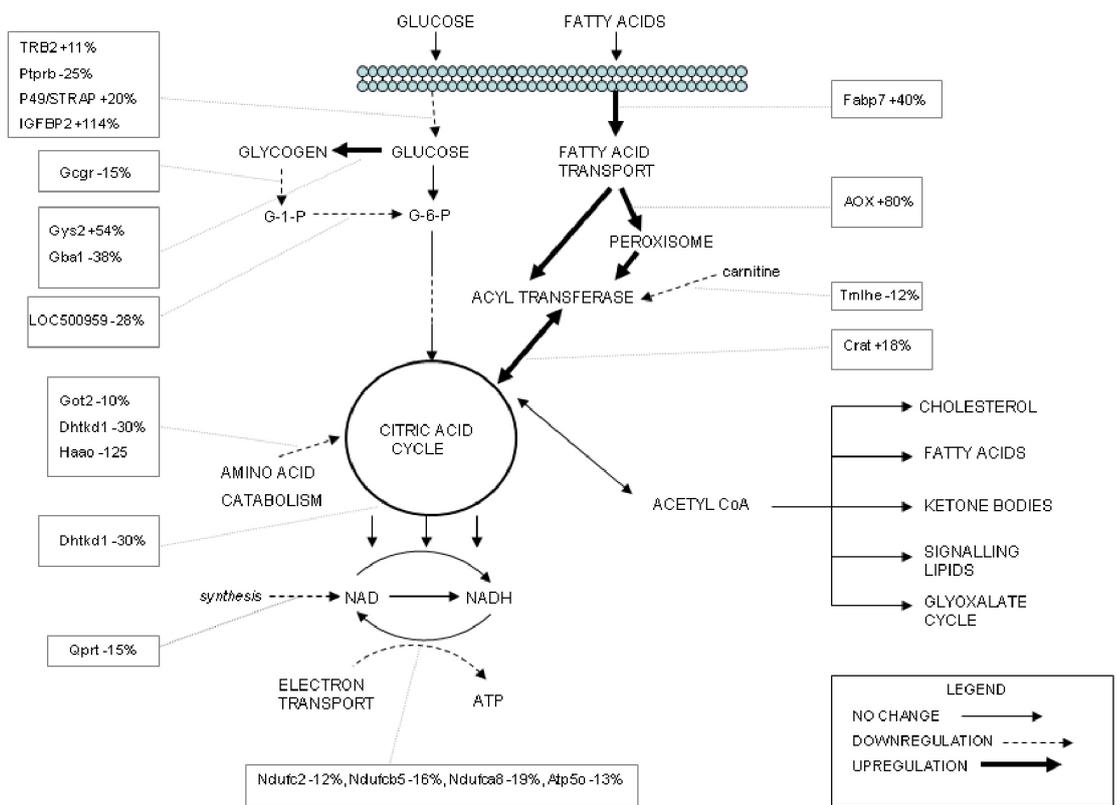
3.3 Discussion

Animals at postnatal day 55 were chosen for the study because they had completed puberty, but had not begun to exhibit the features of the metabolic syndrome phenotype observed later in adulthood and, as shown in the present study, as early as day 110 postnatally [45, 57]. Our results presented show that while there are differences in gene expression in hepatic tissue in day 55 males, there are no observed differences in retroperitoneal white adipose or biceps femoris. This is an intriguing finding, as both fat and skeletal muscle is adversely affected in the adult. This suggests that there is a cascade of biological processes that amplifies over time to lead to the phenotypic abnormalities. Previously Ozanne *et al.* [88] has shown, in a different but related experimental model of maternal low protein intake, that the expression of a limited number of genes in skeletal muscle is minimal early in life and accumulates later in life. It would appear that major changes in gene expression in the liver precede those in other metabolically relevant tissues.

The data for all three tissues is derived from 8 biological replicates in each treatment group with technical replication on each array varying from 20–30 fold. This means that for each gene represented on the array, there are between 160 and 240 measurements of expression level. This permits an accurate estimate of the variance

that is used to derive the final p-value and provides confidence in the detection of small gene expression changes. Thus the changes found in liver are based on many measurements and are highly significant. This is further underscored by the absence of any significant changes in gene expression between the two groups in muscle or adipose tissue subjected to the same analysis. The qRT-PCR analysis of candidate genes (*Ppara* and *GR*) showed very little difference in expression levels between the two groups in the day 55 male offspring. This contrasts with the 10-fold, 3-fold, and 3-fold increase for *Ppara*, *GR*, and *Aox*, respectively, in the day 34 offspring born to nutritionally restricted mothers as reported by Lillycrop *et al.* [82]. Four factors may explain this difference. First, the restricted diet is different (50% protein restriction only) whereas the diet in the study reported here is 30% of normal chow for all nutrients and may impact differently on gene expression. Second, in the Lillycrop *et al.* study the animals were weaned on day 28 and spent only 6 days in the postweaning period compared to 33 days in this present study. Third, the composition of the postweaning diet in the present study is grain-based as opposed to a purified diet containing 50% sucrose in the Lillycrop *et al.* study. Finally, the male and female tissues were pooled in the Lillycrop *et al.* study and there may be gender differences we have not explored. Similar to previous reports, *11 β -hsd2* and *Aox* show increases of between 1.5 to 2-fold in the offspring of undernourished mothers used in this study [82].

Analysis of the gene categories and associated pathways associated with differentially expressed genes (see Appendix C, Table 4) has revealed some relevant changes that potentially impact on the metabolic phenotype. Table 3.5 summarises the relevant categories to emerge from analysis of gene functions (some of which emerged from DAVID and Ingenuity analyses) associated with the list of differentially expressed genes. Figures 3.29A and B summarise the interrelationships between pathways involved in glucose, fat, and energy metabolism and highlights differentially regulated genes that influence these pathways.



B

EXPRESSION CHANGES IN THE INSULIN/IGF SIGNALLING PATHWAY

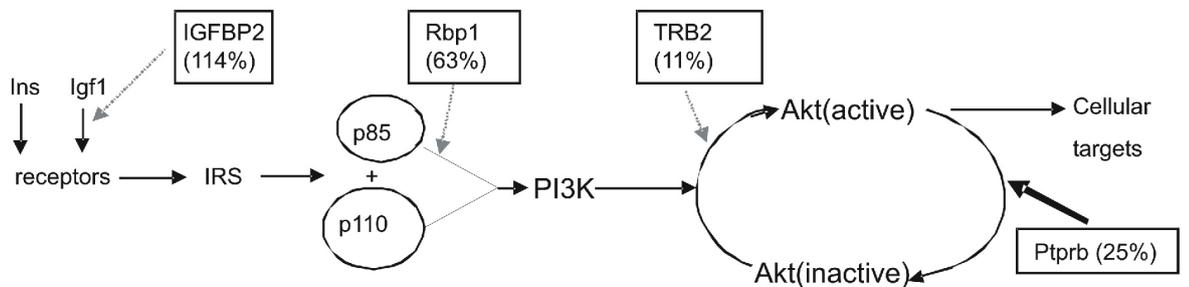


Figure 3.29: (A) Summarises the interrelationships between pathways involved in glucose, fat, and energy metabolism and highlights differentially regulated genes that influence these pathways. (B) A summary of the genes that are altered in signalling through the IGF-1 and PI3K signalling pathways and their relevance to glucose and fat metabolism. Normal arrows indicate no change in gene expression; dashed arrows indicate down regulation; bold arrows indicate up-regulation.

3.3.1 Glucose Metabolism

In utero, the foetus relies mainly on maternal glucose for its energy supply and only in late gestation develops the capacity for glucose storage, mobilization, and gluconeogenesis in preparation for postnatal life. Glucose and glycogen metabolism appears to have developed inappropriately in the livers of UN animals [4, 89]. Many of the changes in hepatic gene expression we detected are relatively modest and this may reflect alterations in hepatic development. Burns *et al.* reported changes in the balance of periportal to periarterial hepatocytes that have different metabolic profiles in the offspring of low protein fed dams [90].

The livers of UN animals show changes in the anomalous expression of RNA for several enzymes involved in glycogenesis, glycogenolysis, and glycolysis. There is a 54% increase in the expression of the glycogen synthase gene (*Gys2*); however, this is counterbalanced by a marked decrease in expression of glucan branching enzyme (*Gbe1*) that is essential for the formation and storage of glycogen. Furthermore, there is a modest decrease in phosphoglucomutase (*Pgm1*) that converts the glucose-1-phosphate resulting from glycogenolysis to glucose-6-phosphate destined for glycolysis. There is also a deficit of triose phosphate isomerase, which could result in an inefficient conversion of fructose 1-6 bisphosphate to glyceraldehyde-3 phosphate in the glycolytic process. The reduced expression of triose phosphate isomerase would therefore, affect the efficiency of glucose conversion to pyruvate and might indicate an increased accumulation of dihydroxyacetone phosphate. This can readily be reduced to glycerol-3-P that is the backbone for phosphatidates and signalling phospholipids. Overproduction of these could disrupt intracellular function. The efficiency of the TCA would be compromised by the reduction in dehydrogenase E1 (DHTKD1). This enzyme catalyses the conversion of α -ketoglutarate to succinyl CoA- a key step in the TCA. Its subsequent conversion to succinate is the only step in glucose catabolism that produces high-energy phosphate (GTP) directly, without the need for oxidative phosphorylation. The GTP can be used for nucleotide-specific metabolic reactions or it can convert ADP to ATP. Production of oxaloacetate further on in the cycle would

also be affected. This would slow the rate of formation of citric acid and the entry into the TCA of acetyl-CoA from oxidative decarboxylation of pyruvate. Supplementation of oxaloacetate from metabolism of branched-chain amino acids would also be reduced. This means that most of the oxaloacetate is required as an intermediate for TCA function and less is available for gluconeogenesis. It is interesting to note, that there is a reduction in the transcripts of other enzymes involved in the catabolism of amino acids that supply intermediates to the TCA cycle. In the foetus, this impasse may be circumvented to an extent by the presence of the glyoxalate cycle in which acetate can be converted to either glucose or succinate. However, this pathway disappears after birth and hence the inability for gluconeogenesis would be exacerbated postnatally.

3.3.2 Insulin and IGF-1 Signalling

The insulin and IGF-1 receptors are closely related structurally and share many post-receptor signalling mechanisms. The livers of UN animals show changes in expression of genes whose products are known to attenuate the PI3 kinase signalling pathway that is common to both receptors. Both *Igfbp2* and *Rbp1* are known to interact with components of the PI3Kinase/Akt signal

A primary function of IGF-1 is to drive postnatal cell proliferation and growth, principally through the MAP kinase signal

3.3.3 Mitochondrial Activity

The array analysis indicates a number of changes in the expression of genes that could affect the efficiency of electron transport and ATP generation in the mitochondrion in liver from UN animals. Transcripts of NADH dehydrogenase subunits (*Ndufc2*, *Ndufb6*, and *Ndufa8*) of complex I of the electron transport chain are all reduced as is quinolate phosphoribosyltransferase (*Qprt*) required for the synthesis of NAD, the electron carrier. In addition, the ATP synthase (*Atp5o*) that generates ATP from the electron gradient is reduced by a similar amount. Changes are also

observed in peptidylprolyl isomerase D (*Ppid*) that controls the mitochondrial permeability transition pore and *Tim14*, a component of the mitochondrial protein import motor. Taken together these findings suggest a degree of mitochondrial dysfunction in the livers of UN animals that could significantly affect the energy balance of the tissue and may be the precursor of later adiposity. Since the liver acts as the initial nutrient sensor, this defect could affect not only its own metabolism but also that of other organs that respond to its systemic signals.

3.3.4 Fat

In this well-characterised model, it has been shown that undernourishment of dams during pregnancy predisposes offspring to later adiposity and hyperinsulinaemia when fed a normal diet [91] but only at an age later than we have undertaken gene expression analysis. In the present study, even as early as day 110, male offspring of UN dams are showing features of metabolic syndrome related to increased adiposity, hyperinsulinemia, and hyperleptinemia. The absence of any differences in fat expression profiles between control and UN animals at day 55 may suggest that the development of adiposity is secondary to early metabolic perturbations in postnatal life. It is interesting to note that at day 35 UN rats show the onset of a sedentary phenotype (and hence reduced energy expenditure) before the presence of obesity [92]. This may reflect early changes in neurological gene expression modifying behaviour.

The elevated expression of RNA for fatty acid binding protein (FABP7) suggests that in livers from UN animals there is an increase in the intracellular trafficking of fatty acids. This is paralleled by increases in expression of *Aox* (from qRT-PCR analysis), which catalyses the first step in the peroxisomal degradation of fatty acids, and carnitine acyltransferase (CRAT) that shuttles fatty acids into the mitochondrial matrix. It is interesting to note that the expression of the first enzyme in the biosynthetic pathway of carnitine, trimethyllysine hydroxylase (*Tmlhe*) is down regulated and this may reflect a homeostatic response to excess fatty acid trafficking. These findings suggest that more energy is derived from β -oxidation of fatty acids in these

animals. However, this can only be achieved if the NADH from β -oxidation regenerates NAD and forms ATP. The observed defects (especially in complex I) indicate a disruption in this process. The provenance of these fatty acids is not clear, since there were no significant differences between treatment groups in the circulating levels of either triglycerides or free fatty acids, although LDL levels are elevated by maternal undernutrition. Other possible sources of fatty acids could be from increased endogenous synthesis or from turnover of cellular components. There is no evidence from the microarray analysis for an increase in enzymes for the latter; indeed the modest decrease in levels of diacylglycerol acyltransferase (DFAT2) in UN animals would indicate a blunting of triglyceride synthesis in liver.

The increased expression of SREBP cleavage activating protein (SCAP) suggests that there may be higher levels of sterol synthesis taking place in livers of undernourished animals. The formation of a transcriptionally active sterol response element binding protein (SREBP) involves the proteolytic action of SCAP [32]. In support of these observations, the level of circulating LDL cholesterol in the UN group is significantly elevated. Interestingly a 60% increase in *11 β -hsd2* transcripts was seen in the livers of UN animals accompanied by a similar increase in the expression levels of *GR*. *11 β -hsd2* appears in late gestation/postnatal tissues in human fetal tissues [93]. Elevated levels of this enzyme in the livers of UN animals would be a mechanism to moderate the effects of high levels of circulating corticosterone that may arise in response to nutritional deprivation/stress.

3.3.5 Ribosomal Proteins and Protein Turnover

An intriguing finding of the analysis is an increase (in the range 7-24%) in transcripts in UN animals of 13 of 80 recognised ribosomal proteins (79 of which are represented on the array by multiple oligonucleotides) that occur in the large and small ribosomal subunits. Eight of the proteins are components of the mitochondrial ribosome. Similar increases in expression are also seen in: *brix* (*Bxdc1*), which is involved in ribosome biogenesis; *NPG1*, a nucleolar GTPase required for the mat-

uration and nuclear export of pre-ribosomes; and *Cebpz*, a gene important in the processing of rDNA involved in the 60s subunit [94]. These findings suggest a greater degree of metabolic flux in ribosomes from UN animals; however it is difficult to understand the significance of these observations given the fact that on the basis of the current annotation only about 16% of the ribosomal proteins show changes in expression levels. It is possible that the increased turnover of ribosomal proteins is contributing to the higher levels of urea observed in the day 55 UN animals.

3.3.6 Effects on Gene Transcription

It is notable that a group of genes that influence transcription show upregulation to varying extents. Some (such as the *Gtf* and *Med* factors) have a general effect on the transcriptional apparatus. The most notable increase is in the *Rpb1* gene known to be expressed highly in the liver, particularly in hepatic stellate cells [95, 96]. In liver, *Rpb1* has an important role in the production of retinoid derivatives that activate the RAR and RXR retinoic acid receptors [97]. RXR can heterodimerise with the PPAR proteins to form active nuclear receptors that target genes involved in fat metabolism and energy homeostasis; for example *Crat* and *Aox* are shown to be upregulated in the UN offspring. Thus elevated or abnormal *Rpb1* expression may have a major role in establishing the metabolic syndrome phenotype.

Two other genes show significant changes in gene expression that may impact hepatic function. The first is *Col14a1* (a member of the collagen gene family) that was originally called undulin [98, 99]. This gene is expressed in hepatic stellate cells that have a role in the formation of the hepatic extra cellular matrix [100] and shows a 20% increase in gene expression. This may be a marker of some underlying pathology as *Col14a1* has been associated with the rearrangement of connective tissue occurring in hepatic fibrosis [98]. It is of note that *Rpb1* expression changes have also been associated with hepatic fibrosis [96] and thus the gene may be involved in the regulation of *Col14a1*. The second gene is *Adra1b* (adrenergic receptor alpha 1b showing a 79% upregulation) that has been associated with carbohydrate metabolism

in the liver [101] and disruption of glucose homeostasis when the gene is inactivated [102, 103]. Given that the role of the receptor is to stimulate glycogenolysis and the process of glycogen formation is impaired in UN animals, the significance of increased expression is not clear. This may reflect an irreversible change in response to the intrauterine nutritional state or a postnatal compensatory mechanism to the altered metabolic profile set *in utero*.

3.4 Conclusions

Studies to date have primarily focused on a candidate gene approach in animals, where the classical features of the metabolic syndrome in the programmed phenotype are already evident. There is a relative paucity of data on gene expression, either by candidate or an array-based approach, on key tissues related to the metabolic syndrome in animals at an age preceding development of the metabolic phenotype. The gene expression changes in appropriate pathways in the livers of the male offspring of maternally undernourished dams at day 55 would suggest that these animals may be predisposed to a persistent and perturbed ability to coordinate fat and carbohydrate metabolism with a shift to a use of fatty acids as an energy source. We have also shown that these animals proceed to develop a phenotype similar to that of the metabolic syndrome, as early as postnatal day 110. Array analysis of tissues at the day 110 time point was beyond the scope of the present trial. However, although the present study cannot *directly* correlate the observed gene expression changes at day 55 with the phenotype at day 110, it provides a clear evidence of disturbed hepatic function in a number of key genes related to lipid oxidation and mitochondrial function at a pre-phenotypic age. There are parallels with the observations of Koves *et al.* [89] and Sparks *et al.* [104], where diet induced obesity in adult life induces oxidative stress associated with increased β -oxidation of fat metabolism, impaired switching of carbohydrate substrates, depletion of TCA cycle intermediates and decreased expression of components of the electron transport chain. They propose that mitochondrial dysfunction may be driving the pre-diabetic insulin resistant state.

The observations in our rat model indicate that the liver is irreversibly programmed to respond to a nutritionally restricted environment and that this persists into early adulthood (day 55). At this age, similar changes are not observed in adipose tissue or skeletal muscle, suggesting that the liver can meet the immediate energy requirements of these peripheral tissues and that it manifests metabolic abnormalities in advance of the full metabolic syndrome phenotype. These data have uncovered potential candidate genes and pathways that when perturbed lead to the development of the metabolic syndrome in older animals and therefore, provide a focus for more detailed gene-specific studies.

It is clear that developmental changes are occurring in these animals due to the difference in maternal diet during pregnancy that is in turn affecting metabolic pathways later in life. However, it is yet to be established exactly what is the mechanism of this developmental programming. Gluckman *et al.* suggests epigenetic means. Epigenetic changes have been found in humans affected by the Dutch Hunger Winter and have also been associated with specific aspects of metabolic disease [65]. The details of the window in which these changes occur and whether they are primary changes or simply a result of a cascade of events remain unclear. In Chapter 5, this thesis will explore how methylation may be involved in the metabolic syndrome as a result of developmental insults and will investigate the phenomenon with a preliminary study. Additional work is being done by collaborators at the Liggins Institute and it is hoped that in the near future more answers will be available.

Analysis of Gene Function				
Metabolic Area	Symbol	Gene Identification	% Change Expression	Function
Carbohydrate and Glucose Metabolism	Gys2	glycogen synthase	54	stimulation of glucose storage
	Gbe1	glucan 1,4 α branching enzyme	-38	impairment of glycogen storage
	Pgm1	phosphoglucomutase	-14	increases glycogenesis over glycolysis
	Tpi	triose phosphate isomerase	-28	attenuation of glycolysis and energy production
	Gcgr	glucagon receptor	-15	attenuation of glycogenolysis
	Trib3	Tribbles homolog 3	11	impairment of insulin/IGF signalling at Akt [105]
	Ptprb	protein tyrosine phosphatase receptor type B	25	impairment of insulin/IGF signalling at Atk [106]
	Rbp1	Retinol binding protein 1	63	impairment of insulin/IGF signalling at PI3Kinase [107]
	Igfbp2	insulin-like growth factor binding protein 2	114	impaired IGD binding to receptor
	p49/STRAP	SRF-dependent transcription regulation associated protein	10	interference with GLUT4 biosynthesis and recycling [108]

Metabolic Area	Symbol	Gene Identification	% Change Expression	Function
Carbohydrate and Glucose Metabolism	TRIP3	thyroid hormone interactor protein 3 (TRIP3)	6	activator of NHF-4 α dependent effects on glucose metabolism [109]
	Dhtkd1	dehydrogenase E1 and transketolase	30	catalyses α ketoglutarate \longrightarrow succinylCoA in TCA
Fat and Lipid Metabolism	Crat	carnitine acyltransferase	18	increased transport of acylCoA from cytosol to mitochondrion
	Tmhle	trimethyllysine hydroxylase	-12	biosynthesis of carnitine
	Dgat2	diacylglycerol acyltransferase	-12	decreased synthesis of triglycerides
	Fabp7	fatty acid binding protein 7	40	increased uptake and intracellular transport of fatty acids
	Scap	SREBP cleavage activating protein (Scap)	20	increased hepatic lipid synthesis [110]
	Igfbp2	IGF binding protein 2	114	protection against obesity and insulin resistance [111]
	Ptprb	protein tyrosine phosphatase receptor type B	25	contributes to hepatic leptin resistance [112]
	Stard5	steroidogenic acute regulatory protein	-12	intracellular cholesterol transport

Metabolic Area	Symbol	Gene Identification	% Change Expression	Function
Fat and Lipid Metabolism	Cyp4f4	cytochrome P450	-28	involved in cholesterol, steroid and leukotriene synthesis
	Cyp4f6	cytochrome P450	-18	involved in cholesterol, steroid and leukotriene synthesis
	Atp8b1	ATPase class II type 8B	16.5	phosphatidyl serine and threonine transferase
	Cidea	cell death inducing DNA fragmentation factor α -like effector	-28	shown to regulate energy expenditure and lipolysis [113, 114, 115, 116]
	Ebp	phenylethylamine Ca ²⁺ antagonism (emopamil) binding protein	20	isomerase involved in the conversion of lanosterol to cholesterol
Amino Acid Metabolism	Got2	glutamate/oxaloacetate transaminase	-10	aspartate aminotransferase in a catabolism and malate shuttle
	Dhtkd1	dehydrogenaseE1 and transketolase domain containing 1	-30	2-ketoglutarate dehydrogenase activity
	Hpd	4-hydroxyphenylpyruvate oxidase	16	participates in tyrosine catabolism

Metabolic Area	Symbol	Gene Identification	% Change Expression	Function
Amino Acid Metabolism	Hao	3-hydroxyanthranilate dioxygenase	-12	participates in tryptophan catabolism
	Zfand2a	arsenite inducible RNA associated protein	40	adapts proteasome to counteract stress-induced proteotoxicity [117]
Protein Turnover	Tmprss6	transmembrane serine protease 6	-15	
	Capn7	calpain 7	-15	
	Cpn2	carboxypeptidase N83	-58	
	Capsn1	calpain small subunit 1	-21	
	Fam108b1	Cgi67 serine protease	-2.8	
	Serpinc6b	serine(or cysteine) peptidase inhibitor,clade B, member 6	25	
	Cand1	cullin associated/neddylated dissociated 1	15	involved in targeted protein proteosomal degradation [118]
	Dcun1d5	defective in cullin neddylation1, domain 5 (DCN1)	13	involved in targeted protein proteosomal degradation [119]
	Txndc12	Thioredoxin domain containing 12 (Txndc12)	20.6	protein folding, defence against oxidative stress
	Ubf1	ubiquitin family domain containing 1 (Ubf1)	14	involved in targeted protein proteosomal degradation

Metabolic Area	Symbol	Gene Identification	% Change Expression	Function
Protein Turnover	Eif4a1	translation initiation factor 4A1	20	
Ribosomes	Rpl7a	ribosomal protein L7a (mitochondrial)	24	
	Rps19	ribosomal protein S19	22	
	Rpl14	ribosomal protein L14 (mitochondrial)	20	
	Rps26	ribosomal protein S26 (mitochondrial)	19	
	Rp23a	ribosomal protein 23a (mitochondrial)	19	
	Rpl27a	ribosomal protein 27a (mitochondrial)	18	
	Rpl21	ribosomal protein L21 (mitochondrial)	17	
	Rpl31	ribosomal protein L31	17	
	Rpl39	ribosomal protein L39	15	
	Rpl6	ribosomal protein L6	14	
	Rps17	ribosomal protein S17 (mitochondrial)	14	
	Rpl24	ribosomal protein L24	14	

Metabolic Area	Symbol	Gene Identification	% Change Expression	Function
Ribosomes	Rpl13	ribosomal protein L13 (mitochondrial)	7	
	NPG1	autoantigen NPG1	16	required for the maturation/nuclear export of pre-ribosomes
	Bxdc1	box domain containing 1	11	control of ribosome biogenesis [120]
	Criz1	charged amino acid leucine zipper 1	26	required for 18s rRNA processing
	Rrp15	ribosomal protein processing homolog 15 (Rrp15)	17	
	Cebpz	CCAAT/enhancer binding protein zeta	17.4	Ribosomal RNA processing [94]
	Dnase13	DNase 1-like 3	40	fragments DNA during apoptosis
Apoptosis	Ppid	peptidyl prolyl isomerase D (Cyclophilin D)	20	anti-apoptotic. Blocks mitochondrial permeability
	Ei24	etoposide induced mRNA	17	may suppress cell growth by inducing apoptosis
	Prkch	protein kinase C eta	14	anti-apoptotic. Potent activator of Raf1

Metabolic Area	Symbol	Gene Identification	% Change Expression	Function
Apoptosis	Sgpl1	sphingosine phosphate lyase	18	apoptotic. Enhances stress-induced ceramide release [121]
	Aatf	apoptosis antagonising factor	13	anti-apoptotic nuclear phosphoprotein transcription factor
	Cidea	cell death inducing DNA fragmentation factor α -like effector	-28	apoptotic factor which induces DNA fragmentation [113, 114, 115, 116]
	SP16	serpin 9b protease inhibitor	27	regulates apoptosis by inhibiting caspases 8 and 10
Mitochondrial and Electron Transport Chain	LOC499612	NADH dehydrogenase (ubiquinone) complex unknown	-11	uncharacterised component of electron transport chain
	Ndufc2	NADH dehydrogenase (ubiquinone) 1 complex unknown 2	-12	uncharacterised component of electron transport chain
	Ndufc6	NADH dehydrogenase (ubiquinone) 1 beta sub-complex 6	-16	component of Complex I of electron transport chain

Metabolic Area	Symbol	Gene Identification	% Change Expression	Function
Mitochondrial and Electron Transport Chain	Ndufa8	NADH dehydrogenase (ubiquinone) 1 alpha subcomplex 8	-19	component of Complex I of electron transport chain
	p45/STRAP	SRF-dependent transcription regulated associated factor	-10	reported to interfere with Complex I assembly
	BC088177 - Qprt	quinolate phosphoribosyl transferase	-29	participates in NAD biosynthesis
	Qprt	quinolate phosphoribosyl transferase	-15	participates in NAD biosynthesis
	Atp5o	ATP synthase H ion transporting	-14	part of the mitochondrial ATP synthesis complex
	Atpi	ATPase inhibitor	-12	prevents ATP hydrolysis during electron flux
	Ppid	peptidyl prolyl isomerase D (Cyclophilin)	-20	part of mitochondrial transition pore
	Tim14	translocase complex subunit Tim14	-29	component of the mitochondrial protein import motor
	Haa0	3-hydroxyanthranilate dioxygenase	-12	provides quinolate for NAD synthesis

Metabolic Area	Symbol	Gene Identification	% Change Expression	Function	
Transcription	MED6	mediator complex subunit 6	9.3	part of the RNA polymerase II transcription unit	
	Gtf2a2	general transcription factor IIa 2	12.5	part of the preinitiation transcription complex	
	Gtf2h1	general transcription factor II H, polypeptide 1	3	part of the preinitiation transcription complex	
	MED21	mediator complex subunit 21	11	part of the RNA polymerase II transcription unit	
	Rbp1	retinoic acid binding protein	63	PRXR formation and PPAR activation	
	DDX52	DEAD box polypeptide 52	18.7	ATP-dependent RNA helicase	
	Ccdc12	Ccdc coiled-coil domain containing 12	14	involved in mRNA splicing	

Table 3.5: Functional Analysis of Differentially Expressed Genes. This table summarises the interesting categories to emerge from analysis of gene functions. Significantly differentially expressed genes from table 3.3 were literature searched individually for functional relevance to metabolic syndrome. Citations are included in the table. Some of the functional information emerged from DAVID and Ingenuity analyses, but no network connection can be assumed from this table. All genes were researched independently.

Chapter 4

Leptin Reversal Study

4.1 Introduction

As previously described, adult offspring of rats subjected to undernutrition during pregnancy develop symptoms of the metabolic syndrome. In Chapter 3, we investigated gene expression changes in the muscle, fat, and liver tissue of such male offspring at 55 days old. At this age, the rats have not yet begun to exhibit the metabolic syndrome phenotype.

Vickers *et al.* published two papers [57, 58] in which the same model of maternal undernutrition was used to investigate the effects of neonatal leptin treatment on the metabolic phenotype of adult offspring. Results for female offspring and male offspring differed significantly and were published separately. This chapter will focus on the results for the female study and all experiments in this chapter utilise those samples. The postnatal (day 3-13) leptin treatment in the female offspring of undernourished mothers resulted in a slowing of neonatal weight gain and normalised caloric intake, locomotor activity, body weight, fat mass, fasting plasma glucose, insulin, and leptin concentrations in programmed offspring. These data were taken in adult life (day 170) and contrast data from saline-treated offspring of undernourished mothers who only developed these features on a postweaning high fat diet. Neonatal leptin had no demonstrable effects on the adult offspring of normally fed mothers. These results suggest that developmental metabolic programming is potentially reversible by an intervention late in the phase of developmental plasticity. The

complete normalisation of the programmed phenotype by neonatal leptin treatment implies that leptin has effects that reverse the prenatal adaptations resulting from relative fetal undernutrition.

This chapter investigates the molecular mechanism of leptin reversal in adult female rats. Female rats can breed as early as 30 days of age and are considered middle age at approximately 270 days of age. In this study we used liver samples from female rats that were 170 days old from the study described above. The samples were analysed using microarray gene expression assays. The results do not reveal molecular evidence for a leptin reversal. However, they show that the interaction between a combination of the three treatments (prenatal diet, leptin treatment and postnatal diet) do have significant metabolic consequences. These changes in gene expression affect pathways involved in the immune response, circadian rhythm, transport, and metabolism. This chapter begins with a review of leptin, the leptin receptor, and their role in energy balance, causes of obesity, and the genetics of obesity, reproduction, and puberty. In addition, the biochemical effects that leptin has on appetite, behaviour, circadian rhythm, and immune response are considered. This background information will be followed by a description of experimental design, data quality, and a presentation and discussion of the results. The discussion aims to unravel the complexity of the experiment, interpret the results, and suggest future avenues of investigation.

4.1.1 Leptin

Leptin is an adipocyte specific protein that functions to sense and regulate body energy stores in humans and in animals. It is largely synthesised and secreted in white adipose tissue; however, low concentrations are also synthesised and secreted by the hypothalamus, pituitary, skeletal muscle and bone, arterial endothelium, intestines, foetus, testes, and placenta. The leptin in these tissues may play a role in the immune system and reproduction [122]. Leptin is remarkably similar across species.

The initial view was that leptin was an anti-obesity hormone, preventing the

storage of excess adipose tissue by feeding back to the hypothalamus to reduce food intake and increase energy expenditure. However, obesity in humans is often accompanied by increased levels of leptin in a state of leptin resistance [57]. Ahima *et al.* suggested that the physiological response to decreasing leptin concentration with starvation may be the dominant role of leptin. His study showed that falling leptin concentration is a critical signal that initiates the neuroendocrine response to starvation, potentially increasing survival by limiting procreation, increasing stress steroids, and decreasing thyroid thermogenesis [123]. Studies investigating leptin expression levels in the foetus and the placenta suggest that leptin may also have a broader range of actions, particularly during growth and development [124].

Early studies investigating the role of leptin were done using the *ob/ob* and *db/db* mouse models. Each of these models has a single distinct gene mutation. The *ob* mutation is on mouse chromosome 6 and the *db* is on mouse chromosome 4. The classic experiments involving parabiosis [125] of *ob/ob* mice with lean wild-type (WT) mice resulted in suppression of feeding and weight loss in the *ob/ob* mice. In contrast, the parabiosis of WT or *ob/ob* mice with *db/db* mice caused dramatic weight loss in the WT and *ob/ob* mice, but the *db/db* mice continued to gain weight. This study suggested that the *ob* locus encoded a circulating ‘satiety’ factor (discovered later to be leptin [126]). The mutation in the *ob/ob* model, leads to a non-functional leptin protein and mice with the *ob/ob* mutation became obese on a normal ad-libitum diet. These mice have low levels of circulating leptin and reduced metabolic rates and body temperature. Injection of exogenous leptin to *ob/ob* mice reduces hyperphagia and weight gain while increasing circulating leptin concentrations, physical activity, and energy expenditure. Humans that have a leptin deficiency due to a genetic mutation exhibit a similar phenotype including congenital obesity and endocrine abnormalities [127]. Leptin treatment in these people decreases energy intake and results in a dramatic fat loss. However, very few human obesity cases are due to this genetic mutation that reduces leptin production.

4.1.2 Leptin Receptor

Leptin interacts directly with the leptin receptor (*Ob-R* or *LEPR*). *Ob-R* is encoded by the *db* gene and binds leptin. The *Ob-R* gene has been shown to have six splice variants. The long form, *OB-RB*, has a long cytoplasmic region containing several motifs required for signal transduction and it is essential for the weight-reducing effects of leptin [122]. It is abundant in the hypothalamus, but it is also found throughout the body. The parabiosis experiments suggested that the *db/db* mice were defective in their ability to respond to the satiety factor, leptin, and subsequently obesity in the *db/db* mice has been linked to a defect of *Ob-R* [123]. Clément *et al.* [128] reported an example of human obesity due to a homozygous mutation in *Ob-R* that resulted in a truncated leptin receptor lacking the transmembrane and intracellular domains. This mutation resulted in morbid obesity and infertility, similar to that of the *db-db* mouse. This mutation is rare, and the affected offspring were children of first cousins. Very few obesity cases are due to defects in *Ob-R*. Considine *et al.* suggests obesity may more commonly be caused by defects in one of the isoforms of *Ob-R* or in one of the specific pathways activated by the isoforms [129].

4.1.3 Genetics of Obesity

Although genetic mutations in the leptin gene and leptin receptor gene are rare causes of obesity, genome-wide association studies have led to the discovery of novel genes in which differential expression levels have been linked to obesity. There are a number of additional rare cases occurring in families with rare genetic mutations. Particular polymorphisms seem to affect certain populations more than others. The Trp64Arg variant in the Beta-3 adrenergic receptor gene has been the subject of 60 independent studies with varying results. There seems to be a significant effect in Asian populations which does not seem to hold true for other populations [130].

Further investigation has shown that obesity is often related to expression changes that affect pathways involved in leptin signalling and brain feedback mechanisms for

controlling diet and weight gain. These pathways that are essential for energy balance may hold the explanation for the more common obesity cases. An example of this is the Jak-Stat pathway. The leptin receptor is important for the activation of the Jak (janus kinase) -Stat (signal transducers and activators of transcription) pathway (see figure 4.1). Mice that have a mutation in the *Stat3* binding site of *Ob-R* have much the same phenotype as db/db mice including hyperphagia and severe early onset obesity, but retained fertility. This suggests that *Stat3* is important for the regulation of feeding behaviour, but is not required to regulate other functions of leptin including reproduction and growth [131]. Two genes involved in negative-regulation of leptin-Jak-Stat signalling are *Socs3* (suppressor of cytokine signalling 3) and *Ptn1* (protein phosphatase nonreceptor type 1). Mice with reduced expression of either one of these genes are lean and exhibit leptin sensitivity, but how their function might be related to human obesity is still unknown [131].

4.1.4 Energy Balance: Interactions of Leptin and the Brain

The cause of obesity is often blamed on human decisions to engage in behaviours including lack of exercise and increased food intake. It is interesting to investigate the biochemistry of energy balance, appetite and satiety. The role of leptin in these affects feedback to the brain and affects behaviour decisions toward food and even exercise. Changes in pathways affecting energy balance and behaviour may underpin the difficulty that obese individuals have in losing weight and keeping it off [132]. The control of energy balance is orchestrated by the hypothalamus via signalling pathways, several of which are affected by changing leptin levels.

The arcuate nucleus (ARC) is an important site in the hypothalamus for regulating leptin's effect on energy balance [133]. Abundant expression of the long form of the leptin receptor (LEPR-B) has been found in the ARC and other areas of the hypothalamus including the ventromedial hypothalamus (VMH), the dorsomedial (DMH), and the paraventricular nucleus (PVN). Figure 4.2 shows a three-armed interaction between the brain and peripheral tissues that are involved in the home-

ostatic pathway. The three arms divide the afferent (blue), central (brown) and efferent (white) pathways. The VMH (and the ARC) receives peripheral peptide signals related to adiposity (leptin), metabolism (insulin), hunger (ghrelin) and satiety (peptide YY₃₋₃₆). The second arm consists of neurons that translate this signal to the PVN and lateral hypothalamic area. These areas then integrate the signals to alter caloric intake and energy expenditure. The third arm consists of signals from the CNS (efferent signals) via the autonomic nervous system. The sympathetic nervous system (SNS) promotes energy expenditure and the vagus nerve promotes energy storage. Leptin is part of the afferent pathway (signalling the CNS) and the SNS and vagus nerve are part of the efferent pathway, which receive these signals and then send signals to the rest of the body. Insulin is part of both afferent and efferent pathways [134]. Leptin conveys a signal of peripheral energy sufficiency to neurons (afferent signal) expressing the leptin receptor. Leptin triggers a permissive signal to the CNS for the initiation of high-energy processes, such as puberty and pregnancy [134]. A drop in leptin is interpreted by the hypothalamus as a diminished energy reserve and invokes the starvation response that is shown in figure 4.1. Inadequate transport of leptin into the CNS could be a contributing factor to obesity. The effects of insulin and leptin on the central nervous system parallel each other (i.e. mice with a knocked out insulin receptor become hyperphagic, obese and infertile) [134].

Leptin has also been shown to have an important role in hypothalamus development. Bouret *et al.* [135] found that leptin is important for the development of hypothalamic circuits. The outgrowth of fibres projecting from the ARC is slower in the absence of leptin and some of the projections never catch up. The connection between the arcuate nucleus and the parvocellular part of the PVN is one such connection. Figure 4.3 shows a focused figure of the brain that includes all of the areas where the leptin receptor expressing neurons are located and those areas that are part of the associated signalling network. In particular, this figure shows the arcuate nucleus, where leptin resistance is predominantly detected [136]. Leptin stimulates POMC (pro-opiomelanocortin) and CARTPT (cocaine and amphetamine related transcript prepropeptide) gene expression in neurons located in the ARC. A

cascade of events leads to neurons that express MC4R (melanocortin receptor 4). The signal transduction pathways involved in this process interact with various brain centres to coordinate eating behaviour. The identification of human mutations occurring in ligands and receptors of this pathway have revealed how this pathway is involved in the regulation of food intake, energy expenditure, lipid and carbohydrate metabolism, and reproductive, thyroid, and immune function [131]. In particular, it has been shown that PI3k signalling is required for leptin's effects on feeding. PI3k signalling enables leptin to depolarise POMC neurons and in effect, suppress food intake [137].

It is hypothesised each human has a different leptin 'set-point'. The set-point is a concentration of leptin, which conveys a message of energy sufficiency to the hypothalamus. Leptin levels drop quickly during short-term fasting (12hours/overnight) and can decline faster than body fat stores are depleted, but still activate the starvation response. In the case of obesity, where this hypothetical set-point is dysfunctional, leptin levels are increased, but do not trigger reduced food intake or lipolysis. If energy intake declines again (as in the case of dieting), the starvation response will be triggered again, resulting in a cycle of overeating [134].

This cycle can be exacerbated by the effects of leptin on the 'hedonic pathway'. The hedonic pathway regulates pleasurable and motivating responses to stimuli. In figure 4.3, several components and regulators of the mesolimbic dopamine system are shown in brown. This is the neural mechanism by which leptin can affect food reward. The relationship between insulin, leptin, reward and obesity is summarised in figure 4.4. Food intake is responsive to the hedonic pathway. Susceptibility to addictive behaviour and pleasurable response increases after food deprivation and can be measured by dopamine release. Obesity results in decreased density of D_2 dopamine receptors [134]. Other neural pathways that input into the regulation of food reward may also undermine leptin's influence preventing it from deterring an affected individual from consuming something palatable [136].

This section highlights the complexity involved in the regulation of energy balance and the effects that errors in brain feedback mechanisms could potentially have on

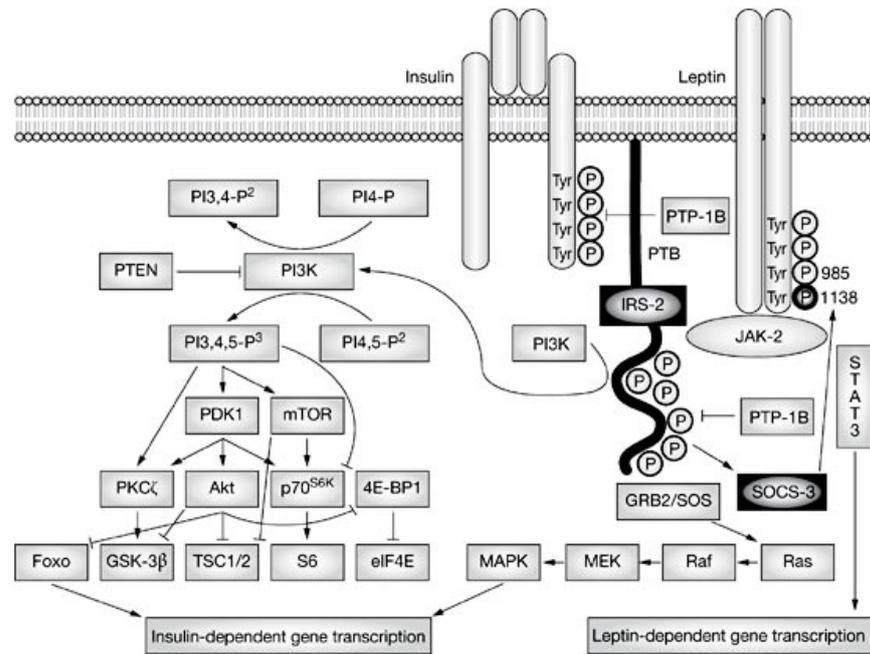


Figure 4.1: This figure was adapted from Lustig *et al.* [134] and depicts the overlap between insulin and leptin signalling pathways (shown in black) in the ventromedial hypothalamic neuron. Insulin stimulates the insulin receptor substrate 2-phosphatidylinositol 3 kinase pathway, whereas leptin stimulates the Janus kinase 2-signal transduction and transcription 3 pathway; however, both the insulin receptor and the leptin receptor recruit the low-abundance-message second messenger insulin receptor substrate 2. Lack of available insulin receptor substrate 2 for the leptin receptor following hyperinsulinemia could result in defective leptin signal transduction. Alternatively, insulin induction of suppressor of cytodine signalling 3 could inactivate the leptin receptor, through dephosphorylation of tyrosine 1138.

eating behaviour. A vast number of biochemical pathways are involved in maintaining energy homeostasis.

4.1.5 Leptin and the Circadian Rhythm

Leptin has been shown to have an effect on the circadian rhythm. A nocturnal rise in leptin occurs in ad-libitum fed rodents and can be prevented by fasting. Fasting can shift the peak plasma leptin level from nocturnal to diurnal. Leptin levels also peak at night in humans, and this pattern is thought to be mediated by insulin levels. This leptin rhythm seems to be blunted with aging, and associated with an increase in visceral adiposity and insulin resistance [138]. A study by Kaneko *et al.* [139] found

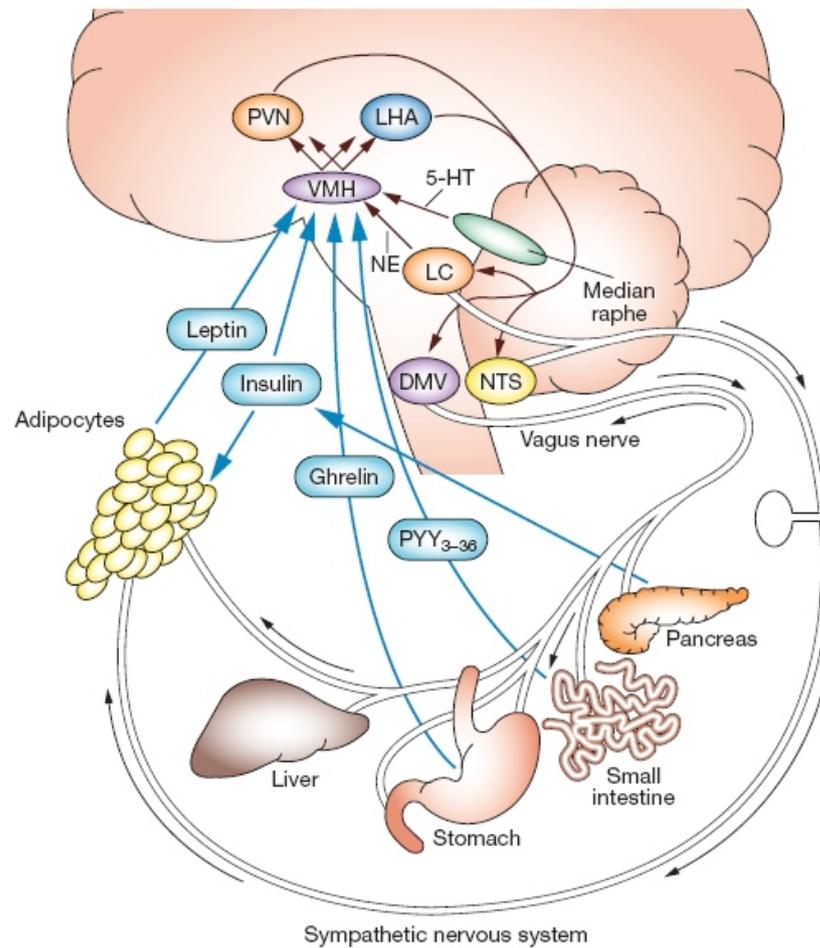


Figure 4.2: This figure shows the homeostatic pathway of energy balance adapted from Lustig *et al.*[134]. It shows the afferent(blue), central(brown) and efferent(white) pathways. The hormones insulin, leptin, ghrelin, and peptide YY_{3-36} send afferent signals relating to short-term energy metabolism and energy sufficiency to the ventromedial hypothalamus (VMH). The VMH sends anorexigenic and orexigenic signals to the melanocortin 4 receptor in the paraventricular nucleus (PVN) and lateral hypothalamic area (LHA). These signals lead to efferent signals via the locus coeruleus and the nucleus tractus solitarius, which activates the sympathetic nervous system and causes adipocytes to undergo lipolysis, or via the dorsal motor nucleus of the vagus, which activates the vagus nerve and causes energy storage, both by increasing pancreatic insulin secretion, and (in rodents) by increasing adipose-tissue sensitivity to insulin.

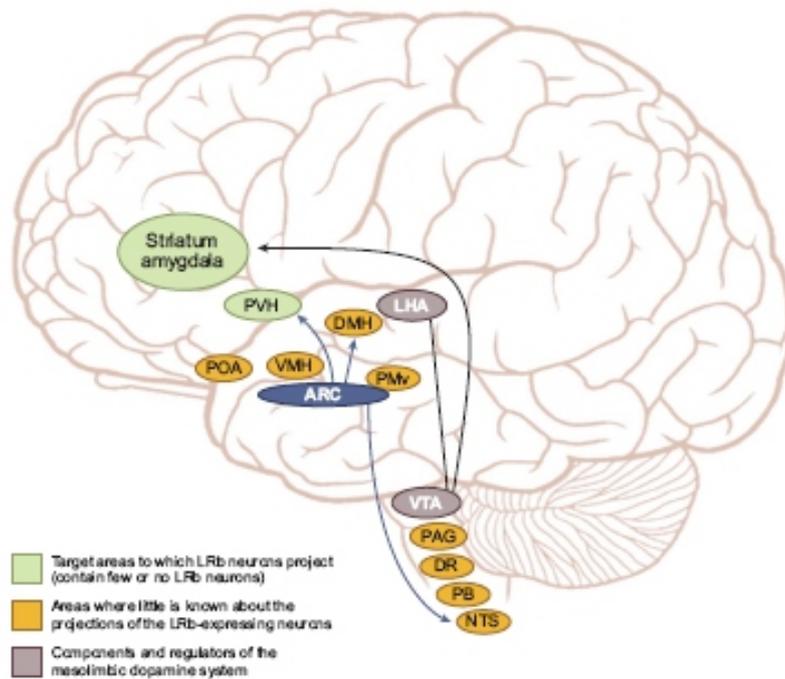


Figure 4.3: This figure was adapted from Myers *et al.* and shows the network of leptin receptor (LRb) expressing neurons. Blue, yellow and brown bubbles indicate regions containing LRb expressing neurons. Yellow bubbles are areas where little is known about the projections of these LRb expressing neurons. The arrows show projection patterns. Light green bubbles indicate areas where the LRb neurons project but do not contain any. Components and regulators of the mesolimbic dopamine system are shown in brown bubbles. ARC, arcuate nucleus; PVH, paraventricular hypothalamic nucleus; VMH, ventromedial hypothalamic nucleus; DMH, dorsomedial hypothalamic nucleus; LHA, lateral hypothalamic area; PMv, ventral premammillary nucleus; POA, preoptic area; VTA, ventral tegmental area; PAG, periaqueductal gray; DR, dorsal raphe; PB, parabrachial nucleus; NTS, nucleus of the solitary tract.[136]

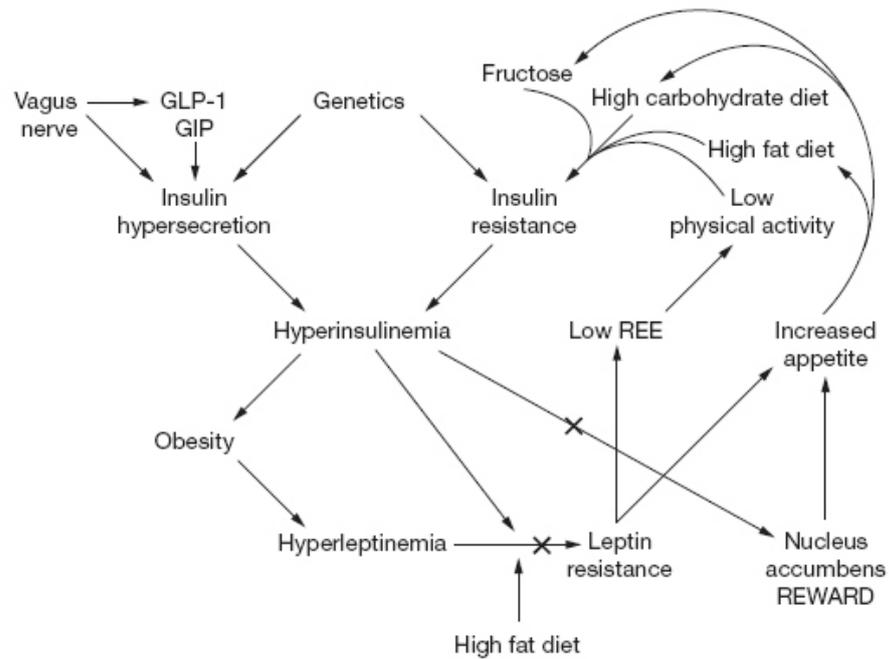


Figure 4.4: This figure was adapted from Lustig et al. [134] depicts the role of hyperinsulinemia in the dysfunction of the energy balance pathway. Various factors can lead to hyperinsulinemia, including vagus nerve mediated insulin hypersecretion or hepatic and/or skeletal muscle insulin resistance. Hyperinsulinemia can interfere with leptin signal transduction in the hypothalamus, promoting leptin resistance. This interference can cause resting energy expenditure to decrease and appetite to increase, promoting further weight gain. Hyperinsulinemia can also cause reduced dopamine uptake, which can lead to increased reward associated with eating food, again promoting further weight gain. Hyperinsulinemia can convert homeostatic and hedonic pathways from negative-feedback to feed-forward mechanisms.

that gene expression of core clock genes were down regulated in the caudal brainstem nucleus of the solitary tract (NTS) in obese mice. This perturbation of clock genes may have an effect on energy homeostasis as well as glucose and lipid metabolism. From an evolutionary standpoint it is sensible that the circadian rhythm would be involved in the regulation of hunger as historically finding food was easier and safer during the day. If metabolic processes have adjusted to this cyclical hunger cycle, one can imagine that a disturbance in a clock gene could have a cascade affect on other processes involved in energy balance.

4.1.6 Leptin During Pregnancy, Lactation and Puberty

Hyperinsulinemia and insulin resistance are both symptoms of pregnancy and puberty. Leptin levels are elevated during pregnancy and seem to be regulated by factors other than body fat content, including sex hormones such as progesterone. Leptin levels are regulated to insure the weight accrual necessary for pregnancy occurs. Variable concentrations of immunoreactive leptin are also present in human milk and the concentration is correlated with maternal adiposity [140]. It has not been determined what effect this exposure to leptin in milk in human infants might have on long-term infant growth, development and adiposity. Similarly, growth hormones during puberty increase leptin levels to insure necessary weight gain [134].

4.1.7 Leptin Differences in the Sexes

In several species, a sexual dimorphism for leptin has been shown. In humans, females have higher leptin levels than males. The absolute peak leptin concentrations are 2 times higher in women than in men [138]. This may be attributed to the fact that females have greater amounts of subcutaneous fat. Additionally, males tend to have a higher percentage of fat-free mass and this will further lower leptin concentrations relative to females. Differences can also be attributed to the stimulation of leptin by oestrogen in females and suppression of leptin by testosterone in males. The gender difference may be influenced by the amount of leptin released or removed per unit

time, which may suggest that women are more resistant to leptin feedback than men. This may explain the female's greater susceptibility to eating disorders and body weight regulation [123].

This sexual dimorphism was mentioned by Vickers *et al.* 2000 [45], 'the programmed phenotype is expressed only in the presence of the high-fat postweaning diet, whereas in males it was previously shown that programming could manifest independently of nutrition'. It was also seen by Vickers in the study discussed in this chapter, [57], and in a similar study of male rats at day 110 [58]. Neonatal leptin treatment promoted obesity in AD male offspring, particularly when given a postweaning HF diet. Whereas in UN males, leptin prevented diet-induced obesity if males were fed a chow postweaning diet. This contrasted with the data in females in which neonatal leptin treatment had no significant effect on body composition or metabolism, regardless of the postweaning diet. Neonatal leptin treatment protected UN females from becoming obese on both a HF and a standards chow diet [58]. Clegg *et al.* suggests male rats are more sensitive to the anorectic effects of insulin, whereas females are more sensitive to changes in serum leptin levels [141]. This may be indicative of altered leptin set-points rather than leptin resistance.

4.1.8 Overview

This background highlights many of the major pathways that may be affected by a leptin treatment. It has provided some insight into the relationship between the feedback mechanism to brain in relationship to food and energy balance. It has highlighted the involvement of leptin in energy homeostasis, feeding behaviour, fat storage, hunger and behaviour toward food. It has mentioned the difference between leptin in males and females. This introduction gives a platform in which to consider the complexity of the results. Many pathways may have been affected by the treatments in this study and therefore additional studies would need to be conducted to tease out the details. However, this study provides a vast amount of food for thought and definitely reveals the importance of the hypothalamus.

4.1.9 Initial Study

The aim of the leptin reversal study conducted by Vickers *et al.* [57] was to establish whether neonatal leptin treatment can alleviate postnatal obesity and the associated metabolic syndrome that occurs in the offspring of undernourished mothers. Methods are described in that paper, but will also be summarised here. The study design is illustrated in figure 4.5.

The study utilised the same model of undernourishment as the study in Chapter 3. Virgin Wistar rats (age 100 ± 5 days) were time mated using a rat oestrous cycle monitor to assess the stage of oestrous of the animals prior to introducing the male. After confirmation of mating, rats were housed individually in standard rat cages with free access to water. All rats were kept in the same room with constant temperature maintained at $25\text{ }^{\circ}\text{C}$ and a 12-h light: 12-h darkness cycle.

Animals were randomly assigned to one of two nutritional groups: a) undernutrition (30% of ad-libitum) of a standard diet throughout gestation (UN group), b) standard diet ad-libitum throughout gestation (AD group). Food intake and maternal weights were recorded daily until the end of pregnancy. After birth, pups were weighed and litter size was adjusted to 8 pups per litter to assure adequate and standardised nutrition until weaning. A minimum of six litters per group were used. At postnatal day 3 female AD and UN pups were randomised to receive either saline or recombinant rat leptin ($2.5\text{ }\mu\text{g/g}\cdot\text{d}$) for 10 days by subcutaneous injection. Pups from undernourished mothers were cross-fostered onto dams that had received AD feeding throughout pregnancy. All animals were fed ad-libitum until weaning (day 22) and then animals were weight matched and put on either a standard rat chow diet (CHOW group) or high-fat diet (45% kcal as fat; HF group). Weight matching was done to ensure that the mean starting weight was similar for the two post natal diets. In each litter of 8 pups, there were four males and four females. Pairs of females were randomly placed on either the chow or HF diet. Mean weight of the cages were checked to ensure that the offspring placed into the HF group were not statistically different from those allocated to the chow diet. The animals were culled and tissue

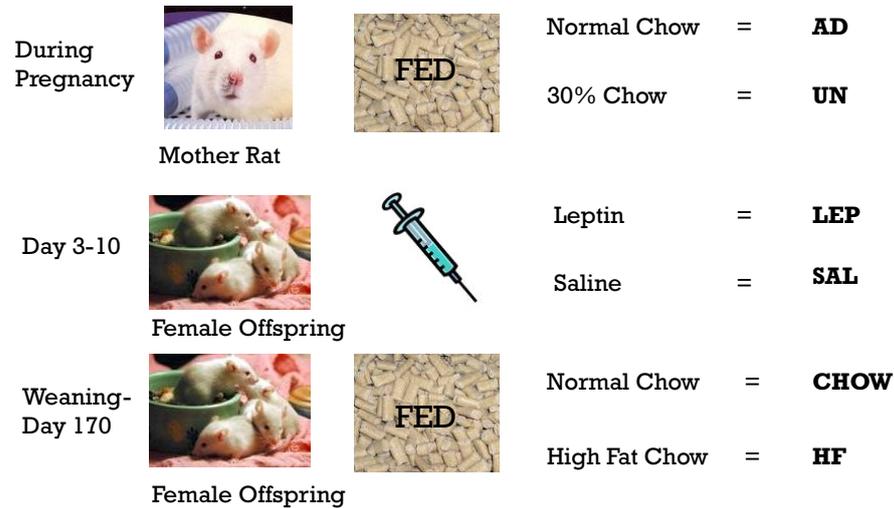


Figure 4.5: Design of Leptin Reversal Study. This figure summarises the study design. During pregnancy rats were fed either ad libitum a standard chow diet (AD) or 30% of a standard chow diet (UN). The female offspring were given either leptin (LEP) or saline (SAL) injections from day 3-10, and then from weaning offspring were fed either ad libitum a standard chow diet (CHOW) or a high fat chow diet (HF).

samples were collected at postnatal day 170. The results of the study are described in detail by Vickers *et al.* [57], but are summarised here.

Phenotypic Assessment

Physiological measurements were taken as described in [57]. These values are summarised in table 4.1. The values for insulin, leptin and c-peptide are shown in figure 4.6 taken from Vickers *et al.*.

The study demonstrates that all the measured metabolic consequences of maternal undernutrition were reversed by a period of neonatal leptin treatment in female rats. Neonatal leptin treatment resulted in reduced pup weight gain for both AD and UN offspring. The high fat diet after weaning significantly increased weight gain with a marked amplification of the weight gain in the UN pups. The leptin treatment normalised the diet-induced weight gain in the UN animals to match the AD animals. These results indicate that leptin induced a phenotype reversal and possibly a reversal of epigenotype. The female infant rats that had not been prenatally nutritionally programmed had no phenotypic effects, but those who had been nutritionally programmed did not develop the metabolic phenotype even when placed on a high fat

diet after weaning. The findings indicate that, at least in the rat, there is an early postnatal window during which the process can be reversed. This animal model resembles the increased weight gain and metabolic abnormalities seen in humans born small for gestational age (SGA), thus the mechanisms responsible for the effects in rats may be similar in humans.

The results of this study indicate that developmental adaptations during fetal life can be reversed by interventions in the neonatal period. The goal of this chapter is to use the samples from Vickers *et al.* [57] to investigate the molecular mechanisms that were affected by the prenatal diet, the leptin treatment and the postweaning diet. We hypothesised that gene expression and methylation studies would provide information on mechanisms and may enable us to pinpoint the postnatal age after which leptin intervention may be ineffective.

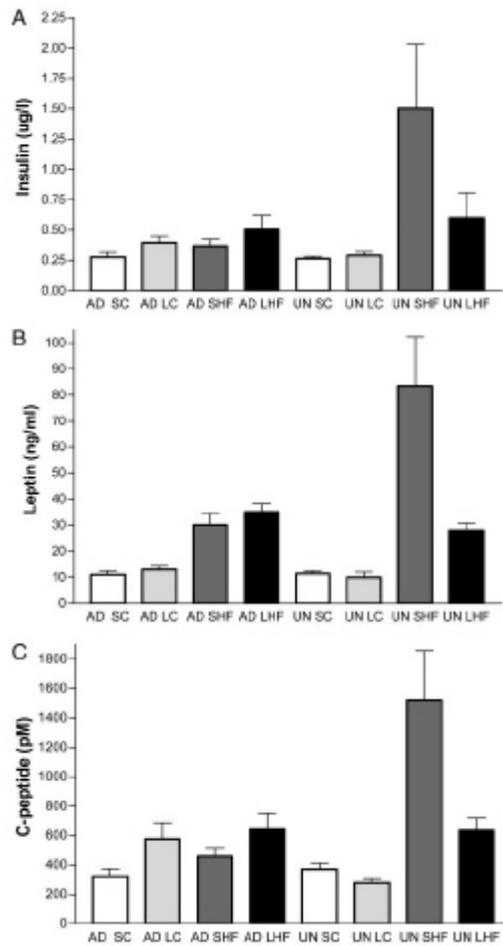


Figure 4.6: (A) Fasting plasma insulin concentrations at day 170 in AD and UN animals on either a chow (C) or high-fat (HF) diet after neonatal saline or leptin treatment. Neonatal leptin treatment normalised fasting plasma insulin concentrations in UN leptin-treated HF (UNLHF) animals compared with saline-treated animals ($P < 0.05$ for effect of programming and diet; programming \times treatment, programming \times diet, and treatment \times diet interactions, $P < 0.05$) (B) Fasting plasma leptin concentrations at day 170. Neonatal leptin treatment normalised fasting plasma leptin concentrations in UNLHF animals compared with saline-treated animals ($P < 0.05$ for effect of diet and treatment; programming \times diet, programming \times treatment, treatment \times diet, and programming \times treatment \times diet interactions, all $P < 0.05$). (C) Fasting plasma C-peptide concentrations at day 170. Neonatal leptin treatment normalised fasting plasma C-peptide concentrations in UNLHF animals compared with saline-treated animals ($P < 0.05$ for effect of programming; $P < 0.0001$ for effect of diet; programming \times treatment, programming \times diet, and treatment \times diet interactions, $P < 0.05$). Data were analysed by three-way factorial ANOVA, mean \pm SEM; $n = 8$ per group. Result from Vickers *et al.* [57]

Measurement	ASC	ASHF	ALC	ALHF	USC	USHF	ULC	ULHF
Body Weight (grams)	335	370.25	330.25	375.25	322.63	387.25	285.5	332.38
Body Length (mm)	218.88	220.13	218.25	215.88	204.75	212.75	203.25	209.25
Total Fat (% of body weight)	30.21	40.56	33.89	42.3	35.55	52.55	28.38	42.64
Leptin (ng/ml)	11.09	26.58	16.83	37.58	25.93	69.05	11.63	28.43
Insulin (ng/ml)	0.27	0.38	0.38	0.51	0.34	1.84	0.30	0.66
C-Peptide (pg/ml)	319	444.88	551.75	658.13	470.88	1398.5	320.63	611.38
Total Ghrelin (fmol/ml)	4588.38	3051.25	4066.75	2428.63	4090.38	2926.25	3901.63	2862.38
Liver Weight (% of body weight)	2.52	2.24	2.51	2.37	2.52	2.33	2.46	2.25
Liver Glycerol	3.68	4.13	3.70	4.40	4.83	6.17	3.98	4.55
Liver Triglycerides	19.84	37.37	24.08	39.44	27.76	48.81	33.42	43.32
Plasma Triglycerides	1.19	0.96	1.01	0.89	0.86	1.05	1.34	0.71
Plasma Glycerol	0.28	0.29	0.27	0.31	0.31	0.36	0.28	0.32
Plasma Free Fatty Acids (mmol/l)	0.91	0.87	0.83	0.77	0.79	1.03	0.95	0.83
Kidney Weight (% of body weight)	0.69	0.64	0.62	0.61	0.60	0.55	0.63	0.63
Adrenal Weight (% of body weight)	0.02	0.02	0.03	0.02	0.02	0.02	0.03	0.02
Heart Weight (% of body weight)	0.31	0.30	0.31	0.31	0.28	0.29	0.33	0.31
Spleen Weight (% of body weight)	0.22	0.21	0.21	0.20	0.22	0.21	0.23	0.20

Table 4.1: Phenotypic measurements relevant to metabolic syndrome measured for each of the eight animals in each treatment group at day 170.

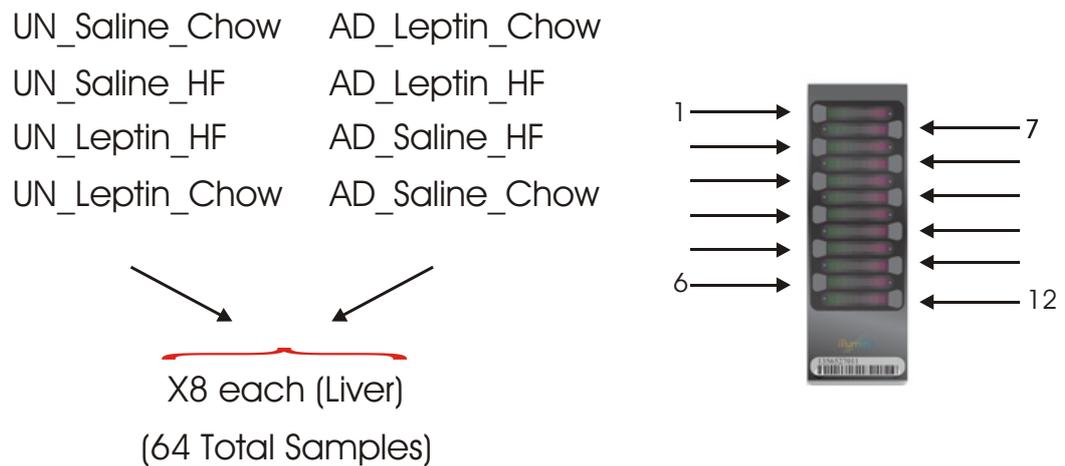


Figure 4.7: Microarray Design. This study included eight separate treatment groups. For each of these groups there were eight biological replicates. RNA extracted from the liver of each of these replicates was hybridised on an Illumina Rat-Ref12 chip. Each chip has 12 separate arrays; 6 chips were used in total.

4.2 Results

4.2.1 Illumina Microarray

On two separate visits to the Liggins Institute, RNA was extracted from eight biological replicates in each of eight treatment groups for both liver and skeletal muscle tissue. Extracted RNA samples were shipped to the University of Cambridge on dry ice along with the remnant phenol-interphase mix for subsequent DNA extraction. Only the liver samples were hybridised to the Illumina chip.

Hybridisations

RNA extracted from liver was hybridised to Illumina microarrays as described in Chapter 2. Extractions, amplifications, and hybridisations were randomised. The microarray study design is shown in figure 4.7.

Data Quality Control

The Illumina system incorporates a number of internal controls for estimating hybridisation quality. These controls are described in more detail in Chapter 2. Figure 4.9 shows the details of the five graphs for the internal controls across all

64 arrays. In figure 4.9A the low, medium and high intensities are shown across all arrays. Ideally, arrays should have signal intensities in the same range. It is expected that there be a linear increase of signal between the three intensities. In figure 4.9B low stringency of the data is analysed by comparing the signal intensity of two mismatches to the signal intensity of a perfect match. The perfect matches should have a 3–4 fold higher signal as seen in this data. No signal intensity indicates the hybridisation failed. Figure 4.9C shows the high stringency control of the data. For high quality data it is expected that the biotin signal would be 3–4 fold lower than the high stringency signal as seen here. Figure 4.9D shows the data for the negative control. The background signal should be approximately 70. This value is accounted for in the normalisation algorithms. The noise signal should be very low for high quality data as seen in this data. Figure 4.9E compares the signal from the housekeeping genes to the signal from all genes. As housekeeping genes are always expressed they should have a much higher signal compared with all genes. The overall quality of this data set was generally high.

To compare the samples in more detail the data was exported from BeadStudio and imported into R Bioconductor. Here the raw data was plotted in several graphs (pairwise, MA, density plot and box plot), then normalised and replotted to compare pre- and post-normalisation. In addition, the arrays were clustered using hierarchical clustering. The hierarchical clustering is shown in figure 4.8. The samples appear to be distributed evenly, not clustering on biological replicate or similar treatments. The sample for UNSALCHOW 28.0.0 appears to be an outlier.

Figures 4.10 to 4.17 show pairwise comparisons for each of the eight treatment groups. A tight diagonal line is expected for replicates, as this indicates low variability between samples. Figures 4.18 and 4.19 show plots of all samples before normalisation. These reveal that there is variation in intensity between arrays and therefore a normalisation is necessary.

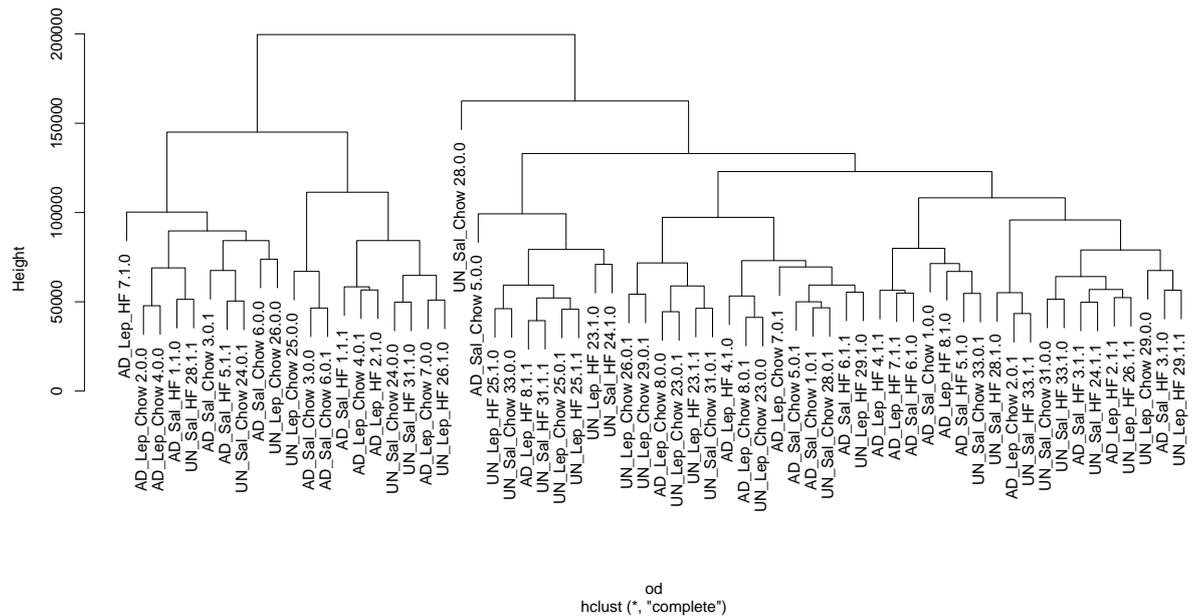


Figure 4.8: This figure shows the hierarchical clustering for all 64 arrays before normalisation. Arrays do not seem to cluster on any extraneous factor, nor does one of the treatments cause enough difference for them to cluster together. UNSALCHOW 28.0.0 appears to be an outlier.

Normalisation/Differential Expression

Following the quality control assessment, all samples were confirmed to be of good quality and the data was filtered for genes present on at least one array. The original 22,517 genes were filtered down to 10,774, by this low stringency filtering method. Data was then transformed with a variance stabilising transformation and normalised with a quantile normalisation (as done in Chapter 3 and described in Chapter 2). Figures 4.20 and 4.21 show a boxplot and density plot of all samples after normalisation.

Due to the complexity of the study design, several methods of analysis were implemented to find the most biologically relevant method appropriate for interpreting the data. Initially, the FSPMA method was used (as described in Chapter 2). This method utilises a mixed model ANOVA in R and can be rerun to look at the effects of different treatments. FSPMA was run three times, once for each of the treatments (maternal diet, leptin treatment, and postweaning diet). The lists of differentially expressed genes were subjected to a FDR adjustment and genes with a p-value <

Summary of Illumina Internal Controls For Leptin Data

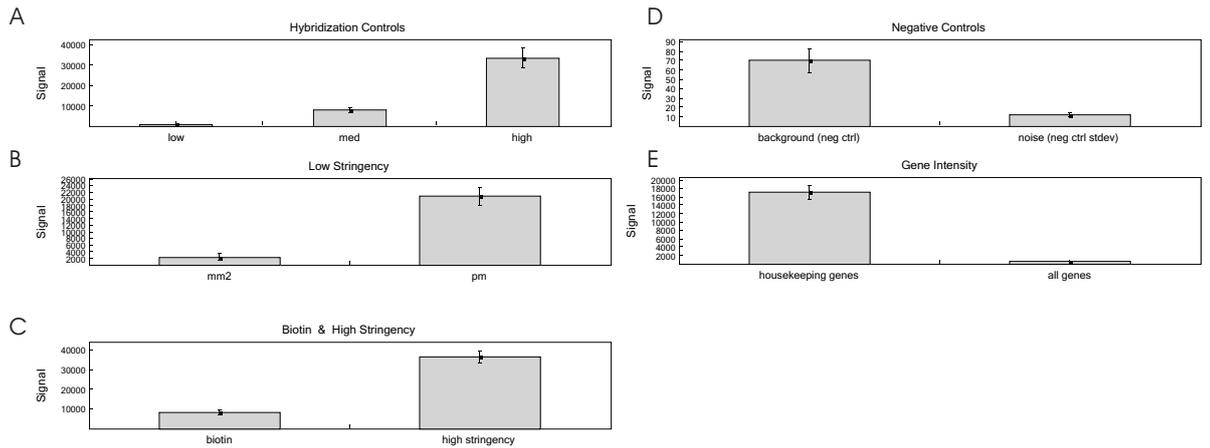


Figure 4.9: Illumina Internal Controls. This figure shows the details of the internal quality control tools for the Illumina platform. (A) Low, medium, and high intensities across arrays. Ideally, arrays should have signal intensities in the same range. It is expected that there be a linear increase of signal between the three intensities. (B) High stringency of the data. For high quality data it is expected that the biotin signal would be 3–4 fold lower than the high stringency signal as shown for this data. (C) Low stringency of the data by comparing the signal intensity of two mismatches to the signal intensity of a perfect match. The perfect matches should have a 3–4 fold higher signal as shown for this data. No signal intensity indicates the hybridisation failed. (D) Data for the negative control. The background signal should be approximately 70. This value is accounted for in the normalisation algorithms. The noise signal should be very low for high quality data. (E) Comparison of the signal from the housekeeping genes to the signal from all genes. As housekeeping genes are always expressed they should have a much higher signal compared with all genes. This shows that this data was generally high quality.

Pairwise Plot for Sample Correlation for ADSALCHOW Samples

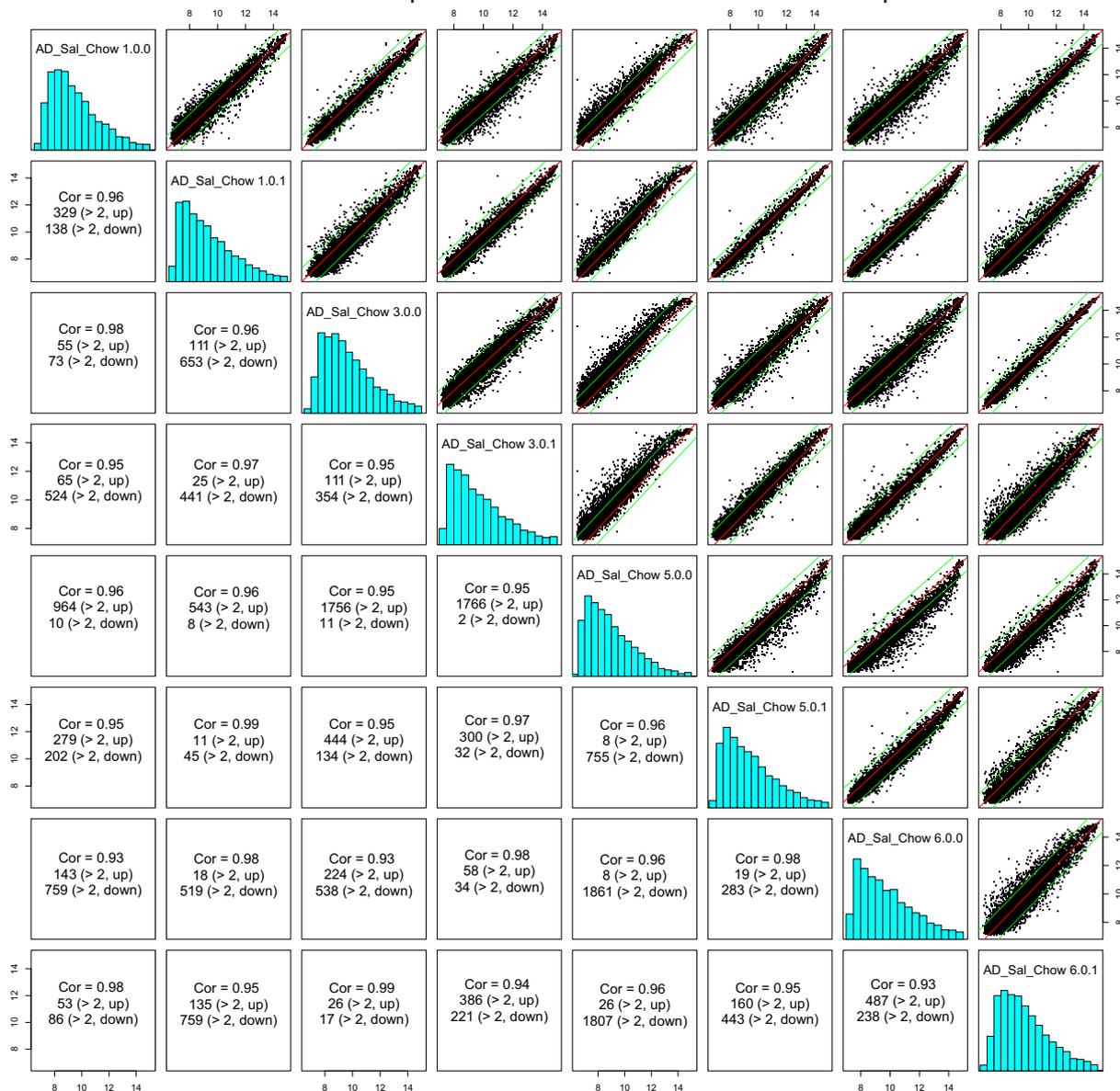


Figure 4.10: Pairwise Comparison of ADSALCHOW Samples Before Normalisation. This plot compares all eight ADSALCHOW biological replicates against each other. These replicates should be very similar and should be tightly distributed on the $x=y$ axis, as these samples are. The correlations for these comparisons are ≥ 0.93 .

Pairwise Plot for Sample Correlation for UNSALCHOW Samples

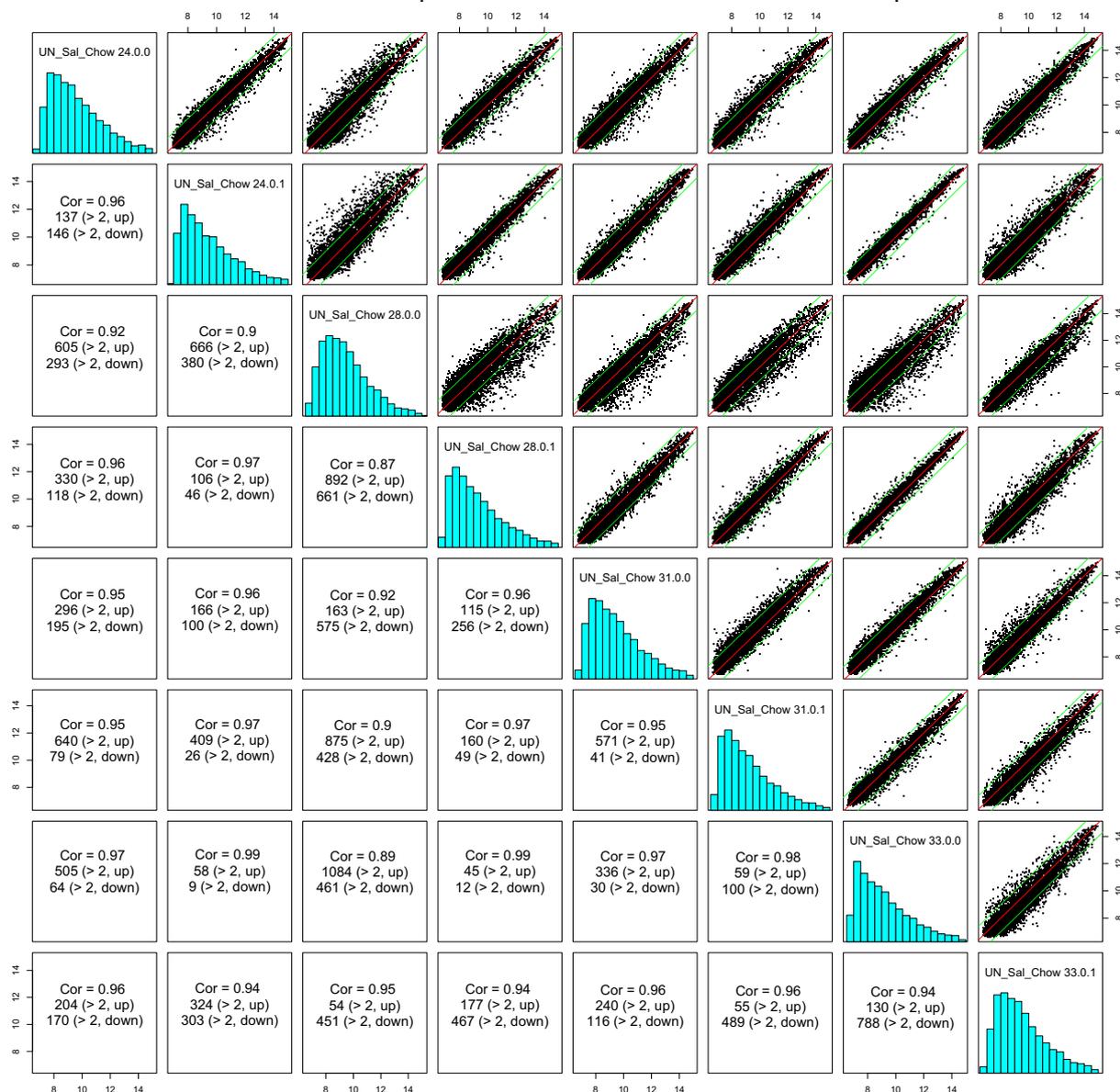


Figure 4.11: Pairwise Comparison of UNSALCHOW Samples Before Normalisation. This plot compares all eight UNSALCHOW biological replicates against each other. These replicates should be very similar and should be tightly distributed on the $x=y$ axis, as these samples are. The comparisons show higher variability of sample 28.0.0. All other comparisons have a correlation ≥ 0.94 .

Pairwise Plot with Sample Correlation for ADLEPCHOW Samples

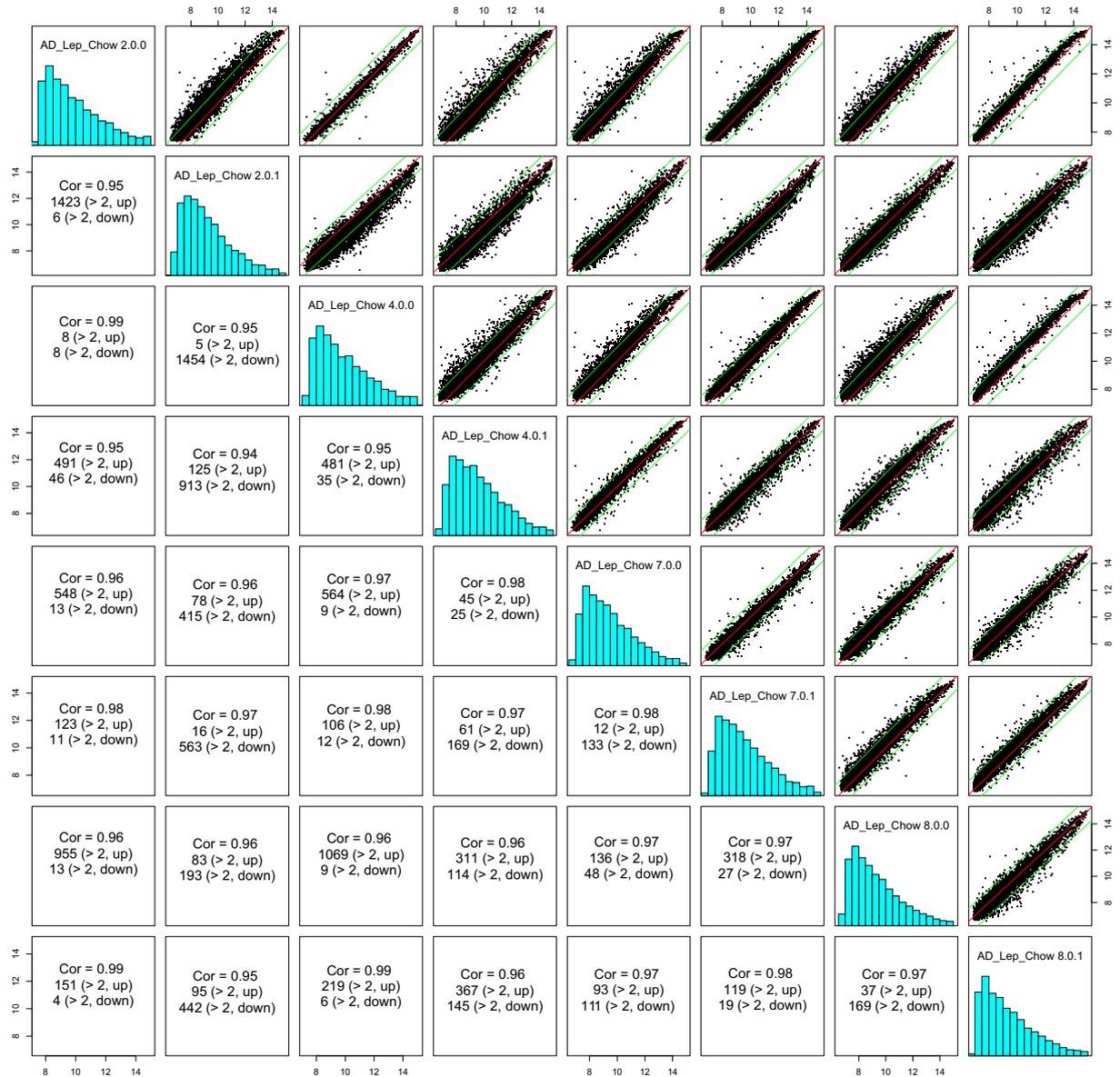


Figure 4.12: Pairwise Comparison of ADLEPCHOW Samples Before Normalisation. This plot compares all eight ADLEPCHOW biological replicates against each other. These replicates should be very similar and should be tightly distributed on the $x=y$ axis, as these samples are. The correlations for these comparisons are ≥ 0.94 .

Pairwise Plot for Sample Correlation for UNLEPCHOW Samples

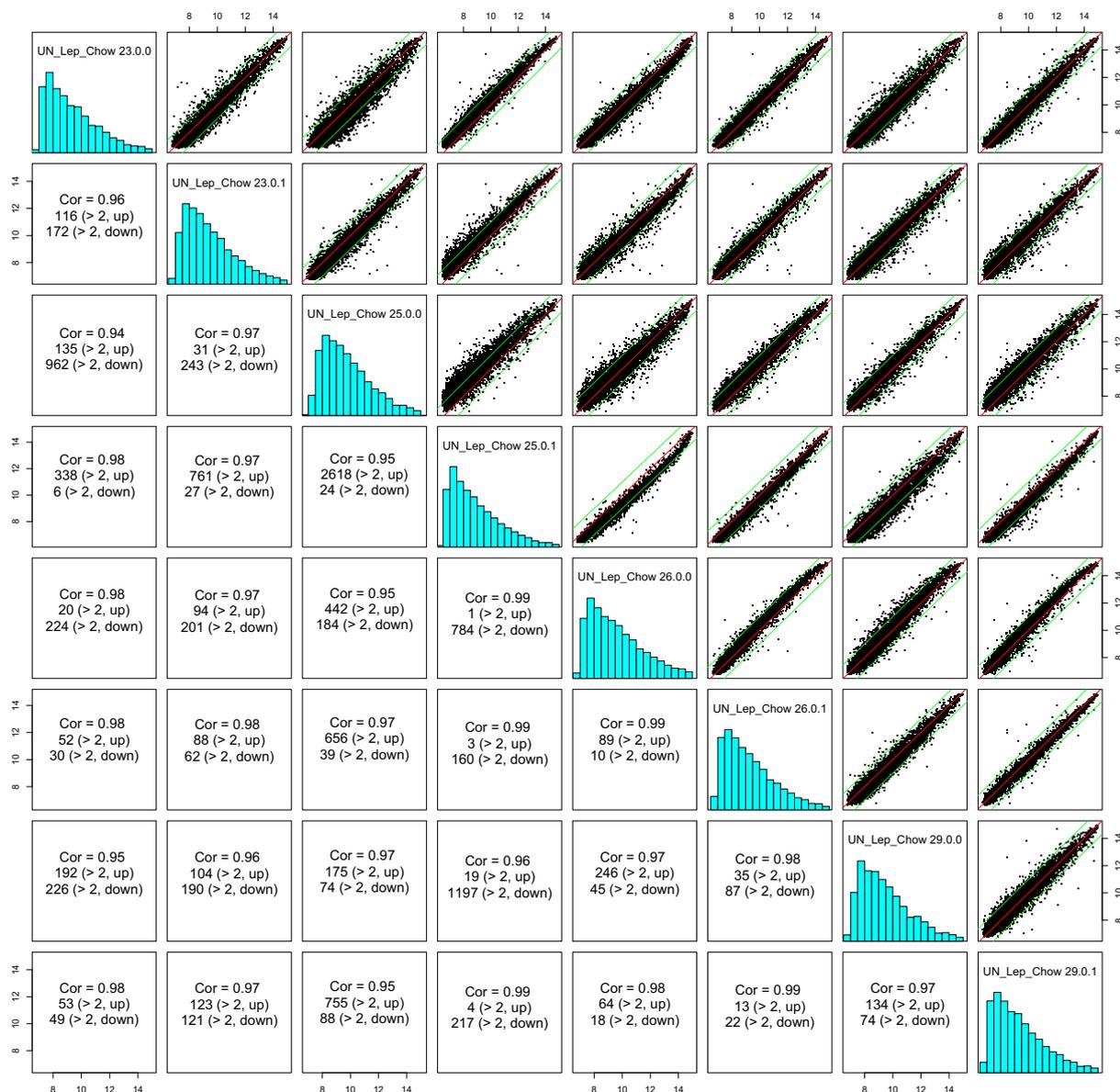


Figure 4.13: Pairwise Comparison of UNLEPCHOW Samples Before Normalisation. This plot compares all eight UNLEPCHOW biological replicates against each other. These replicates should be very similar and should be tightly distributed on the x=y axis, as these samples are. The correlations for these comparisons are ≥ 0.94 .

Pairwise Plot for Sample Correlation for ADSALHF Samples

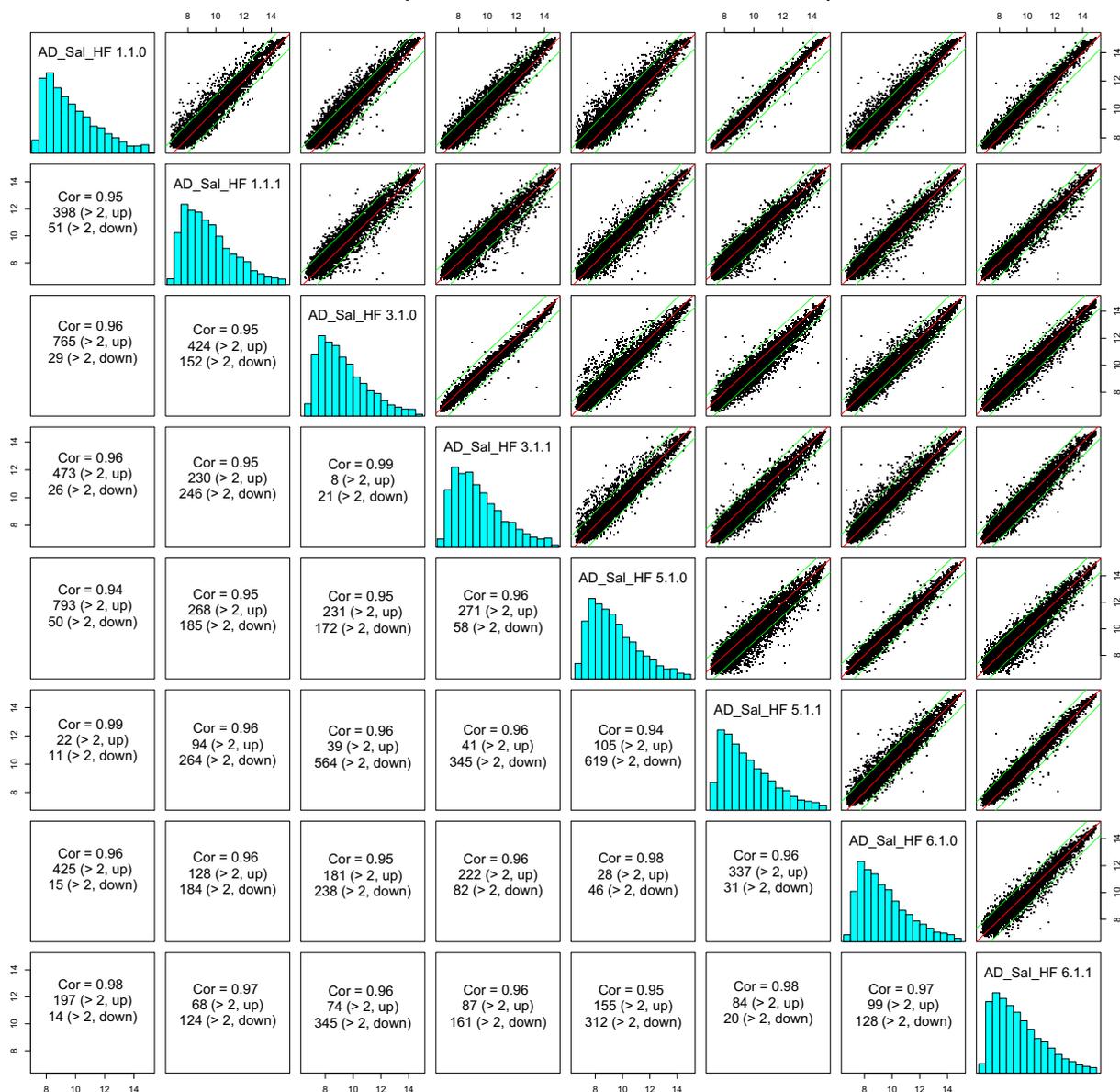


Figure 4.14: Pairwise Comparison of ADSALHF Samples Before Normalisation. This plot compares all eight ADSALHF biological replicates against each other. These replicates should be very similar and should be tightly distributed on the $x=y$ axis, as these samples are. The correlations for these comparisons are ≥ 0.94 .

Pairwise Plot for Sample Correlation for UNSALHF Samples

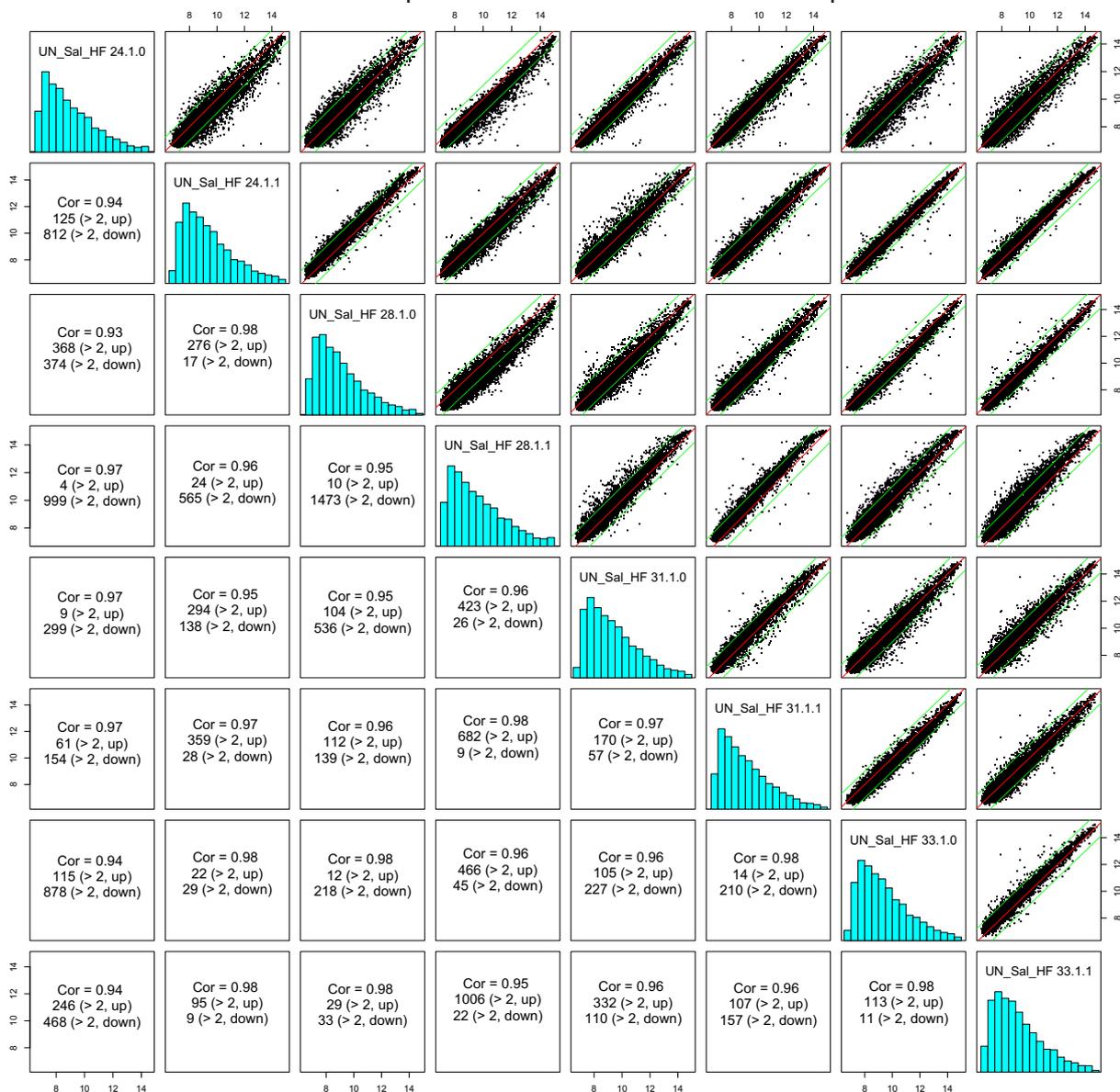


Figure 4.15: Pairwise Comparison of UNSALHF Samples Before Normalisation. This plot compares all eight UNSALHF biological replicates against each other. These replicates should be very similar and should be tightly distributed on the $x=y$ axis, as these samples are. The correlations for these comparisons ≥ 0.93 .

Pairwise Plot for Sample Correlation for ADLEPHF Samples

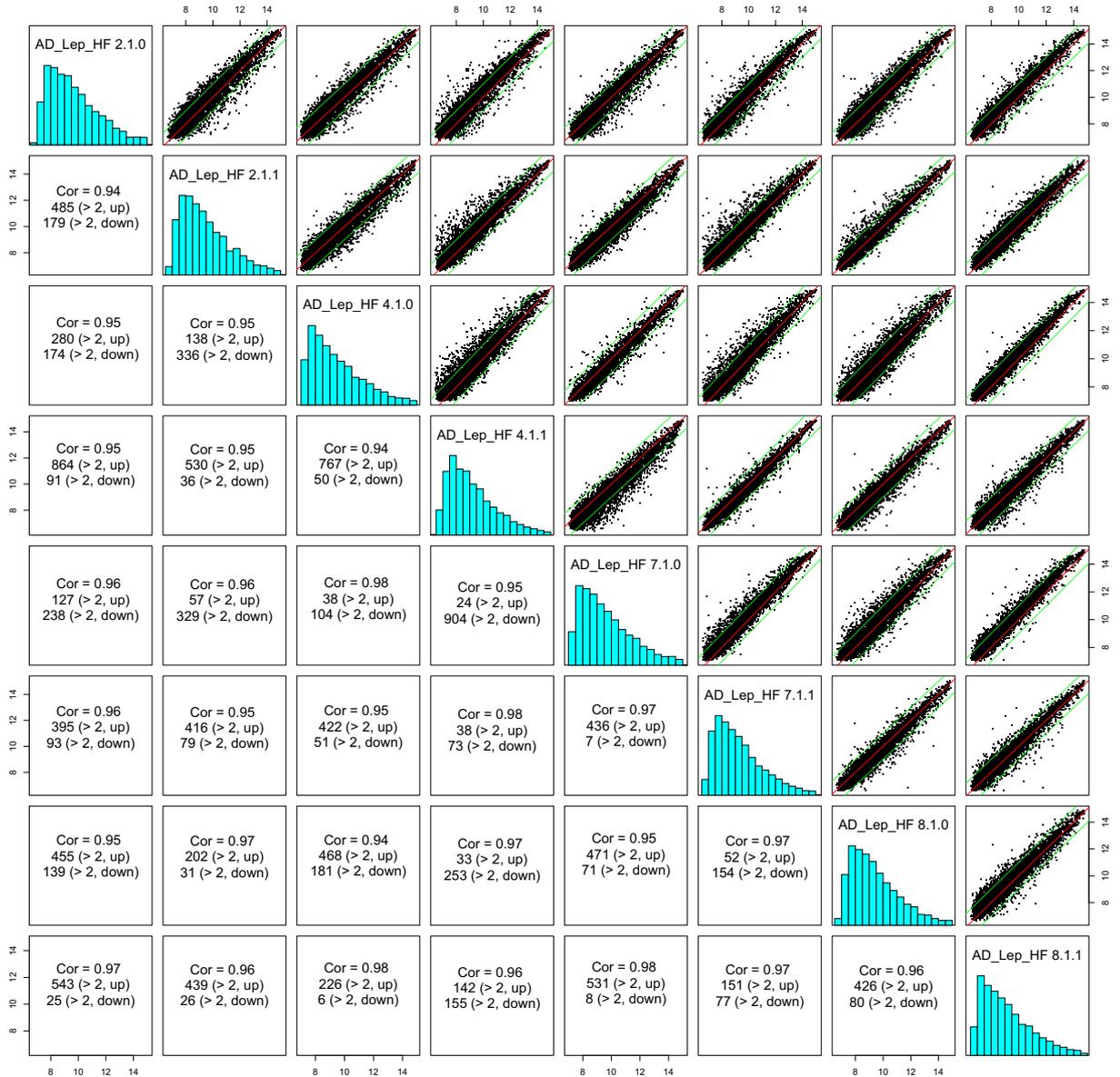


Figure 4.16: Pairwise Comparison of ADLEPHF Samples Before Normalisation. This plot compares all eight ADLEPHF biological replicates against each other. These replicates should be very similar and should be tightly distributed on the $x=y$ axis, as these samples are. The correlations for these comparisons ≥ 0.94 .

Pairwise Plot for Sample Correlation for UNLEPHF Samples

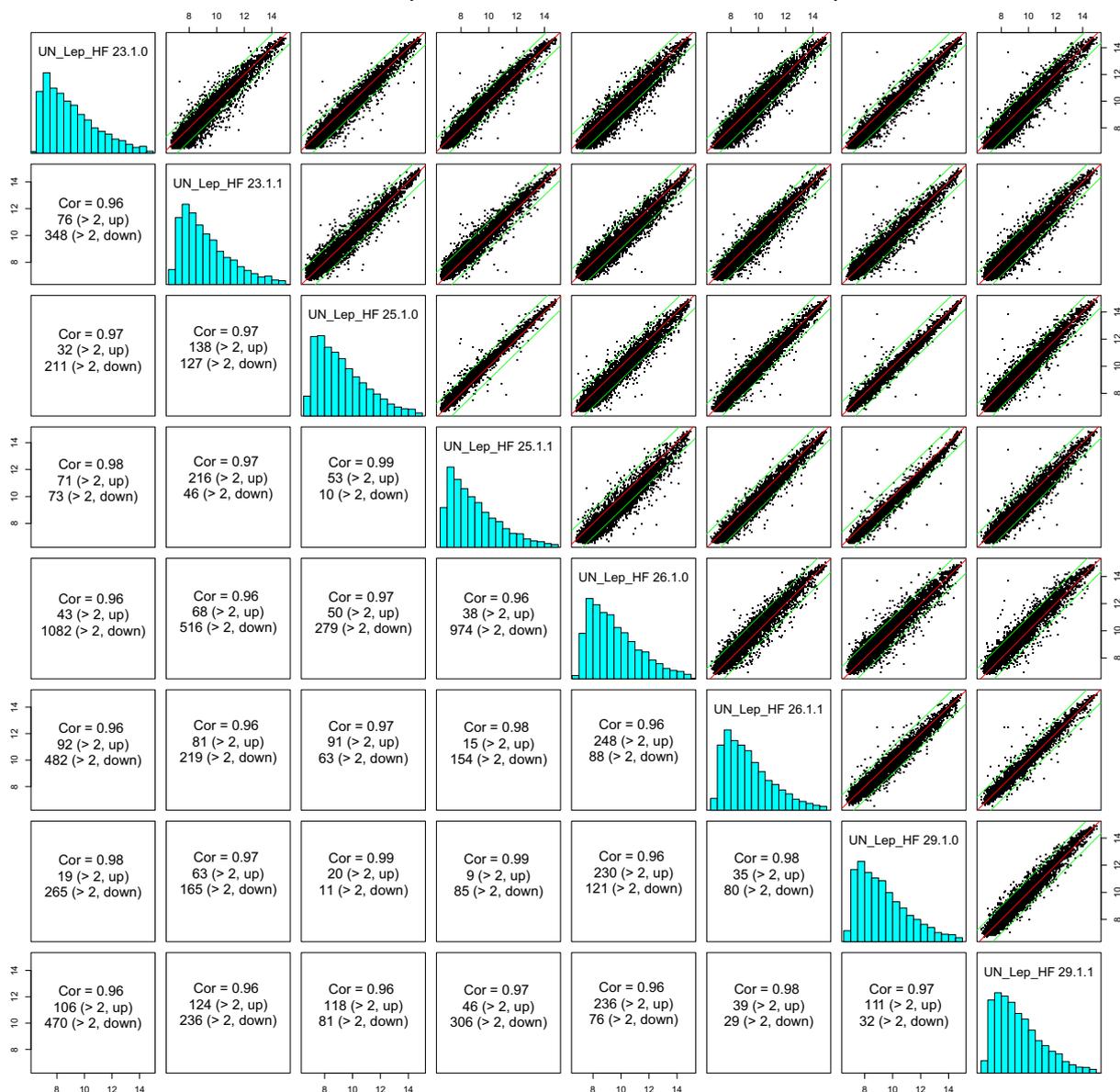


Figure 4.17: Pairwise Comparison of UNLEPHF Samples Before Normalisation. This plot compares all eight UNLEPHF biological replicates against each other. These replicates should be very similar and should be tightly distributed on the $x=y$ axis, as these samples are. The correlations for these comparisons ≥ 0.96 .

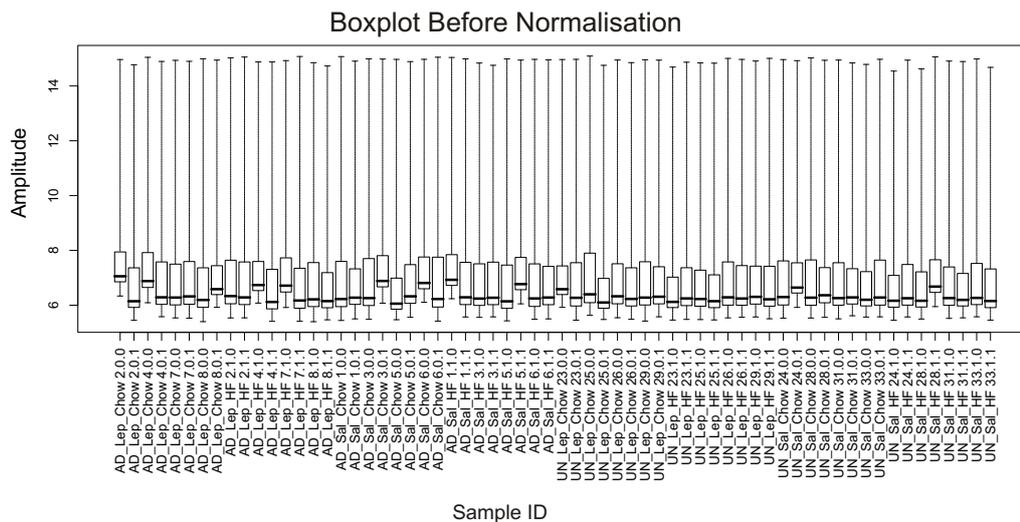


Figure 4.18: Boxplot Before Normalisation. The box itself contains the middle 50% of the data points. The line in the box indicates the median value of the data. If the median line within the box is not equidistant from the hinges, then the data is skewed. The ends of the vertical lines or ‘whiskers’ indicate the minimum and maximum data values. The plot compares the amplitude of all expression values on each array. This shows that the majority of the signal is at lower expression levels. It also shows that there is some variability between each array. This analysis was done in Lumi.

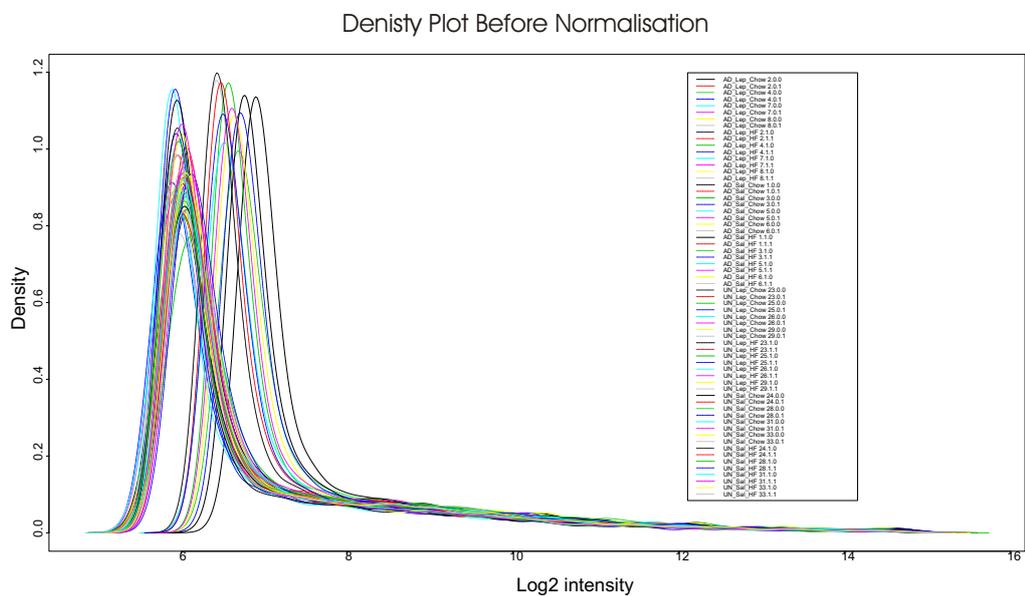


Figure 4.19: Histogram of Density of Log2 Intensities For All Arrays. This is a histogram of each array showing the density of intensities (log2). This plot shows that the majority of intensity values are low and that there is significant variation between samples. This analysis was done in Lumi.

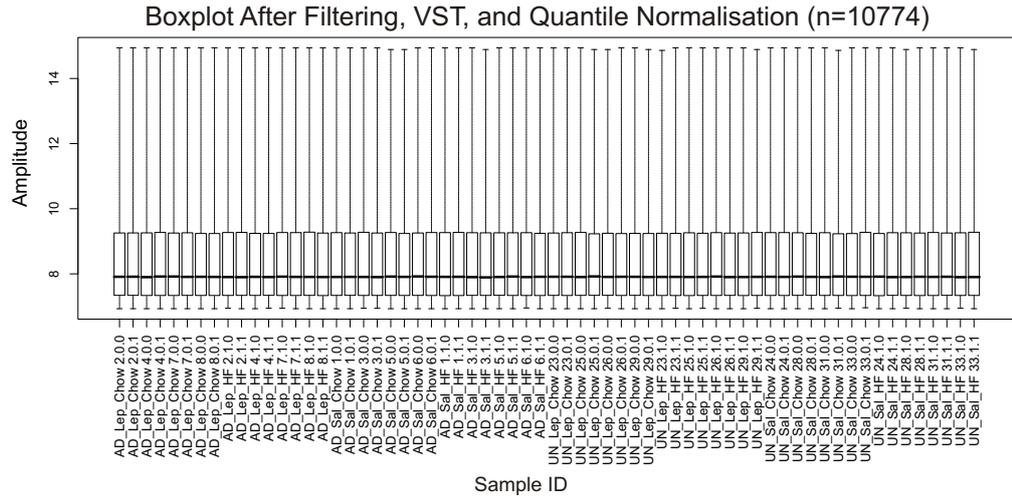


Figure 4.20: Boxplot After Quantile Normalisation. This is a boxplot of the data from each array. The plot compares the amplitude of all expression values on each array. This shows that the majority of the signal is at lower expression levels. This plot shows that there is no variability between the arrays. The normalisation has effectively adjusted the expression values to remove the variability. This analysis was done in Lumi.

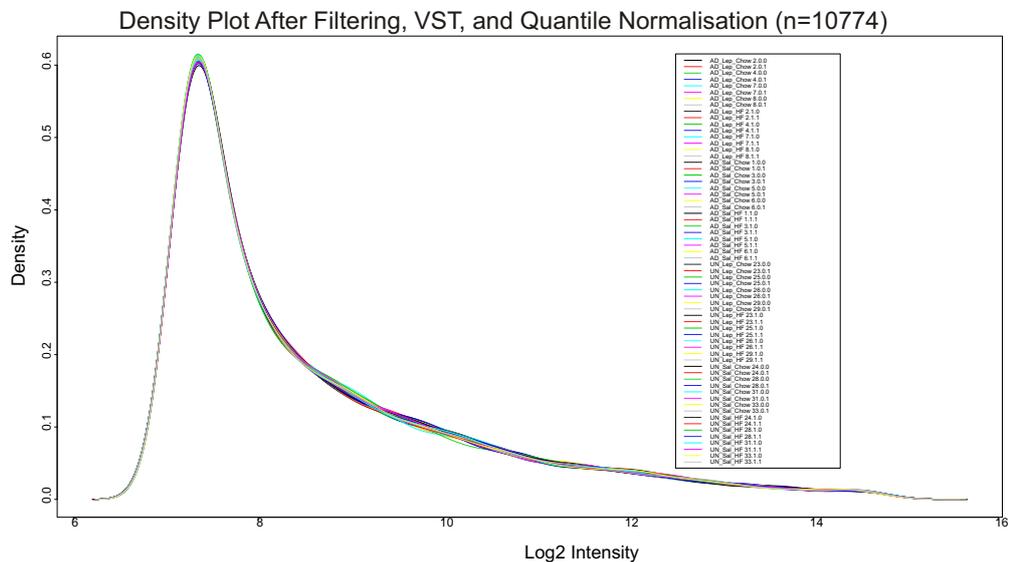


Figure 4.21: Histogram of Density of Log2 Intensities For All Arrays After Quantile Normalisation. This is a histogram of each array showing the density of intensities (\log_2). This plot shows that the majority of intensity values are low and that there is no variation between samples. The normalisation has effectively adjusted the expression values to remove the variability. This analysis was done in Lumi.

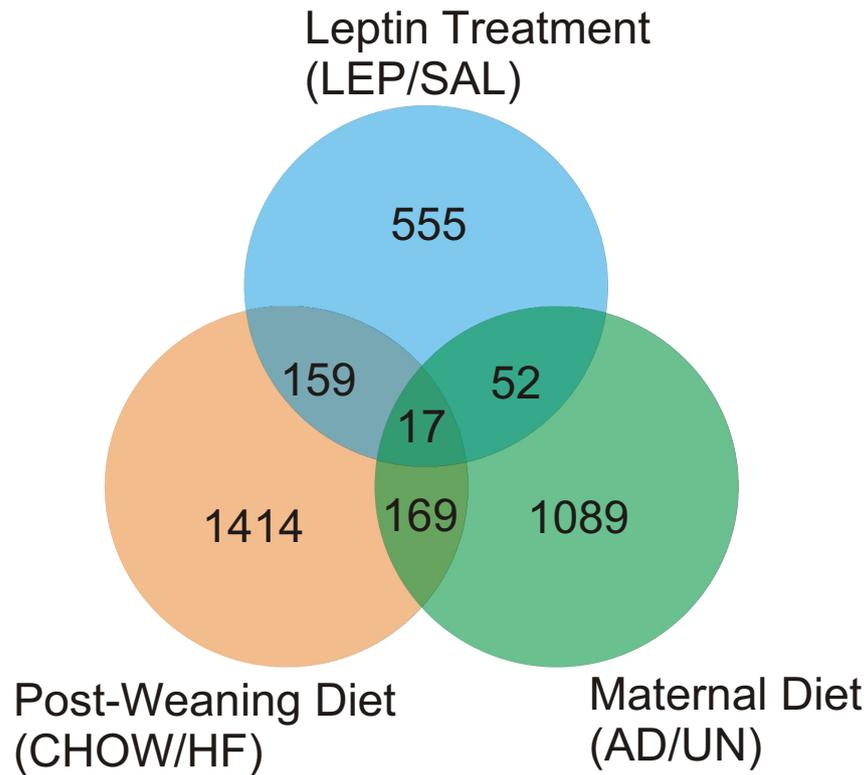


Figure 4.22: This figure shows the results of three FSPMA analyses. FSPMA uses a mixed model ANOVA to compute p-values based on a particular effect. The analysis was run three times, once for each of the treatments (maternal diet, leptin treatment, and postweaning diet). This Venn diagram shows the number of significant (p-value<0.05) genes in each list and the overlap between each list.

0.05 were selected as statistically significant for further investigation. The lists of genes affected by maternal diet, leptin treatment and postweaning diet had 1,089 genes, 555 genes and 1,414 genes respectively. There was some overlap between these lists as shown on the venn diagram in figure 4.22, including 17 genes that appeared on all three lists.

The second analysis involved pair-wise comparisons performed in Limma as utilised in Chapter 3. To compare the 8 different treatment groups against each other resulted in 28 separate comparisons. The number of genes for each of these comparisons is shown in figure 4.23. Of these comparisons, 13 revealed one or more significantly expressed genes. The combined list revealed a total of 160 significantly (p-value < 0.05) differentially expressed genes. Of the 160 genes, 55 had also been revealed by

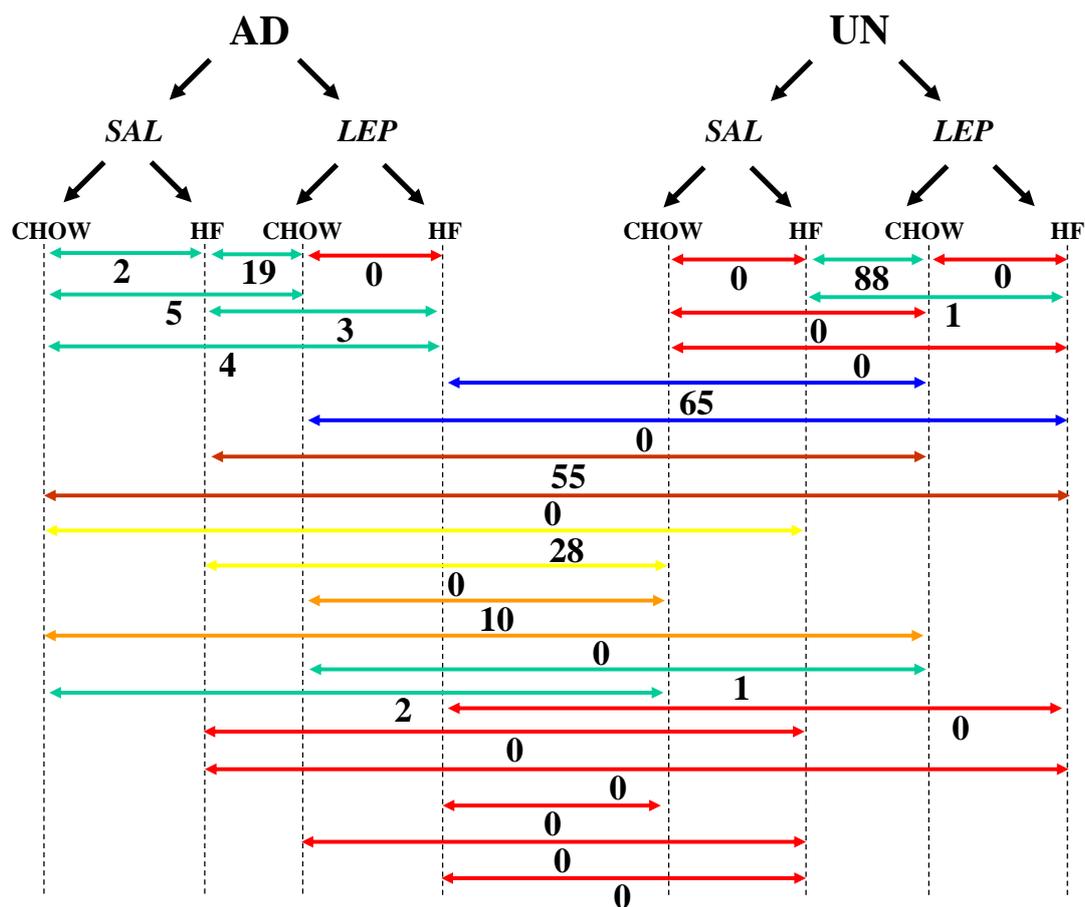


Figure 4.23: This figure shows the results of the pairwise comparisons between each combination of treatment groups. The groups are shown at the top and the number shown below the arrow connecting two treatment groups indicates the number of genes that were significantly, differentially expressed for that comparison.

the FSPMA analysis

A final analysis was done in KDE. It was a simple one-way ANOVA with an FDR multiple testing correction. This revealed 99 significant genes. Many of the genes found in the other two analyses (especially FSPMA) revealed genes that had very small expression differences, thus we added a filtering regime to the one-way ANOVA in an effort to remove outliers and focus on genes with larger fold-changes. The filtering removed individual outliers which were > 1.96 standard deviations from the mean of the 8 biological replicates. Genes were excluded from ANOVA if they were detected in fewer than 16 of the 64 samples. In this list, we found 79 of the 99 genes from the first one-way ANOVA and 139/160 genes found in the pair-wise comparisons. Using this method, 334 genes were found with a fold-change between the maximum

and minimum values that was less than 25% after averaging the biological replicates and 127 genes were found when the fold-change difference was less than 50% after averaging the biological replicates. By eliminating outliers in this way, the error bars were tightened and a less stringent FDR correction could be used. It also allowed us to focus on genes with the largest fold-change differences. Gene lists for all analyses can be found in Appendix C (FSPMA results, Tables 7-9; Pairwise results, Tables 10-22; one-way ANOVA results, Table 23).

Hierarchical Clustering

Hierarchical clustering was performed on the 127 genes with an ANOVA p-value less than 5% and at least a 50% change in gene expression between the highest and lowest values. The data was normalised to a median value of 0 to allow clustering. The hierarchical clustering dendrogram is shown in figure 4.24. The hierarchical clustering analysis separated genes into 10 clusters based on the expression patterns between treatment groups. By plotting the expression values in a dendrogram (see figure 4.24), generalisations can be made about the expression patterns in each cluster. In this figure, each row represents a gene and each column a treatment group. Cluster 0 (the one closest to the top) consists of three genes. These appear to have gene expression induced by a HF diet in the offspring of AD mothers and repressed by a HF diet in the offspring of UN mothers. There also may be some interaction with leptin. Cluster 1 has 10 genes and is induced by a HF diet in the offspring of AD mothers and is repressed by a HF diet in the offspring of UN mothers. Cluster 2 has 6 genes and is induced by the combination of leptin treatment and a postnatal chow diet in the offspring of both AD and UN mothers. Clusters 3 and 4 have only 3 genes and 2 genes respectively; therefore, they are difficult to interpret. Cluster 5 has 28 genes and the gene expression is induced by leptin. Leptin induction in this cluster is reduced in the UNLEPHF group. Cluster 6 has 10 genes and gene expression is induced by UN and repressed by AD with a possible leptin effect. Clusters 7 and 8 have only one gene each; they are also difficult to interpret. Cluster 9 has 13 genes and gene expression is repressed by a HF diet in the offspring of AD mothers and

the expression is induced by a HF diet in offspring of UN mothers. Cluster 10 has 49 genes and gene expression is repressed by leptin in general, however, the leptin repression is less effective in the UNLEPHF treatment group and more effective in the UNLEPCHOW treatment group. This can also be described as the interaction of two effects: repression by leptin, and induction by a HF diet, specifically in the offspring of UN mothers.

Pathway Analysis

Initially Pathway Analysis was performed on the FSPMA gene lists, as these are larger and could potentially reveal trends based on the effect of an individual treatment. Onto-Express [78] was used to find biological processes, molecular functions and cellular components associated with each gene list. The results of this analysis are shown in figures 4.25-4.27. These pie charts represent all the genes involved in biological processes, molecular functions or cellular components. The gray area shows the proportion of biological processes, molecular functions or cellular components not represented in the gene list, while the coloured wedges represent the categories that are most enriched in the gene list. A similar analysis was also done in DAVID, which revealed the most enriched categories for the list of genes affected by maternal diet to be related to protein binding, localisation and transport; those for the leptin treatment to be immune response, nucleobase metabolism and transport; and those for the postweaning diet to be related to cytoplasm, catalytic activity and metabolism (cellular lipid, lipid, alcohol, sterol, and steroid metabolism specifically).

Once the other analyses had been done and the dendrogram for the one-way ANOVA had been produced, a pathway analysis was performed on each of the clusters to determine if they represented genes involved in related biological processes, molecular functions or cellular components. Table 4.2 shows the key biological processes and kegg pathways that were significantly enriched for each cluster based on DAVID pathway analysis. During this process, it also became interesting to compare the overlap between the lists produced in the three separate analyses. Table 4.3 shows the 83 genes that overlapped between the FSPMA analysis and the filtered one-way

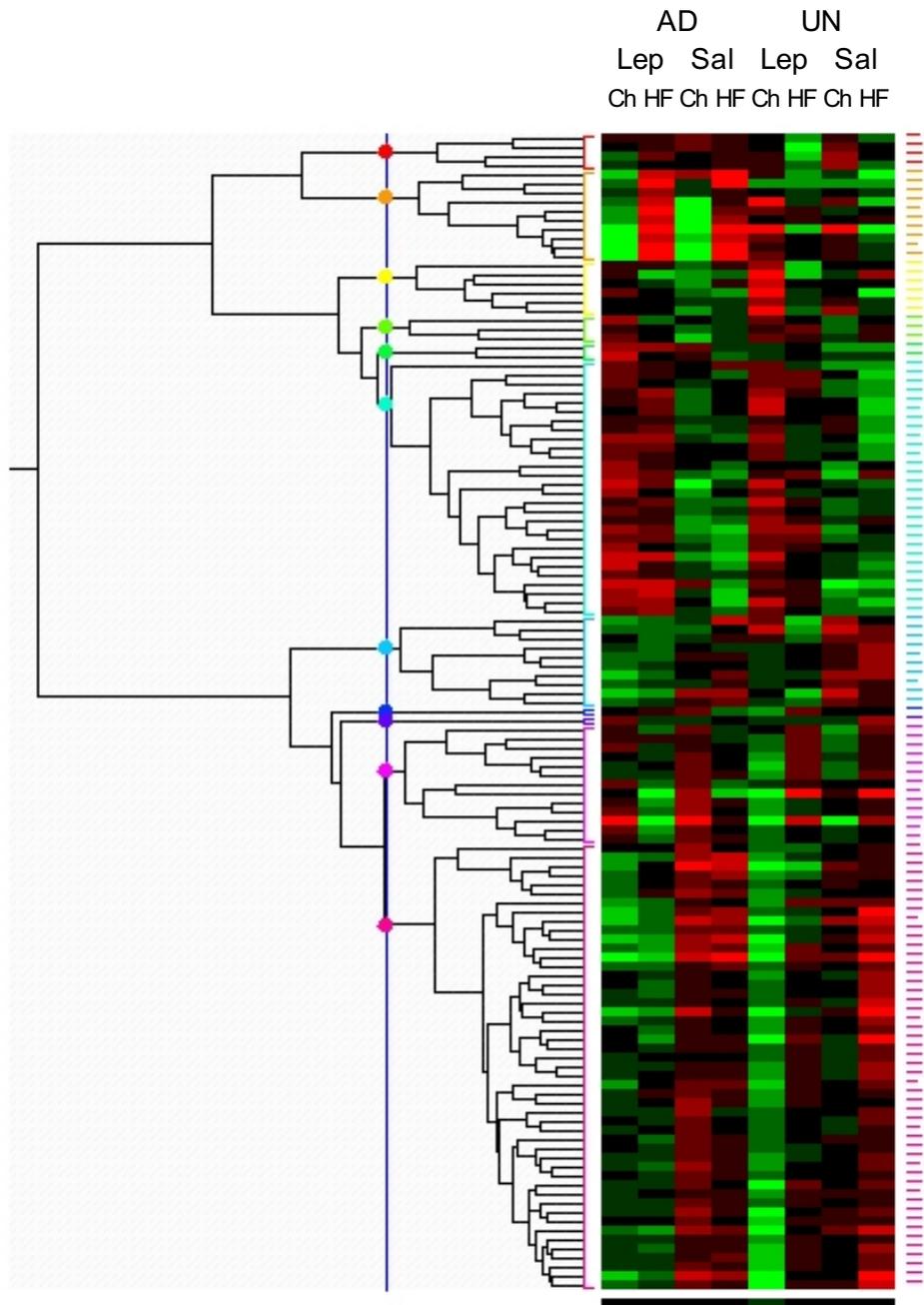


Figure 4.24: This is a heatmap showing the results of hierarchical clustering of the data normalised to a median value of 0. Each treatment group is represented in a separate column as shown on the top. The ten clusters are shown with coloured brackets on the left side.

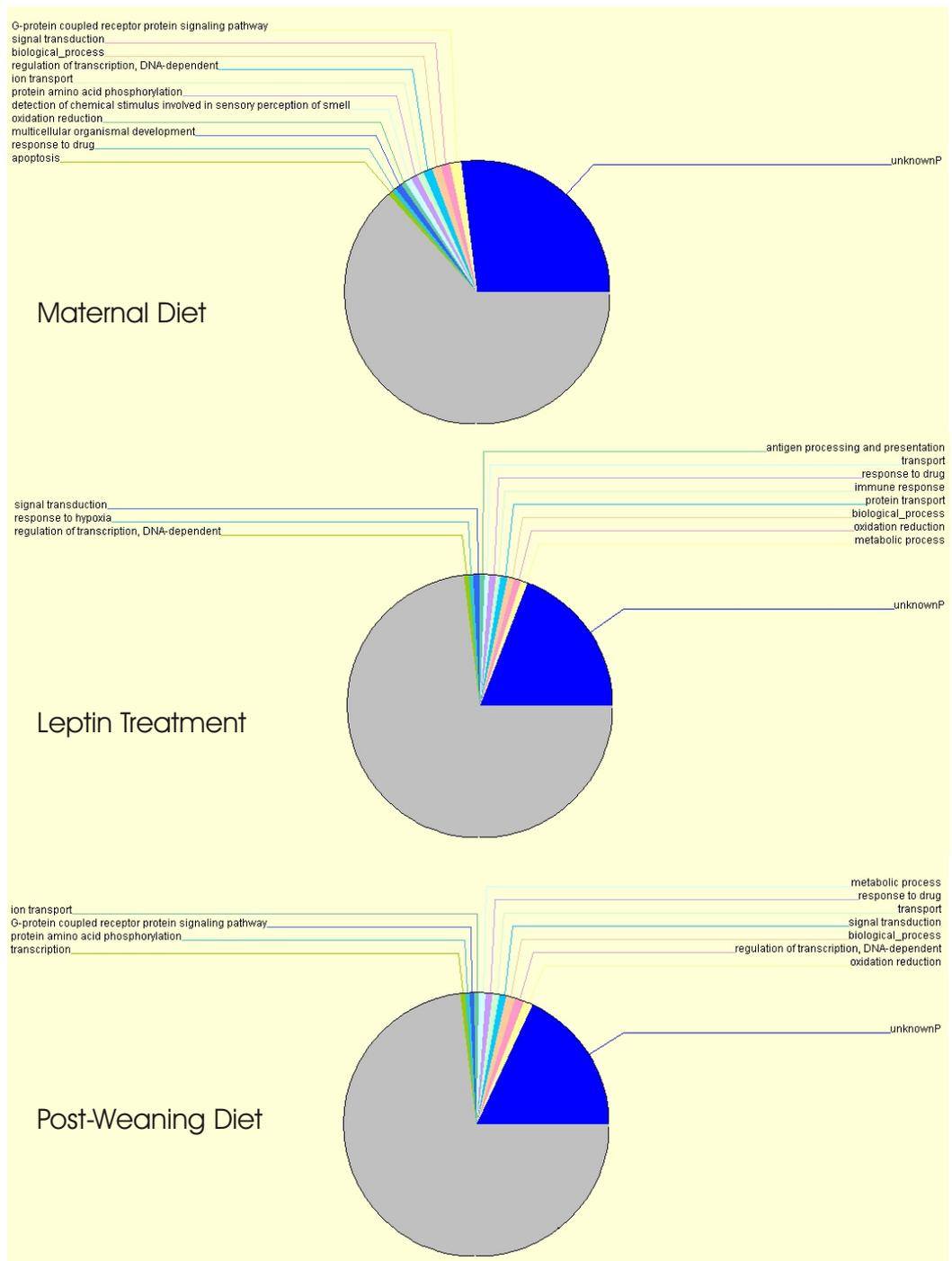


Figure 4.25: This pie chart shows the biological processes most enriched in each of the gene lists produced with the FSPMA analysis. The circle represents all the genes involved in biological processes. The gray area represents the biological processes not represented in the gene list. The coloured wedges show the proportion of genes involved in a particular biological process. This figure was made using Onto-Express [78].

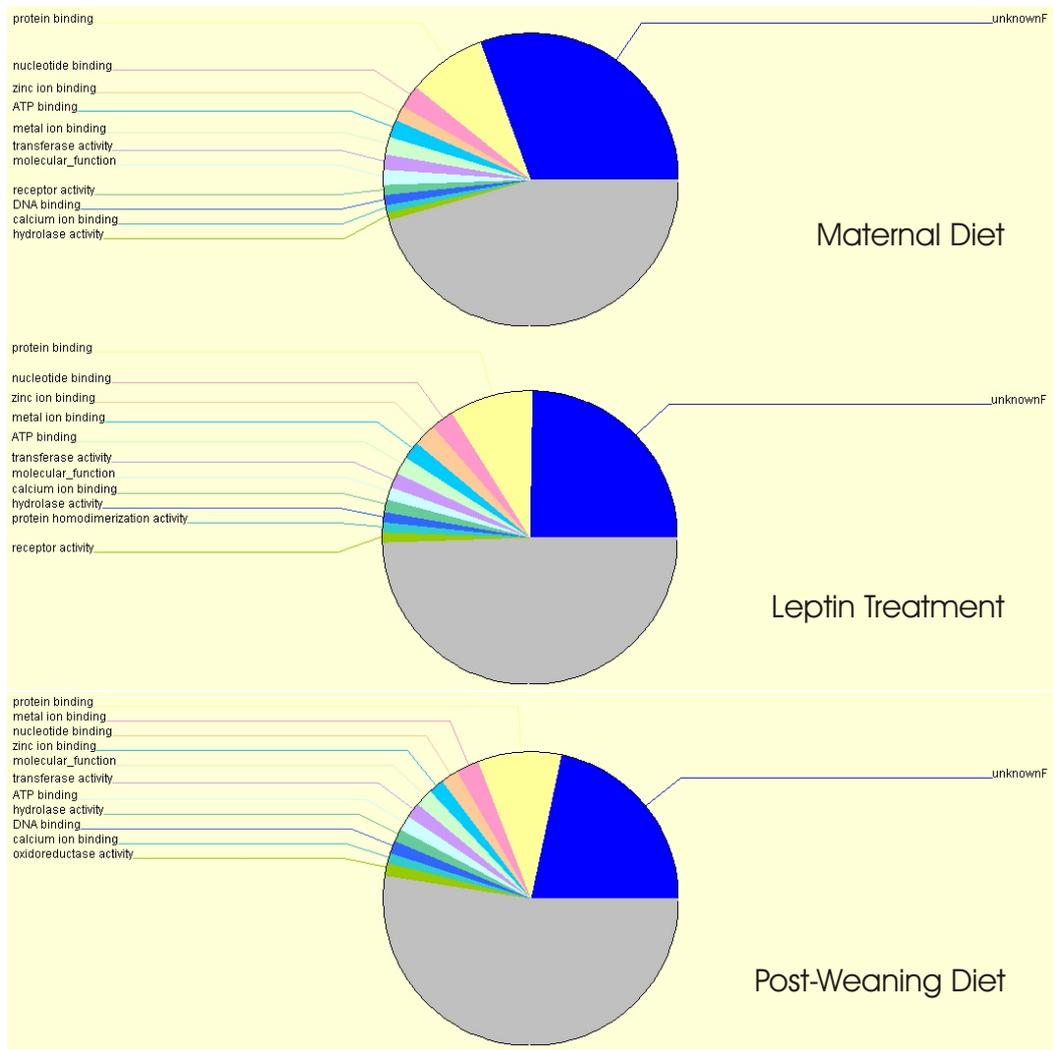


Figure 4.26: This pie chart shows the molecular functions most enriched in each of the gene lists produced with the FSPMA analysis. The circle represents all the genes involved in molecular functions. The gray area represents the molecular functions not represented in the gene list. The coloured wedges show the proportion of genes involved in a particular molecular function. This figure was made using Onto-Express [78].

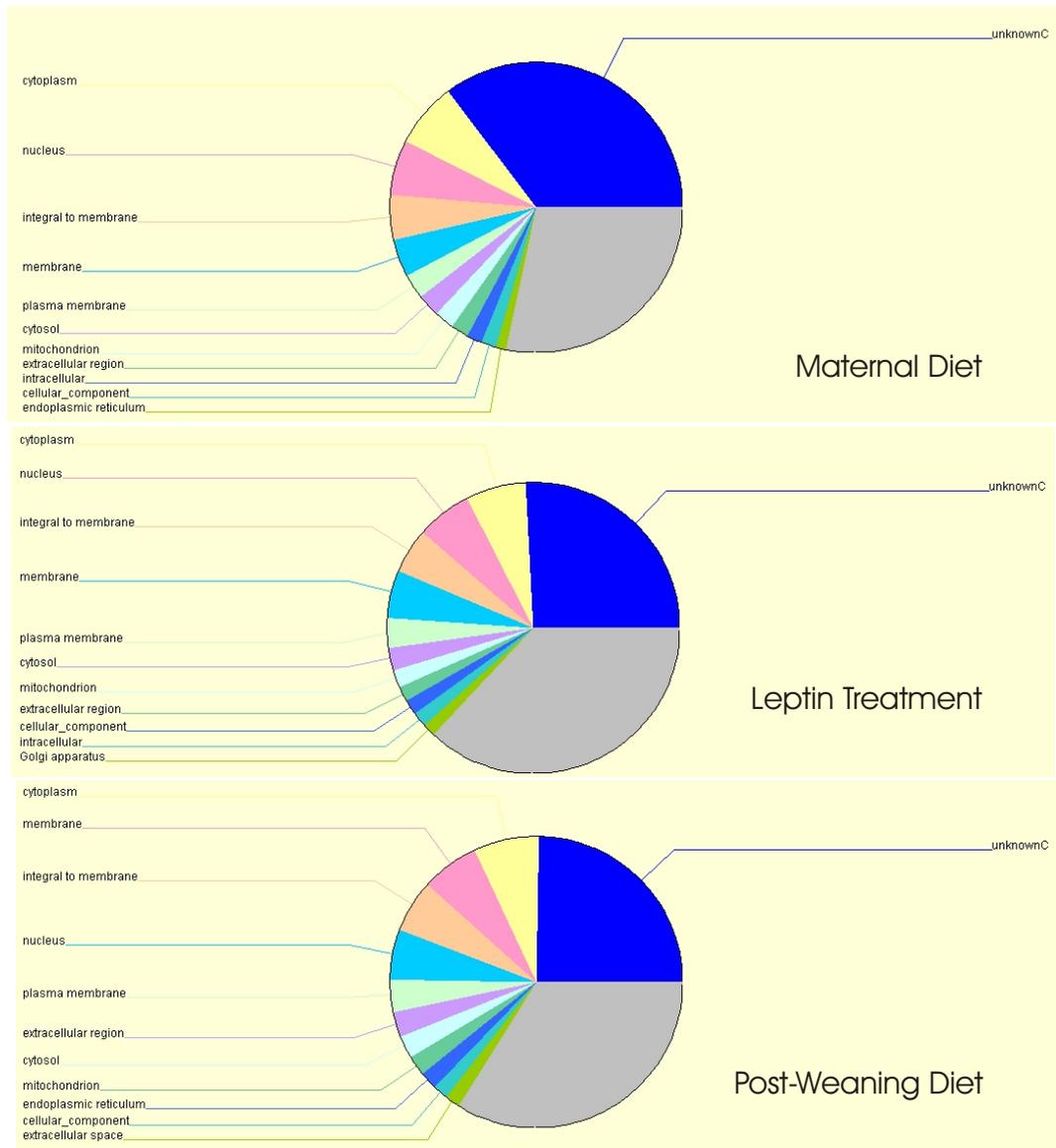


Figure 4.27: This pie chart shows the cellular components most enriched in each of the gene lists produced with the FSPMA analysis. The circle represents all the genes involved in cellular components. The gray area represents the cellular components not represented in the gene list. The coloured wedges show the proportion of genes involved in a particular cellular component. This figure was made using Onto-Express [78].

ANOVA analysis listed by cluster. This table has a column to indicate which of the three FSPMA lists the gene was present in (M for maternal diet, L for leptin treatment and/or D for postweaning diet). Another column shows which, if any, pair-wise comparisons the gene was present in.

Cluster	Number of genes	Biological Process/Kegg Pathway
Cluster0	4	Not found in pathway software
Cluster1	10	Antigen processing and presentation immune response
Cluster2	6	Regulation of biological processes
Cluster3	3	Transport of lipid (1gene)
Cluster4	2	Not found in pathway software
Cluster5	28	Metabolic processes (carboxylic acid, organic acid, amino acid etc), cholesterol biosynthetic process kegg pathways: carbon fixation, biosynthesis of steroids alanine and aspartate metabolism, arginine and proline metabolism,
Cluster6	10	Lipid transport, transport, digestion, establishment of localisation,response to drug kegg pathway: PPAR signalling pathway
Cluster7	1	Proteolysis kegg pathway: Metabolism of xenobiotics by cytochrome P450
Cluster8	1	Not found in pathway software
Cluster9	13	Response to external stimulus, humoral immune response, defence response, immunoglobulinn mediated immune response kegg pathway: complement and coagulation cascades
Cluster10	49	Metabolic processes (glucan,glycogen,polysaccharide, and more) kegg pathway: circadian rhythm

Table 4.2: This table lists the number of genes found in each of the hierarchical clustering clusters and the main biological processes and kegg pathways that are affected by the genes in the particular cluster found using DAVID [77].

Cluster	Symbol	Definition	Treatments	Pair-Wise Comparisons
Cluster0	Clic2	chloride intracellular channel 2	DM	ASCvsUSC
	Hrasls3	HRAS like suppressor 3	M	
Cluster1	Igfbp2	insulin-like growth factor binding protein 2	D	ALHFvsULC
	RT1-149	RT1 class I, T24, gene 4	L	ALCvsASHF,ALCvsASC,ALHFvsASC
	RT1-A3	RT1 class I, A3	L	
	RT1-Ba	RT1 class II, locus Ba	DL	ALHFvsASC
	RT1-CE15	RT1 class I, CE15	L	
	RT1-M6-2	RT1 class I, M6, gene 2	L	ALCvsASHF,ALCvsASC
	Sds	serine dehydratase	M	
	Sparc1	SPARC-like 1 (hevin)	M	
	Tnfsf13	tumour necrosis factor (ligand) superfamily, member 13	M	
Cluster2	Cgref1	cell growth regulator with EF hand domain 1	MDL	
	Colq	collagen-like tail subunit (single strand of homotrimer) of asymmetric acetylcholinesterase	DM	ALHFvsULC

Cluster	Symbol	Definition	Treatments	Pair-Wise Comparisons
	Hspb1	heat shock 27kDa protein 1	DL	ULCvsUSHF
Cluster4	Csad	cysteine sulfinic acid decarboxylase	DM	
Cluster5	Asl	argininosuccinate lyase	D	ULCvsUSHF
	Cd63	CD63 antigen	DL	ULCvsUSHF,ASHFvsULC,ASCvsUSHF,ALCvsUSHF
	Csrp2	cysteine and glycine-rich protein 2	MDL	ULCvsUSHF,ASCvsUSHF
	Cyp3a13	cytochrome P450, family 3, subfamily a, polypeptide 13	D	ALHFvsULC
	Dhcr7	7-dehydrocholesterol reductase	D	ASHFvsULC,ALCvsASHF
	Ftl1	ferritin light chain 1	D	
	Got1	glutamic-oxaloacetic transaminase 1, soluble (aspartate aminotransferase 1)	D	
	Hnrpab	heterogeneous nuclear ribonucleoprotein A/B	D	ALHFvsULC,ULCvsUSHF
	LOC316130	similar to mannose-6-phosphate receptor binding protein 1	DL	ULCvsUSHF,ASHFvsULC,ASCvsUSHF,ALCvsUSHF
	LOC499654	similar to Ac2-125	D	
	Me1	malic enzyme 1	D	

Cluster	Symbol	Definition	Treatments	Pair-Wise Comparisons
	Nnmt_pred	nicotinamide N-methyltransferase (predicted)	D	ASCvsUSHF
	Orn1	orosomuroid 1	D	ASHFvsULC,ALCvsASHF
	Prkcdbp	protein kinase C, delta binding protein	D	ULCvsUSHF,ASHFvsULC,ASCvsUSHF,ALCvsASHF,ASCvsASHF
	Rnf125_pred	ring finger protein 125 (predicted)	DL	
	Slc16a13	solute carrier family 16, member 13 (monocarboxylic acid transporter 13)	DL	
	Slc22a5	solute carrier family 22 (organic cation/carnitine transporter), member 5	D	
	Tmod1	tropomodulin 1	DL	ULCvsUSHF,ASHFvsULC,ALCvsASHF
	Tor3a	torsin family 3, member A	DM	
	Trpm6_pred	transient receptor potential cation channel, subfamily M, member 6 (predicted)	D	

Cluster	Symbol	Definition	Treatments	Pair-Wise Comparisons
	Yc2	glutathione S-transferase Yc2 subunit	DM	
Cluster6	Abcg5	ATP-binding cassette, sub-family G (WHITE), member 5	D	ALCvsASHF
	Afp	alpha-fetoprotein	MDL	ULCvsUSHF,ASCvsUSHF,ALCvsUSHF
	Ctsh	cathepsin H	DM	
	Fabp2	fatty acid binding protein 2, intestinal	DM	
	Fabp7	fatty acid binding protein 7, brain	D	ALCvsASHF
	Hdc	histidine decarboxylase	D	
	Serpina7	serine (or cysteine) peptidase inhibitor, clade A (alpha-1 antipeptidase, antitrypsin), member 7	D	
	Synj2	synaptojanin 2	DM	ULCvsUSHF
Cluster9	C2	complement component 2	L	
	Ift1	interferon-induced protein with tetratricopeptide repeats 1	M	

Cluster	Symbol	Definition	Treatments	Pair-Wise Comparisons
	Mlc1_pred	megalencephalic leukoencephalopathy with subcortical cysts 1 homolog (human) (predicted)	L	ULCvsUSHF
	Sod2	superoxide dismutase 2, mitochondrial, nuclear gene encoding mitochondrial protein	L	
	Ugt2b10_pred	UDP glycosyltransferase 2 family, polypeptide B10 (predicted)	D	ALHFvsULC
Cluster10	Aadac	arylamidase (esterase)	D	ALHFvsULC,ULCvsUSHF,ASHFvsULC,ASCvsUSHF
	Acox2	acyl-Coenzyme A oxidase 2, branched chain	D	ALHFvsULC,ASHFvsULC
	Ak3l1	adenylate kinase 3-like 1	DL	ALHFvsULC,ULCvsUSHF,ASHFvsULC
	Avpr1a	arginine vasopressin receptor 1A	DL	ULCvsUSHF
	Bche	butyrylcholinesterase	DL	ALHFvsULC,ULCvsUSHF,ASHFvsULC,ASCvsUSHF
	Bhlhb2	basic helix-loop-helix domain containing, class B2	D	ALHFvsULC,ULCvsUSHF,ASHFvsULC

Cluster	Symbol	Definition	Treatments	Pair-Wise Comparisons
	Btg1	B-cell translocation gene 1, anti-proliferative	DL	ULCvsUSHF
	Ca3	carbonic anhydrase 3	DL	ALHFvsULC,ULCvsUSHF
	Car14_pred	carbonic anhydrase 14 (predicted)	DL	ALHFvsULC,ULCvsUSHF,ASHFvsULC
	Cyp2b3	cytochrome P450IIB3	D	ALHFvsULC
	Dclre1a_pred	DNA cross-link repair 1A, PSO2 homolog (S cerevisiae) (predicted)	D	
	Enpp2	ectonucleotide pyrophosphate/phosphodiesterase 2	MDL	ALHFvsULC,ULCvsUSHF,ASHFvsULC
	Gck	glucokinase	MDL	ULCvsUSHF,ASCvsUSHF
	Gclc	glutamate-cysteine ligase, catalytic subunit	D	ALHFvsULC
	Klf15	Kruppel-like factor 15	DM	ASHFvsULC
	Kmo	kynurenine 3-monooxygenase (kynurenine 3-hydroxylase)	DL	ALHFvsULC,ULCvsUSHF

Cluster	Symbol	Definition	Treatments	Pair-Wise Comparisons
	LOC497901	similar to Liver-expressed antimicrobial peptide 2 precursor (LEAP-2)	D	ALHFvsULC,ULCvsUSHF
	LOC500080	similar to solute carrier family 13 (sodium/sulfate symporters), member 4	DL	ALHFvsULC,ULCvsUSHF,ASHFvsULC,ASCvsUSHF
	Olr59	olfactory receptor 59	DL	ULCvsUSHF
	Otc	ornithine carbamoyltransferase, nuclear gene encoding mitochondrial protein	DL	ALHFvsULC,ULCvsUSHF,ASHFvsULC
	Per2	period homolog 2 (Drosophila)	DM	ALHFvsULC,ULCvsUSHF,ASHFvsULC
	Polg2-pred	polymerase (DNA directed), gamma 2, accessory subunit (predicted)	D	ALHFvsULC
	Ppp1r3b	protein phosphatase 1, regulatory (inhibitor) subunit 3B	DL	ULCvsUSHF,ASHFvsULC,ASCvsUSHF
	Ptprf	protein tyrosine phosphatase, receptor type, F	D	
	Pygl	liver glycogen phosphorylase	DL	ALHFvsULC,ULCvsUSHF,ASHFvsULC

Cluster	Symbol	Definition	Treatments	Pair-Wise Comparisons
	Rarb	retinoic acid receptor, beta, transcript variant 2	DL	ULCvsUSHF
	Serpina3m	serine (or cysteine) proteinase inhibitor, clade A, member 3M	DL	ALHFvsULC,ULCvsUSHF,ASHFvsULC
	Sez6	seizure related 6 homolog (mouse)	DL	ALHFvsULC,ULCvsUSHF
	Slc40a1	solute carrier family 39 (iron-regulated transporter), member 1	DM	ALHFvsULC,ASHFvsULC
	Stac3-pred	SH3 and cysteine rich domain 3 (predicted)	D	ULCvsUSHF
	Tdo2	tryptophan 2,3-dioxygenase	D	ALHFvsULC,ULCvsUSHF,ASHFvsULC
	Thrsp	thyroid hormone responsive	D	ALHFvsULC
	Xpnpep2	X-prolyl aminopeptidase (aminopeptidase P) 2, membrane-bound	DM	ALHFvsULC

Table 4.3: This is a list of 83 genes that were found in the filtered one-way ANOVA and the FSPMA analysis. The genes are listed by cluster (hierarchical clustering performed on the filtered one-way ANOVA gene list). The Treatments column indicates which FSPMA analysis the gene showed up in (maternal diet (M), leptin treatment (L), or postweaning diet (D)) and the final column lists the pair-wise comparisons the gene was significant in.

4.2.2 qRT-PCR for Microarray Validation

Quantitative RT-PCR analysis of selected genes was performed in order to confirm the expression data derived from array analysis. Eight genes (*Igfbp2*, *Rt1-149*, *Rt1-ba*, *Rt1-m6-2*, *Orm1*, *Ca3*, *Gck*, and *Per2*) found to be significantly, differentially expressed on the Illumina microarray were selected for qRT-PCR validation. qRT-PCR was performed as described in Chapter 2 on the pooled RNA samples of the biological replicates in each treatment group. Cyclophilin was used as an invariant control for normalisation. The individual CT values are shown for the cyclophilin technical replicates in figure 4.28. All except one of the replicates vary less than 4% from the mean of all replicates. The microarray and qRT-PCR data for each treatment group for each gene are shown in figures 4.29-4.32.

4.2.3 Comparison of Treatment Groups

The genes selected for qRT-PCR are also shown in figures 4.33-4.36. These graphs aim to compare the effects of the leptin treatment across the x-axis and the postweaning diet along the y-axis.

Genes/Pathways Highlighted in Other Studies

Gluckman *et al.* [91] also used the liver tissue from this initial leptin reversal study for molecular analysis, but utilising a candidate gene approach. They examined the gene expression of *11 β -hsd2*, *Ppara α* , *Pepck*, and *GR*. They found a significant effect of leptin in UN animals that had been fed a high-fat postweaning diet for *11 β -hsd2*, *Ppara α* , and *Pepck* (results shown in figure 4.37A). We repeated these experiments for comparison using the same primers and a similar protocol. We obtained different results that showed no significant change between treatment groups (results shown in figure 4.37B).

Figure 4.38 shows the array signals for the five genes from Gluckman *et al.* [91] plotted with leptin treatment on the x-axis and the postweaning diet on the y-axis to determine the effect of the treatment on the gene expression.

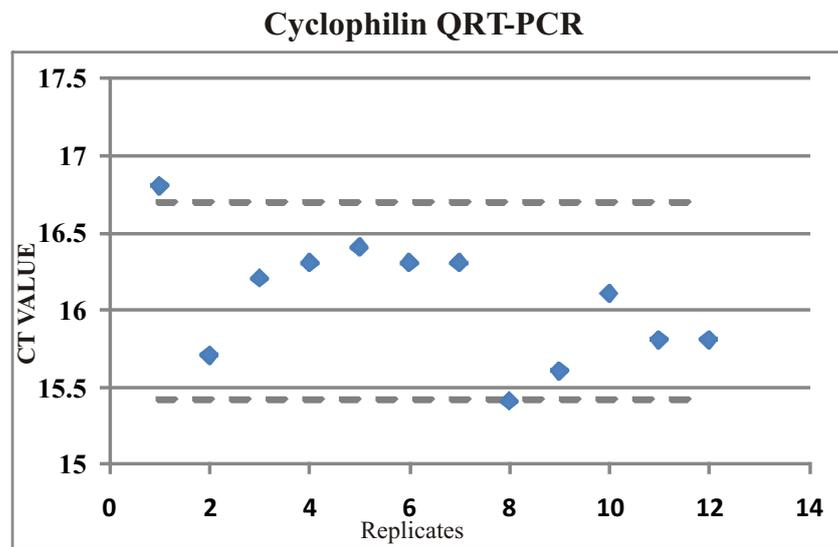


Figure 4.28: This figure shows the 12 technical replicates run with primers for cyclophilin. This data served as the invariant control for the other qRT-PCR data presented in this chapter. The lines represent 4% of the mean.

Comparison of Day 170 Female Saline Data to Day 55 Male Undernourished Data

Although the samples used in this chapter and in Chapter 3 are very different and were not chosen for comparison, a brief analysis was done out of interest. The comparison was between the ASC and USC samples from this chapter and all the AD and UN samples from Chapter 3. For review, the two studies are from different cohorts that went through experimental protocols with slight variations. The two groups are different sexes and the samples are from different time points.

Two separate analyses were done; a simple pairwise comparison between each of the combinations of the 4 groups (AD-Day170, UN-Day170, AD-Day55 and UN-Day55) and a three way ANOVA with interaction term. Figure 4.39 shows the number of genes significantly, differentially expressed in each of the pairwise comparisons. The number in brackets shows the genes unique to that comparison. Although it is impossible to say whether the genes are being up or downregulated due to sex differences or age differences, it does highlight the need to do a more detailed study looking at different developmental windows (see Appendix C Tables 24-27 for genelists).

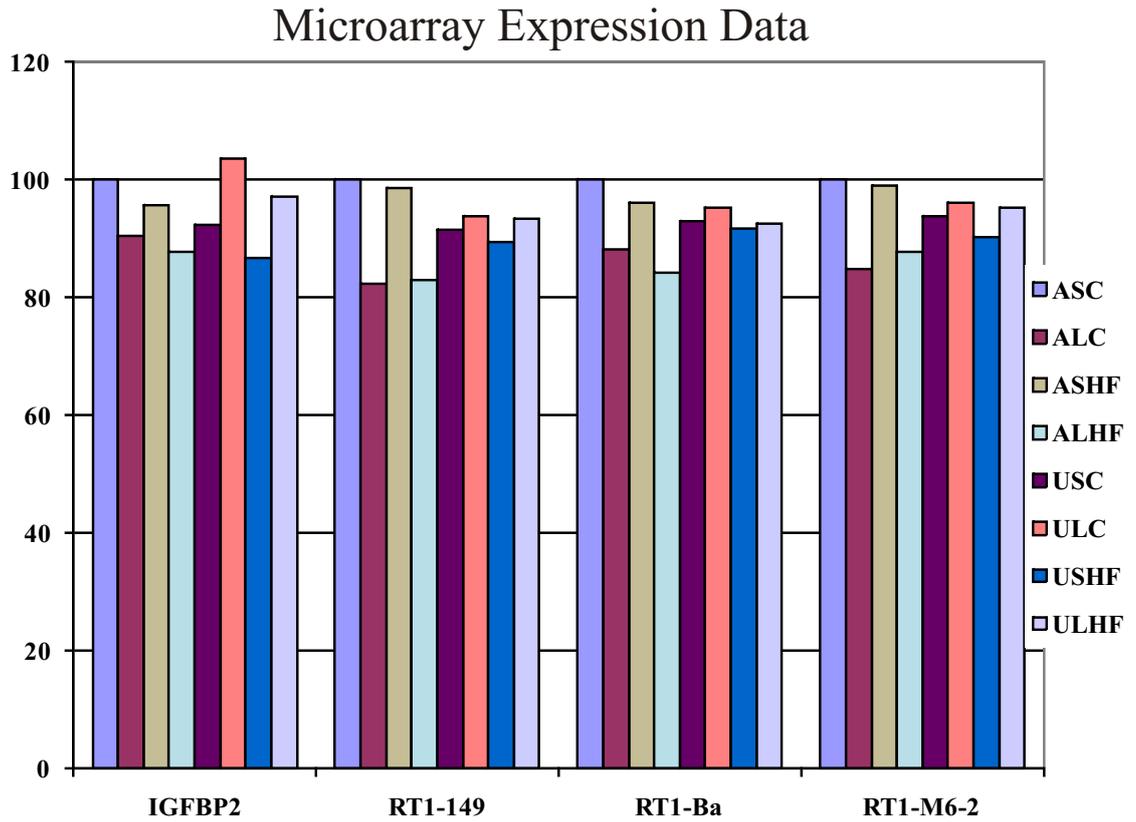


Figure 4.29: This figure shows four of the genes (*Igfbp2*, *Rt1-149*, *Rt1-ba*, and *Rt1-M6-2*) that were chosen from the list of significantly, differentially expressed genes on the Illumina microarray. For each gene 8 biological replicates for each of the 8 treatment groups were hybridised and analysed. The data here is shown as a % of ASC expression.

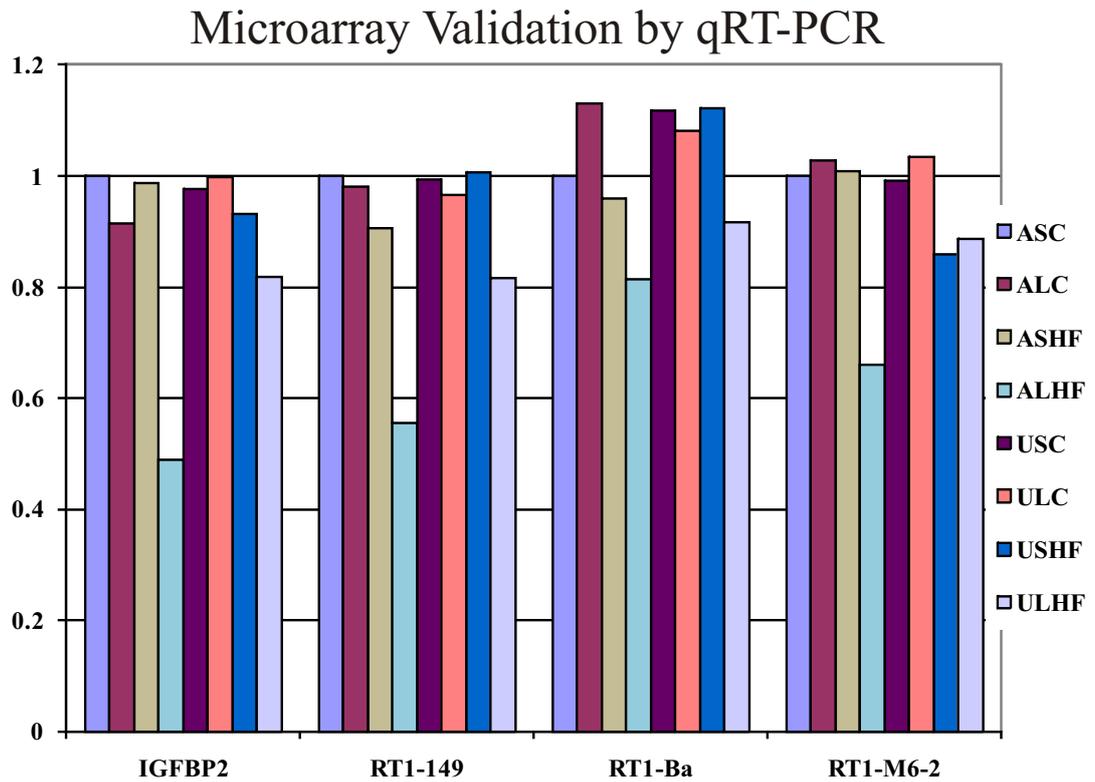


Figure 4.30: This figure shows four of the genes (*Igfbp2*, *Rt1-149*, *Rt1-ba*, and *Rt1-M6-2*) that were chosen from the list of significantly, differentially expressed genes on the Illumina microarray. The RNA from the eight biological replicates was pooled for each of the eight treatment groups. qRT-PCR for the eight samples was run in triplicate with cyclophilin as an invariant control. The data have been normalised to the ASC treatment group at 1.

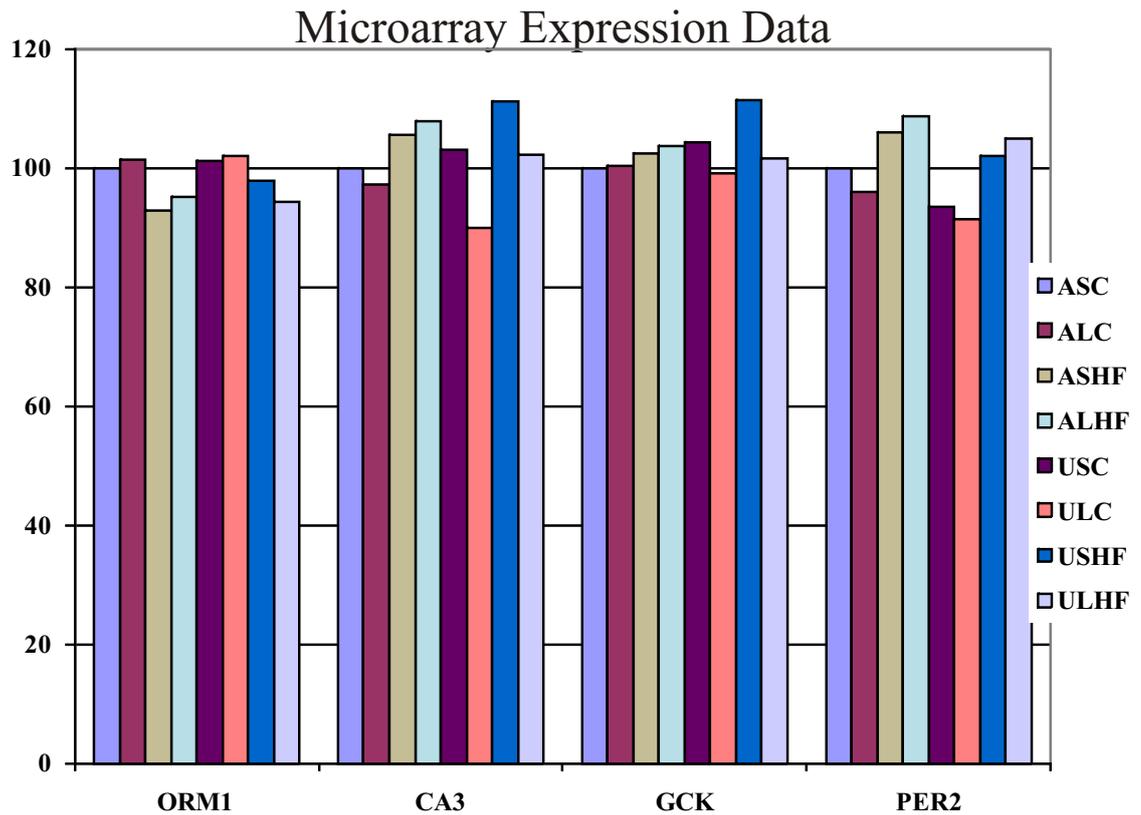


Figure 4.31: This figure shows four of the genes (*Orm1*, *Ca3*, *Gck*, and *Per2*) that were chosen from the list of significantly, differentially expressed genes on the Illumina microarray. For each gene 8 biological replicates for each of the 8 treatment groups were hybridised and analysed. The data here is shown as a % of ASC expression.

Microarray Validation by qRT-PCR

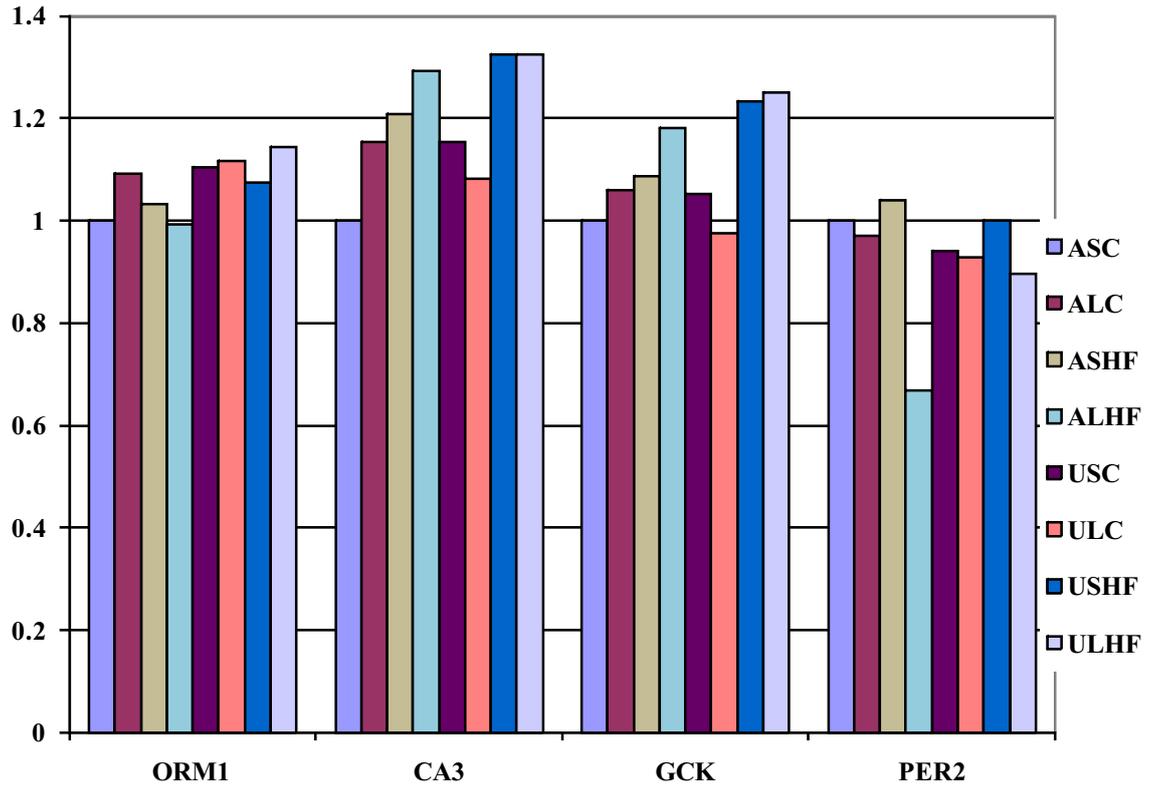


Figure 4.32: that were chosen from the list of significantly, differentially expressed genes on the Illumina microarray. The RNA from the eight biological replicates was pooled for each of the eight treatment groups. qRT-PCR for the eight samples was run in triplicate with cyclophilin as an invariant control. The data have been normalised to the ASC treatment group at 1.

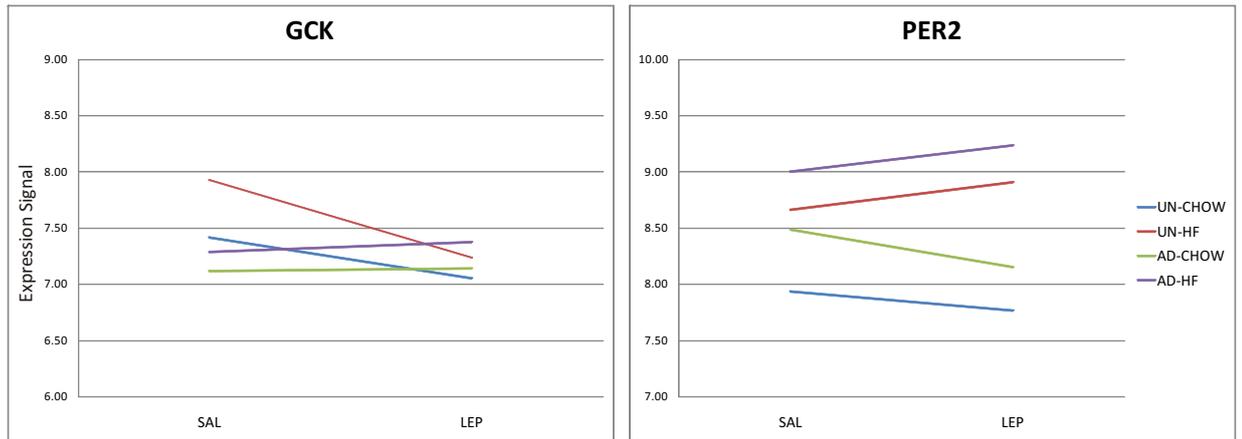


Figure 4.33: The figure on the left shows the array signals for *Gck* and the figure on the right for *Per2*. The x-axis represents the effect of the leptin treatment and the y-axis represents the postweaning diet. The blue and red lines show the undernourished maternal diet (UN) and the green and purple lines show the ad-libitum fed maternal diet (AD).

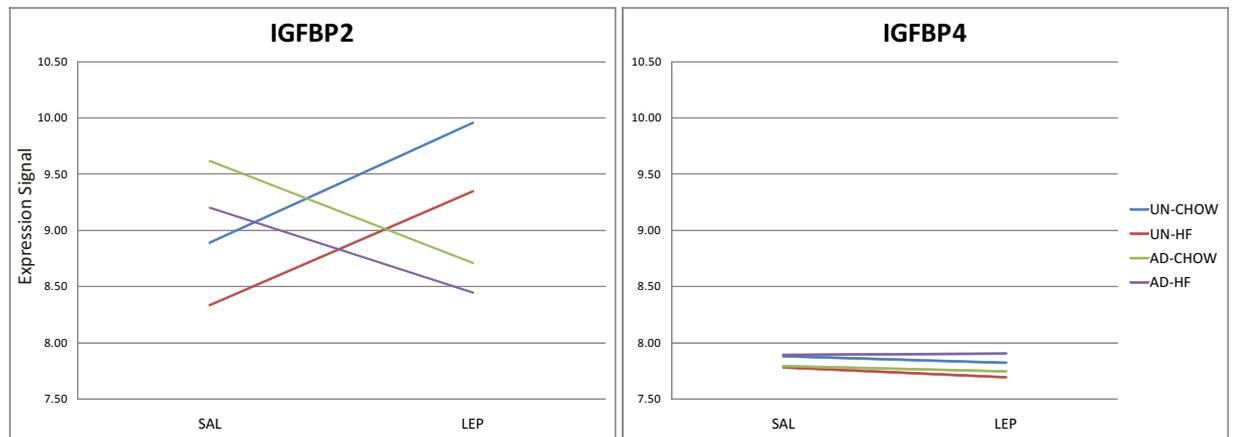


Figure 4.34: The figure on the left shows the array signals for *Igfbp2* and the figure on the right for *Igfbp4*. The x-axis represents the effect of the leptin treatment and the y-axis represents the postweaning diet. The blue and red lines show the undernourished maternal diet (UN) and the green and purple lines show the ad-libitum fed maternal diet (AD).

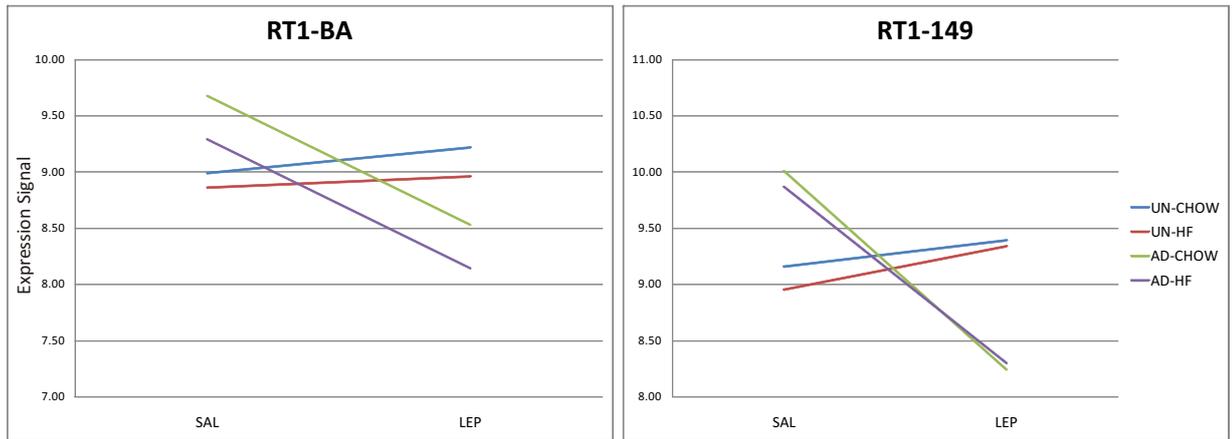


Figure 4.35: The figure on the left shows the array signals for *Rt1-ba* and the figure on the right for *Rt1-149*. The x-axis represents the effect of the leptin treatment and the y-axis represents the postweaning diet. The blue and red lines show the undernourished maternal diet (UN) and the green and purple lines show the ad-libitum fed maternal diet (AD).

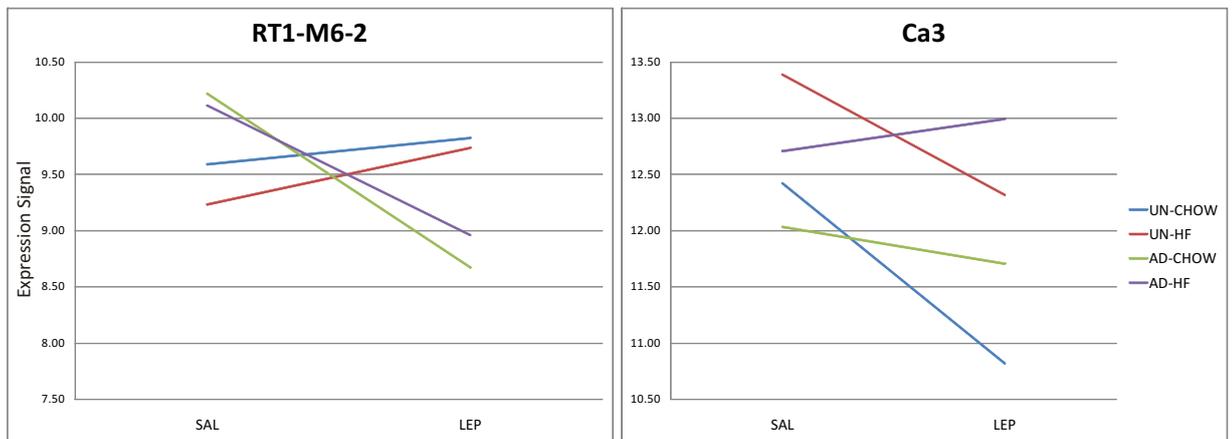


Figure 4.36: The figure on the left shows the array signals for *Rt1-m6-2* and the figure on the right for *Ca3*. The x-axis represents the effect of the leptin treatment and the y-axis represents the postweaning diet. The blue and red lines show the undernourished maternal diet (UN) and the green and purple lines show the ad-libitum fed maternal diet (AD).

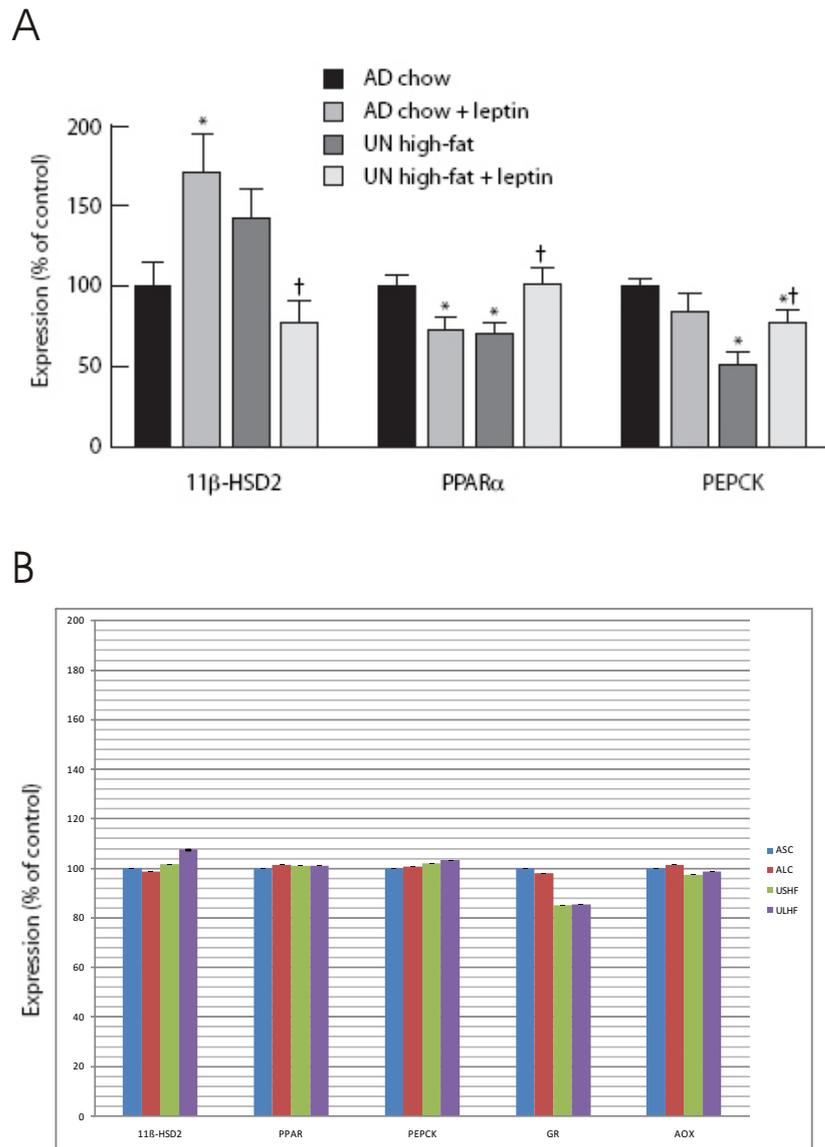


Figure 4.37: (A) This figure was taken from Gluckman *et al.* [91] and shows RT-PCR results for the mRNA expression of *11β-hsd2*, *Pparα*, and *Pepck* for the same adult day 170 female rats as used in this chapter. Data are means \pm SEM for $n=8$ per group; values for gene expression are expressed relative to those of normally nourished (AD maternal diet and CHOW postnatal diet) saline-treated offspring set as 100%. All three genes were found to have a significant $p < 0.05$ effect of leptin treatment in UN high-fat animals (+). (*) indicates data is significantly $p < 0.05$ different from ADSALCHOW. (B) We performed similar qRT-PCR experiments and found no significant differences among treatment groups. In this case, data are means \pm SEM for $n=3$ per group except *PPARα*, which is $n=8$.

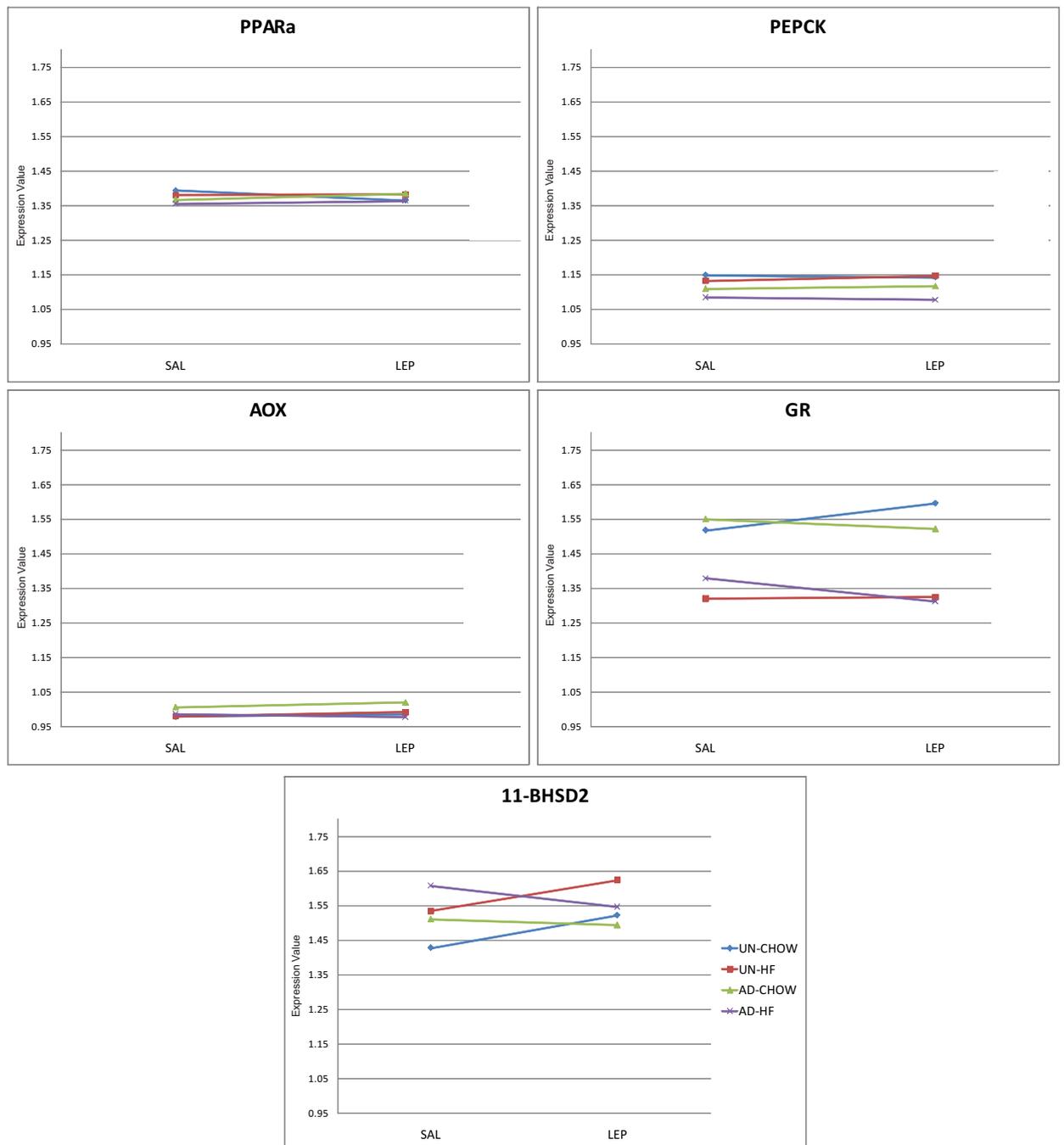


Figure 4.38: This figure shows the array signals for the five genes from Gluckman *et al.* [91] plotted with leptin treatment on the x-axis and the postweaning diet on the y-axis to determine the effect of the treatment on the gene expression. *GR* and *11β-HSD2* are the only plots to show interactions between the treatments.

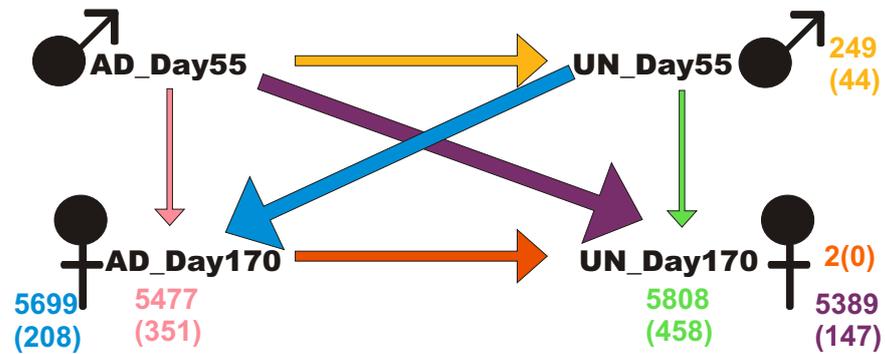


Figure 4.39: This figure shows the results of pairwise comparisons between the AD-SALCHOW and UNSALCHOW data from this chapter with the AD and UN data from Chapter 3. The number of genes that came up as differentially expressed for the comparison ($p < 0.05$) is shown in the colour of the arrow for that comparison. The number shown in brackets is the number of genes unique to that comparison. The effects of sex and time point cannot be separated.

Cluster	Symbol	Biofunction
Cluster_0	Clic2	regulation of cellular processes
Cluster_0	Hrasls3	role in adipogenesis [142]
Cluster_1	Igfbp2	impaired IGD binding to receptor
Cluster_1	RT1-149	histocompatibility complex
Cluster_1	RT1-A3	histocompatibility complex
Cluster_1	RT1-Ba	histocompatibility complex
Cluster_1	RT1-CE15	histocompatibility complex
Cluster_1	RT1-M6-2	histocompatibility complex
Cluster_1	Sds	regulation of cell cycle
Cluster_1	Sparcl1	glycoprotein that binds calcium
Cluster_1	Tnfsf13	important for B cell development
Cluster_2	Cgref1	cell growth; response to stress
Cluster_2	Colq	collagen-like tail subunit
Cluster_2	Hspb1	Heat shock 27kD protein induced by environmental stress
Cluster_4	Csad	involved in taurine biosynthesis
Cluster_5	Asl	involved in urea cycle
Cluster_5	Cd63	cell growth; associated with tumour progression
Cluster_5	Csrp2	regulation of development and cell differentiation
Cluster_5	Cyp3a13	drug metabolism
Cluster_5	Dhcr7	cholesterol metabolism
Cluster_5	Ftl1	regulation of iron storage
Cluster_5	Got1	amino acid metabolism; involved in urea cycle

Cluster_5	Hnrpab	mRNA metabolism and transport
Cluster_5	LOC316130	M6prbp1 -receptor binding protein
Cluster_5	Me1	activity links glycolytic and TCA cycles
Cluster_5	Nnmt_predicted	methyltransferase activity; implicated in cancers
Cluster_5	Orm1	regulation of immune system; inflammatory response
Cluster_5	Prkcdbp	negative regulation of cell cycle; tumour suppressor
Cluster_5	Rnf125_predicted	T-cell activation
Cluster_5	Slc16a13	transport
Cluster_5	Slc22a5	transport; linked to metabolic decompensation
Cluster_5	Tmod1	actin binding; organisation of actin filaments
Cluster_5	Tor3a	nucleotide binding
Cluster_5	Trpm6_predicted	calcium ion and nucleotide binding
Cluster_6	Abcg5	nucleotide binding; transport of dietary cholesterol
Cluster_6	Afp	metal ion binding
Cluster_6	Ctsh	degradation of proteins in lysosome
Cluster_6	Fabp2	binds long-chain fatty acids; maintains energy homeostasis by acting as lipid sensor
Cluster_6	Fabp7	expressed during development; involved in brain development
Cluster_6	Hdc	modulates numerous physiologic processes
Cluster_6	Serpina7	thyroid hormone transport protein
Cluster_6	Synj2	signal transduction
Cluster_9	C2	proteolysis
Cluster_9	Mlc1_predicted	may be involved in transport; associated with neurological disorders
Cluster_9	Sod2	protein binding; mitochondrial
Cluster_9	Ugt2b10_predicted	metabolic processes
Cluster_10	Aadac	lipid mobilisation (regulated diurnally)
Cluster_10	Acox2	involved in degradation of long chain Fas
Cluster_10	Ak3l1	mitochondrial nucleotide mobilisation
Cluster_10	Avpr1a	receptor mediates cell contraction and proliferation, platelet aggregation, release of coagulation factor and glycogenolysis
Cluster_10	Bche	drug metabolism
Cluster_10	Bhlhb2	basic helix-loop-helix domain
Cluster_10	Btg1	regulates cell growth and differentiation

Cluster_10	Ca3	involved in cellular response to stress
Cluster_10	Car14_predicted	carbonic anhydrase
Cluster_10	Cyp2b3	cyt P450 complex
Cluster_10	Delre1a_predicted	involved in blocking transcription, replication, and segregation of DNA
Cluster_10	Enpp2	modulator of cell motility; may induce parturition; stimulates tumour cell motility
Cluster_10	Gck	important in glucose utilisation
Cluster_10	Gclc	rate limiting step in glutathione synthesis
Cluster_10	Klf15	glucose and fat metabolism [143]
Cluster_10	Kmo	required for synthesis of quinolinic acid
Cluster_10	LOC497901	Leap2-may have antimicrobial activity
Cluster_10	LOC500015	Samd9l
Cluster_10	LOC500080	similar to Slc13 member4 - may be involved in lifespan determination [144]
Cluster_10	Olr59	sensory perception of smell
Cluster_10	Otc	involved in the urea cycle
Cluster_10	Per2	regulation of circadian rhythm [145],[139] and [146]
Cluster_10	Polg2_predicted	mitochondrial DNA polymerase gamma subunit
Cluster_10	Ppp1r3b	regulates enzymes (PYGL) involved in glycogen metabolism
Cluster_10	Ptprf	regulates cell growth and differentiation
Cluster_10	Pygl	important in carbohydrate metabolism [147]
Cluster_10	Rarb	regulates gene expression and affects development
Cluster_10	Serpina3m	protein metabolism
Cluster_10	Sez6	cell-cell recognition and neuronal membrane signalling
Cluster_10	Slc40a1	regulation of iron levels
Cluster_10	Stac3_predicted	intracellular signalling
Cluster_10	Tdo2	tryptophan metabolism
Cluster_10	Thrsp	may have a controlling role tumour lipid metabolism
Cluster_10	Xpnpep2	role in inflammatory process and response to injury

Table 4.4: This list includes most of the genes that were found in the filtered one-way ANOVA and the FSPMA analysis. The list includes the cluster as in table 4.3 for cross reference. These are the functions as found in literature search or online database information (genecards, NCBI or wikipedia).

4.3 Discussion

By comparing the transcriptional profiles of liver RNA from 170 day old rats that have been subjected to undernourishment during pregnancy, a leptin treatment and a postnatal high fat diet; this study has highlighted many genes that have been altered due to a combination of these treatments. Pathway analysis has revealed that many of the genes are involved in biological functions that are relevant to metabolism and may reveal primary or secondary effects of the treatments. An understanding of the role of leptin, its signalling to the brain, and the evolution of diet in humans (as discussed in Chapter 1) enables us to hypothesise the reasons for the phenotypes that develop. This study has now provided molecular evidence to back up some of these theories.

In this study, significant gene expression differences were dependent on the interaction between three treatments: maternal diet during pregnancy, leptin treatment and postweaning diet. Initial analysis aimed to separate the three treatments and consider, independently, the genes being influenced by each treatment. This analysis resulted in very large gene lists that included a majority of genes with very small expression changes. The large gene lists, however, were useful for pathway analysis. Figures 4.25 (biological processes), 4.26 (molecular functions), and 4.27 (cellular components) show the breakdown of the results of these analyses. These figures were made using Onto-Express and analysis done in David (see Chapter 2) supported the results. These analyses indicated that genes affected by maternal diet were involved in pathways related to protein binding, localisation, transport, organ development and apoptosis. The presence of organ development seems particularly logical given the hypothesis that a restricted diet during pregnancy causes the foetus to focus resources on brain development rather than organ development. The genes affected by the leptin treatment were related to immune response, nucleobase metabolism and transport. As mentioned in the introduction to this chapter, leptin has been implicated as affecting immune response. Finally, the genes affected by the postweaning diet were related to cytoplasm, catalytic activity, signal transduction and

metabolism (cellular lipid, lipid, alcohol, sterol and steroid metabolism specifically). Pair-wise analysis of treatment groups against each other produced interesting results that have proven challenging to interpret (see figure 4.23). Three of these comparisons stood out as having greater than 50 significantly expressed genes and intriguingly all three of these comparisons involved the UNLEPCHOW treatment group. This is the group that was undernourished during pregnancy, received the leptin treatment and was then put on a standard chow postnatal diet. Vickers *et al.* [57] has claimed that female rats that are undernourished during pregnancy only show the programmed phenotype if they are then subjected to the mismatch of the postnatal HF diet. This treatment group did not show significant phenotypic differences from the control ADSALCHOW treatment group. It is also worth mentioning the three treatment groups that showed the most significant differences to the UNLEPCHOW group. The most differences (88 genes) were found in comparison to the UNSALHF treatment group, which is the programmed phenotype of maternal undernourishment followed by the mis-match of the postnatal HF diet. Interestingly, the fourth largest gene list (28 genes) occurred in the comparison of this programmed phenotype (UNLEPHF) and the normal control treatment (ADSALCHOW). Phenotypically it seems to be the same comparison, but there must be metabolic changes induced by the maternal undernourishment given the larger number of differentially expressed genes in that comparison. Without the HF diet insult, these metabolic changes did not affect the phenotype. The second and third largest gene lists occurred in the comparison against the ADLEPHF (65) and ADSALHF (55) treatment groups. These treatments result in similar phenotypes although the leptin treated rats had slightly higher body fat, locomotor activity, insulin, leptin, and C-peptide levels and slightly lower food intake as would be expected with raised leptin levels, but these measurements were not statistically significant. A large number (51) of the genes occur in at least 2 of the three largest lists.

As described in the results, the gene lists from the different analyses were compared and a table of overlapping genes was created. A close look at table 4.3 shows the agreement of the ‘Treatments’ (this column indicates which gene list that the

gene was significant for in the FSPMA analysis) and the ‘Pair-wise Comparisons’. The lists were sorted by the clusters obtained from the hierarchical clustering of the one-way ANOVA across all eight treatment groups. Table 4.2 lists each cluster and the most enriched biological process/molecular function for each cluster. Table 4.4 then goes into more detail and provides the specific biological function or proposed biological function for most of the 83 genes that overlapped in the different analyses. The largest cluster and the one that seems to include the most intriguing genes is cluster 10. This cluster has 49 genes and gene expression appears to be repressed by leptin in general, however the leptin repression is less effective in the UNLEPHF treatment group and more effective in the UNLEPCHOW treatment group. This can also be described as the interaction of two effects: repression by leptin, and induction by a HF diet specifically in the offspring of UN mothers. The rest of this discussion will focus on the function of these genes and hypotheses on how they might be affecting the observed phenotypes in this study.

4.3.1 Cluster 10

Cluster 10 is the cluster that drew the most interest and includes genes involved in metabolic processes: sensory perception, stress response, regulation of circadian rhythm, transport and signalling. All of these genes were shown to be significantly affected by the postnatal high fat diet in the FSPMA analysis. Although, several were also shown to be affected by the maternal diet and six were shown to be significantly affected by all three treatments. The genes affected by all three treatments were *Gck*, *Per2*, *Klf15*, *Enpp2*, *Xpnpep2*, and *Slc40a1*; four of which are well characterised and are summarised here.

Gck (glucokinase) is a rate-limiting enzyme in glycolysis that regulates the conversion of ATP and D-glucose to ADP and D-glucose 6-phosphate. The gene is found to have reduced expression in patients with type-2 diabetes and in diabetic animals [148]. Insulin stimulates *Gck*, which in turn activates glucose disposal in the liver. Down regulation of *Gck* could, therefore, play a role in reduced insulin sensitivity.

Bogdarina *et al.* [149] found that fasting causes reduced *Gck* expression, but that re-feeding will result in an immediate insulin-induced increase in *Gck*. However, in this study, animals that were exposed to maternal undernutrition during gestation have reduced GCK expression in adulthood compared to control animals. This gene has come up as being differentially expressed in all three treatment groups and figure 4.33 clearly reveals an effect of leptin in conjunction with the maternal UN diet, which is in line with the reduced insulin sensitivity in the UN animals.

Per2 (period homolog 2) is one the core clock genes involved in regulation of the circadian rhythm. The circadian rhythm is the 24-hour cycle by which biological processes in the body are regulated. There are several core clock genes including (*Npas2*, *Bmal1*, *Per1*, *Rev-erba*, and *Per2*) [139]. *Per1* was significantly, differentially expressed due to the maternal diet and the leptin treatment in the FSPMA analysis, however, it was not present as significant in the other two analyses. Changes in the expression of the core clock genes have been tied to obesity. As mentioned at the beginning of this chapter, leptin has been shown to affect the circadian rhythm. Figure 4.33 shows an effect of leptin in both AD and UN fed animals.

Klf15 (Krupple-like Factor 15) is a gene that was significant in the FSPMA analysis as being affected by the postnatal diet and in the pair-wise analysis it came up in the comparison of ASHFvsULC; which indicates it has been affected by all treatments. Figure 4.40 shows the expression signals and emphasises the effect of leptin and the postnatal diet. Gray *et al.* has shown that *Klf15* plays a role in the regulation of gluconeogenesis [143]. *Klf15*^{-/-} mice are capable of gluconeogenesis but are less efficient at using alanine as a source for the process.

Enpp2 is found in adipose tissue and Ferry *et al.* [150] conducted a detailed study into its role in this tissue. They found it is strongly up-regulated during adipocyte differentiation and it seems to be a consequence of the differentiation as it is associated with the accumulation of triglycerides. The expression levels are significantly increased in obese diabetic db/db mice and, therefore, *Enpp2* may be associated with diabetes or obesity. Although this study was done in adipocyte tissue and our study involved liver tissue, it can be assumed that *Enpp2* may also be related

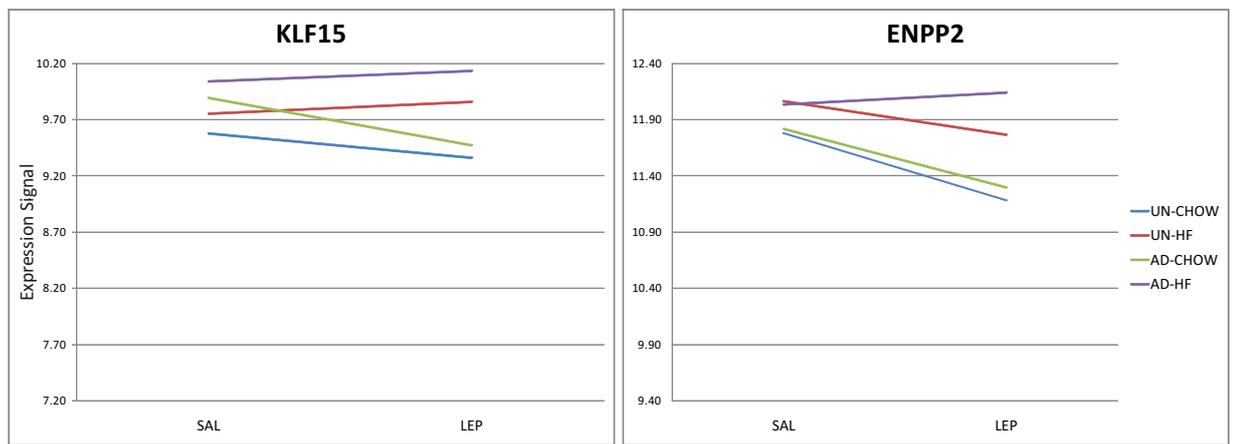


Figure 4.40: The figure on the left shows the array signals for *KLF15* and the figure on the right for *Enpp2*. The x-axis represents the effect of the leptin treatment and the y-axis represents the postweaning diet. The blue and red lines show the undernourished maternal diet (UN) and the green and purple lines show the ad-libitum fed maternal diet (AD).

to obesity in liver tissue [150]. Figure 4.40 shows an effect of leptin and the postnatal diet.

4.3.2 Genes Involved in the Immune Response

Several genes involved in the histocompatibility complex were found to be significantly, differentially expressed. These genes (*RT1-Ba*, *RT1-149*, *RT1-A3*, *RT1-CE15*, *RT1-A1*, *RT1-M62*, *RT1-CE16*, *RT1-CE12*, *RT1-CE7*, *RT1-N2*, *RT1-CE13*, and *Ca3*) were down regulated in animals that had been treated with leptin. This was the case for offspring of both AD and UN mothers. There was more down regulation in those that had subsequently received a high fat postnatal diet. The animals were 170 days of age so increased expression may be due to aging or to repair of liver damage. However, it is intriguing that the data show the down regulation in the treated animals indicating less inflammation and a lower immune response. Vila *et al.* conducted a study of hepatic gene expression on a mouse model of senescence in which the mice experience accelerated aging. They found that at 5 months (approximately 150 days) inflammation is taking place in the liver and HSP1 is upregulated. This rat model also shows an increase in expression of HSPb1 in ULC and less so

in ALC rats. Previous studies have shown that HSP1 decreases in expression with aging; the increase in this mouse model would indicate a counter-anti-inflammatory response [151]. It is possible, the decreased gene expression in the leptin treated rats could be due to errors in signalling pathways that have suppressed the immune system.

4.3.3 Relevant Genes in Other Clusters

IGFBP2 is another relevant gene that was significantly, differentially expressed in all three analyses. The FSPMA analysis revealed it as being affected by the postnatal diet, it was in cluster 1, and was significant in the pairwise comparison of ALHFvsULC. A relative, *Igfbp4*, also showed up in the FSPMA analysis, but not in the pairwise comparison. Both of these genes are expressed in mammary tissue. Igfbps are involved in fetal and neonatal development and *Igfbp2* is predominantly expressed in the liver while *Igfbp4* is only expressed in small amounts in the liver. The expression of both *Igfbp2* and *Igfbp4* is known to increase in aging. IGFs are involved in inhibition of insulin-like activity [152]. Figure 4.34 shows an obvious effect of leptin and the maternal diet for *Igfbp2*, but no effects are seen in *Igfbp4*.

4.3.4 Candidate Genes

The qRT-PCR analysis of candidate genes *11 β -hsd2*, *Pepck*, *Ppar α* , *GR*, and *Aox* showed no significant expression difference between the treatment groups (see figure 4.37). This was in contrast with the results obtained by Gluckman *et. al* [65] for expression changes in *11 β -hsd2*, *Ppar α* , and *Pepck*. Gluckman used the same tissues from the same animals so the difference in results cannot be explained. Figure 4.38 shows the qRT-PCR expression signals for comparison of the effect of each of the treatments. This figure shows some small effect from the leptin treatment on the expression of *GR* and *11 β -hsd2*.

4.3.5 Adipose Tissue

In an unpublished study, (KA Lillycrop, GC Burdge, MH Vickers, MA Hanson, and PD Gluckman) used the adipose tissue samples from the same animals used here and conducted a candidate gene expression study using RT-PCR to determine mRNA concentrations in *Ppara*, *Ppar γ 2*, *Aox*, *Cpt-1* (carnitine:palmitoyl transferase), *Lpl* (lipoprotein lipase), *GR*, *11 β -hsd2*, and leptin. These samples were normalised to ribosomal 18S RNA using the Δ CT method. The results showed that maternal diet and leptin treatment had independent effects on the gene expression and there was no leptin reversal. Similar to their results we found an effect of the maternal diet and postnatal diet on GR mRNA expression, but there was no effect on *Ppara*, *Aox*, or *11 β hsd-2* (or *Pepck*). This was not statistically significant but the comparisons are shown in figures 4.37 and 4.38.

The RNA had been extracted from muscle tissue for the same animals; however, due to time constraints these samples were not used. It would be interesting to see how the expression changes in muscle differ from fat and liver. In addition in future studies it would be of interest to collect brain tissue as this would give insight into the origin of signalling errors.

4.4 Conclusions

The initial goal of this study was to uncover the molecular mechanism of the leptin reversal; however, we did not uncover any evidence of the leptin reversal. This finding was supported by the unpublished work done by Lillycrop *et al.*. Instead we found, the interaction between the maternal diet, leptin treatment, and the postweaning diet could have significant effects on gene expression in a number of metabolically relevant genes. Many of the changes were supported by functional data found in the literature. A comparative analysis of the results of this study and those of the study in Chapter 3 revealed a large number of differentially expressed genes, however, these were confounded due to the difference in sex of animals and time-point of tissue collection. This has highlighted the need to perform a more detailed study

incorporating different tissues taken at different time points to really understand the changes that have occurred. In this study the effect of leptin was associated more with the postweaning chow diet rather than the postweaning high-fat diet. This is an unexpected result and difficult to explain. A simplified study design that looks at the immediate effects of leptin may help unravel the details of the cascade of events that lead to the metabolic syndrome.

Chapter 5

Methylation Study

5.1 Introduction

The work presented in this chapter aimed to investigate whether epigenetic effects through methylation at CpG dinucleotides in promoter sequences, are responsible for gene expression and resulting phenotypic changes in the rat model. We wanted to gain a general impression of the presence, scale, and significance of such methylation change, by exploring a diverse range of genes that were highlighted in the previously described genome-wide expression studies. This chapter includes experiments done by a research student under the author's guidance. Neil Graham was a part II student in 2007-2008, that used DNA previously extracted from the AD/UN liver samples described in Chapter 3, to look for methylation differences in differentially expressed genes with the McrBC assay (see Chapter 2). Neil was able to follow up on 87 genes highlighted in the microarray analysis.

A second research student, Chathika Weerasuriya, was a summer student in 2008, that extracted DNA from the leptin liver samples described in Chapter 4 and preceded to treat them with bisulphite conversion and then, use PCR and sequencing to determine if there were methylation changes. Due to the inability to get high quality DNA from the leptin samples, the bisulphite conversion that summer was largely unsuccessful and no results were produced.

This chapter begins with background information on epigenetics, specifically methy-

lation and how it has been tied to the metabolic syndrome. This is followed by a description and discussion of the results.

5.1.1 Epigenetics

The term epigenetics was first coined by Conrad Waddington [153], to unite the fields of developmental biology and genetics. His definition was general and said simply, ‘All those events which lead to the unfolding of the genetic program for development’. Over the years, this term has taken on the more specific definition to describe the changes in gene function that occur without a change in gene sequence. It refers to modifications that regulate gene activity; the modifications can affect the DNA itself or the proteins that package it (histones), but the DNA sequence does not change [154]. Figure 5.1 illustrates the epigenetic mechanisms that can cause changes in gene activity. One component of epigenetic regulation, imprinting, allows the control of whether the maternal or paternal allele for a particular gene is expressed. For instance, the epigenetic imprinting of IGF2 regulates fetal growth and fetal size depending which of the parents’ alleles is silenced [155]. Methylation occurs depending on the availability of methyl groups and can be affected by nutrition. As discussed in this thesis, changes in gene expression can increase the risk of disease. Changes in components of epigenetic regulation affect gene expression and in turn, can lead to disease. Knowing the epigenotype related to a disease state could help to target treatments for disease.

In mammals, there are two main developmental windows when epigenetic modifications occur. During gametogenesis, complete demethylation occurs, followed by remethylation before fertilisation. A second demethylation event occurs in early embryogenesis, and then in early embryonic life, just after implantation, the methylation is re-established. The imprinted genes are able to escape this second wave of demethylation [157]. Other genes may also be able to escape this second wave of demethylation. This may explain transgenerational effects of undernutrition, as observed in offspring of those that were pregnant during the Dutch Hunger Winter, for

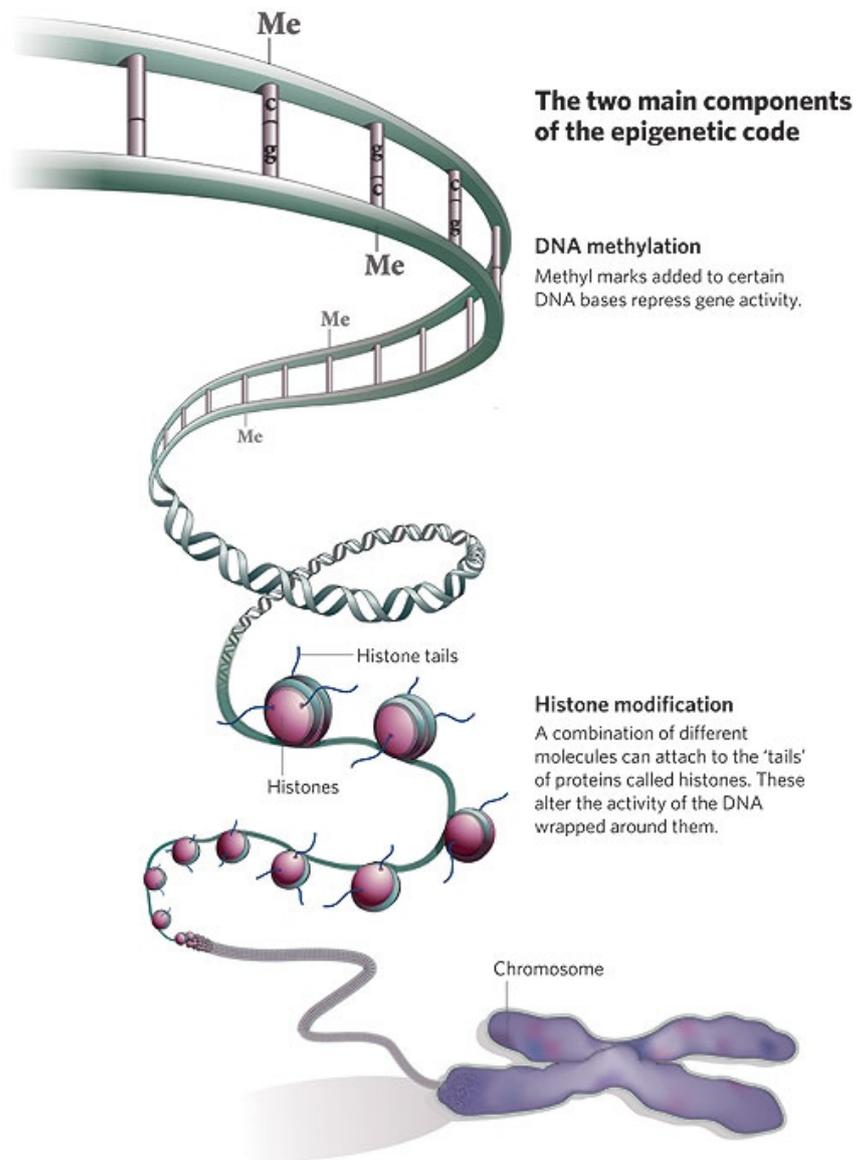


Figure 5.1: Epigenetic Mechanisms. Epigenetic changes are referring to modifications that regulate gene activity without changes in the DNA sequence. Two common mechanisms for epigenetic changes are shown in this figure adapted from Qiu [156]. The modifications can affect the DNA itself (through methylation) or the proteins that package it (histone modification).

example. These events may represent critical windows of time when reprogramming could occur.

Methylation

Methylation is essential for the development of mammals and it has been shown that embryos that cannot maintain normal methylation die after gastrulation [158]. DNA methylation is the addition of a methyl group CH_3 to the fifth carbon position of the cytosine pyrimidine ring via a methyltransferase enzyme. There are three major enzymes involved in the establishment and maintenance of DNA methylation pattern: DNMT 3A and 3B are *de novo* methyltransferases, and DNMT1 is the maintenance DNMT that ensures that methylation patterns are copied throughout each cell division [159]. The majority of DNA methylation in mammals occurs in 5'-CpG-3' dinucleotides, but other methylation patterns do exist. In fact, about 80% of all 5'-CpG-3' dinucleotides in mammalian genomes is methylated, whereas the majority of the 20% that remain unmethylated is within promoters or in the first exons of genes. CpG dinucleotides, the site of almost all methylation in mammals, are underrepresented in DNA. Clusters of CpGs called CpG islands, are often found in association with genes, most often in the promoters and first exons, but also in regions more toward the 3' end [160].

CpG islands often function as strong promoters and have been proposed to function as replication origins. Even though they are generally not methylated, most investigations into the role of DNA methylation in mammals have focused on CpG islands, rather than on the regions in which the majority of methylation is found [161]. There is a growing understanding as to how the methylation signal is interpreted by mammalian cells. The post-synthetic addition of methyl groups to the 5-position of cytosines alters the appearance of the major groove of DNA to which the DNA proteins bind. These epigenetic 'markers' on DNA can be copied after DNA synthesis, resulting in heritable changes in chromatin structure. Methylation of CpG-rich promoters is used by mammals to prevent transcriptional initiation and to ensure the silencing of genes on the inactive X chromosome, imprinted genes, and parasitic

DNAs [161]. Methylation has been shown to block gene transcription by two general mechanisms. The first involves the modification of the cytosine bases. The methylation at a single site within the binding region for a transcription factor can strongly inhibit the binding of the factor to the DNA recognition sequence [162]. The second involves methyl-CpG-binding proteins (MBPs) that recognise methyl CpGs and use transcriptional co-repressor molecules to silence transcription and to modify surrounding chromatin. MBPs provide a link between DNA methylation and chromatin remodelling and modification [163, 164]

Histone Modification

Histone modifications (particularly posttranslational modifications of amino-terminal tail domains) are another important epigenetic mechanism for controlling gene expression [165]. Histone acetylation is associated with active gene transcription while other histone modifications such as the methylation of histone H3 lysine 9 (H3K9) are indicative of condensed and inactive chromatin [166]. Histone modification and DNA methylation work together in a reciprocal relationship to control gene expression. Methylation of histone H3 lysine 9 can be triggered by DNA methylation [167]. DNA methyltransferases have been shown to interact with histone deacetylases, histone methyltransferases, and methyl-cytosine-binding proteins in a complex network [168, 169, 170].

5.1.2 Epigenetic Change, Developmental Programming, and the Metabolic Syndrome

Epigenetic modifications offer the mechanism for the fixed genome to have a flexible means to alter gene regulation that can be transferred through mitosis without changing the DNA sequence. Epigenetics enables an organism to respond to the environment and change its gene expression, offering phenotypic plasticity with a fixed genotype [171, 33]. Waterland and Michels proposed the term ‘Epigenetic Epidemiology’ as the study of the associations between epigenetic variation and the risk of

disease based on the increasing evidence that epigenetic mechanisms play a role in DOHaD [172]. This is similar to genetic epidemiology, but rather than focusing on the role of inherited genes in disease etiology, epigenetic epidemiology also considers epigenetic inheritance, developmental stochasticity, environment influences, and aging. Recent epigenomic advances are accelerating the rate of discovery of human loci where epigenetic regulation can be correlated with early environmental exposures. Linking this to a growing understanding of the mechanisms by which epigenetic change leads to metabolic syndrome is key to the future research in DOHaD.

Wolff *et al.* conducted the first study that demonstrated that the maternal diet during pregnancy could trigger epigenetic changes in the offspring [173]. The study was done on mice with the agouti gene (A^{vy}). The agouti gene is an example of a gene with metastable epialleles. Metastable epialleles are gene loci that can be modified by an epigenetic component in a variable or reversible way. This will result in a range of phenotypes from one genotype. The agouti gene encodes a signalling molecule that promotes the production of yellow phaeomelanin rather than black eumelanin pigment. If the degree of methylation varies, then it can produce a wide distribution in coat colour (from yellow for unmethylated alleles to brown for methylated). Changing the maternal dietary supplementation to include methyl-donor nutrients (i.e. folic acid, vitamin B_{12} , choline, and betaine) shifts the coat colour of the offspring toward the brown phenotype. In addition to the change in coat colour, the hypomethylation of the agouti gene promoter leads to obesity and in a particular strain of mouse it leads to cancer [174]. Figure 5.2 shows mice that have differing levels of methylation of the agouti gene based on the maternal diet they were exposed to during gestation. Other studies revealed that, epigenetic changeability based on early nutrition is a characteristic of metastable epialleles. In humans, population studies have shown folate status is highly correlated with coronary artery disease [175]; however, this has not yet been linked to a change in methylation of affected genes.

Additional studies in rodent models of intrauterine growth restriction have shown that maternal diet during pregnancy leads to epigenetic changes in the offspring. A study on protein-restricted rats found that the promoters of GR and $PPAR\alpha$ were



Figure 5.2: Diet and Methylation. This image, adapted from Waterland *et al.* [174], shows mice that have been exposed to different maternal diets during gestation, which has resulted in differing levels of methylation of the Agouti gene. This has caused changes in the phenotype and has also made the mice prone to obesity and cancer.

hypomethylated in the offspring and were associated with increase gene expression [176]. A study on hypertension used the offspring of protein restricted dams and found decreased methylation in the promoter of the *Agr1b* gene and increased expression [177]. These changes affected the regulation of blood pressure and were associated with the development of hypertension in later life. These findings give evidence to the hypothesis that maternal nutrition may affect the methylation in the offspring and in turn, lead to disease later in life. Methylation changes in offspring of individuals exposed to the Dutch hunger winter has shown that maternal diet can have long-term effects on the epigenome [178]. These individuals had decreased levels of methylation of the *IGF2* gene in adulthood. In addition to nutrition, other environmental factors including maternal behaviour during pregnancy and early life of the offspring can affect the establishment and/or maturation of epigenetic mechanisms, causing persistent changes in gene expression [172].

5.1.3 Available Techniques for Methylation Analysis

Recent advances in genomics have led to the development of multiple new techniques for DNA methylation profiling and the profiling of histone modifications.

These techniques cover a range of resolution and throughput. A very recent review, by Laird *et al.* [179], compares and contrasts these techniques. Measurement of methylation can be made by determining the pattern of methylated target sequences along individual DNA molecules or by finding an average methylation level at a single genomic locus across many DNA molecules. The uneven distribution of methylation in CpG islands creates an extra complication. Simple PCR to look for methylation changes is not possible, as methylation is erased during amplification. This makes it necessary to perform some type of methylation dependent treatment before the amplification process. The three approaches to the methylation treatment are: endonuclease digestion, affinity enrichment, and bisulphite conversion. After one of these approaches has been used to treat the DNA, a variety of techniques are available (i.e. probe hybridisation and sequencing) to reveal the location of the methylated cytosine residues.

Endonuclease Digestion

Restriction enzyme digestion utilises methylation sensitive restriction enzymes for DNA methylation studies. These enzymes are inhibited by a 5-methyl-C in a CpG region of a sequence. This reveals DNA methylation by the pattern of cutting that is determined from the digestion. The most common are *HpaII* and *SmaI*. This is largely because for each of these enzymes there is another enzyme (*MspI* and *XmaI* respectively) that is not inhibited by CpG methylation [179]. This is the method of methylation detection that has been utilised for the data presented in this chapter.

Affinity Enrichment

Chromatin immunoprecipitation is a technique used to identify the location of DNA-binding proteins and epigenetic marks in the genome. Genomic sequences that are methylated or contain another mark of interest are enriched by binding soluble DNA chromatin extracts to an antibody that recognises the mark. ChIP-chip technology uses this technique followed in conjunction with microarray hybridisation and is very useful for identifying histone modifications. Affinity purification is use-

ful for identifying methylation patterns when a methyl-binding protein is used (i.e. MECP2). Techniques that use this approach include MeDIP, mDIP, or mCIP.

Bisulphite Conversion

Sodium bisulphite conversion is a chemical treatment of DNA that turns an epigenetic difference into a genetic difference (unmethylated Cs are converted to Ts by uracil). This enables the detection of DNA methylation using PCR or sequencing. Bisulphite sequencing is the best method for producing base-pair resolution. The Golden Gate BeadArray technology has been adapted to interrogate DNA methylation using bisulphite converted DNA. Illumina has also developed the Infinium platform, which is a more comprehensive approach to DNA methylation analysis.

5.1.4 Summary

The techniques mentioned above can be coupled with array based analysis or sequencing approaches to obtain a high-throughput method for methylation interrogation. An alternative approach to sequencing or hybridisation, involves detection by matrix-assisted laser desorption ionization time-of-flight (MALDI-TOF) mass spectrometry. This requires gene-specific amplification and is best suited to a candidate gene study. Sequenom has developed the EpiTYPER which enables automation for a large number of samples. A number of factors will play into the choice of technique used including number of samples, quality and quantity of DNA, desired coverage, resolution, and cost. Figure 5.3 shows a variety of different techniques and platforms available, the different enzymatic pre-treatment each requires, and the analysis method used.

5.1.5 Study Design

In this thesis, endonuclease digestion has been utilised to look for methylation changes highlighted in the gene expression profiling study. Genes were selected based on the results of the MEEBO expression study and preliminary microarray expression

Main principles of DNA methylation analysis				
Pretreatment	Analytical step			
	Locus-specific analysis	Gel-based analysis	Array-based analysis	NGS-based analysis
Enzyme digestion	<ul style="list-style-type: none"> • <i>HpaII</i>-PCR 	<ul style="list-style-type: none"> • Southern blot • RLGs • MS-AP-PCR • AIMS 	<ul style="list-style-type: none"> • DMH • MCAM • HELP • MethylScope • CHARM • MMass 	<ul style="list-style-type: none"> • Methyl-seq • MCA-seq • HELP-seq • MSCC
Affinity enrichment	<ul style="list-style-type: none"> • MeDIP-PCR 		<ul style="list-style-type: none"> • MeDIP • mDIP • mCIP • MIRA 	<ul style="list-style-type: none"> • MeDIP-seq • MIRA-seq
Sodium bisulphite	<ul style="list-style-type: none"> • MethylLight • EpiTYPER • Pyrosequencing 	<ul style="list-style-type: none"> • Sanger BS • MSP • MS-SNuPE • COBRA 	<ul style="list-style-type: none"> • BiMP • GoldenGate • Infinium 	<ul style="list-style-type: none"> • RRBS • BC-seq • BSPP • WGSBS

AIMS, amplification of inter-methylated sites; BC-seq, bisulphite conversion followed by capture and sequencing; BiMP, bisulphite methylation profiling; BS, bisulphite sequencing; BSPP, bisulphite padlock probes; CHARM, comprehensive high-throughput arrays for relative methylation; COBRA, combined bisulphite restriction analysis; DMH, differential methylation hybridization; HELP, *HpaII* tiny fragment enrichment by ligation-mediated PCR; MCA, methylated CpG island amplification; MCAM, MCA with microarray hybridization; MeDIP, mDIP and mCIP, methylated DNA immunoprecipitation; MIRA, methylated CpG island recovery assay; MMass, microarray-based methylation assessment of single samples; MS-AP-PCR, methylation-sensitive arbitrarily primed PCR; MSCC, methylation-sensitive cut counting; MSP, methylation-specific PCR; MS-SNuPE, methylation-sensitive single nucleotide primer extension; NGS, next-generation sequencing; RLGs, restriction landmark genome scanning; RRBS, reduced representation bisulphite sequencing; -seq, followed by sequencing; WGSBS, whole-genome shotgun bisulphite sequencing.

Figure 5.3: Methylation Techniques. This figure from Laird *et al.* shows a majority of the methylation techniques available. The chart shows which enzymatic pretreatment and which analytical step is used for each of the available techniques.

results for the Illumina expression study, presented in Chapter 3. There were 87 genes selected and analysed according to the pipeline shown in figure 5.4.

5.2 Results

Of the 87 genes analysed in the methylation assay, 12 were significantly differentially expressed on the MEEBO microarray platform and 26 were significantly differentially expressed in the final analysis of the data from the Illumina platform. The full list of 87 genes is shown in table 5.1. The table reveals which chip the gene was found on. Those genes found using the MEEBO chip are labelled ‘M’ and those that were in the final Illumina analysis are labelled ‘I-FL’. The remaining genes, labelled ‘I’ that were found in the preliminary Illumina analysis, are not considered significantly differentially expressed, but are included here, as a significant amount of work did go into locating their CpG islands and determining the methylation status of these genes. The sequence for each of the genes was obtained and promoters were found. The region 5kb upstream of the promoter was analysed in search of CpG islands using the EBI CpG island prediction software. In the instances where no island was detected, the UCSC CpG searching tool was also used with a less strin-

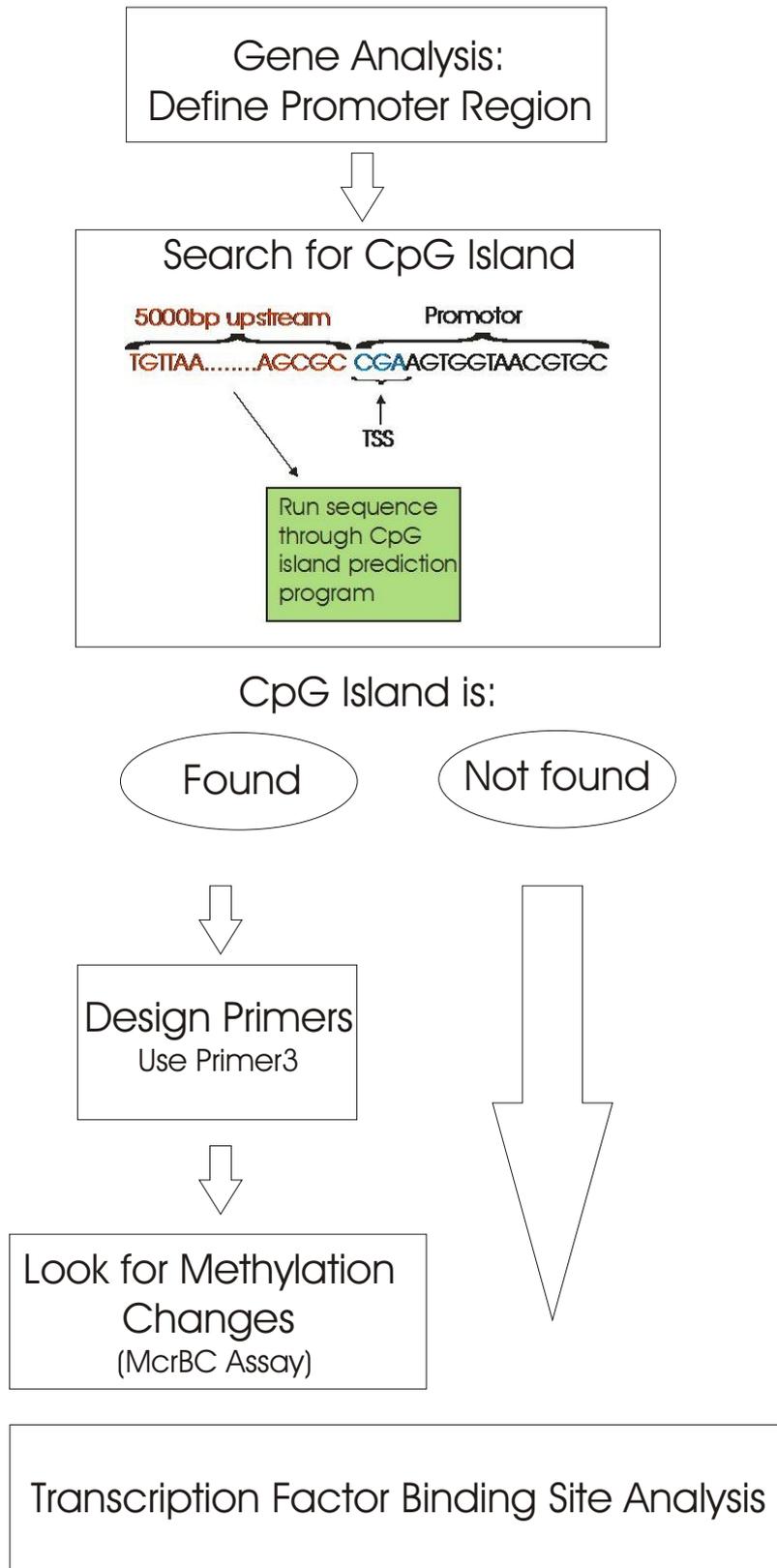


Figure 5.4: Methylation Analysis Pipeline. This figure shows steps followed in this chapter to assess the methylation status of the genes highlighted in the microarray expression studies.

gent search algorithm to provide negative confirmation. One CpG island or more was identified in 56% of the promoter regions; flanking primers were designed using Primer3 software. In cases where two islands were found, primers were designed for both. The primers were then used in the McrBC assay. Separate pools of DNA for the 8 biological replicates from ad-libitum fed (AD) and undernourished (UN) rats were digested by methylation sensitive restriction enzymes. The high frequency cutting enzymes make it possible to assay a large proportion of CpG islands. The enzymes used were: *HpaII* (cuts unmethylated alleles), *McrBC* (cuts methylated alleles), and *MspI* (used as a negative control as it cuts independent of methylation). Digestions of *HpaII* and *McrBC* give four possible results: full, null, incomplete, or composite methylation. Before running the assay, the New England Biolabs NEBcutter tool was used to confirm *HpaII* and *MspI* cutting sites. In addition, GC content was confirmed by searching for the frequency of rare-cutter sites (*NarI*, *NaeI*, *SacII*, *NotI*, *BssHIII*, *EagI*, *SmaI*, *MnII*, and *MluI*). Products were separated by gel electrophoresis, and then the methylation status of the island was determined by the relative band intensities. The findings and details of each gene are shown in table 5.1. In addition, to the gene name and the chip the gene was found on, the table shows the direction of the expression change, whether or not the gene has CpG islands, whether or not primers were designed for methylation analysis, and whether a methylation change was revealed by the McrBC assay. Finally, the promoter sequences of each of the genes were analysed for possible transcription factor binding sites. The assay succeeded for 18 genes and there was little or no apparent methylation difference between the AD and UN samples. CpG islands were generally unmethylated in the AD samples (61%). Of the 18 genes assayed, 5 were from the preliminary Illumina expression analysis and their PCR results will not be included here. Figures 5.5 and 5.6 show the PCR images from the McrBC assay for the 13 genes that were shown as differentially expressed in the MEEBO expression study or the Illumina expression study (final list). Differential methylation was observed in 4 of the 18 successful assays (*Mapk4*, *Tnfrsf13*, *Zfand2a*, and *Atf4*), all of which were demethylated. Two of these genes were found as significantly differentially expressed in one of the microar-

ray platforms. *Mapk4* was differentially expressed on the MEEBO chip, and *Zfand2a* was differentially expressed in the final analysis of the Illumina data. The other two genes (*Tnfsf13* and *Atf4*) came from the preliminary Illumina analysis and are not considered significant.

Gene Symbol	Chip	Exp. Change	CpG islands	Primers Designed	Initial Meth.	Meth. Change
Aass	M	↓	0	N/A	Unk	N/A
Adamts1	I-FL	↑	2	yes	Not	No Change
Adn	I	↑	0	N/A	Unk	Not Examined
Adra1b	I-FL	↓	1	yes	Unk	PCR fail
Amd1	I	↑	2	yes	Not	No Change
Arhgef19_pred	I-FL	↓	3	yes	Mixed	No Change
Arl6ip2_pred	I	↓	1	no	Unk	Not Examined
Atf4	I	↑	1	yes	Mixed	Demethylation (isl1)
Besh3	I-FL	↓	1	no	Unk	N/A
Bhlhb2	I	↓	1	no	Unk	Not Examined
C4-2	I	↑	0	N/A	Unk	Not Examined
Calm3	I	↓	1	no	Unk	Not Examined
Cbx3	M	↑	0	N/A	Unk	N/A
Cdo1	I	↑	1	no	Unk	Not Examined
Ces3	I	↑	0	N/A	Unk	Not Examined
Cidea_pred	I-FL	↓	1	yes	Not	No Change
Cited2	I	↓	1+	no	Unk	Not Examined
Cpa1	I	↑	0	N/A	Unk	Not Examined
Crlz1_pred	I-FL	↑	2	yes	Not	No Change
Csng	M	↑	0	N/A	Unk	N/A
Cyp2c40	I	↓	0	N/A	Unk	N/A
Dhtkd1_pred	I-FL	↓	1	yes	Not	No Change
Dnajb9	I	↓	1	no	Unk	Not Examined
Dnase1l3	I-FL	↑	2	yes	Unk	PCR fail
Eef2k	I	↓	1+	no	Unk	Not Examined
Eif4g2_pred	I	↓	1	no	Unk	Not Examined
Fabp7	I-FL	↑	0	N/A	Unk	N/A
Fcgr2b	I-FL	↑	0	N/A	Unk	N/A
Fxyd6	I-FL	↑	0	N/A	Unk	Not Examined
Galnt11	I	↑	1	yes	Not	No Change
Gng11	I-FL	↑	1	N/A	Unk	N/A

Gene Symbol	Chip	Exp. Change	CpG islands	Primers Designed	Initial Meth.	Meth. Change
Gstp2	I-FL	↑	1	yes	Not	No Change
Igfbp2	I-FL	↑	1	yes	Mixed	No Change
isg12(b)	I	↑	1	no	Unk	Not Examined
LOC299823	I	↑	0	N/A	Unk	Not Examined
LOC301711	I	↓	0	N/A	Unk	Not Examined
LOC303576	I	↑	0	N/A	Unk	Not Examined
LOC303861	I-FL	↓	0	N/A	Unk	N/A
LOC314964	I	↓	1	no	Unk	Not Examined
LOC361061	I	↑	0	N/A	Unk	Not Examined
LOC363151	I-FL	↑	1	yes	Unk	PCR fail
LOC365214	I	↓	0	N/A	Unk	Not Examined
LOC366485	I	↑	1+	no	Unk	Not Examined
LOC366941	I	↑	0	N/A	Unk	Not Examined
LOC498731	I	↑	0	N/A	Unk	Not Examined
LOC500322	I	↑	1+	no	Unk	Not Examined
Magel2	M	↑	0	N/A	Unk	N/A
Mapk4	M	↓	2	yes	Mixed	Demethylation (isl2)
MGC105601	I-FL	↑	0	N/A	Unk	N/A
MGC94010	I-FL	↑	0	N/A	Unk	N/A
Mgst1	I	↑	0	N/A	Unk	Not Examined
Mx1	I	↑	0	N/A	Unk	Not Examined
Mybph	I	↑	0	N/A	Unk	Not Examined
Myl2	M	↓	0	N/A	Unk	N/A
Npm1	I-FL	↓	2	yes	Not	No Change
Olfr802	M	↑	0	N/A	Unk	N/A
Paip1_pred	I	↓	1	no	Unk	Not Examined
Pck1	I	↑	0	N/A	Unk	Not Examined
Pde4d	I	↓	1	no	Unk	Not Examined
Pdia5	M	↓	0	N/A	Unk	N/A
Phka1	I	↓	1	no	Unk	Not Examined
Plekhb1	I	↑	0	N/A	Unk	Not Examined
Plvap	I	↑	1	no	Unk	Not Examined
Pmp22	I	↑	0	N/A	Unk	Not Examined
Ppap2a	I	↑	0	N/A	Unk	Not Examined
Psip1	M	↓	1	yes	Unk	PCR fail

Gene Symbol	Chip	Exp. Change	CpG islands	Primers Designed	Initial Meth.	Meth. Change
Ptprb_pred	I-FL	↑	0	N/A	Unk	N/A
Rabepk	M	↑	1	yes	Unk	PCR fail
Rbp1	I-FL	↑	1	yes	Unk	PCR fail
RGD1306512_pred	I-FL	↓	1	yes	Unk	PCR fail
RGD1308082	I-FL	↓	1	yes	Not	No Change
Zfand2a	I-FL	↑	1	yes	Mixed	Demethylation
RGD1311155_pred	I	↑	0	N/A	Unk	N/A
Rpl10l_pred	I	↓	1	yes	Mixed	No Change
Rpl23a_pred	I	↑	1+	no	Unk	Not Examined
Sara2	I	↓	1	no	Unk	Not Examined
Serpinb6	I-FL	↑	1	yes	Not	No Change
Sgpl1	I-FL	↑	0	N/A	Unk	N/A
Slfn3	I	↑	0	N/A	Unk	Not Examined
Tceal8	I	↓	0	N/A	Unk	Not Examined
Tcf15_pred	I	↓	1	no	Unk	Not Examined
Thbs4	I	↓	0	N/A	Unk	Not Examined
Tmpo	I	↓	1	no	Unk	Not Examined
Tnfsf13	I	↑	1	yes	Mixed	Demethylation
Ubl3_pred	I	↓	1	no	Unk	Not Examined
Usf2	M	↑	1	yes	Unk	PCR fail
XM_978865	M	↓	1	yes	Not	No Change

Table 5.1: Table of Methylation. This table shows the 87 genes that were chosen from the microarray study presented in Chapter 3. The promoter regions of the genes were characterised, CpG islands were identified, primers were designed and the McrBC assay was performed. This table shows the results of each process of the pipeline for each gene including: the chip it was found on (MEEBO ‘M’ or Illumina ‘I’) and whether it made it to the final Illumina gene list ‘I-FL’, the direction of expression change found in microarray, the number of CpG islands, whether primers were designed for CpG Islands, whether the McrBC assay was performed and successful and if this assay showed a methylation change.

Bisulphite Sequencing

Additional work attempting to use bisulphite sequencing for methylation analysis was attempted using these DNA samples. After multiple attempts, bisulphite conversion was finally successful, but problems with sequencing caused the termination of this avenue of investigation. Primers designed for these assays are listed in Appendix A.

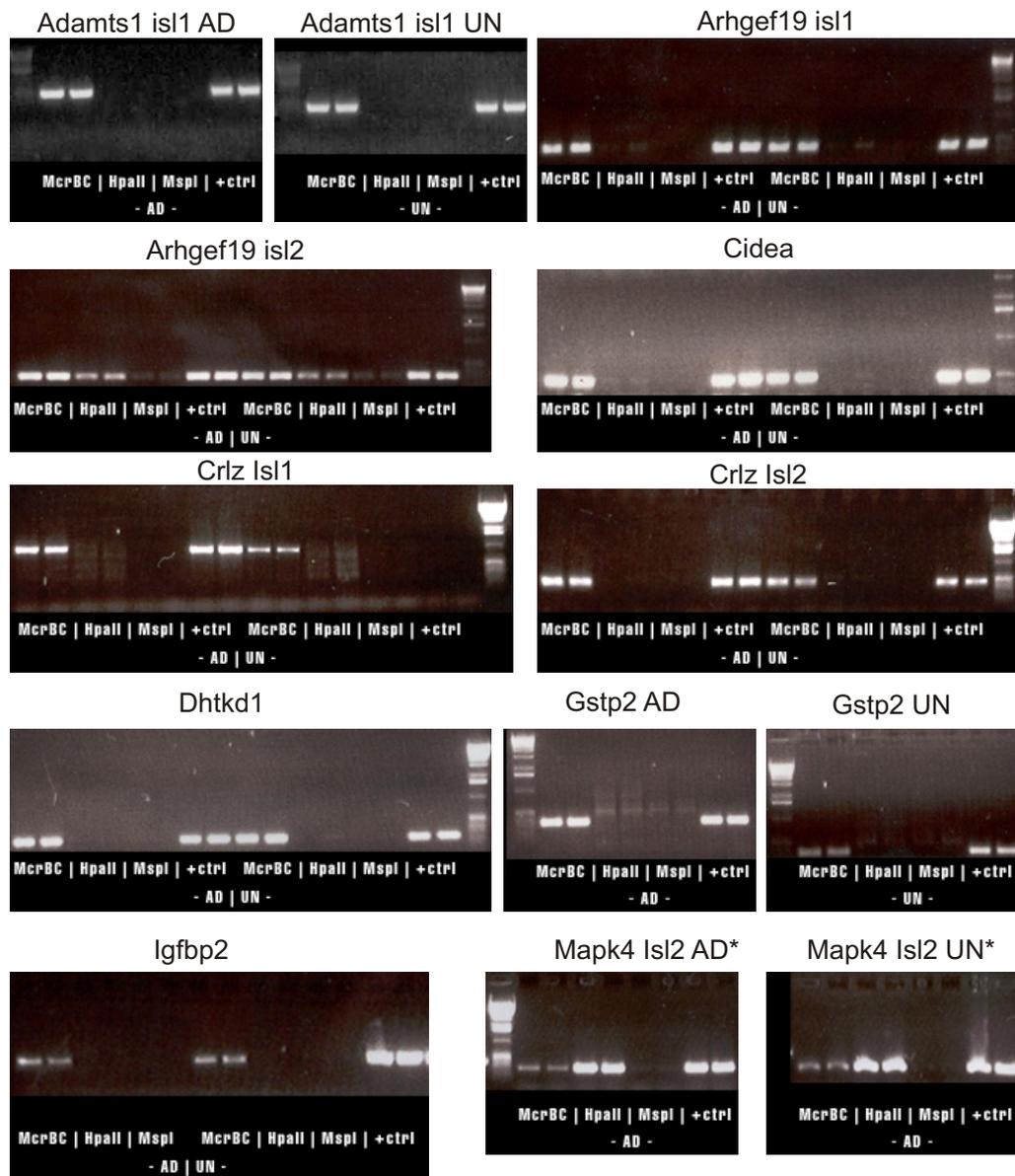


Figure 5.5: PCR results for 8 of the 13 genes that were significantly differentially expressed in the microarray study and were successfully assayed using McrBC. Some genes were assayed twice for different CpG islands. Those that show a methylation change due to the nutrition treatment *in utero* are indicated with an *.

5.3 Discussion

This chapter aimed to investigate the methylation status of genes that were found to be differentially expressed in the livers of day 55 rats exposed to maternal under-nourishment during gestation (as presented in Chapter 3). This study was conducted by a part II student during the preliminary stages of the analysis of the Illumina

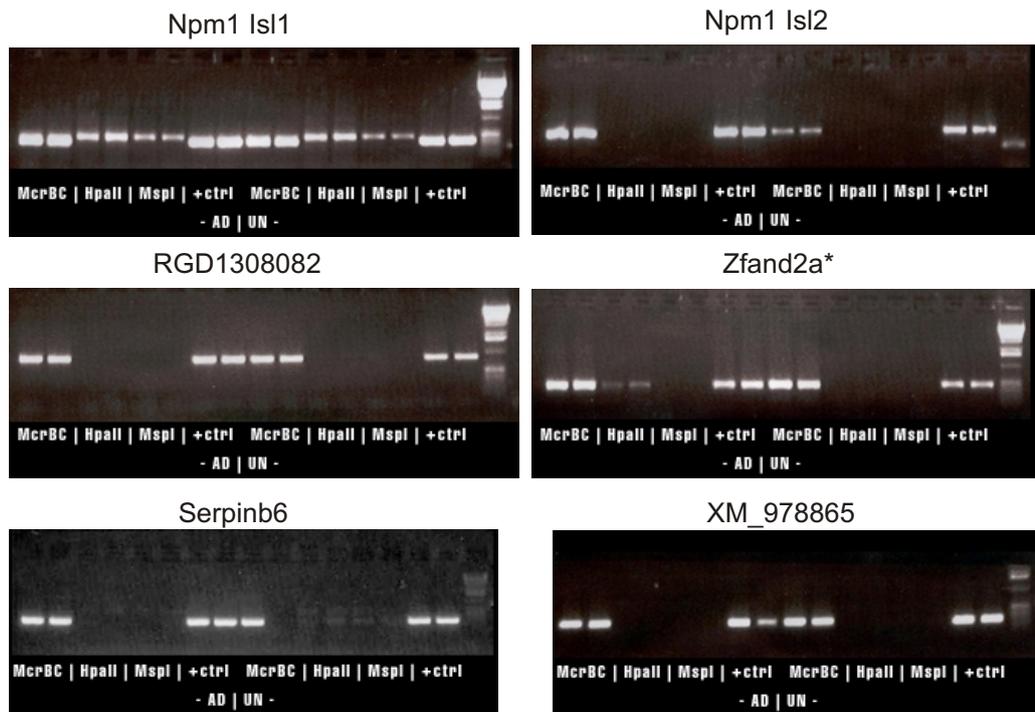


Figure 5.6: PCR results for 5 of the 13 genes that were significantly differentially expressed in the microarray study and were successfully assayed using McrBC. Some genes were assayed twice for different CpG islands. Those that show a methylation change due to the nutrition treatment *in utero* are indicated with an *.

BeadArray data set. For this reason, many of the genes that were assayed for this chapter did not show up in the final microarray statistical analysis. Of the 87 genes analysed, only 26 were in the final list of differentially expressed genes presented in Chapter 3. Of these 26, only one showed a change in methylation. This gene was *Zfand2a*. Another 12 included in the methylation analysis were genes found to be differentially expressed on the MEEBO microarray platform. Of these 12, one showed a methylation change in the McrBC assay. That gene was *Mapk4*. The discussion below will elaborate more on the function of *Zfand2a* and *Mapk4* and their potential role in the metabolic syndrome. Due to the timing of this study, many genes that did show up in the final microarray analysis were not assayed. Therefore, this chapter does not represent a thorough analysis of the methylation status of the genes implicated in Chapter 3, but is a preliminary study and a test of the McrBC assay and bisulphite sequencing as methods for future use.

5.3.1 *Zfand2a*

Zfand2a was found to be upregulated by 40% on the Illumina microarray. This is an arsenite inducible RNA associated protein that adapts proteosomes to counteract stress-induced proteotoxicity [117]. This transcript is located on chromosome 12. The CpG island search revealed one island spanning the transcription start site. Separation of PCR products on an agarose revealed the presence of bands in the HpaII AD digests, absent in the corresponding HpaII UN position, see figure 5.6. Given that HpaII cuts DNA at unmethylated positions, this indicates the presence of some methylated CpGs in the AD pool, and thus, failed digestion, permitting PCR amplification of the fragment. Absence of visible bands in the HpaII UN position thus suggests an increase in the proportion of unmethylated CpGs. It would appear that a subtle change in the direction of demethylation has taken place.

An analysis to characterise likely sites for transcription factor binding was performed. The CpG island spans the transcription start site and four potential binding sites were found (CREB, E2F, c-Ets, and TATA). It is proposed, these sites, playing a role in the binding of the basal transcription machinery, may be responsible for mediating the repressive effects on transcription of generalised promoter methylation. Results of transcription factor binding site analysis are shown in Appendix C Figure 1. Upregulation of *Zfand2a* may indicate an alteration in oxidative stress due to the effects of undernutrition. Sequence similarity also suggests *Zfand2a* may be involved in nucleic acid interactions.

5.3.2 *Mapk4*

Mitogen-Activated Protein Kinase 4 (*Mapk4*) is located on chromosome 18 and gene ontology entries suggest a role in the regulation of cell cycle processes and thus, this gene product may be important in producing long-term pathogenic phenotypes implicated in the metabolic syndrome. The CpG island prediction software revealed two CpG islands. Primers flanking the first island (the long transcript) failed to amplify efficiently. However, in island 2, following separation of PCR products on an

agarose gel (see figure 5.5), the McrBC UN band is more intense than the corresponding AD band. As McrBC digests methylated sites, this change indicates decreased methylation due potentially to undernutrition *in utero*. This decreased methylation would suggest increased expression of (*Mapk4*); however, the Illumina gene expression assay highlighted it as being downregulated. This result is paradoxical and further investigation is necessary to clarify what is going on. Computational analysis to show likely transcription factor binding sites within the CpG island showing potential differential methylation were performed. The nucleotides in the promoter sequence comprising the CpG island in the short transcript were selected and TF binding site analysis was performed on the reverse complement. One possible transcription factor binding site (CREB) was found in the CpG island that could be affecting the expression of downstream genes. Bisulphite sequencing or protein assays would enable verification. Results of transcription factor binding site analysis are shown in Appendix C Figure 1.

5.4 Summary

A number of studies have investigated the methylation status of genes associated with metabolism that have previously been highlighted as candidate genes for the metabolic syndrome. Fujiki *et al.* found the promoter of the *Ppar γ 2* gene is hypermethylated in pre-adipocytes, then it is progressively demethylated, as the differentiation into adipocytes progresses and as mRNA expression increases [180]. In addition, they found decreased *Ppar γ* mRNA and increased methylation of the *Ppar γ* promoter in the visceral adipose tissue of a diabetic mouse model compared to the wild-type mice. This suggests the methylation of the promoter region of *PPAR γ* influences the expression of the downstream gene.

A study by Plagemann *et al.*, investigated the effects of prenatal and neonatal overfeeding on methylation patterns in hypothalamic promoter regions of genes involved in the regulation of food intake and body weight [181]. The study found nutritional intake affects the methylation pattern and as a result, the ‘set-point’ is

altered in the gene promoter of genes involved in body weight regulation. This ties in closely with the introduction to Chapter 4.

In this study, that majority of genes showed no change in methylation status. As the sites tested were all upstream of a CpG island, the results are consistent with studies that show CpG islands are largely unmethylated and methylation changes tend to occur in intergenic regions [182]. In addition, work on *Pdx1* has shown that histone modifications precede methylation changes. Progressive histone modifications accompanied the gradual decrease in *Pdx1* expression [183]. The increased histone modifications occur as glucose homeostasis deteriorates and oxidative stress increases. This suggests that the expression changes could be the result of a cascade of events and that testing for methylation changes would not necessarily reveal the mechanism for decreased expression.

A more detailed epigenetic investigation into the methylation changes and histone modifications occurring in the tissues from Chapter 4, in addition, to an expression study of the hypothalamus of rats similarly undernourished during pregnancy, would be interesting to tie together the interaction of brain and peripheral organs along with the interaction of methylation, histone modification, and expression changes.

The work in this chapter barely scratched the surface of interesting avenues of exploration in the understanding of the involvement of methylation in developmental programming and the later occurrence of the metabolic syndrome. This chapter was intended to be a preliminary study to see if interesting methylation differences could be uncovered in the genes highlighted in the gene profiling study from Chapter 3. The methylation work done revealed that most genes showed no change in methylation status. There are three possible explanations for this result. First, the McrBC assay is not as sensitive as a sequencing based approach and therefore by using this assay subtle changes in methylation status may be overlooked. The assay is not sensitive enough to detect methylation differences among cell types. Liver cells may show a heterogeneous pattern of methylation between cell types which would have been lost when RNA was extracted. Second, this thesis has been investigating gene expression profiles as a consequence of the change in maternal diet during pregnancy. The

changes we may be observing could be due to a secondary change, which are a result of methylation changes in upstream primary regulators. Third, we are looking at animals that are 55 days old. While the majority of epigenetic modifications occur early in life it is possible that a key change has occurred in a window after day 55. Subsequent studies should investigate methylation changes at a variety of time points. It would also be beneficial to use another methylation technique to interrogate more genes highlighted due to their gene expression profile or alternatively a genome-wide methylation approach could be taken. A genome-wide approach might highlight methylation changes in genes upstream of the genes found in this thesis, revealing a cascade of changes in the affected pathways.

Chapter 6

microRNA Expression Study

6.1 Introduction

The work presented in this chapter aimed to investigate the role of miRNAs in gene expression regulation as a consequence of maternal undernutrition in the rat model (see Chapter 3). We investigated miRNA expression changes in the liver samples of 55 day old offspring of mothers that had been on a restricted diet or an ad-libitum diet during pregnancy. This chapter begins with a review of microRNA biogenesis and function. This background information is followed by a description of experimental design, data quality and a presentation and discussion of the results.

6.1.1 microRNA

MicroRNAs (miRNAs) are small (18-25 nucleotide) noncoding RNAs that regulate gene expression in plants and animals by targeting mRNAs for cleavage or translational repression. Recent studies have shown that miRNA regulation involves a complex system of positive and negative post-transcriptional control that is only just now being unravelled [184]. The miRNA target region (seed sequence) is located at the 5' end from bases 2 to 8. This region is important in identifying miRNAs. Complex computational algorithms have been designed to predict miRNAs by looking for these seed sequences. These algorithms predict that each miRNA can potentially bind to 200 targets and estimate that miRNAs control the expression of one-third

of human mRNAs. The official miRNA database (miRBase) currently lists over 600 human miRNAs [185]; however, computational prediction estimates there could be over 1,000 [186]. Overall, miRNAs control the expression of hundreds or possibly thousands of genes in a given cell type; therefore, they are likely to influence nearly every genetic pathway [184]. Given this, it is not surprising that microRNA gene regulation has been shown to play a role in many biological functions including human development, cellular differentiation, adaptation to environment, oncogenesis, host cell interactions with pathogens, and has also been linked to human diseases such as cancer, metabolic syndrome and aging [187, 184]. In addition, because of its tissue specificity miRNAs are a potential target as biomarkers for human disease.

miRNA Biogenesis

The complex process of miRNA biogenesis in the human cell is shown in figure 6.1. miRNA biogenesis involves two processing steps: the first takes place in the nucleus and the next in the cytoplasm, where the final mature, single stranded miRNA is produced. miRNA genes can be found in intergenic regions and in exons or introns of other genes. These miRNA genes are transcribed by both RNA polymerase II (most often) and III (occasionally) into primary transcripts (pri-miRNAs that are 100-1000s of nucleotides long) [188]. These pri-miRNAs contain Cap structures and poly(A)-tails and can encode sequences for multiple miRNA genes. Two RNA type-III endonucleases (Drosha and Dicer) coordinate the maturation of the miRNAs. Drosha initiates the processing of the pri-miRNA by cleaving it at the stem loop. It is then called pre-miRNA, and it is transported to the cytoplasm by Exportin-5/Ran-GTP complex. There it is processed by Dicer into the short (≈ 22 nucleotide) mature miRNA duplexes. These miRNA duplexes are then incorporated into the miRNA-induced silencing complex (miRISC, also known as the miRNP complex), where one strand is eliminated and one remains in the complex. There are occasions where both sides of the miRNA-miRNA duplex are retained and associate with the miRISC to target distinct subsets of mRNA for down-regulation [184]. miRNAs then affect gene expression by annealing to the 3'-UTR (untranslated region) of target

genes and cause mRNA degradation or repression of translation [189]. According to the genomic region where a miRNA resides, miRNAs can be grouped into several categories: intronic miRNAs in protein-coding genes; exonic miRNAs in non-coding genes; intronic miRNAs in non-coding genes.

The Complexity of miRNA Transcription

Little detail is known on exactly how miRNAs are transcribed and what promoter elements regulate their transcription. A few regulatory factors have been identified that bind directly to miRNA promoter elements and control their expression. Specifically, the oncogene C-myc and the tumour-suppressor gene p53 are capable of positive and negative control of miRNA transcription. In addition, subsets of miRNA genes reside within CpG islands and are under epigenetic control. Transcription of miRNAs miR-148a, miR-34b/c and let-7a-3 are dependent on their methylation status [184]. Recent studies estimate that 5-10% of mammalian miRNAs are epigenetically regulated. In a study by Brueckner *et al.* [190], the let-7a-3 loci was hypomethylated in a subset of patients with lung adenocarcinoma compared to patients with normal lung tissue, suggesting that a lack of epigenetic control could contribute to the progression of cancer. Another layer of complexity is added with recent work indicating the bidirectional transcription of miRNA genes. *Drosophila* mir-307 and the mammalian mir-338 genes are both transcribed in the sense and anti-sense directions.

miRNAs and Methylation

In plants, small interfering RNAs have been known to direct methylation as a mechanism of transcriptional control. Recent studies have provided evidence that RNA-directed DNA methylation in mammals may also occur [191]. siRNAs and miRNAs are considered to be closely related and many of the enzymes involved in the RNA interference pathway (RNAi) are also processing pathways of siRNAs and miRNAs[165]. Recent evidence also suggests that they affect histone modifications [192]. Given that siRNAs and miRNAs are considered closely related miRNAs could also be involved in controlling DNA methylation and histone modifications. A re-

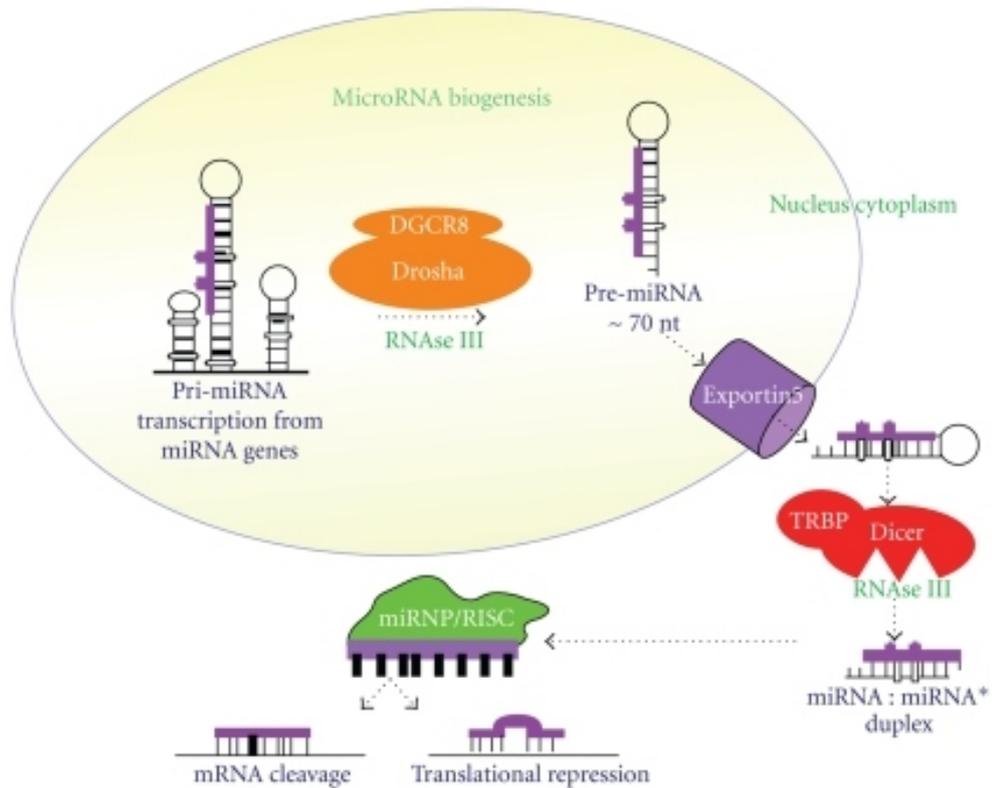


Figure 6.1: miRNA biogenesis. miRNAs are transcribed by polymerase II or polymerase III to primary transcripts (pri-miRNA). Two RNA type-III endonucleases, Drosha and DGCR8, process pri-miRNA by cutting it at the bottom of its stem loop. It is then exported, as pre-miRNA, to the cytoplasm by export factor-5 (exportin-5). The pre-miRNA is then processed by Dicer to generate an approximately 22 nucleotide long miRNA duplex. One strand of the now mature miRNA is then permanently incorporated into the RNA-induced silencing complex (RISC) and will bind to the 3'-UTR of a target mRNA. This can lead to silencing of the transcript via mRNA degradation or translational repression. This figure has been adapted from Poy *et al.* [189]

cent study showed that miR-165 and miR-166 were required for the methylation of the *PHABULOSA (PHB)* gene in *Arabidopsis*. These two miRNAs react with the newly processed *PHB* gene to change chromatin of the gene's template [193]. These results present the mechanism that would allow miRNAs to control gene expression in addition to the RNAi pathway. Similar findings in mammalian cells have yet to be shown. MiRNAs may also regulate chromatin structure by regulating key histone modifiers. A study in mice has shown that miR-140 can target histone deacetylase 4 [194]. There is still a lot to understand about miRNAs and their role in gene expression. The mounting evidence points to miRNAs as key players in the epigenetic control of gene expression and it is worth investigating whether this is a mechanism involved in the programming of the metabolic syndrome.

microRNAs in the Metabolic Syndrome

MicroRNAs have important regulatory roles in a variety of biological processes including several related to metabolic processes including adipocyte differentiation, metabolic integration, insulin resistance and appetite regulation. These are highlighted in figure 6.2 [185]. A role in energy metabolism was first shown in a drosophila study that implicated miR-14 in fat metabolism. Subsequently, miRNAs have been shown to have functional roles in all organs directly related to the metabolism of glucose (pancreatic islet, liver, skeletal muscle, adipose tissue, and brain). All of these tissues have a unique miRNA expression pattern; therefore, miRNA may contribute to specific tissue functions by regulating a set of unique target genes [189]. The liver miRNA expression profile is dominated by a single sequence, miR-122. This miRNA has been implicated in cholesterol and lipid metabolism, and in hepatitis C virus replication. This miRNA is expressed in human and rodent liver tissue with estimates of 50,000-80,000 copies per cell [195]. Recent findings have identified miRNAs as having a principal role in the production and secretion of insulin [185, 189]. miR-375 is specific to pancreatic islet cells and has been shown to play a key role in blood glucose homeostasis through its regulation of β -cell function. miR-124a, let-7b and miR-30d are also found in pancreatic islet cells. The two former are important

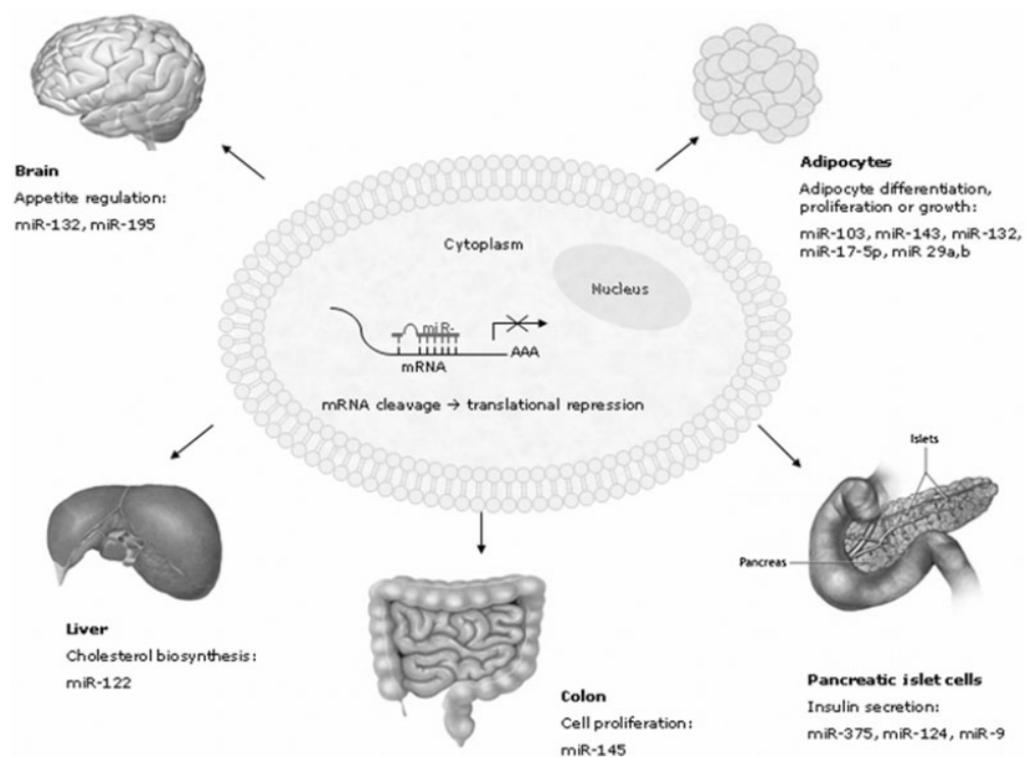


Figure 6.2: This figure summarises the known roles of particular miRNAs in specific metabolically related tissues (brain, liver, muscle, adipocyte, and pancreatic islet). This figure was adapted from Heneghan *et al.* [185].

ribo-regulators of blood glucose and the latter has been shown to influence insulin transcription [185]. In addition, a number of miRNAs (including miR-103, miR-143, and miR-132) have been found in adipose tissue that are involved in adipocyte differentiation, proliferation and growth, and insulin resistance [185]. Studies in *Drosophila* flies have revealed two miRNAs (miR-14 and miR-278) that regulate lipid metabolism in body fat [185] and another (miR-1) is essential for proper muscle function [195].

miRNAs as Biomarkers

miRNA offers the opportunity to locate unique biomarkers of metabolic health and disease because of its tissue specific expression and association with clinicopathologic variables. The recent discovery of their presence in circulation makes them a potentially non-invasive biomarker option. Reasons for their release into circulation need to be investigated thoroughly, in addition to their exact roles in gene expression. However, they would be an excellent tool for diagnosing metabolic syndrome [185].

6.1.2 Study Design

In this study, the same samples were used as in Chapter 3. These samples included eight biological replicates from the AD nutritional group and eight biological replicates from the UN nutritional group. The samples are rat samples, but due to the unavailability of rat miRNA expression arrays the samples were cross-hybridised to mouse arrays. The Illumina miRNA Profiling Assay is similar to the BeadArray in that there are 12 arrays on each chip, and 2 chips were used for this study. Each array incorporates the mouse microRNA panel which contains 419 sequences. As miRNAs are conserved between species it was reasonable to use the mouse array to interrogate rat RNA samples. The experimental design is the same as that illustrated in figure 3.14. Some samples were run a second time to try and achieve a better quality hybridisations.

6.2 Results

Total RNA was extracted from the rat livers using a protocol to ensure that small RNAs were not filtered out (see Chapter 2). Eight biological replicates for each of the two treatment groups (AD and UN) were hybridised to the Illumina MicroRNA Profiling Assay using the Mouse microRNA Panel. Amplifications and hybridisations were randomised and were performed by Kerry Cline at Cambridge Genomic Services in the Department of Pathology. Cross hybridisations of rat samples to mouse chips resulted in the detection of an average of 64% of the features across all arrays. This result was not as good as anticipated, but was reproducible when samples were re-hybridised. Samples 4.0, 101.0, 106.0, and 102.1 were re-amplified and hybridised a second time due to initially questionable quality control results (these repeated sample are called 4.0.2, 101.0.2, 106.0.2, and 102.1.0). The second hybridisations for each of these samples were remarkably consistent (as seen most clearly in the MA plots). Ultimately only 4.0 and 4.0.2 were removed as outliers for the final differential expression analysis. This resulted in a total of 18 arrays in the analysis.

6.2.1 Data Quality Control

The Illumina system incorporates a number of internal controls for estimating hybridisation quality. Figure 6.3 shows the details of the eight graphs for the internal controls across all 20 arrays. The results of the quality control assessment are not as good as recommended by Illumina, but this is most likely due to the cross-hybridisation of rat RNA to a mouse array. Figure 6.3A shows the data for the negative controls. The background and noise signals should be low. This data shows values less than 700 for the noise, which is slightly higher than ideal. Figure 6.3B shows the PAP controls comparing the housekeeping genes with all miRNA expression. The housekeeping genes should be much higher than the other genes as they are always expressed. Here the values are quite similar with large error bars. Figure 6.3C compares the mismatches versus the perfect matches. As expected the perfect matches are greater than mismatches. Figure 6.3D is a contamination control. The

positive represents the high signals and the negative represents background. Figure 6.3E shows the overall miRNA intensity. Figure 6.3F shows the high, medium and low annealing controls. Figure 6.3G shows the perfect match versus mismatch extension controls. Figure 6.3H shows the hybridisation controls, the expression levels should be high. Some of the controls (particularly the PAP controls) gave less than ideal results, but overall the quality is acceptable. In addition to the Illumina Internal Controls, an R library, `arrayQualityMetrics` [196], was used to assess the quality of the data before and after normalisation. This programme produces a quality control report and a number of figures for data visualisation (MA plots, density plots, boxplots, and a heatmap). The report includes information on which samples appear to be outliers. The different metrics the program utilises are meant to assess individual array quality (MA plot), homogeneity between arrays (density and boxplots), between array comparison (heatmap) and variance mean dependency (plot of standard deviation versus rank of the mean). The images from this analysis are shown in figures 6.4-6.11. A visual inspection of these images shows that sample 4.0 and 4.0.2 appear to be outliers. The array quality metrics program marked sample 4.0 as an outlier in two metrics (heatmap and MA plot).

6.2.2 Normalisation

Following the quality control assessment, this data was normalised in much the same way as the mRNA expression data in Chapter 3. The data was filtered for genes present on at least one array. Data was then transformed with a variance stabilising transformation and normalised with a quantile normalisation (as described in Chapter 2). This was all done using the `Lumi` library in R. The `arrayQualityMetrics` library in R was also used post-normalisation to assess the quality of the data. At this stage, the programme suggested the removal of one outlier (sample AD 4.0). This sample failed in all three of the programmes metrics. Sample 4.0.2 failed in 2 out of 3 metrics (MA plot and heatmap) and samples 106.1 and 102.0 failed in 1 out of 3 of the metrics (boxplot/density plot). Due to this assessment and the visual inspection of all

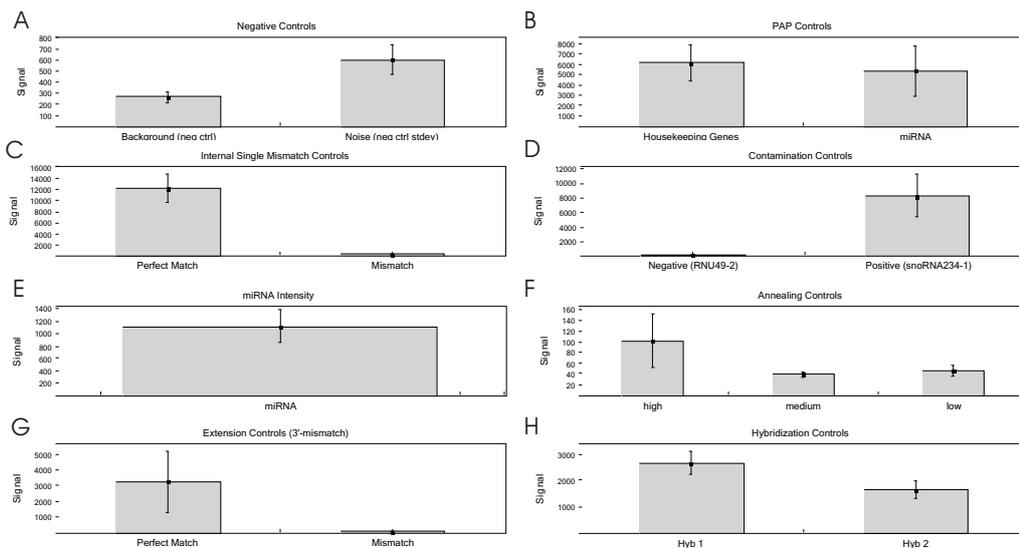


Figure 6.3: Illumina BeadStudio Internal Controls. This figure shows the internal quality control tools for the Illumina platform. (A) Negative Controls. Arrays should have low expression values for the negative controls (background and noise). (B) PAP Controls. This represents the housekeeping genes and miRNA and expression values should be positive. (C) Internal Single Mismatch Controls. Perfect matches should be greater than mismatches. (D) Contamination Controls. Positive = High Signal; Negative = Background (E) miRNA Intensity. The expression should be higher than background. (F) Annealing Controls. The Higher T_m MSO > Lower T_m MSO (G) Extension Controls. Perfect Match < Mismatch (H) Array Hybridization Controls. Expression levels should be high.

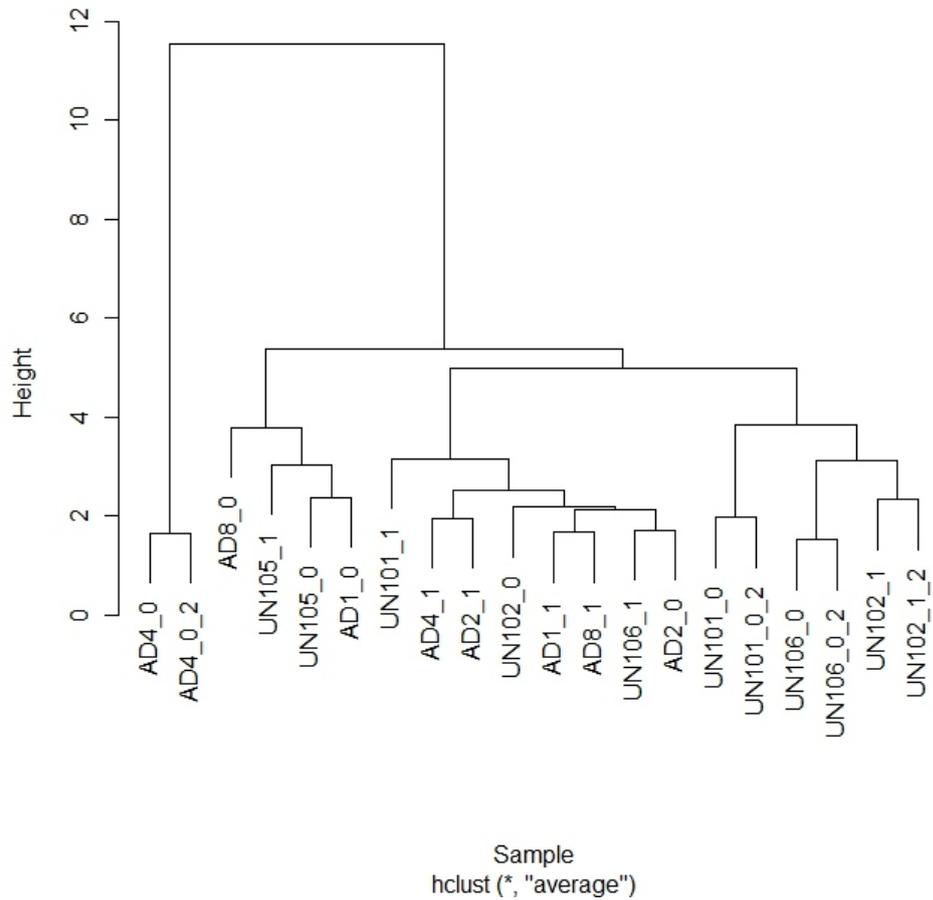


Figure 6.4: Hierarchical Clustering of Raw miRNA Data. This plot shows the hierarchical clustering of the raw miRNA data. The AD and UN samples have not separated clearly, which is likely to be due to the high level of variation between replicates. The two samples 4.0 and 4.0.2 which cluster far away from the rest were determined to be outliers and were removed from the final analysis.

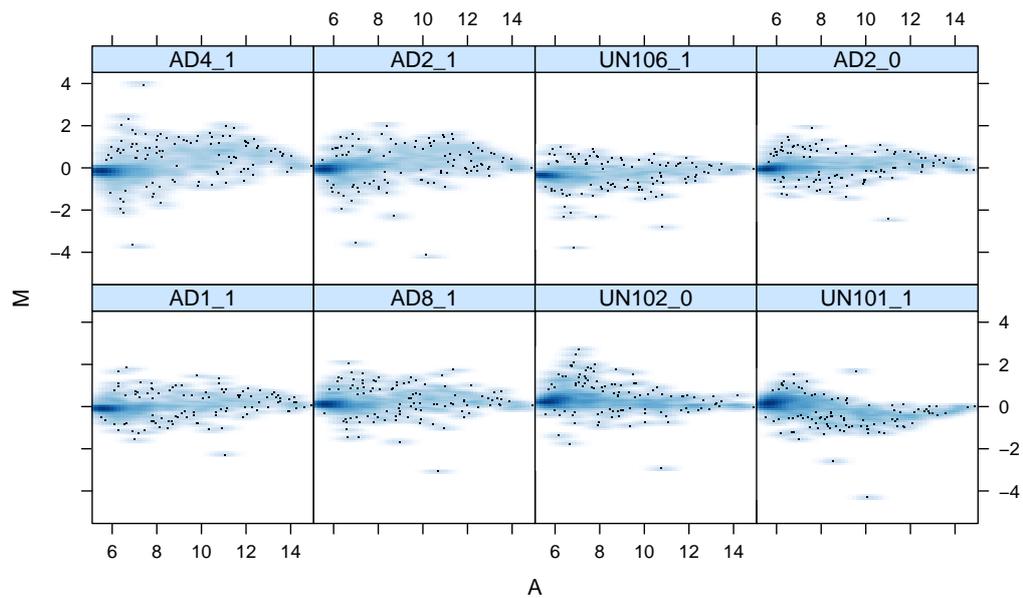


Figure 6.5: MA Plot of Raw miRNA Data. This plot shows MA plots for eight of the samples before normalisation. MA is defined as $M = \log_2 I_1 - \log_2 I_2$; $A = \frac{1}{2}(\log_2 I_1 + \log_2 I_2)$. The mass of the distribution is expected to be concentrated along the $M=0$ axis, and there should be no trend in the mean of M as a function of A . You can see that the majority of data is concentrated at the lower end of the A scale, which is expected as smaller expression values are more common.

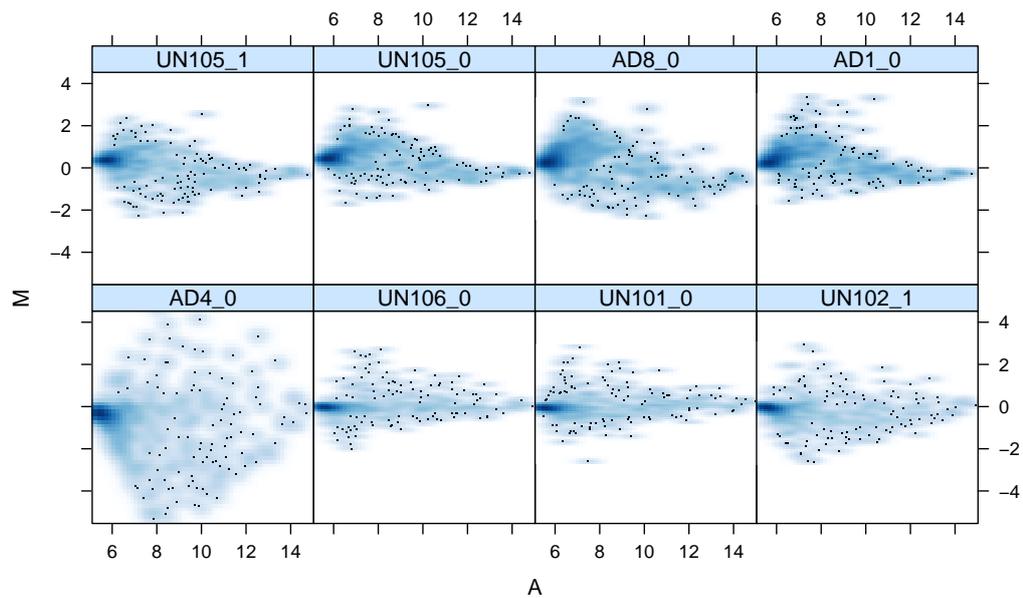


Figure 6.6: MA Plot of Raw miRNA Data. This plot shows MA plots for eight of the samples before normalisation. MA is defined as $M = \log_2 I_1 - \log_2 I_2$; $A = \frac{1}{2}(\log_2 I_1 + \log_2 I_2)$. The mass of the distribution is expected to be concentrated along the $M=0$ axis, and there should be no trend in the mean of M as a function of A . You can see that the majority of data is concentrated at the lower end of the A scale, which is expected as smaller expression values are more common.

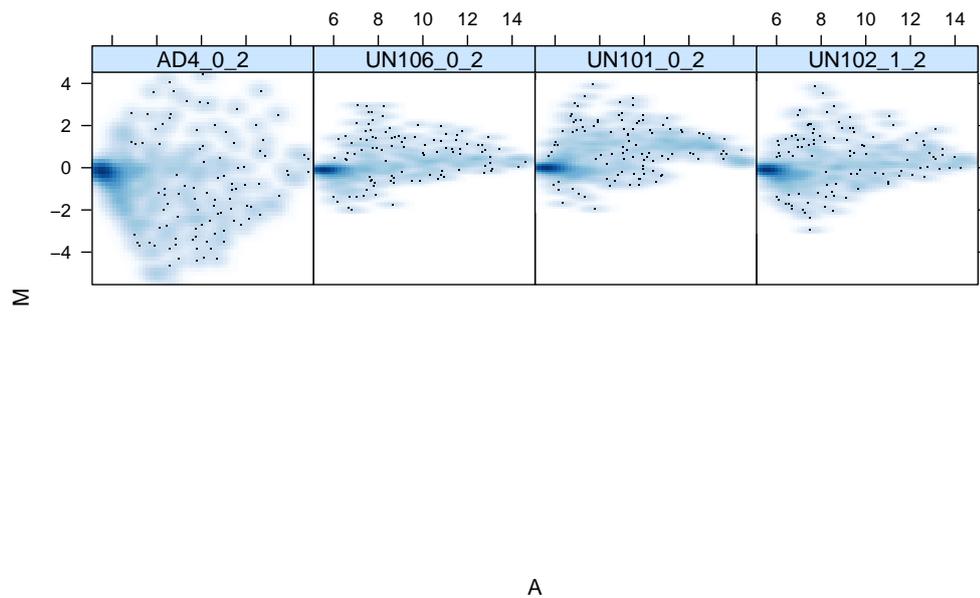


Figure 6.7: MA Plot of Raw Re-done miRNA Data. This plot shows MA plots for four samples before normalisation that were re-amplified and re-hybridised. MA is defined as $M = \log_2 I_1 - \log_2 I_2$; $A = \frac{1}{2}(\log_2 I_1 + \log_2 I_2)$. The mass of the distribution is expected to be concentrated along the $M=0$ axis, and there should be no trend in the mean of M as a function of A . You can see that the majority of data is concentrated at the lower end of the A scale, which is expected as smaller expression values are more common.

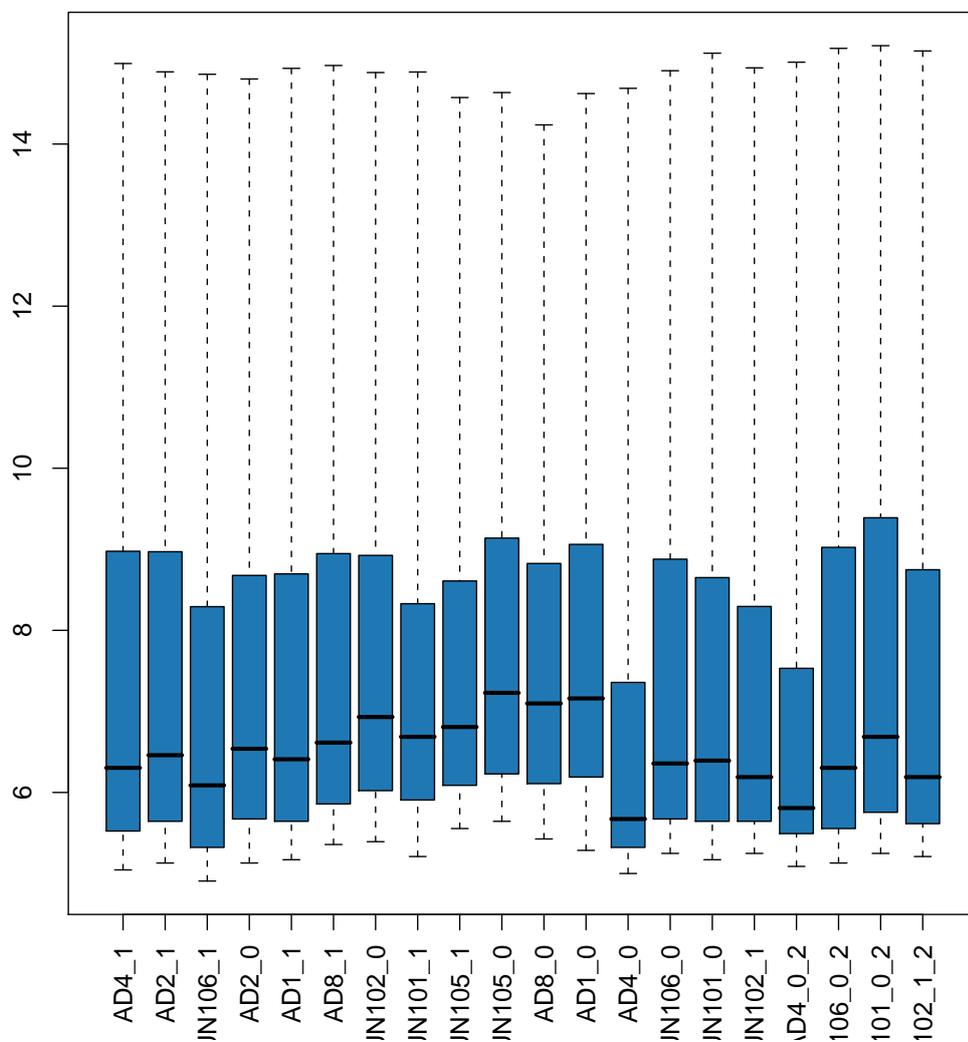


Figure 6.8: Boxplot of Raw miRNA Data. This figure presents boxplots of the $\log_2(\text{Intensities})$. Each box represents one array. Differences in box size and y position indicate increased variation between arrays.

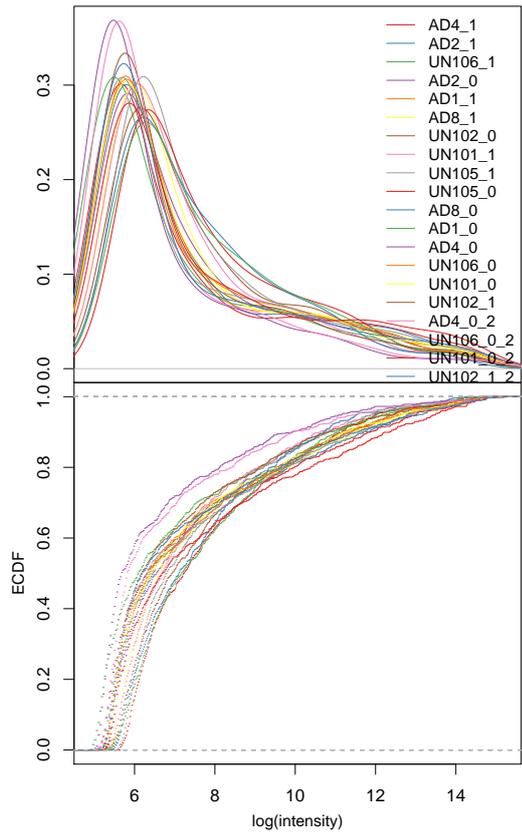


Figure 6.9: Density Plot of Raw miRNA Data. This graph presents density estimates (histograms) of the data. Each line represents a separate array. Lines that are very different to the others should be investigated as outliers.

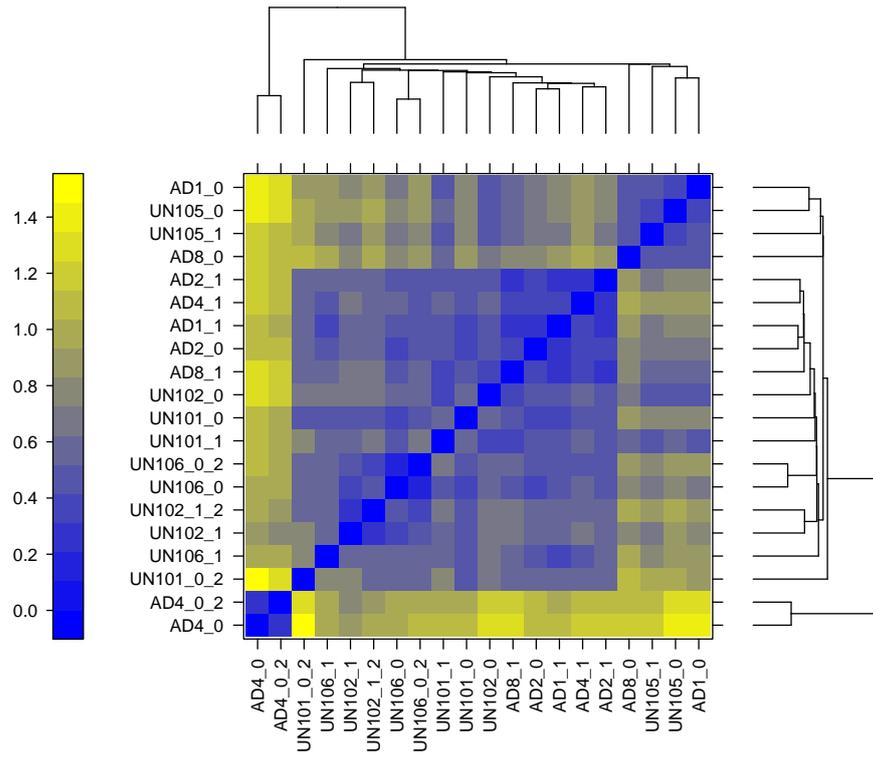


Figure 6.10: Heatmap of Raw miRNA Data. This heatmap shows the distance between arrays based on hierarchical clustering. This plot can help to detect outlier arrays.

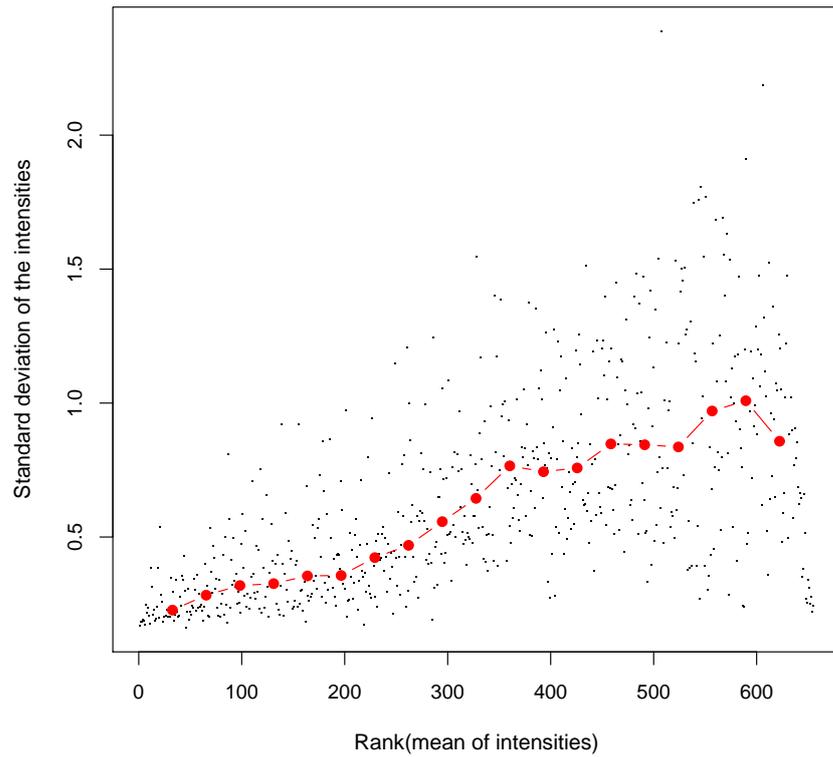


Figure 6.11: Mean SD of Raw miRNA Data. This plot shows the empirical standard deviation of the intensities for each feature of all the arrays on the y-axis versus the rank of the mean of intensities of the arrays on the x-axis. The red dots, connected by lines, show the running median of the standard deviation. After VSN normalisation, this should be approximately horizontal. There should be no substantial trend.

plots samples 4.0 and 4.0.2 were removed from the analysis. Data was re-normalised with data for the remaining 18 arrays before proceeding to the differential expression analysis. Figures 6.12-6.16 show the normalised data. The Limma library in R was used for pairwise comparison using a linear model to compute p-values and produce a gene list. An FDR adjustment was computed and miRNA genes with a p-value < 0.05 were selected as statistically significant for further investigation.

6.2.3 Differential Expression

This study aimed to highlight miRNA that might be involved in the development of the metabolic syndrome due to changes in nutrition during gestation. The results; however, revealed no miRNAs that were significantly differentially expressed across all eight biological replicates. Due to the cross hybridisation of rat RNA onto a mouse miRNA array, only 64 % of the features were detected. Closer inspection of the data revealed a significant amount of variation among these biological replicates. In an effort to look at the data in more detail, two approaches to data analysis were taken. The first involved a manual/candidate gene approach to look for the genes implicated in pathways involved in the metabolic syndrome and determine if any of them had potentially interesting expression profiles. The candidate miRNAs that are used in this analysis were mentioned in the introduction to this chapter as being metabolically relevant. Unfortunately, none of the miRNAs that were highlighted in the literature had an interesting expression profile in this study. Overall, they all have very small fold changes and very high p-values. As the candidate gene method did not pull out any interesting results, a second method for analysing the data involved filtering the data for those that were detected in at least 14 of the 18 samples. This reduced the list from miRNAs to 213. Further filtering involved taking the p-value before the FDR adjustment and selecting the samples with a p-value < 0.05 . This produced a filtered list of 44 miRNAs. Two of these miRNAs, mir-335-5p and 451, have a two-fold change of increased expression. All 44 miRNAs are listed in table 6.1.

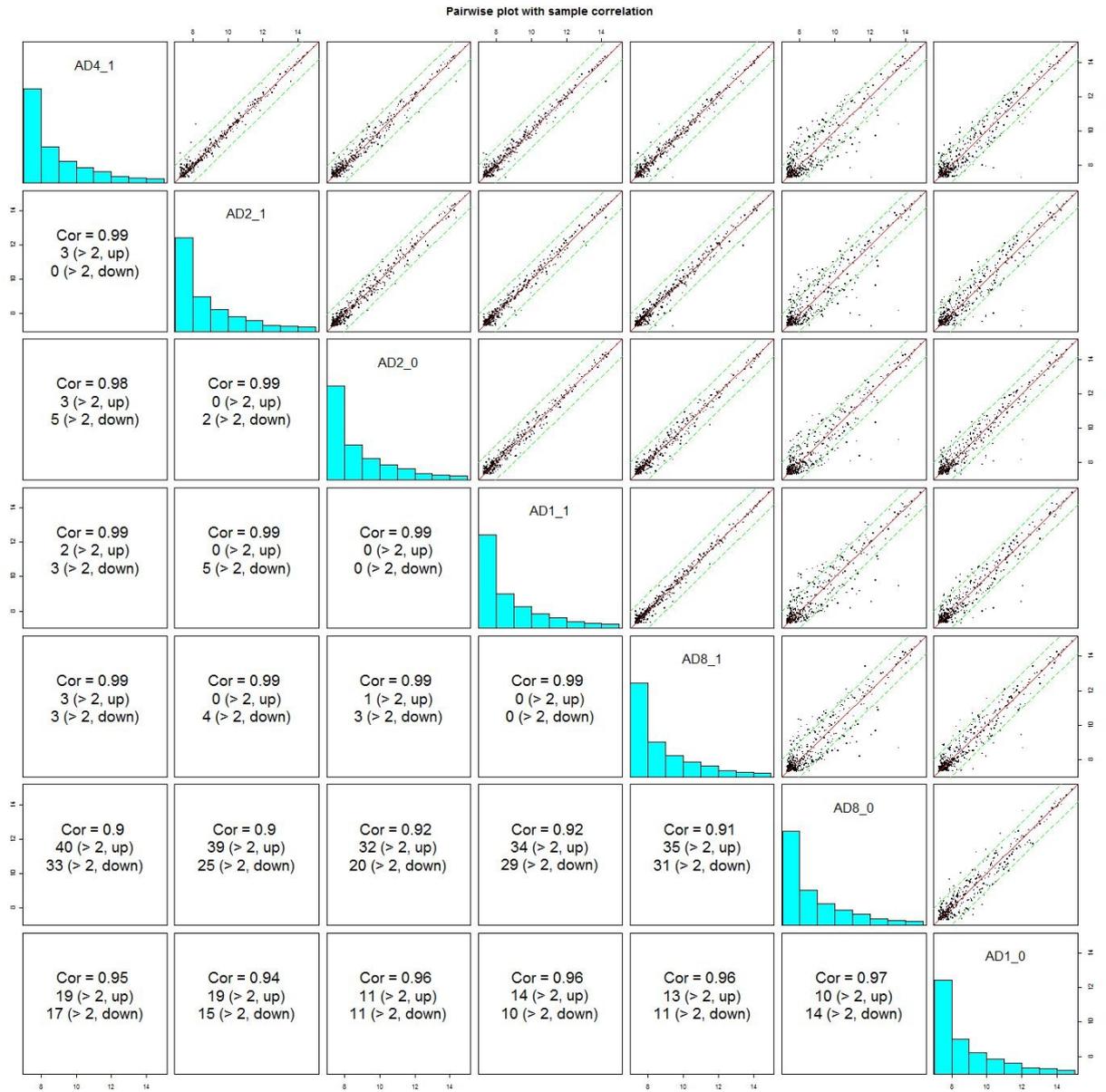


Figure 6.12: Pairwise Comparison of AD Liver miRNA Samples After Normalisation. This plot compares all seven AD liver biological replicates against each other after normalisation and removal of outliers. These replicates should be very similar and should be tightly distributed on the $x=y$ axis, as these samples are. The correlation for all comparisons ≥ 0.90 .

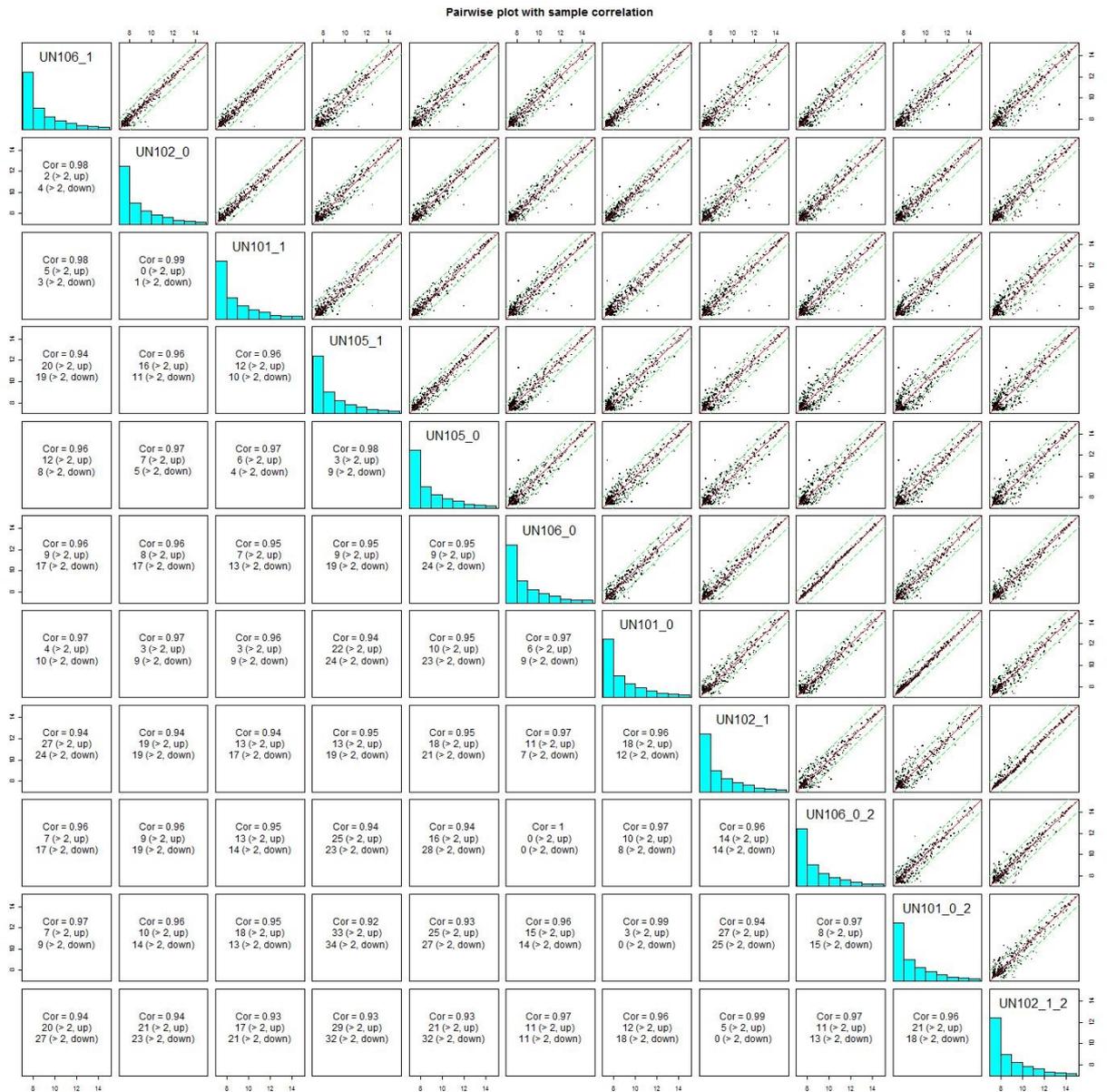


Figure 6.13: Pairwise Comparison of UN Liver miRNA Samples After Normalisation. This plot compares all ten UN liver biological replicates and technical replicates against each other after normalisation and removal of outliers. These replicates should be very similar and should be tightly distributed on the $x=y$ axis, as these samples are. The correlation for all comparisons ≥ 0.93 .

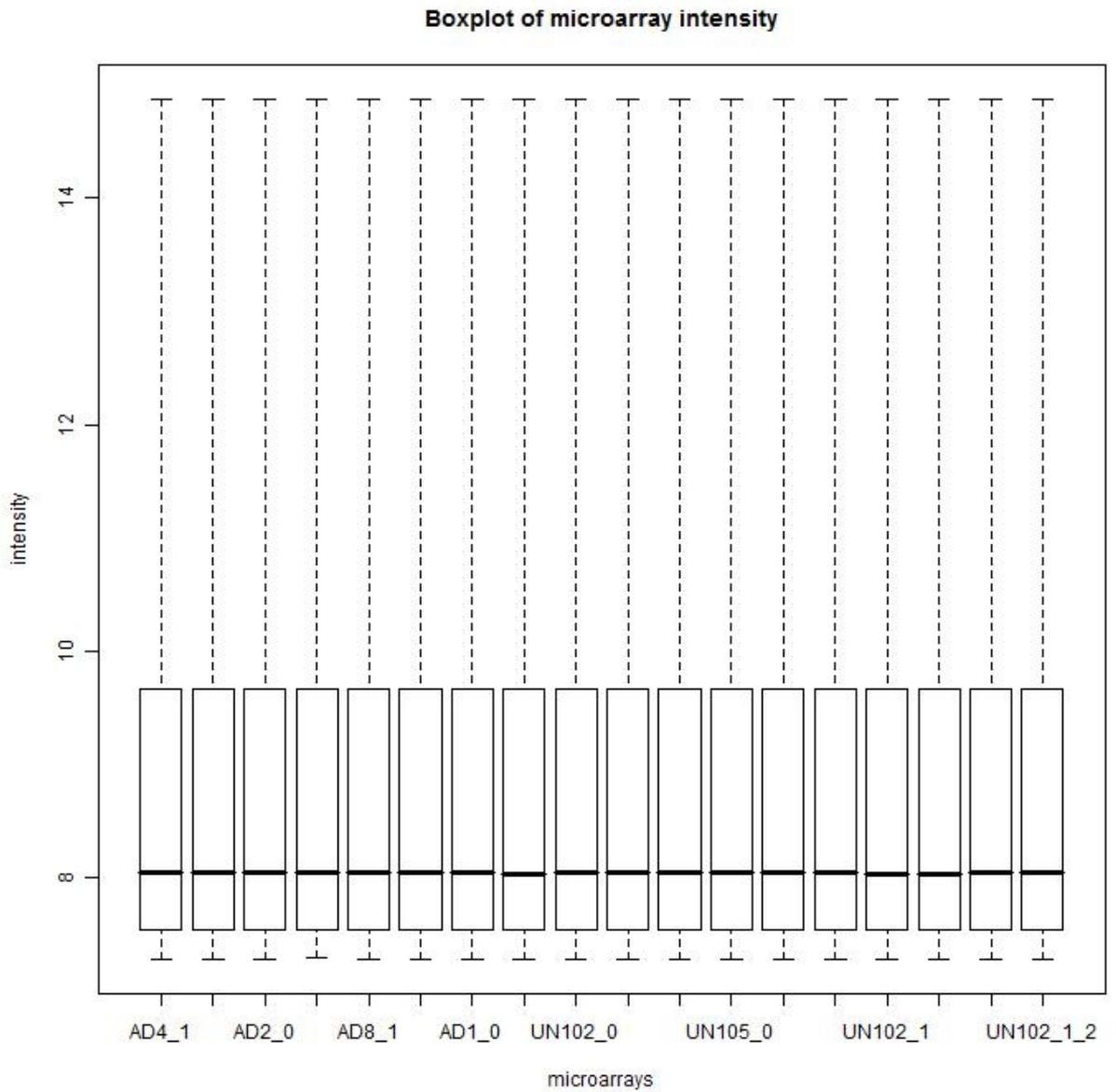


Figure 6.14: Boxplot After Quantile Normalisation. This is a boxplot of the data from each array. The plot compares the amplitude of all expression values on each array. This shows that the majority of the signal is at lower expression levels. This plot shows that there is no variability between the arrays. The normalisation has effectively adjusted the expression values to remove the variability. This analysis was done in Lumi.

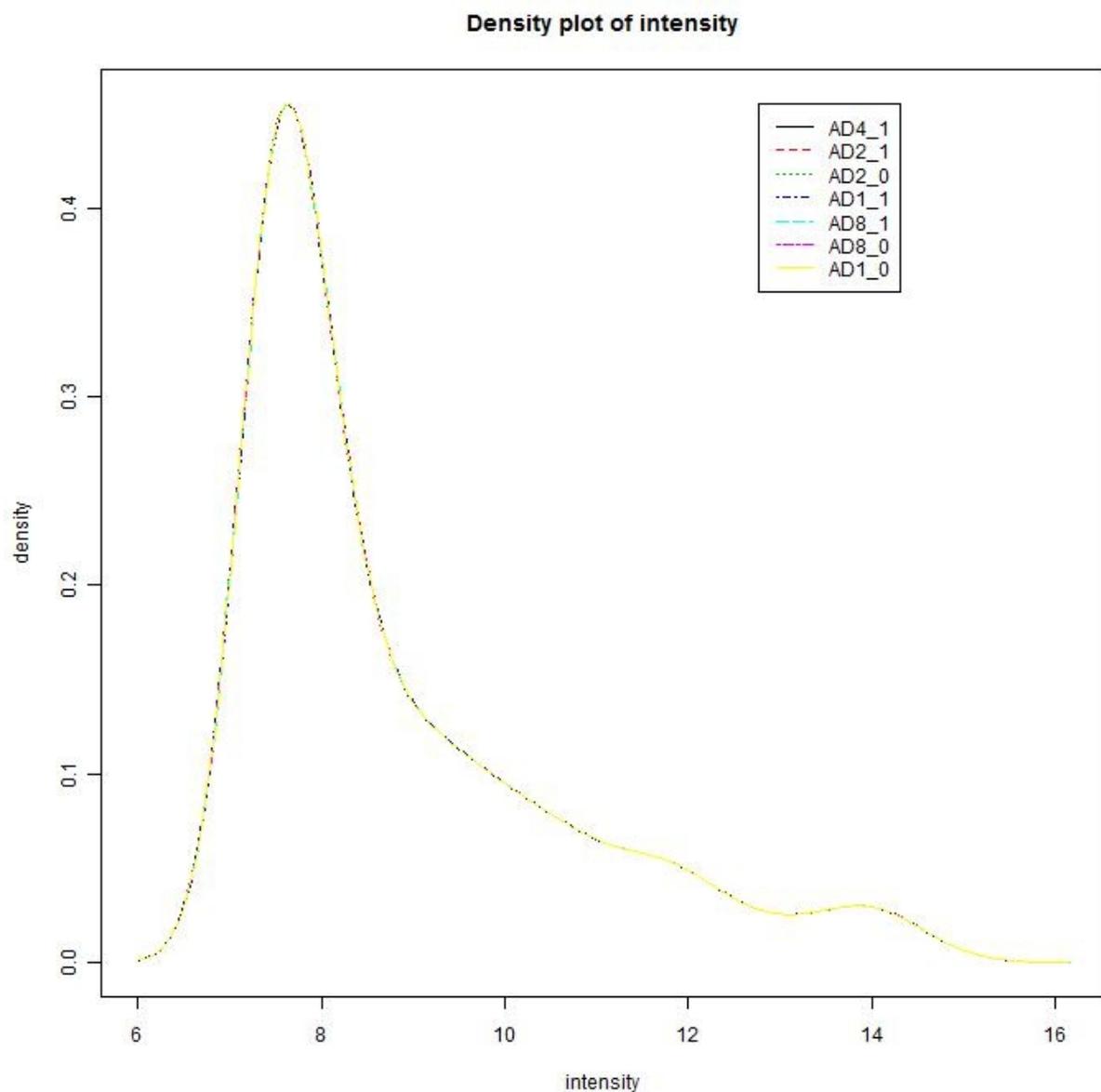


Figure 6.15: Density Plot Histogram of Density of Log2 Intensities for AD miRNA Arrays After Quantile Normalisation. This is a histogram of each array showing the density of intensities (log2). This plot shows that the majority of intensity values are low and that there is no variation between samples. The normalisation has effectively adjusted the expression values to remove the variability. This analysis was done in Lumi.

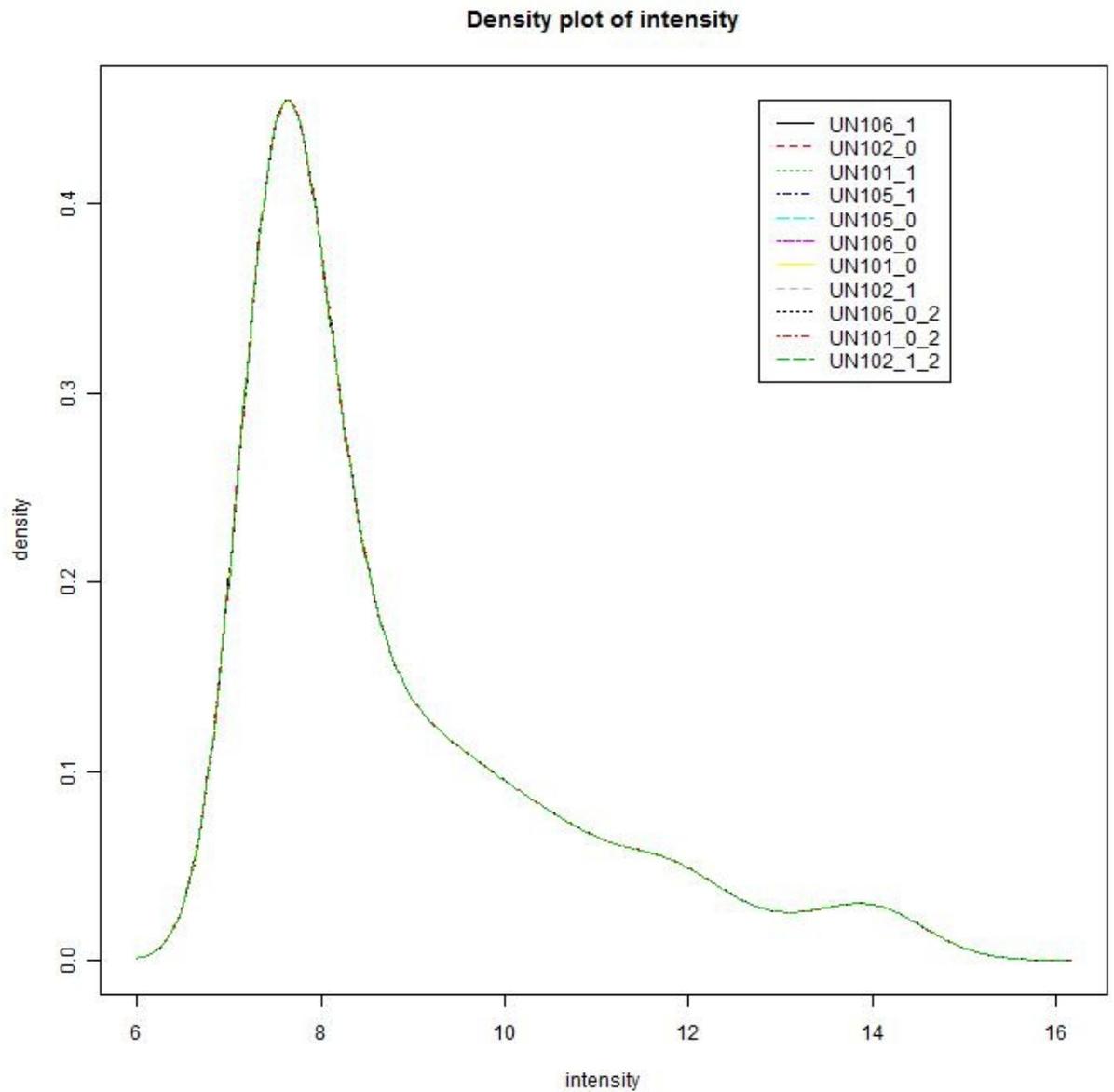


Figure 6.16: Density Plot Histogram of Density of Log2 Intensities for UN miRNA Arrays After Quantile Normalisation. This is a histogram of each array showing the density of intensities (log2). This plot shows that the majority of intensity values are low and that there is no variation between samples. The normalisation has effectively adjusted the expression values to remove the variability. This analysis was done in Lumi.

TargetID	logFC	P-Value	Detect
mmu-miR-335-5p	0.97	0.00	18
mmu-miR-451	0.91	0.00	18
mmu-miR-203	0.66	0.00	18
mmu-miR-499	0.65	0.01	18
mmu-miR-200a	0.56	0.04	18
mmu-miR-194	0.55	0.01	18
mmu-miR-126-5p	0.53	0.03	18
mmu-miR-29c	0.52	0.00	18
mmu-miR-16	0.51	0.01	18
mmu-miR-223	0.50	0.03	18
mmu-miR-100	0.42	0.02	18
mmu-miR-182	0.42	0.03	14
mmu-miR-20a	0.40	0.01	18
mmu-miR-130a	0.40	0.03	18
mmu-miR101b:9.1	0.40	0.01	18
mmu-miR-30b	0.40	0.03	18
mmu-miR-7a	0.38	0.03	18
mmu-miR-142-5p	0.37	0.02	18
mmu-miR-26a	0.35	0.01	18
mmu-miR-350	0.32	0.03	18
mmu-miR-546	0.31	0.04	18
mmu-miR-450a-5p	0.30	0.03	18
mmu-miR-146a	0.27	0.05	18
mmu-miR-540-3p	0.27	0.05	14
mmu-miR-22	0.26	0.02	18
mmu-miR-190b	0.21	0.00	15
mmu-miR-293	0.20	0.00	15
mmu-miR-26b	0.20	0.01	18
mmu-let-7g	0.17	0.02	18
mmu-let-7c	-0.14	0.04	18
mmu-miR-693-5p	-0.15	0.01	18
mmu-miR-469	-0.24	0.03	18
mmu-miR-23a	-0.30	0.02	18
mmu-miR-150	-0.32	0.01	18
mmu-miR-682	-0.33	0.04	18
mmu-let-7d*	-0.41	0.05	18

1=1

TargetID	logFC	P-Value	Detect
mmu-let-7e	-0.42	0.02	18
mmu-miR-669c	-0.42	0.02	18
mmu-miR-381	-0.44	0.02	18
mmu-miR-92a	-0.48	0.01	18
mmu-miR-18a	-0.54	0.01	18
mmu-miR-326	-0.57	0.03	18
mmu-miR-207	-0.64	0.02	18
solexa-103-3961	-0.77	0.01	18

Table 6.1: Table of miRNA Results. This table shows the 44 miRNAs that had a p-value < 0.05 and were detected in 14/18 arrays.

6.3 Discussion

This chapter aimed to highlight miRNAs that have been significantly differentially expressed in the livers of male day 55 rats that were exposed to maternal undernutrition during gestation. The study involved the cross hybridisation of rat samples to a mouse array that resulted in the hybridisation of 64% of the features. In this type of experiment, a multiple testing correction is recommended to eliminate a majority of potential false positives. However, in this study the multiple testing correction resulted in no significantly, differentially expressed genes. Therefore, out of interest, a less stringent analysis was done by removing the multiple testing correction, but still ensuring the feature was detected in at least 14/18 arrays. This analysis revealed 44 miRNAs with a p-value < 0.05 . This represents 10% of the total number of features on the array, which is less than would be expected by chance. The fact that only two of these showed a nearly 2-fold upregulation in the UN livers (miR-335-5p and miR-451) and both were found to be either directly or indirectly associated with the metabolic syndrome in very recent literature makes further investigation into their involvement worthwhile. In addition, miR-let-7e was shown by Ingenuity Pathways Analysis to target Mapk4, which was highlighted as being down regulated in Chapter 4 and demethylated in Chapter 5. These results are not consistent and follow up studies are important to elucidate exactly what is going on.

6.3.1 miR-335 Implicated in Lipid Metabolism

A literature search revealed that miR-335 has been associated with lipid metabolism in the livers of genetically obese mice [197]. Nakanishi *et al.* conducted a study using 63 day old male ob/ob (leptin deficient) mice, db/db (leptin-receptor deficient) mice, and KKAY (lethal agouti gene) mice. They used a miRNA microarray similar to the one used here but on a different platform (HOKKADO System) to compare db/db mice to wild-type (WT). They found miR-335 to be significantly increased in the db/db mice. They used real-time PCR to check the expression levels of miR-335 in multiple tissues in WT mice and found high expression in brain, lung, heart, and white adipose tissue (WAT). This figure is shown in figure 6.17. They then used real-time PCR to check the liver expression levels in ob/ob, db/db, and KKAY mice. All of these mice had significantly elevated body weight, liver weight, hepatic triglyceride concentration, and hepatic cholesterol concentration compared to the WT. The PCR revealed significant increase of liver miR-335 in the three obesity/diabetic mouse models. These results are shown in figure 6.18. The authors also investigated miR335 expression in white adipose tissue and miR-335 involvement in adipose differentiation and hypothesised that miR-335 may have a role in regulating lipid metabolism. Li *et al.* [198] also found miR-335 to be upregulated (4-fold) in ob/ob mice compared to WT on another miRNA microarray platform (CapitalBio Mammalian miRNA Array). This study used male and female \approx 90 day old mice. Another study using a human breast cancer cell line implicated miR-335 as a metastasis suppressor due to significant down regulation [199]. This publication revealed that miR-335 suppresses metastasis and migration through targeting of *Sox4*, the progenitor cell transcription factor and tenascin C, extracellular matrix component. How exactly miRNA-335 functions in the metabolic syndrome is yet to be determined.

miR-335 Gene Homology to Mest

A more thorough investigation into miR-335 revealed significant homology to two transcripts. One is involved in the inner ear and the other, Mest (also Peg1

or Zpeg1), has been linked to fat storage. Mest (mesoderm-specific transcript) is a paternally expressed imprinted transcript that is important during development. It is a member of the α/β fold hydrolase superfamily (along with lipases, acyltransferases, and esterases) [200] and is localised in the endoplasmic reticulum. A microarray study by Takahashi *et al.* [201] found that Mest is significantly expressed in the white-adipose tissue of obese mice. Additional experiments revealed that the expression was related to adipocyte size rather than diabetes or increased body weight. These results suggest that Mest is involved in the formation of adipose tissue and the determination of adipose cell size. Later work by Nikonova *et al.* [202], has led to the hypothesis that MEST controls developmental adipose tissue expansion at the onset of a positive energy balance by regulating adipocyte hypertrophy. The work showed that adipose tissue expansion always accompanies increased Mest expression, but increased Mest expression is not always associated with increased fat in the diet. This suggests that Mest is not triggered by the diet itself, but rather other cues that indicate the need for increased lipid storage. A more detailed investigation into the control of Mest in relation to miR-335 and the combined effects on the metabolic syndrome may offer more insight into the mechanisms of the metabolic syndrome

6.3.2 miR-451 Expression in Human Cancer Cells and Possible Association with Adipocyte Differentiation

Zhu *et al.* [203] conducted an miRNA expression study using multidrug resistant human cancer cell lines and found that miR-451 along with miR-27a were found to be upregulated in the human ovarian cancer cell line and the cervix carcinoma cell line compared to the respective parental cell lines. These miRNAs were shown to be involved in activating the expression of P-glycoprotein, the *MDR1* gene that has been implicated in cancer cell resistance to a broad range of chemotherapeutics. Zhu *et al.* hypothesised that miR-27a and miR-451 may be potential targets for a therapeutic strategy to control multidrug resistance in cancer cells. In a very recent publication by Sang Yun Kim *et al.* [204], miR-27a was shown to be a negative regulator of

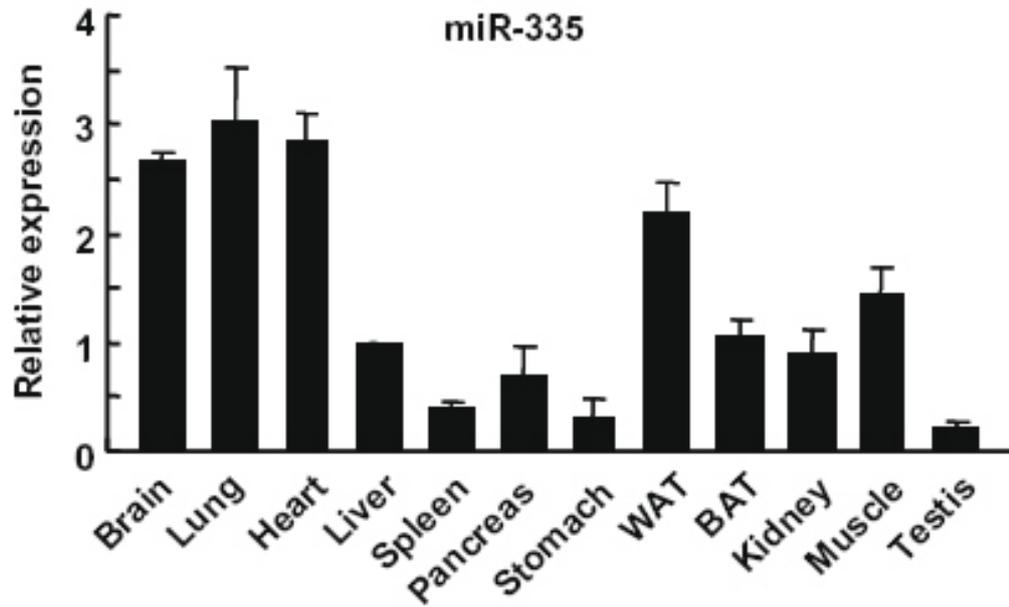


Figure 6.17: This figure was adapted from Nakanishi *et al.* [197] and shows the expression of miRNA-335 in a variety of tissues in WT mice. This shows that highest expression was observed in brain, lung, heart, and white adipose tissue (WAT).

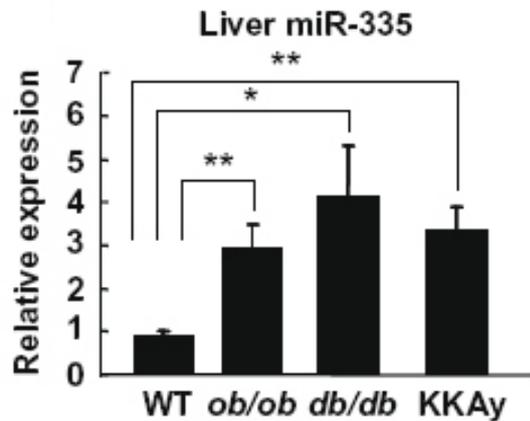


Figure 6.18: This figure was adapted from Nakanishi *et al.* [197] and shows the expression of miRNA in the livers of *ob/ob* (leptin deficient), *db/db* (leptin-receptor deficient), and *KKAY* (lethal agouti gene) mice compared to the WT. * = p-value <0.05; ** = p-value <0.01

adipocyte differentiation by suppressing PPAR γ expression. This study compared expression of miRNAs in white adipose tissue of obese mice and lean mice. miR-27a was downregulated in obese mice suggesting it might play a role in the suppression of adipocyte differentiation through the control of PPAR γ . As miR-27a was shown to have a similar expression profile as miR-451 in human cancer cell lines, it would be interesting to investigate whether it also plays a role in adipocyte differentiation in white adipose tissue.

6.4 Summary

This study aimed to investigate the role of miRNA in metabolic syndrome by comparing the expression profiles of livers of day 55 male rats exposed to maternal undernutrition during gestation to those whose mothers had been fed normally. Two of these miRNAs (miR-335 and miR-451) were up regulated by two-fold and were found to be either directly or indirectly associated with the metabolic syndrome in very recent literature. MiR-335 has been shown more than once, to be upregulated in the livers of obese/diabetic mice. By association with miR-27a, miR-451 might be involved in aspects of lipid metabolism in adipose tissue, but exactly what function the upregulation found in this study might be associated with in the liver is unknown. The next step in this investigation would be to use an alternative method to verify the expression changes suggested by this miRNA study. Further work into the miRNA expression in different tissues, at different time points would also be informative.

Chapter 7

General Discussion and Future Work

7.1 Conclusions

This thesis aimed to investigate the molecular mechanisms responsible for the developmental programming of the metabolic syndrome due to the exposure to maternal nutritional insults. This was the first time that a global gene expression technique was published in the DOHaD field [80]. Microarray technology was used to detect changes in gene expression in target tissues, between offspring of control and undernourished mothers, to obtain a broader picture of the cellular functions and genetic pathways that may be implicated in the metabolic syndrome. Multiple studies have utilised the candidate gene approach to look for gene expression changes and methylation differences due to changes in maternal nutrition. However, a global approach makes it possible to find novel genes that have not been implicated in the metabolic syndrome previously. Preliminary analyses of methylation status and miRNA expression were also performed in an effort to determine the mechanism of programming and the role of miRNA in gene expression regulation, as related to the metabolic syndrome.

7.1.1 Gene Expression Profiling Male 55 day old Rats Exposed to Maternal Undernutrition *in utero*

The global gene expression approach was successful in revealing that at a young age (day 55) and prior to the development of the metabolic syndrome phenotype, the livers of male rats exposed to maternal undernutrition during gestation were already showing significant changes in metabolism compared to the control. The study showed no differentially expressed genes in skeletal muscle and white adipose tissue, which indicated the liver was being affected before changes in fat and muscle were occurring. The hypothesis drawn from this was that the animals had been metabolically programmed to favour fat as an energy source. This had resulted in mitochondrial dysfunction, which initially affects hepatic function, but would potentially lead to the development of the metabolic syndrome later in life. The phenotypic data was compared to that of older male rats (day 110) that had developed the symptoms of the metabolic syndrome. In a future study, an investigation into the expression levels in the day 110 male rats would likely support this hypothesis and reveal more information into the pathways being affected.

7.1.2 Gene Expression Profiling Female 170 day old Rats Exposed to Maternal Undernutrition *in utero*, Postnatal Leptin and postweaning High Fat Diet

In this thesis, a later time point (170 days) in female rats was investigated. Unfortunately, because of the difference in sex, a direct comparison could not be made between the two studies; however, this second study involved a postnatal leptin treatment and a postweaning high-fat diet. Phenotypic data from this study had indicated that the postnatal leptin treatment resulted in a reversal of the metabolic syndrome phenotype. Here, we aimed to find the molecular mechanism responsible for the reversal. Molecular evidence for a reversal of the developmental programming was not revealed; however, it was clear that the interaction between treatments did result

in significantly, differentially expressed genes. Intriguingly, the genes involved were related to immune function, regulation of the circadian rhythm, and metabolism. The interpretation was complex due to the interaction of the three treatments. A follow up study would be useful to look into the highlighted genes in more detail, and obtain a clearer picture of how the diet mismatch and the leptin treatment are affecting the metabolism of these 170 day old female rats. In addition, a simplified study design to observe immediate effects of leptin would help to separate the effects of the post-natal diet.

7.1.3 Methylation Assay on Male 55 day old Rats Exposed to Maternal Undernutrition *in utero*

In an effort to investigate the mechanism of fetal programming, a preliminary study aimed to determine whether epigenetic mechanisms (specifically DNA methylation) might be responsible for the observed transcriptional changes. This study was conducted by a part II student during the preliminary stages of the analysis of the Illumina BeadArray data set. For this reason, many of the genes that were analysed did not show up in the final analysis. Genes that were identified as significantly, differentially expressed in the livers of 55 day old rats exposed to maternal undernourishment during gestation were analysed to locate the promoter regions and to find any nearby CpG islands. Primers were designed and methylation specific PCR was used to look for methylation differences. Of the 87 genes analysed, only 26 were in the final list of differentially expressed genes presented in Chapter 3. Of these 26, only one showed a change in methylation. This gene was *Zfand2a*. Another 12 included in the methylation analysis were genes that were found to be differentially expressed on the MEEBO microarray platform. Of these 12, one showed a methylation change in the McrBC assay. That gene was *Mapk4*. At least one potential transcription factor binding site was found in the CpG islands of both of these genes. Further investigation verifying the methylation differences and its effect on downstream gene expression changes, would clarify the involvement in the metabolic syndrome. In

addition, it would be of use to conduct a global methylation study on the day 55 or day 170 samples or a similar study, to highlight methylation changes in other genes that may be affecting the genes highlighted in this thesis, due to a cascade of events.

7.1.4 MiRNA Expression Profiling of Male 55 day old Rats Exposed to Maternal Undernutrition *in utero*

The liver samples from this study were also used on a miRNA expression array to highlight any miRNAs that might be involved in the differential regulation of gene expression in the day 55 male offspring of undernourished mothers. Initially, this study revealed no significant results. A closer look at the individual data points revealed a large amount of variation between individuals that was over-shadowing any potential differences between treatment groups. By using a less stringent statistical test (removing the multiple testing correction) and also, selecting for probes that were detected on 14/18 arrays, 44 miRNAs were revealed as being differentially expressed (p-value < 0.05). Two of these had a nearly 2-fold upregulation in the UN livers. Both of these miRNAs were found to be either directly or indirectly associated with the metabolic syndrome in very recent literature. MiR-335 has been shown, more than once to be upregulated in the livers of obese/diabetic mice and has significant homology to a transcript that has been shown to be involved in lipid storage (Mest). By association with miR-27a, miR-451 might be involved in aspects of lipid metabolism in adipose tissue. However, the function of the miR-451 upregulation in the liver as found in this study is unknown. This data suggests, it would be worthwhile to follow up on miR-335 and miR-451 to verify the expression changes found in this study. In addition, if the variation between animals could be controlled, a global miRNA study on a rat chip may reveal more candidate miRNAs that are differentially expressed.

7.2 Future Directions

These findings provide a number of interesting genes and pathways for further studies and also highlight the need to conduct a thorough study in multiple tissues at different time-points to pinpoint the window of developmental plasticity. Overall, more investigation into the molecular mechanisms of the metabolic syndrome is necessary to understand the pathways involved and to potentially reveal new treatments, biomarkers, or preventive health advice. Studies designed to effectively highlight important developmental windows and studies in a variety of tissues (i.e. pancreas and hypothalamus), in addition to those studied here, will also help to clarify our understanding of the pathways involved.

Appendix A

Primer Lists

Sequence Name	Sequence 5'to3'
Tpi1_F_Exp	GCCCAGGAAGTACACGAGAA
Tpi1_R_Exp	CAGGCTACGCAGGAAGGTAG
Acad9_F_Exp	GGCCTCTCCAACACCATGTA
Acad9_R_Exp	GTGTTCCCCAGACGACAGTT
Slc37a4_F_Exp	GCTCCAGCAAAAATGAAAGC
Slc37a4_R_Exp	TGCAGCTAAACTACCCAGCA
GCK_F_Exp	AGTATGACCGGATGGTGGAT
GCK_R_Exp	CCGTGGAACAGAAGGTTCTC
HPRT_F_Exp	CTCATGGACTGATTATGGACAGGAC
HPRT_R_Exp	GCAGGTCAGCAAAGAACTTATAGCC
BETA ACTIN_F_Exp	GATTACTGCCCTGGCTCCTA
BETA ACTIN_R_Exp	TCATCGTACTCCTGCTTGCT
CA3_F_Exp	TTCTGAAGATAGGACGGGAG
CA3_R_Exp	ATGGGCTCTTTCAGTAGCAG
ORM1_F_Exp	TTTAACCTGACAGATGAGAACC
ORM1_R_Exp	GCACTTATCCTTTGTCCAGTC
PER2_exp_F	GATCCTGTACATCTCCAACCA
PER2_exp_R	CCTGAGTGAAAGAATCTAAGCC

Table A.1: Sequences of primers used for expression assays as described in chapter 3.

Sequence Name	Sequence 5'to3'
IGFBP2_exp_F	CCTCAAACAGTGCAAGATGTC
IGFBP2_exp_R	TGCTCGTTGTAGAAGAGATGG
IGFBP4_exp_F	CAGACCTCTGACAAGGATGAG
IGFBP4_exp_R	CCACAACCTTCATCTTGCTC
RT1-Ba_exp_F	CTACCAACAAGGTTCCCTGAG
RT1-Ba_exp_R	TCGTTAGAAGGGATGAAGGTG
RT1-M6.2_exp_F	TTCTACCCTGCTGACATCAC
RT1-M6.2_exp_R	GTTGCTGGTTTCCATCTCAG
Ng22_exp_F	CACACTCCGTTATCACACTG
Ng22_exp_R	TGTAGGCATTACGGTTGAGG
RT1-149_exp_F	CTCCTTCATCCACTGATTCCA
RT1-149_exp_R	TGTCTTCATGCTTCACAATCTG

Table A.2: Sequences of primers used for expression assays as described in chapter 4.

Sequence Name	Sequence 5'to3'
GR_F_EXP	GGAGAATTATGACCACACTCAAC
GR_R_EXP	GCAGTAGGTAAGGAGATTCTCAA
PPARa_F_EXP	CTGGTCAAGCTCAGGACACA
PPARa_R_EXP	AAACGGATTGCATTGTGTGA
PPARg_F_EXP	TGCAGATTACAATGATGAC
PPARg_R_EXP	TCGATATCACTGGAGATC
AOX_F_EXP	CCAATCACGCAATAGTTCTGG
AOX_R_EXP	CGCTGTATCGTATGGCGAT
11B-HSD2_F_EXP	TGGCCACTGTGTTGGATTT
11B-HSD2_R_EXP	ATCGGCCACTACCATGTTG
PPARa_NEW_F_EXP	CGGGTCATACTCGCAGGAAAG
PPARa_NEW_R_EXP	TGGCAGCAGTGGGAAGAATCG
PEPCK_F_EXP	AGCTGCATAATGGTCTGG
PEPCK_R_EXP	GAACCTGGAGTTGAATGC
Cyclophilin_F_EXP	TTCGGTCGCGTCTGCTTCGA
Cyclophilin_R_EXP	GCCAGGACCTGTATGCTTCA

Table A.3: Sequences of primers used for expression assays as found in Lillycrop *et al.* [82]. These assays are described in chapters 3 and 4.

Sequence Name	Sequence 5'to3'
Igfbp2_f.m	AGGGAGTGGTCTCCAAAAGG
Igfbp2_r.m	CCGCTGAGCTACGAGTTTCT
Serpib6_f.m	GACTGAGTCGCCTTGTGGTC
Serpib6_r.m	CAGGCACSCCTGGACAAGAAA
LOC363151_f.m	CAGCCAACATGCAAAGTGGAG
LOC363151_r.m	CTCCTCCTCCAGCCTGACT
Gstp2_f.m	TGTTTTGTCCCCCAGAACTC
Gstp2_r.m	CCAAAAATGAACCCAGCACT
RGD1310991_f.m	TCAGCCTAGGCCAGAGATGT
RGD1310991_r.m	CAGCACTCCTGACCCTGTC
Adamts1_isl1_f.m	CACCGTCTGGAGGGTGAA
Adamts1_isl1_r.m	TTTCGGAGCTCTCAGTCTGC
Adamts1_isl2_f.m	CTGACCCCAAAGGGACTTCT
Adamts1_isl2_r.m	GGGCTCCAATGTGGCTATAA
Atf4_isl1_f.m	AACCTCTGGTGGCTCTTCC
Atf4_isl1_r.m	CAGGCTCTGCTGCCTCTAAT
Atf4_isl2_f.m	CATTTCTGCTTGCTCTGTGG
Atf4_isl2_r.m	TTGCACAAGATGGAGGCTTA
Galnt11_f.m	GACAACTCGGGCCTCACA
Galnt11_r.m	CTGGACCAACTACCAAACCTTC
Crlz1_isl1_f.m	TTCTCTGACCCCAGGACAAG
Crlz1_isl1_r.m	GGCTGAGACTAGCCTGGACA
Crlz1_isl2_f.m	CCGAGTGGAAGTTTGGTTGT
Crlz1_isl2_r.m	TGTTGGTGCTTGACCAATGT
Amd1_isl1_f.m	AAGCAGGCAAACCTCCT
Amd1_isl1_r.m	CTTCGTACCATCCCAAGGTG
Amd1_isl2_f.m	ATTTTCCCGGCTATTTTCGT
Amd1_isl2_r.m	CTTCATCTGCAACCAAGCTG
Rbp1_f.m	TCTTTCTAGGCTGGGGAGGT
Rbp1_r.m	GCCCTCTAGTTGGCAGCA

Table A.4: Sequences of primers used for methylation assays described in chapter 5.

Sequence Name	Sequence 5'to3'
Tnfsf13_f_m	GACCACGACCAGCAGGAG
Tnfsf13_r_m	TGTGGGAGGGCTCAACATAC
Adra1b_f_m	GACCAAGGCACCTCAGCTAC
Adra1b_r_m	GGGTGTATGAACCCGGATG
Npm1_isl1_f_m	GGTAGGCCTCGCCTCACT
Npm1_isl1_r_m	ACTGACTGGAAGGGAGAGCTT
Npm1_isl2_f_m	TCCCTTCCAGTCAGTTACCG
Npm1_isl2_r_m	GCCTTCGAGCAAGTCAAATC
Cidea_f_m	AGCCCCCAGAAGAGAAAAAC
Cidea_r_m	AGGGGGACTATAGCCGTCTG
RGD1308082_f_m	GACCGACCGCAAATAGAAA
RGD1308082_r_m	CTCACCCACTTCAGGGTCTC
Dhtkd1_f_m	CCATTCTTCCACGCGTCT
Dhtkd1_r_m	GAAGAGGCCCTTTGGTC
Arhgef19_isl1_f_m	CATCTCCATTTAGGGCTGA
Arhgef19_isl1_r_m	GTGCCTAGTGGGTCTTCCAC
Arhgef19_isl2_f_m	CATCCGAGACCCCTAAGTCA
Arhgef19_isl2_r_m	CCCGTGGAAAACCTGGAGAC
Rpl10l_f_m	TTTCTGCCAGCAAAGTCCTC
Rpl10l_r_m	GTAGAGTGACCCGGAGGTTG
Usf2_isl1_f_m	CACCTTCCTGCAGCTCAAC
Usf2_isl1_r_m	CCCATGGACATGCTGGAC
Usf2_isl2_f_m	GCGCTGATTTTGGGACAA
Usf2_isl2_r_m	CCGAGGATCTGGGAAACAG
Rabepk_f_m	TCCGCCACAGTCTTTTCTTC
Rabepk_r_m	GGAGACCGAGCGAATTAGAA
Psip1_f_m	TCTTGCGGAAGATGAGGTCT
Psip1_r_m	AGCATCCCTGCCTCAGGT
Mapk4_isl1_f_m	AGCGGGCAGCTAGAGAAAA
Mapk4_isl1_r_m	ACGCAGGAGGTGCAGAAG

Table A.5: Sequences of primers used for methylation assays described in chapter 5 continued.

Sequence Name	Sequence 5'to3'
Mapk4.isl2.f.m	GGACAATATAAGCCACGCTGA
Mapk4.isl2.r.m	GACTGCATCGCCAGTGTCTA
XM_978865.f.m	AGCTGTGCTTGTAGGGGAAG
XM_978865.r.m	CAAGCACTCAGACCCCAGAT
Acad9_F_Meth	ACATTACATACAGCAATGCTTGG
Acad9_R_Meth	AAAGGTGGGAGCAGAGAACA
Tpi1_F_Meth	TGTCCCTAGGCCACCATCTA
Tpi1_R_Meth	TTGAAAAGGTGCCCTCAAAC
Slc37a4_F_Meth	GCAGGGAATATTTCTTTAGTTTTCTG
Slc37a4_R_Meth	GTAAGGCACCATGGCAAAAG

Table A.6: Sequences of primers used for methylation assays as described in chapter 5 continued.

Sequence Name	Sequence 5'to3'
GR_F_METH	CGTCTTGTTCACCCACT
GR_R_METH	CCTTGCAGTTGCCGACAG
PPARa_F_METH	TGTGTCTCGTTCTGAACCG
PPARa_R_METH	TCCACCCACCTCACTGTC
PPARg_F_METH	CGACTGTGAGGAGCAAGG
PPARg_R_METH	CCCAGGTCTCTTCTTCAG
Hexokinase_F_METH	GAACCTGGACAGGTGTAGGAGAATC
Hexokinase_R_METH	AGCACTAGTGTGTCCCCTGTCC
PPARg2_F_METH	GTCTCTGCTCTGGTAATTC
PPARg2_R_METH	AAGGCTTGTGGTCATTGAG

Table A.7: Sequences of primers designed for methylation assays as found in Lillycrop *et al.* [82].

Sequence Name	Sequence 5'to3'
IGFBP2_F_BISEQ	TTGGTTT TAGAAGGATTGAATTTTTT
IGFBP2_R_BISEQ	ACTTTACTCCCTTCAACCTAAAATC
ISL1_F_BISEQ	TGTTAGTTTTAGGAGATGTATTTTTTT
ISL1_R_BISEQ	AATTAAACAATTAATAACAACCCACAA
EHMT2_F_BISEQ	GGGGGTAAAAATGTTTAAAAGGTT
EHMT2_R_BISEQ	ACAAACA ACTAAAAACCCAAAAC
MAPK4_F_BISEQ	GTGGAGTTTAGTTTTTTTGGT
MAPK4_R_BISEQ	CTTTAAAAAACCTCCAACCCTTAC
ATF4_F_BISEQ	TGGATTGATAGGATTGGATTAGGTA
ATF4_R_BISEQ	CTACCACAAAACAAACAAAATAAC
TNFSF13_F_BISEQ	ATGTGTTTAGAAAGGGGTATGTTG
TNFSF13_R_BISEQ	TTCACTAACCTCAAAAACCTTAAT
GCK1_F_BISEQ	GAATTTTATAGAAGAGTTTAGAATGTTTTGG
GCK1_R_BISEQ	CACACCTTATAATATCCATAACCATCTC
GCK2_F_BISEQ	GGGTGTTAGGGTAGTTAGAGGATTTG
GCK2_R_BISEQ	CCTAACTCCTAAAACCACCTATTAC
Crat_biseq_F	TTTAGTAGGGAAATGAGTGTTGGTTT
Crat_biseq_R	ATATCCAAATCTAACCAATATCCTTAAACA
Tpi_biseq_F	TTAGTTATGTTTGTGAAATATTTGGGTT
Tpi_biseq_R	AAACTCAAACATCCCACCTTAATAA

Table A.8: Sequences of primers designed for bisulphite sequencing assays as described in chapter 5.

Appendix B

Publications and Presentations

Publications

- T.J. MORRIS, M. VICKERS, P. GLUCKMAN, S. GILMOUR and N. AFFARA (2009) *Transcriptional Profiling of Rats Subjected to Gestational Undernourishment: Implications for Developmental Variations in Metabolic Traits*. PLoS ONE 4(9):e7271. doi:10.1371/journal.pone.0007271

Presentations

- T. J. MORRIS, N. AFFARA, S. GILMOUR, M. VICKERS, and P. GLUCKMAN. *Microarray Analysis to Investigate the Mechanisms of Metabolic Syndrome*. Poster presentation. 17th Annual International Conference on Intelligent Systems from Molecular Biology & 8th European Conference on Computational Biology June 2009: Stockholm, Sweden.
- T. J. MORRIS, N. AFFARA, S. GILMOUR, M. VICKERS, and P. GLUCKMAN. *A Global Approach to Epigenetic Change in an Undernourished Rat Model*. Poster presentation. 5th International Conference on Developmental Origins of Health and Disease (Dohad) November 2007: Perth, Western Australia.
- T. J. MORRIS. *A Gene Expression Study of Fetal Programming in Undernourished Wistar Rats*. Poster presentation. Sanger PhD Symposium April 2007: Cambridge, UK.

Appendix C

Supplementary Data

Supplementary Data including gene lists, tables, and transcription factor analysis results as cited throughout this thesis are included on the accompanying DVD.

Bibliography

- [1] G. M. REAVEN and Y. D. CHEN. *Role of insulin in regulation of lipoprotein metabolism in diabetes*. *Diabetes Metab Rev* **4**(7), 639–52 (—1988—).
- [2] L. POSTON. *Influences of maternal nutritional status on vascular function in the offspring*. *Curr Drug Targets* **8**(8), 914–22 (—2007—).
- [3] J. D. MCGARRY. *Banting lecture 2001: dysregulation of fatty acid metabolism in the etiology of type 2 diabetes*. *Diabetes* **51**(1), 7–18 (—2002—).
- [4] S. E. OZANNE and C. N. HALES. *Pre- and early postnatal nongenetic determinants of type 2 diabetes*. *Expert Rev Mol Med* **4**(24), 1–14 (—2002—).
- [5] D. J. BARKER. *The developmental origins of well-being*. *Philos Trans R Soc Lond B Biol Sci* **359**(1449), 1359–66 (—2004—).
- [6] D. J. BARKER. *The developmental origins of adult disease*. *J Am Coll Nutr* **23**(6 Suppl), 588S–595S (—2004—).
- [7] P. D. GLUCKMAN and M. A. HANSON. *Living with the past: evolution, development, and patterns of disease*. *Science* **305**(5691), 1733–6 (—2004—).
- [8] T. D. PRICE, A. QVARNSTROM and D. E. IRWIN. *The role of phenotypic plasticity in driving genetic evolution*. *Proc Biol Sci* **270**(1523), 1433–40 (—2003—).
- [9] P. D. GLUCKMAN, M. A. HANSON and H. G. SPENCER. *Predictive adaptive responses and human evolution*. *Trends Ecol Evol* **20**(10), 527–33 (—2005—).

- [10] P. BATESON, D. BARKER, T. CLUTTON-BROCK, D. DEB, B. D'UDINE, R. A. FOLEY, P. GLUCKMAN, K. GODFREY, T. KIRKWOOD, M. M. LAHR, J. MCNAMARA, N. B. METCALFE, P. MONAGHAN, H. G. SPENCER and S. E. SULTAN. *Developmental plasticity and human health*. Nature **430**(6998), 419–21 (—2004—).
- [11] P. D. GLUCKMAN, M. A. HANSON and MYILIBRARY. *Mismatch why our world no longer fits our bodies* (—2006—). 2006019107 GBA647553 013469769 (CaONFJC)cis5931326 450959 75317 MIL [electronic resource] : Peter Gluckman and Mark Hanson. 25 cm. Title from e-book title screen (viewed October 15, 2007). Includes bibliographical references (p. 213-266) and index. Electronic reproduction. UK : MyiLibrary, 2007. Connect to MyiLibrary resource.
- [12] S. E. OZANNE and C. N. HALES. *Lifespan: catch-up growth and obesity in male mice*. Nature **427**(6973), 411–2 (—2004—).
- [13] H. VAN REMMEN, Z. GUO and A. RICHARDSON. *The anti-ageing action of dietary restriction*. Novartis Found Symp **235**, 221–30; discussion 230–3 (—2001—).
- [14] C. N. HALES and D. J. BARKER. *Type 2 (non-insulin-dependent) diabetes mellitus: the thrifty phenotype hypothesis*. Diabetologia **35**(7), 595–601 (—1992—).
- [15] J. R. SPEAKMAN. *Obesity: the integrated roles of environment and genetics*. J Nutr **134**(8 Suppl), 2090S–2105S (—2004—).
- [16] J. C. WELLS. *Flaws in the theory of predictive adaptive responses*. Trends Endocrinol Metab **18**(9), 331–7 (—2007—).
- [17] J. R. SPEAKMAN. *Thrifty genes for obesity, an attractive but flawed idea, and an alternative perspective: the 'drifty gene' hypothesis*. Int J Obes (Lond) **32**(11), 1611–7 (—2008—).
- [18] J. V. NEEL. *Diabetes mellitus: a "thrifty" genotype rendered detrimental by "progress"?*. Am J Hum Genet **14**, 353–62 (—1962—).

- [19] D. J. BARKER. *Maternal nutrition, fetal nutrition, and disease in later life.* Nutrition **13**(9), 807–13 (—1997—).
- [20] C. N. HALES and D. J. BARKER. *The thrifty phenotype hypothesis.* Br Med Bull **60**, 5–20 (—2001—).
- [21] I. C. McMILLEN and J. S. ROBINSON. *Developmental origins of the metabolic syndrome: prediction, plasticity, and programming.* Physiol Rev **85**(2), 571–633 (—2005—).
- [22] S. O’RAHILLY and I. S. FAROOQI. *Human obesity as a heritable disorder of the central control of energy balance.* Int J Obes (Lond) **32 Suppl 7**, S55–61 (—2008—).
- [23] P. D. TAYLOR and L. POSTON. *Developmental programming of obesity in mammals.* Exp Physiol **92**(2), 287–98 (—2007—).
- [24] D. J. BARKER. *In utero programming of chronic disease.* Clin Sci (Lond) **95**(2), 115–28 (—1998—).
- [25] A. FORSDAHL. *Are poor living conditions in childhood and adolescence an important risk factor for arteriosclerotic heart disease?.* Br J Prev Soc Med **31**(2), 91–5 (—1977—).
- [26] A. C. RAVELLI, J. H. VAN DER MEULEN, R. P. MICHELS, C. OSMOND, D. J. BARKER, C. N. HALES and O. P. BLEKER. *Glucose tolerance in adults after prenatal exposure to famine.* Lancet **351**(9097), 173–7 (—1998—).
- [27] A. C. RAVELLI, J. H. VAN DER MEULEN, C. OSMOND, D. J. BARKER and O. P. BLEKER. *Obesity at the age of 50 y in men and women exposed to famine prenatally.* Am J Clin Nutr **70**(5), 811–6 (—1999—).
- [28] G. P. RAVELLI, Z. A. STEIN and M. W. SUSSER. *Obesity in young men after famine exposure in utero and early infancy.* N Engl J Med **295**(7), 349–53 (—1976—).

- [29] W. C. KNOWLER, D. J. PETTITT, M. F. SAAD, M. A. CHARLES, R. G. NELSON, B. V. HOWARD, C. BOGARDUS and P. H. BENNETT. *Obesity in the Pima Indians: its magnitude and relationship with diabetes*. Am J Clin Nutr **53**(6 Suppl), 1543S–1551S (—1991—).
- [30] D. R. McCANCE, D. J. PETTITT, R. L. HANSON, L. T. JACOBSSON, W. C. KNOWLER and P. H. BENNETT. *Birth weight and non-insulin dependent diabetes: thrifty genotype, thrifty phenotype, or surviving small baby genotype?*. BMJ **308**(6934), 942–5 (—1994—).
- [31] R. SIMMONS. *Developmental origins of adult metabolic disease: concepts and controversies*. Trends Endocrinol Metab **16**(8), 390–4 (—2005—).
- [32] J. GOLDING. *The Avon Longitudinal Study of Parents and Children (ALSPAC)—study design and collaborative opportunities*. Eur J Endocrinol **151** Suppl **3**, U119–23 (—2004—).
- [33] M. J. WARNER and S. E. OZANNE. *Mechanisms involved in the developmental programming of adulthood disease*. Biochem J **427**(3), 333–47 (—2010—).
- [34] O. VARGA, M. HARANGI, I. A. OLSSON and A. K. HANSEN. *Contribution of animal models to the understanding of the metabolic syndrome: a systematic overview*. Obes Rev (—2009—).
- [35] J. A. ARMITAGE, I. Y. KHAN, P. D. TAYLOR, P. W. NATHANIELSZ and L. POSTON. *Developmental programming of the metabolic syndrome by maternal nutritional imbalance: how strong is the evidence from experimental models in mammals?*. J Physiol **561**(Pt 2), 355–77 (—2004—).
- [36] A. SNOECK, C. REMACLE, B. REUSENS and J. J. HOET. *Effect of a low protein diet during pregnancy on the fetal rat endocrine pancreas*. Biol Neonate **57**(2), 107–18 (—1990—).
- [37] S. E. OZANNE. *Metabolic programming in animals*. Br Med Bull **60**, 143–52 (—2001—).

- [38] S. E. OZANNE, N. D. MARTENSZ, C. J. PETRY, C. L. LOIZOU and C. N. HALES. *Maternal low protein diet in rats programmes fatty acid desaturase activities in the offspring*. *Diabetologia* **41**(11), 1337–42 (—1998—).
- [39] P. R. SHEPHERD, N. J. CROWTHER, M. DESAI, C. N. HALES and S. E. OZANNE. *Altered adipocyte properties in the offspring of protein malnourished rats*. *Br J Nutr* **78**(1), 121–9 (—1997—).
- [40] C. J. PETRY, M. W. DORLING, D. B. PAWLAK, S. E. OZANNE and C. N. HALES. *Diabetes in old male offspring of rat dams fed a reduced protein diet*. *Int J Exp Diabetes Res* **2**(2), 139–43 (—2001—).
- [41] D. S. FERNANDEZ-TWINN, A. WAYMAN, S. EKIZOGLU, M. S. MARTIN, C. N. HALES and S. E. OZANNE. *Maternal protein restriction leads to hyperinsulinemia and reduced insulin-signaling protein expression in 21-mo-old female rat offspring*. *Am J Physiol Regul Integr Comp Physiol* **288**(2), R368–73 (—2005—).
- [42] A. GAROFANO, P. CZERNICHOW and B. BREANT. *In utero undernutrition impairs rat beta-cell development*. *Diabetologia* **40**(10), 1231–4 (—1997—).
- [43] A. GAROFANO, P. CZERNICHOW and B. BREANT. *Beta-cell mass and proliferation following late fetal and early postnatal malnutrition in the rat*. *Diabetologia* **41**(9), 1114–20 (—1998—).
- [44] A. GAROFANO, P. CZERNICHOW and B. BREANT. *Effect of ageing on beta-cell mass and function in rats malnourished during the perinatal period*. *Diabetologia* **42**(6), 711–8 (—1999—).
- [45] M. H. VICKERS, B. H. BREIER, W. S. CUTFIELD, P. L. HOFMAN and P. D. GLUCKMAN. *Fetal origins of hyperphagia, obesity, and hypertension and postnatal amplification by hypercaloric nutrition*. *Am J Physiol Endocrinol Metab* **279**(1), E83–7 (—2000—).

- [46] L. GAMBLING, S. DUNFORD, D. I. WALLACE, G. ZUUR, N. SOLANKY, S. K. SRAI and H. J. MCARDLE. *Iron deficiency during pregnancy affects postnatal blood pressure in the rat.* J Physiol **552**(Pt 2), 603–10 (—2003—).
- [47] B. T. FELT and B. LOZOFF. *Brain iron and behavior of rats are not normalized by treatment of iron deficiency anemia during early development.* J Nutr **126**(3), 693–701 (—1996—).
- [48] J. N. GORSKI, A. A. DUNN-MEYNELL, T. G. HARTMAN and B. E. LEVIN. *Postnatal environment overrides genetic and prenatal factors influencing offspring obesity and insulin resistance.* Am J Physiol Regul Integr Comp Physiol **291**(3), R768–78 (—2006—).
- [49] C. M. BONEY, A. VERMA, R. TUCKER and B. R. VOHR. *Metabolic syndrome in childhood: association with birth weight, maternal obesity, and gestational diabetes mellitus.* Pediatrics **115**(3), e290–6 (—2005—).
- [50] M. A. BROWN, G. PASSARIS and M. A. CARLTON. *Pregnancy-induced hypertension and acute fatty liver of pregnancy: atypical presentations.* Am J Obstet Gynecol **163**(4 Pt 1), 1154–6 (—1990—).
- [51] S. SRINIVASAN, G. R. AMBLER, L. A. BAUR, S. P. GARNETT, M. TEPSA, F. YAP, G. M. WARD and C. T. COWELL. *Randomized, controlled trial of metformin for obesity and insulin resistance in children and adolescents: improvement in body composition and fasting insulin.* J Clin Endocrinol Metab **91**(6), 2074–80 (—2006—).
- [52] S. A. BAYOL, R. MACHARIA, S. J. FARRINGTON, B. H. SIMBI and N. C. STICKLAND. *Evidence that a maternal "junk food" diet during pregnancy and lactation can reduce muscle force in offspring.* Eur J Nutr **48**(1), 62–5 (—2009—).
- [53] A. M. SAMUELSSON, P. A. MATTHEWS, M. ARGENTON, M. R. CHRISTIE, J. M. MCCONNELL, E. H. JANSEN, A. H. PIERSMA, S. E. OZANNE, D. F.

- TWINN, C. REMACLE, A. ROWLERSON, L. POSTON and P. D. TAYLOR. *Diet-induced obesity in female mice leads to offspring hyperphagia, adiposity, hypertension, and insulin resistance: a novel murine model of developmental programming*. *Hypertension* **51**(2), 383–92 (—2008—).
- [54] P. NIVOIT, C. MORENS, F. A. VAN ASSCHE, E. JANSEN, L. POSTON, C. REMACLE and B. REUSENS. *Established diet-induced obesity in female rats leads to offspring hyperphagia, adiposity and insulin resistance*. *Diabetologia* **52**(6), 1133–42 (—2009—).
- [55] M. J. NYIRENDA, R. CARTER, J. I. TANG, A. DE VRIES, C. SCHLUMBOHM, S. G. HILLIER, F. STREIT, M. OELLERICH, V. W. ARMSTRONG, E. FUCHS and J. R. SECKL. *Prenatal programming of metabolic syndrome in the common marmoset is associated with increased expression of 11beta-hydroxysteroid dehydrogenase type 1*. *Diabetes* **58**(12), 2873–9 (—2009—).
- [56] D. S. GARDNER, A. A. JACKSON and S. C. LANGLEY-EVANS. *Maintenance of maternal diet-induced hypertension in the rat is dependent on glucocorticoids*. *Hypertension* **30**(6), 1525–30 (—1997—).
- [57] M. H. VICKERS, P. D. GLUCKMAN, A. H. COVENY, P. L. HOFMAN, W. S. CUTFIELD, A. GERTLER, B. H. BREIER and M. HARRIS. *Neonatal leptin treatment reverses developmental programming*. *Endocrinology* **146**(10), 4211–6 (—2005—).
- [58] M. H. VICKERS, P. D. GLUCKMAN, A. H. COVENY, P. L. HOFMAN, W. S. CUTFIELD, A. GERTLER, B. H. BREIER and M. HARRIS. *The effect of neonatal leptin treatment on postnatal weight gain in male rats is dependent on maternal nutritional status during pregnancy*. *Endocrinology* **149**(4), 1906–13 (—2008—).
- [59] S. YURA, H. ITOH, N. SAGAWA, H. YAMAMOTO, H. MASUZAKI, K. NAKAO, M. KAWAMURA, M. TAKEMURA, K. KAKUI, Y. OGAWA and S. FUJII. *Role*

of premature leptin surge in obesity resulting from intrauterine undernutrition. Cell Metab **1**(6), 371–8 (—2005—).

- [60] A. ALEIXANDRE DE ARTINANO and M. MIGUEL CASTRO. *Experimental rat models to study the metabolic syndrome.* Br J Nutr **102**(9), 1246–53 (—2009—).
- [61] J. PETRIK, J. M. PELL, E. ARANY, T. J. McDONALD, W. L. DEAN, W. REIK and D. J. HILL. *Overexpression of insulin-like growth factor-II in transgenic mice is associated with pancreatic islet cell hyperplasia.* Endocrinology **140**(5), 2353–63 (—1999—).
- [62] C. E. BERTRAM and M. A. HANSON. *Animal models and programming of the metabolic syndrome.* Br Med Bull **60**, 103–21 (—2001—).
- [63] R. A. WATERLAND and C. GARZA. *Potential mechanisms of metabolic imprinting that lead to chronic disease.* Am J Clin Nutr **69**(2), 179–97 (—1999—).
- [64] M. E. SYMONDS, S. P. SEBERT, M. A. HYATT and H. BUDGE. *Nutritional programming of the metabolic syndrome.* Nat Rev Endocrinol **5**(11), 604–10 (—2009—).
- [65] P. D. GLUCKMAN, M. A. HANSON, T. BUKLIJAS, F. M. LOW and A. S. BEEDLE. *Epigenetic mechanisms that underpin metabolic and cardiovascular diseases.* Nat Rev Endocrinol **5**(7), 401–8 (—2009—).
- [66] S. DRAGHICI. *Data analysis tools for DNA microarrays.* Chapman & Hall/CRC mathematical biology and medicine series. Chapman & Hall/CRC, Boca Raton (—2003—).
- [67] T. K. KARAKACH, R. M. FLIGHT and P. D. WENTZELL. *Bootstrap method for the estimation of measurement uncertainty in spotted dual-color DNA microarrays.* Anal Bioanal Chem **389**(7-8), 2125–41 (—2007—).
- [68] K. L. GUNDERSON, S. KRUGLYAK, M. S. GRAIGE, F. GARCIA, B. G. KERMANI, C. ZHAO, D. CHE, T. DICKINSON, E. WICKHAM, J. BIERLE,

- D. DOUCET, M. MILEWSKI, R. YANG, C. SIEGMUND, J. HAAS, L. ZHOU, A. OLIPHANT, J. B. FAN, S. BARNARD and M. S. CHEE. *Decoding randomly ordered DNA arrays*. *Genome Res* **14**(5), 870–7 (—2004—).
- [69] K. KUHN, S. C. BAKER, E. CHUDIN, M. H. LIEU, S. OESER, H. BENNETT, P. RIGALT, D. BARKER, T. K. MCDANIEL and M. S. CHEE. *A novel, high-performance random array platform for quantitative gene expression profiling*. *Genome Res* **14**(11), 2347–56 (—2004—).
- [70] R. C. GENTLEMAN, V. J. CAREY, D. M. BATES, B. BOLSTAD, M. DETTLING, S. DUDOIT, B. ELLIS, L. GAUTIER, Y. GE, J. GENTRY, K. HORNIK, T. HOTHORN, W. HUBER, S. IACUS, R. IRIZARRY, F. LEISCH, C. LI, M. MAECHLER, A. J. ROSSINI, G. SAWITZKI, C. SMITH, G. SMYTH, L. TIERNEY, J. Y. YANG and J. ZHANG. *Bioconductor: open software development for computational biology and bioinformatics*. *Genome Biol* **5**(10), R80 (—2004—).
- [71] G. K. SMYTH. *Linear models and empirical bayes methods for assessing differential expression in microarray experiments*. *Stat Appl Genet Mol Biol* **3**, Article3 (—2004—).
- [72] P. DU, W. A. KIBBE and S. M. LIN. *lumi: a pipeline for processing Illumina microarray*. *Bioinformatics* **24**(13), 1547–8 (—2008—).
- [73] S. M. LIN, P. DU, W. HUBER and W. A. KIBBE. *Model-based variance-stabilizing transformation for Illumina microarray data*. *Nucleic Acids Res* **36**(2), e11 (—2008—).
- [74] B. M. BOLSTAD, R. A. IRIZARRY, M. ASTRAND and T. P. SPEED. *A comparison of normalization methods for high density oligonucleotide array data based on variance and bias*. *Bioinformatics* **19**(2), 185–93 (—2003—).
- [75] P. SYKACEK, R. A. FURLONG and G. MICKLEM. *A friendly statistics package for microarray analysis*. *Bioinformatics* **21**(21), 4069–70 (—2005—).

- [76] L. WERNISCH, S. L. KENDALL, S. SONEJI, A. WIETZORREK, T. PARISH, J. HINDS, P. D. BUTCHER and N. G. STOKER. *Analysis of whole-genome microarray replicates using mixed models*. *Bioinformatics* **19**(1), 53–61 (—2003—).
- [77] J. DENNIS, G., B. T. SHERMAN, D. A. HOSACK, J. YANG, W. GAO, H. C. LANE and R. A. LEMPICKI. *DAVID: Database for Annotation, Visualization, and Integrated Discovery*. *Genome Biol* **4**(5), P3 (—2003—).
- [78] S. DRAGHICI, P. KHATRI, R. P. MARTINS, G. C. OSTERMEIER and S. A. KRAWETZ. *Global functional profiling of gene expression*. *Genomics* **81**(2), 98–104 (—2003—).
- [79] Y. YAMADA, H. WATANABE, F. MIURA, H. SOEJIMA, M. UCHIYAMA, T. IWASAKA, T. MUKAI, Y. SAKAKI and T. ITO. *A comprehensive analysis of allelic methylation status of CpG islands on human chromosome 21q*. *Genome Res* **14**(2), 247–66 (—2004—).
- [80] T. J. MORRIS, M. VICKERS, P. GLUCKMAN, S. GILMOUR and N. AFFARA. *Transcriptional profiling of rats subjected to gestational undernourishment: implications for the developmental variations in metabolic traits*. *PLoS One* **4**(9), e7271 (—2009—).
- [81] M. LIEBERMAN, A. D. MARKS, C. M. SMITH, D. B. MARKS and C. M. M. B. M. B. SMITH. *Marks' essential medical biochemistry*. Lippincott Williams and Wilkins, Philadelphia, Pa. ; London (—2007—).
- [82] K. A. LILLYCROP, E. S. PHILLIPS, A. A. JACKSON, M. A. HANSON and G. C. BURDGE. *Dietary protein restriction of pregnant rats induces and folic acid supplementation prevents epigenetic modification of hepatic gene expression in the offspring*. *J Nutr* **135**(6), 1382–6 (—2005—).
- [83] S. M. WOODALL, B. H. BREIER, B. M. JOHNSTON and P. D. GLUCKMAN. *A model of intrauterine growth retardation caused by chronic maternal under-*

- nutrition in the rat: effects on the somatotrophic axis and postnatal growth.* J Endocrinol **150**(2), 231–42 (—1996—).
- [84] P. GLUCKMAN, A. BEEDLE, M. HANSON and M. H. VICKERS. *Leptin reversal of the metabolic phenotype - evidence for the role of developmental plasticity in the development of the metabolic syndrome.* Hormone Research **67**(Suppl 1), 115–120 (—2007—).
- [85] D. GRANNER and S. PILKIS. *The genes of hepatic glucose metabolism.* J Biol Chem **265**(18), 10173–6 (—1990—).
- [86] C. E. BERTRAM and M. A. HANSON. *Prenatal programming of postnatal endocrine responses by glucocorticoids.* Reproduction **124**(4), 459–67 (—2002—).
- [87] L. SEDOVA, O. SEDA, L. KAZDOVA, B. CHYLIKOVA, P. HAMET, J. TREMBLAY, V. KREN and D. KRENOVA. *Sucrose feeding during pregnancy and lactation elicits distinct metabolic response in offspring of an inbred genetic model of metabolic syndrome.* Am J Physiol Endocrinol Metab **292**(5), E1318–24 (—2007—).
- [88] S. E. OZANNE, G. S. OLSEN, L. L. HANSEN, K. J. TINGEY, B. T. NAVE, C. L. WANG, K. HARTIL, C. J. PETRY, A. J. BUCKLEY and L. MOSTHAFSEEDORF. *Early growth restriction leads to down regulation of protein kinase C zeta and insulin resistance in skeletal muscle.* J Endocrinol **177**(2), 235–41 (—2003—).
- [89] T. R. KOVES, J. R. USSHER, R. C. NOLAND, D. SLENTZ, M. MOSEDALE, O. ILKAYEVA, J. BAIN, R. STEVENS, J. R. DYCK, C. B. NEWGARD, G. D. LOPASCHUK and D. M. MUOIO. *Mitochondrial overload and incomplete fatty acid oxidation contribute to skeletal muscle insulin resistance.* Cell Metab **7**(1), 45–56 (—2008—).
- [90] S. P. BURNS, M. DESAI, R. D. COHEN, C. N. HALES, R. A. ILES, J. P. GERMAIN, T. C. GOING and R. A. BAILEY. *Gluconeogenesis, glucose han-*

- dling, and structural changes in livers of the adult offspring of rats partially deprived of protein during pregnancy and lactation. J Clin Invest* **100**(7), 1768–74 (—1997—).
- [91] P. D. GLUCKMAN, K. A. LILLYCROP, M. H. VICKERS, A. B. PLEASANTS, E. S. PHILLIPS, A. S. BEEDLE, G. C. BURDGE and M. A. HANSON. *Metabolic plasticity during mammalian development is directionally dependent on early nutritional status. Proc Natl Acad Sci U S A* **104**(31), 12796–800 (—2007—).
- [92] M. H. VICKERS, B. H. BREIER, D. MCCARTHY and P. D. GLUCKMAN. *Sedentary behavior during postnatal life is determined by the prenatal environment and exacerbated by postnatal hypercaloric nutrition. Am J Physiol Regul Integr Comp Physiol* **285**(1), R271–3 (—2003—).
- [93] P. M. STEWART, B. A. MURRY and J. I. MASON. *Type 2 11 beta-hydroxysteroid dehydrogenase in human fetal tissues. J Clin Endocrinol Metab* **78**(6), 1529–32 (—1994—).
- [94] P. MILKEREIT, O. GADAL, A. PODTELEJNIKOV, S. TRUMTEL, N. GAS, E. PETFALSKI, D. TOLLERVEY, M. MANN, E. HURT and H. TSCHOCHNER. *Maturation and intranuclear transport of pre-ribosomes requires Noc proteins. Cell* **105**(4), 499–509 (—2001—).
- [95] K. UCHIO, B. TUCHWEBER, N. MANABE, G. GABBIANI, J. ROSENBAUM and A. DESMOULIERE. *Cellular retinol-binding protein-1 expression and modulation during in vivo and in vitro myofibroblastic differentiation of rat hepatic stellate cells and portal fibroblasts. Lab Invest* **82**(5), 619–28 (—2002—).
- [96] S. LEPREUX, P. BIOULAC-SAGE, G. GABBIANI, V. SAPIN, C. HOUSSET, J. ROSENBAUM, C. BALABAUD and A. DESMOULIERE. *Cellular retinol-binding protein-1 expression in normal and fibrotic/cirrhotic human liver: different patterns of expression in hepatic stellate cells and (myo)fibroblast subpopulations. J Hepatol* **40**(5), 774–80 (—2004—).

- [97] O. ZIOUZENKOVA and J. PLUTZKY. *Retinoid metabolism and nuclear receptor responses: New insights into coordinated regulation of the PPAR-RXR complex*. FEBS Lett **582**(1), 32–8 (—2008—).
- [98] S. MILANI, C. GRAPPONE, G. PELLEGRINI, D. SCHUPPAN, H. HERBST, A. CALABRO, A. CASINI, M. PINZANI and C. SURRENTI. *Undulin RNA and protein expression in normal and fibrotic human liver*. Hepatology **20**(4 Pt 1), 908–16 (—1994—).
- [99] D. R. GERECKE, X. MENG, B. LIU and D. E. BIRK. *Complete primary structure and genomic organization of the mouse Col14a1 gene*. Matrix Biol **22**(7), 595–601 (—2004—).
- [100] Z. KMIEC. *Cooperation of liver cells in health and disease*. Adv Anat Embryol Cell Biol **161**, III–XIII, 1–151 (—2001—).
- [101] G. KUNOS and E. J. ISHAC. *Mechanism of inverse regulation of alpha 1- and beta-adrenergic receptors*. Biochem Pharmacol **36**(8), 1185–91 (—1987—).
- [102] R. BURCELIN, M. ULDRY, M. FORETZ, C. PERRIN, A. DACOSTA, M. NENNIGER-TOSATO, J. SEYDOUX, S. COTECCHIA and B. THORENS. *Impaired glucose homeostasis in mice lacking the alpha1b-adrenergic receptor subtype*. J Biol Chem **279**(2), 1108–15 (—2004—).
- [103] N. BEGIN-HEICK. *Liver beta-adrenergic receptors, G proteins, and adenylyl cyclase activity in obesity-diabetes syndromes*. Am J Physiol **266**(6 Pt 1), C1664–72 (—1994—).
- [104] L. M. SPARKS, H. XIE, R. A. KOZA, R. MYNATT, M. W. HULVER, G. A. BRAY and S. R. SMITH. *A high-fat diet coordinately downregulates genes required for mitochondrial oxidative phosphorylation in skeletal muscle*. Diabetes **54**(7), 1926–33 (—2005—).

- [105] K. DU, S. HERZIG, R. N. KULKARNI and M. MONTMINY. *TRB3: a tribbles homolog that inhibits Akt/PKB activation by insulin in liver*. *Science* **300**(5625), 1574–7 (—2003—).
- [106] F. G. HAJ, J. M. ZABOLOTNY, Y. B. KIM, B. B. KAHN and B. G. NEEL. *Liver-specific protein-tyrosine phosphatase 1B (PTP1B) re-expression alters glucose homeostasis of PTP1B^{-/-}mice*. *J Biol Chem* **280**(15), 15038–46 (—2005—).
- [107] E. F. FARIAS, C. MARZAN and R. MIRA-Y LOPEZ. *Cellular retinol-binding protein-I inhibits PI3K/Akt signaling through a retinoic acid receptor-dependent mechanism that regulates p85-p110 heterodimerization*. *Oncogene* **24**(9), 1598–606 (—2005—).
- [108] I. LISINSKI, H. MATSUMOTO, D. R. YVER, A. SCHURMANN, S. W. CUSHMAN and H. AL-HASANI. *Identification and characterization of p49/STRAP as a novel GLUT4-binding protein*. *Biochem Biophys Res Commun* **344**(4), 1179–85 (—2006—).
- [109] H. IWAHASHI, K. YAMAGATA, I. YOSHIUCHI, J. TERASAKI, Q. YANG, K. FUKUI, A. IHARA, Q. ZHU, T. ASAKURA, Y. CAO, A. IMAGAWA, M. NAMBA, T. HANAFUSA, J. MIYAGAWA and Y. MATSUZAWA. *Thyroid hormone receptor interacting protein 3 (trip3) is a novel coactivator of hepatocyte nuclear factor-4alpha*. *Diabetes* **51**(4), 910–4 (—2002—).
- [110] M. MATSUDA, B. S. KORN, R. E. HAMMER, Y. A. MOON, R. KOMURO, J. D. HORTON, J. L. GOLDSTEIN, M. S. BROWN and I. SHIMOMURA. *SREBP cleavage-activating protein (SCAP) is required for increased lipid synthesis in liver induced by cholesterol deprivation and insulin elevation*. *Genes Dev* **15**(10), 1206–16 (—2001—).
- [111] S. B. WHEATCROFT, M. T. KEARNEY, A. M. SHAH, V. A. EZZAT, J. R. MIELL, M. MODO, S. C. WILLIAMS, W. P. CAWTHORN, G. MEDINA-

- GOMEZ, A. VIDAL-PUIG, J. K. SETHI and P. A. CROSSEY. *IGF-binding protein-2 protects against the development of obesity and insulin resistance*. *Diabetes* **56**(2), 285–94 (—2007—).
- [112] N. T. LAM, S. D. COVEY, J. T. LEWIS, S. OOSMAN, T. WEBBER, E. C. HSU, A. T. CHEUNG and T. J. KIEFFER. *Leptin resistance following over-expression of protein tyrosine phosphatase 1B in liver*. *J Mol Endocrinol* **36**(1), 163–74 (—2006—).
- [113] A. GUMMESSON, M. JERNAS, P. A. SVENSSON, I. LARSSON, C. A. GLAD, E. SCHELE, L. GRIPETEG, K. SJOHOLM, T. C. LYSTIG, L. SJOSTROM, B. CARLSSON, B. FAGERBERG and L. M. CARLSSON. *Relations of adipose tissue CIDEA gene expression to basal metabolic rate, energy restriction, and obesity: population-based and dietary intervention studies*. *J Clin Endocrinol Metab* **92**(12), 4759–65 (—2007—).
- [114] I. DAHLMAN, M. KAAMAN, H. JIAO, J. KERE, M. LAAKSO and P. ARNER. *The CIDEA gene V115F polymorphism is associated with obesity in Swedish subjects*. *Diabetes* **54**(10), 3032–4 (—2005—).
- [115] Z. ZHOU, S. YON TOH, Z. CHEN, K. GUO, C. P. NG, S. PONNIAH, S. C. LIN, W. HONG and P. LI. *Cidea-deficient mice have lean phenotype and are resistant to obesity*. *Nat Genet* **35**(1), 49–56 (—2003—).
- [116] M. BIBIKOVA, Z. LIN, L. ZHOU, E. CHUDIN, E. W. GARCIA, B. WU, D. DOUCET, N. J. THOMAS, Y. WANG, E. VOLLMER, T. GOLDMANN, C. SEIFART, W. JIANG, D. L. BARKER, M. S. CHEE, J. FLOROS and J. B. FAN. *High-throughput DNA methylation profiling using universal bead arrays*. *Genome Res* **16**(3), 383–93 (—2006—).
- [117] A. STANHILL, C. M. HAYNES, Y. ZHANG, G. MIN, M. C. STEELE, J. KALININA, E. MARTINEZ, C. M. PICKART, X. P. KONG and D. RON. *An arsenite-inducible 19S regulatory particle-associated protein adapts proteasomes to proteotoxicity*. *Mol Cell* **23**(6), 875–85 (—2006—).

- [118] K. W. MIN, M. J. KWON, H. S. PARK, Y. PARK, S. K. YOON and J. B. YOON. *CAND1 enhances deneddylation of CUL1 by COP9 signalosome*. *Biochem Biophys Res Commun* **334**(3), 867–74 (—2005—).
- [119] T. KURZ, N. OZLU, F. RUDOLF, S. M. O’ROURKE, B. LUKE, K. HOFMANN, A. A. HYMAN, B. BOWERMAN and M. PETER. *The conserved protein DCN-1/Dcn1p is required for cullin neddylation in C. elegans and S. cerevisiae*. *Nature* **435**(7046), 1257–61 (—2005—).
- [120] F. EISENHABER, C. WECHSELBERGER and G. KREIL. *The Brix domain protein family – a key to the ribosomal biogenesis pathway?*. *Trends Biochem Sci* **26**(6), 345–7 (—2001—).
- [121] U. REISS, B. OSKOUIAN, J. ZHOU, V. GUPTA, P. SOORIYAKUMARAN, S. KELLY, E. WANG, J. MERRILL, A. H. and J. D. SABA. *Sphingosine-phosphate lyase enhances stress-induced ceramide generation and apoptosis*. *J Biol Chem* **279**(2), 1281–90 (—2004—).
- [122] A. A. VENNER, M. E. LYON and P. K. DOYLE-BAKER. *Leptin: a potential biomarker for childhood obesity?*. *Clin Biochem* **39**(11), 1047–56 (—2006—).
- [123] R. S. AHIMA and S. Y. OSEI. *Leptin signaling*. *Physiol Behav* **81**(2), 223–41 (—2004—).
- [124] N. HOGGARD, J. G. MERCER, D. V. RAYNER, K. MOAR, P. TRAYHURN and L. M. WILLIAMS. *Localization of leptin receptor mRNA splice variants in murine peripheral tissues by RT-PCR and in situ hybridization*. *Biochem Biophys Res Commun* **232**(2), 383–7 (—1997—).
- [125] D. L. COLEMAN. *Obese and diabetes: two mutant genes causing diabetes-obesity syndromes in mice*. *Diabetologia* **14**(3), 141–8 (—1978—).
- [126] Y. ZHANG, R. PROENCA, M. MAFFEI, M. BARONE, L. LEOPOLD and J. M. FRIEDMAN. *Positional cloning of the mouse obese gene and its human homologue*. *Nature* **372**(6505), 425–32 (—1994—).

- [127] C. T. MONTAGUE, I. S. FAROOQI, J. P. WHITEHEAD, M. A. SOOS, H. RAU, N. J. WAREHAM, C. P. SEWTER, J. E. DIGBY, S. N. MOHAMMED, J. A. HURST, C. H. CHEETHAM, A. R. EARLEY, A. H. BARNETT, J. B. PRINS and S. O'RAHILLY. *Congenital leptin deficiency is associated with severe early-onset obesity in humans*. *Nature* **387**(6636), 903–8 (—1997—).
- [128] K. CLEMENT, C. VAISSE, N. LAHLOU, S. CABROL, V. PELLOUX, D. CASUTO, M. GOURMELEN, C. DINA, J. CHAMBAZ, J. M. LACORTE, A. BASDEVANT, P. BOUGNERES, Y. LEBouc, P. FROGUEL and B. GUY-GRAND. *A mutation in the human leptin receptor gene causes obesity and pituitary dysfunction*. *Nature* **392**(6674), 398–401 (—1998—).
- [129] R. V. CONSIDINE, E. L. CONSIDINE, C. J. WILLIAMS, T. M. HYDE and J. F. CARO. *The hypothalamic leptin receptor in humans: identification of incidental sequence polymorphisms and absence of the db/db mouse and fa/fa rat mutations*. *Diabetes* **45**(7), 992–4 (—1996—).
- [130] I. S. FAROOQI and S. O'RAHILLY. *Genetic factors in human obesity*. *Obes Rev* **8 Suppl 1**, 37–40 (—2007—).
- [131] I. S. FAROOQI and S. O'RAHILLY. *Mutations in ligands and receptors of the leptin-melanocortin pathway that lead to obesity*. *Nat Clin Pract Endocrinol Metab* **4**(10), 569–77 (—2008—).
- [132] R. H. LUSTIG. *Childhood obesity: behavioral aberration or biochemical drive? Reinterpreting the First Law of Thermodynamics*. *Nat Clin Pract Endocrinol Metab* **2**(8), 447–58 (—2006—).
- [133] R. COPPARI, M. ICHINOSE, C. E. LEE, A. E. PULLEN, C. D. KENNY, R. A. MCGOVERN, V. TANG, S. M. LIU, T. LUDWIG, J. CHUA, S. C., B. B. LOWELL and J. K. ELMQUIST. *The hypothalamic arcuate nucleus: a key site for mediating leptin's effects on glucose homeostasis and locomotor activity*. *Cell Metab* **1**(1), 63–72 (—2005—).

- [134] R. H. LUSTIG. *The neuroendocrinology of childhood obesity*. *Pediatr Clin North Am* **48**(4), 909–30 (—2001—).
- [135] S. G. BOURET, S. J. DRAPER and R. B. SIMERLY. *Formation of projection pathways from the arcuate nucleus of the hypothalamus to hypothalamic regions implicated in the neural control of feeding behavior in mice*. *J Neurosci* **24**(11), 2797–805 (—2004—).
- [136] M. G. MYERS, M. A. COWLEY and H. MUNZBERG. *Mechanisms of leptin action and leptin resistance*. *Annu Rev Physiol* **70**, 537–56 (—2008—).
- [137] K. W. WILLIAMS, M. M. SCOTT and J. K. ELMQUIST. *From observation to experimentation: leptin action in the mediobasal hypothalamus*. *Am J Clin Nutr* **89**(3), 985S–990S (—2009—).
- [138] J. LICINIO, A. B. NEGRAO, C. MANTZOROS, V. KAKLAMANI, M. L. WONG, P. B. BONGIORNO, P. P. NEGRO, A. MULLA, J. D. VELDHUIS, L. CEARNAL, J. S. FLIER and P. W. GOLD. *Sex differences in circulating human leptin pulse amplitude: clinical implications*. *J Clin Endocrinol Metab* **83**(11), 4140–7 (—1998—).
- [139] K. KANEKO, T. YAMADA, S. TSUKITA, K. TAKAHASHI, Y. ISHIGAKI, Y. OKA and H. KATAGIRI. *Obesity alters circadian expressions of molecular clock genes in the brainstem*. *Brain Res* **1263**, 58–68 (—2009—).
- [140] K. L. HOUSEKNECHT, M. K. MCGUIRE, C. P. PORTOCARRERO, M. A. MCGUIRE and K. BEERMAN. *Leptin is present in human milk and is related to maternal plasma leptin concentration and adiposity*. *Biochem Biophys Res Commun* **240**(3), 742–7 (—1997—).
- [141] D. J. CLEGG, L. M. BROWN, S. C. WOODS and S. C. BENOIT. *Gonadal hormones determine sensitivity to central leptin and insulin*. *Diabetes* **55**(4), 978–87 (—2006—).

- [142] S. HUMMASTI and P. TONTONAZ. *The peroxisome proliferator-activated receptor N-terminal domain controls isotype-selective gene expression and adipogenesis*. Mol Endocrinol **20**(6), 1261–75 (—2006—).
- [143] S. GRAY, B. WANG, Y. ORIHUELA, E. G. HONG, S. FISCH, S. HALDAR, G. W. CLINE, J. K. KIM, O. D. PERONI, B. B. KAHN and M. K. JAIN. *Regulation of gluconeogenesis by Kruppel-like factor 15*. Cell Metab **5**(4), 305–12 (—2007—).
- [144] A. M. PAJOR. *Molecular properties of the SLC13 family of dicarboxylate and sulfate transporters*. Pflugers Arch **451**(5), 597–605 (—2006—).
- [145] H. ANDO, T. TAKAMURA, N. MATSUZAWA-NAGATA, K. R. SHIMA, T. ETO, H. MISU, M. SHIRAMOTO, T. TSURU, S. IRIE, A. FUJIMURA and S. KANEKO. *Clock gene expression in peripheral leucocytes of patients with type 2 diabetes*. Diabetologia **52**(2), 329–35 (—2009—).
- [146] A. Y. SO, T. U. BERNAL, M. L. PILLSBURY, K. R. YAMAMOTO and B. J. FELDMAN. *Glucocorticoid regulation of the circadian clock modulates glucose homeostasis*. Proc Natl Acad Sci U S A **106**(41), 17582–7 (—2009—).
- [147] M. TOMIHIRA, E. KAWASAKI, H. NAKAJIMA, Y. IMAMURA, Y. SATO, M. SATA, M. KAGE, H. SUGIE and K. NUNOI. *Intermittent and recurrent hepatomegaly due to glycogen storage in a patient with type 1 diabetes: genetic analysis of the liver glycogen phosphorylase gene (PYGL)*. Diabetes Res Clin Pract **65**(2), 175–82 (—2004—).
- [148] M. H. JIANG, J. FEI, M. S. LAN, Z. P. LU, M. LIU, W. W. FAN, X. GAO and D. R. LU. *Hypermethylation of hepatic Gck promoter in ageing rats contributes to diabetogenic potential*. Diabetologia **51**(8), 1525–33 (—2008—).
- [149] I. BOGDARINA, H. C. MURPHY, S. P. BURNS and A. J. CLARK. *Investigation of the role of epigenetic modification of the rat glucokinase gene in fetal programming*. Life Sci **74**(11), 1407–15 (—2004—).

- [150] G. FERRY, E. TELLIER, A. TRY, S. GRES, I. NAIME, M. F. SIMON, M. RODRIGUEZ, J. BOUCHER, I. TACK, S. GESTA, P. CHOMARAT, M. DIEU, M. RAES, J. P. GALIZZI, P. VALET, J. A. BOUTIN and J. S. SAULNIER-BLACHE. *Autotaxin is released from adipocytes, catalyzes lysophosphatidic acid synthesis, and activates preadipocyte proliferation. Up-regulated expression with adipocyte differentiation and obesity.* J Biol Chem **278**(20), 18162–9 (—2003—).
- [151] L. VILA, N. ROGLANS, M. ALEGRET, A. CAMINS, M. PALLAS, R. M. SANCHEZ, M. VAZQUEZ-CARRERA and J. C. LAGUNA. *Hepatic gene expression changes in an experimental model of accelerated senescence: the SAM-P8 mouse.* J Gerontol A Biol Sci Med Sci **63**(10), 1043–52 (—2008—).
- [152] S. RAJARAM, D. J. BAYLINK and S. MOHAN. *Insulin-like growth factor-binding proteins in serum and other biological fluids: regulation and functions.* Endocr Rev **18**(6), 801–31 (—1997—).
- [153] H. K. WADDINGTON. *Fetal salvage in abruptio placentae.* Am J Obstet Gynecol **73**(4), 816–21 (—1957—).
- [154] S. E. OZANNE and M. CONSTANCIA. *Mechanisms of disease: the developmental origins of disease and the role of the epigenotype.* Nat Clin Pract Endocrinol Metab **3**(7), 539–46 (—2007—).
- [155] S. H. ZEISEL. *Epigenetic mechanisms for nutrition determinants of later health outcomes.* Am J Clin Nutr **89**(5), 1488S–1493S (—2009—).
- [156] J. QIU. *Epigenetics: unfinished symphony.* Nature **441**(7090), 143–5 (—2006—).
- [157] W. CUTFIELD, P. HOFMAN, M. MITCHELL and I. MORISON. *Could Epigenetics Play a Role in the Developmental Origins of Health and Disease.* Pediatr Res **61**(5), 68R–75R (—2007—).

- [158] R. JAENISCH. *DNA methylation and imprinting: why bother?*. Trends Genet **13**(8), 323–9 (—1997—).
- [159] R. J. KLOSE and A. P. BIRD. *Genomic DNA methylation: the mark and its mediators*. Trends Biochem Sci **31**(2), 89–97 (—2006—).
- [160] F. LARSEN, G. GUNDERSEN, R. LOPEZ and H. PRYDZ. *CpG islands as gene markers in the human genome*. Genomics **13**(4), 1095–107 (—1992—).
- [161] P. A. JONES and D. TAKAI. *The role of DNA methylation in mammalian epigenetics*. Science **293**(5532), 1068–70 (—2001—).
- [162] F. WATT and P. L. MOLLOY. *Cytosine methylation prevents binding to DNA of a HeLa cell transcription factor required for optimal expression of the adenovirus major late promoter*. Genes Dev **2**(9), 1136–43 (—1988—).
- [163] J. BOYES and A. BIRD. *DNA methylation inhibits transcription indirectly via a methyl-CpG binding protein*. Cell **64**(6), 1123–34 (—1991—).
- [164] B. HENDRICH and A. BIRD. *Identification and characterization of a family of mammalian methyl-CpG binding proteins*. Mol Cell Biol **18**(11), 6538–47 (—1998—).
- [165] J. C. CHUANG and P. A. JONES. *Epigenetics and microRNAs*. Pediatr Res **61**(5 Pt 2), 24R–29R (—2007—).
- [166] A. H. PETERS, J. E. MERMOUD, D. O’CARROLL, M. PAGANI, D. SCHWEIZER, N. BROCKDORFF and T. JENUWEIN. *Histone H3 lysine 9 methylation is an epigenetic imprint of facultative heterochromatin*. Nat Genet **30**(1), 77–80 (—2002—).
- [167] L. JOHNSON, X. CAO and S. JACOBSEN. *Interplay between two epigenetic marks. DNA methylation and histone H3 lysine 9 methylation*. Curr Biol **12**(16), 1360–7 (—2002—).

- [168] X. NAN, H. H. NG, C. A. JOHNSON, C. D. LAHERTY, B. M. TURNER, R. N. EISENMAN and A. BIRD. *Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex*. *Nature* **393**(6683), 386–9 (—1998—).
- [169] F. FUKS, P. J. HURD, D. WOLF, X. NAN, A. P. BIRD and T. KOUZARIDES. *The methyl-CpG-binding protein MeCP2 links DNA methylation to histone methylation*. *J Biol Chem* **278**(6), 4035–40 (—2003—).
- [170] F. FUKS, W. A. BURGERS, A. BREHM, L. HUGHES-DAVIES and T. KOUZARIDES. *DNA methyltransferase Dnmt1 associates with histone deacetylase activity*. *Nat Genet* **24**(1), 88–91 (—2000—).
- [171] B. E. BERNSTEIN, A. MEISSNER and E. S. LANDER. *The mammalian epigenome*. *Cell* **128**(4), 669–81 (—2007—).
- [172] R. A. WATERLAND and K. B. MICHELS. *Epigenetic epidemiology of the developmental origins hypothesis*. *Annu Rev Nutr* **27**, 363–88 (—2007—).
- [173] G. L. WOLFF, R. L. KODELL, S. R. MOORE and C. A. COONEY. *Maternal epigenetics and methyl supplements affect agouti gene expression in Avy/a mice*. *FASEB J* **12**(11), 949–57 (—1998—).
- [174] R. A. WATERLAND and R. L. JIRTLE. *Early nutrition, epigenetic changes at transposons and imprinted genes, and enhanced susceptibility to adult chronic diseases*. *Nutrition* **20**(1), 63–8 (—2004—).
- [175] W. Y. TANG and S. M. HO. *Epigenetic reprogramming and imprinting in origins of disease*. *Rev Endocr Metab Disord* **8**(2), 173–82 (—2007—).
- [176] K. A. LILLYCROP, E. S. PHILLIPS, A. A. JACKSON, M. A. HANSON and G. C. BURDGE. *Dietary protein restriction of pregnant rats induces and folic acid supplementation prevents epigenetic modification of hepatic gene expression in the offspring*. *J Nutr* **135**(6), 1382–6 (—2005—).

- [177] I. BOGDARINA, S. WELHAM, P. J. KING, S. P. BURNS and A. J. CLARK. *Epigenetic modification of the renin-angiotensin system in the fetal programming of hypertension*. *Circ Res* **100**(4), 520–6 (—2007—).
- [178] B. T. HEIJMANS, E. W. TOBI, A. D. STEIN, H. PUTTER, G. J. BLAUW, E. S. SUSSER, P. E. SLAGBOOM and L. H. LUMEY. *Persistent epigenetic differences associated with prenatal exposure to famine in humans*. *Proc Natl Acad Sci U S A* **105**(44), 17046–9 (—2008—).
- [179] P. W. LAIRD. *Principles and challenges of genome-wide DNA methylation analysis*. *Nat Rev Genet* **11**(3), 191–203 (—2010—).
- [180] K. FUJIKI, F. KANO, K. SHIOTA and M. MURATA. *Expression of the peroxisome proliferator activated receptor gamma gene is repressed by DNA methylation in visceral adipose tissue of mouse models of diabetes*. *BMC Biol* **7**, 38 (—2009—).
- [181] A. PLAGEMANN, T. HARDER, M. BRUNN, A. HARDER, K. ROEPKE, M. WITTRICK-STAAAR, T. ZISKA, K. SCHELLONG, E. RODEKAMP, K. MELCHIOR and J. W. DUDENHAUSEN. *Hypothalamic proopiomelanocortin promoter methylation becomes altered by early overfeeding: an epigenetic model of obesity and the metabolic syndrome*. *J Physiol* **587**(Pt 20), 4963–76 (—2009—).
- [182] J. CHEONG, Y. YAMADA, R. YAMASHITA, T. IRIE, A. KANAI, H. WAKAGURI, K. NAKAI, T. ITO, I. SAITO, S. SUGANO and Y. SUZUKI. *Diverse DNA methylation statuses at alternative promoters of human genes in various tissues*. *DNA Res* **13**(4), 155–67 (—2006—).
- [183] J. H. PARK, D. A. STOFFERS, R. D. NICHOLLS and R. A. SIMMONS. *Development of type 2 diabetes following intrauterine growth retardation in rats is associated with progressive epigenetic silencing of Pdx1*. *J Clin Invest* **118**(6), 2316–24 (—2008—).

- [184] K. BREVING and A. ESQUELA-KERSCHER. *The complexities of microRNA regulation: mirandering around the rules*. Int J Biochem Cell Biol (—2009—).
- [185] H. M. HENEGHAN, N. MILLER and M. J. KERIN. *Role of microRNAs in obesity and the metabolic syndrome*. Obes Rev pages 1–8 (—2009—).
- [186] E. BEREZIKOV, V. GURYEV, J. VAN DE BELT, E. WIENHOLDS, R. H. PLASTERK and E. CUPPEN. *Phylogenetic shadowing and computational identification of human microRNA genes*. Cell **120**(1), 21–4 (—2005—).
- [187] S. D. BOYD. *Everything you wanted to know about small RNA but were afraid to ask*. Lab Invest **88**(6), 569–78 (—2008—).
- [188] G. M. BORCHERT, W. LANIER and B. L. DAVIDSON. *RNA polymerase III transcribes human microRNAs*. Nat Struct Mol Biol **13**(12), 1097–101 (—2006—).
- [189] M. N. POY, M. SPRANGER and M. STOFFEL. *microRNAs and the regulation of glucose and lipid metabolism*. Diabetes Obes Metab **9 Suppl 2**, 67–73 (—2007—).
- [190] B. BRUECKNER, R. GARCIA BOY, P. SIEDLECKI, T. MUSCH, H. C. KLIEM, P. ZIELENKIEWICZ, S. SUHAI, M. WIESSLER and F. LYKO. *Epigenetic reactivation of tumor suppressor genes by a novel small-molecule inhibitor of human DNA methyltransferases*. Cancer Res **65**(14), 6305–11 (—2005—).
- [191] E. H. BAYNE and R. C. ALLSHIRE. *RNA-directed transcriptional gene silencing in mammals*. Trends Genet **21**(7), 370–3 (—2005—).
- [192] L. HE and G. J. HANNON. *MicroRNAs: small RNAs with a big role in gene regulation*. Nat Rev Genet **5**(7), 522–31 (—2004—).
- [193] N. BAO, K. W. LYE and M. K. BARTON. *MicroRNA binding sites in Arabidopsis class III HD-ZIP mRNAs are required for methylation of the template chromosome*. Dev Cell **7**(5), 653–62 (—2004—).

- [194] L. TUDDENHAM, G. WHEELER, S. NTOUNIA-FOUSARA, J. WATERS, M. K. HAJIHOSSEINI, I. CLARK and T. DALMAY. *The cartilage specific microRNA-140 targets histone deacetylase 4 in mouse cells*. FEBS Lett **580**(17), 4214–7 (—2006—).
- [195] J. KRUTZFELDT and M. STOFFEL. *MicroRNAs: a new class of regulatory genes affecting metabolism*. Cell Metab **4**(1), 9–12 (—2006—).
- [196] A. KAUFFMANN, R. GENTLEMAN and W. HUBER. *arrayQualityMetrics—a bioconductor package for quality assessment of microarray data*. Bioinformatics **25**(3), 415–6 (—2009—).
- [197] N. NAKANISHI, Y. NAKAGAWA, N. TOKUSHIGE, N. AOKI, T. MATSUZAKA, K. ISHII, N. YAHAGI, K. KOBAYASHI, S. YATOH, A. TAKAHASHI, H. SUZUKI, O. URAYAMA, N. YAMADA and H. SHIMANO. *The up-regulation of microRNA-335 is associated with lipid metabolism in liver and white adipose tissue of genetically obese mice*. Biochem Biophys Res Commun **385**(4), 492–6 (—2009—).
- [198] S. LI, X. CHEN, H. ZHANG, X. LIANG, Y. XIANG, C. YU, K. ZEN, Y. LI and C. Y. ZHANG. *Differential expression of microRNAs in mouse liver under aberrant energy metabolic status*. J Lipid Res **50**(9), 1756–65 (—2009—).
- [199] S. F. TAVAZOIE, C. ALARCON, T. OSKARSSON, D. PADUA, Q. WANG, P. D. BOS, W. L. GERALD and J. MASSAGUE. *Endogenous human microRNAs that suppress breast cancer metastasis*. Nature **451**(7175), 147–52 (—2008—).
- [200] M. HOLMQUIST. *Alpha/Beta-hydrolase fold enzymes: structures, functions and mechanisms*. Curr Protein Pept Sci **1**(2), 209–35 (—2000—).
- [201] M. TAKAHASHI, Y. KAMEI and O. EZAKI. *Mest/Peg1 imprinted gene enlarges adipocytes and is a marker of adipocyte size*. Am J Physiol Endocrinol Metab **288**(1), E117–24 (—2005—).
- [202] L. NIKONOVA, R. A. KOZA, T. MENDOZA, P. M. CHAO, J. P. CURLEY and L. P. KOZAK. *Mesoderm-specific transcript is associated with fat mass*

expansion in response to a positive energy balance. FASEB J **22**(11), 3925–37 (—2008—).

[203] H. ZHU, H. WU, X. LIU, B. R. EVANS, D. J. MEDINA, C. G. LIU and J. M. YANG. *Role of MicroRNA miR-27a and miR-451 in the regulation of MDR1/P-glycoprotein expression in human cancer cells.* Biochem Pharmacol **76**(5), 582–8 (—2008—).

[204] S. Y. KIM, A. Y. KIM, H. W. LEE, Y. H. SON, G. Y. LEE, J. W. LEE, Y. S. LEE and J. B. KIM. *miR-27a is a negative regulator of adipocyte differentiation via suppressing PPARgamma expression.* Biochem Biophys Res Commun **392**(3), 323–8 (—2010—).