Genetic and Functional Investigation of Sequence Variation Associated with the Related Phenotypes: Birth Weight, Obesity and Type 2 Diabetes (T2D)

By

Laura Katherine Towns

Thesis submitted in fulfilment of the requirement
For the degree of
Doctor of Philosophy (PhD)

Faculty of Life Sciences, School of Human Sciences
London Metropolitan University

In Collaboration with
Department of Genomics of Common Disease
Imperial College, London

June 2013

Acknowledgements

I am eternally grateful to my supervisors Dr Una Fairbrother and Dr Andrew Walley for their invaluable advice, patience and encouragement over the last five years, without which completion of my thesis would not have been possible.

The guidance and assistance provided by Mrs Marlene Attard, Dr Stephen Clark, Dr Petros Takousis, Dr Afshan Siddiq and Dr Christine Nunn for their technical support was also greatly appreciated. Similarly, I am indebited to Mr Brian Whiting, for his time, patience and help in his artistic expertice.

In addition, special thanks are extended to members of academic and technical staff and fellow PhD researchers at London Metropolitan University and Imperial College London for their advice, words of encouragement, tea and muffins.

I would also like to thank London Metropolitan University and Imperial College London for providing me with the opportunity, facilities and resources to complete my thesis.

Finally, I express my deepest gratitude to my Mum, Dad, Sister, Aileen and Brother-in-Law, Ed and friends, Andrea, Becky, Emily, Steph, Tasha and Sarah for their love, support and for keeping me sane!

Conference Publications

European Society for Human Genetics, Amsterdam June 2011. *HHEX* is associated with T2D in a South Asian Population. L K Towns, M Attard, A J Walley and U L Fairbrother, European Journal Human Genetics: Vol 19, Supplement 1, Section: Complex Traits and Polygenic Disorders P09.109.

European Society for Human Genetics, Gutenberg June 2010. *HHEX-IDE* genotype is associated with birth weight in South Asians. L K Towns, M Attard, A J Walley and U L Fairbrother.European Journal Human Genetics: Vol, 18, Supplement 1, Section: Complex Traits and Polygenic Disorders P09.065.

British Society for Human Genetics, Warwick September 2009. *HHEX-IDE* genotype is a risk factor for IUGR in South Asians. L K Towns, M Attard, A J Walley and U L Fairbrother. Journal of Medical Genetics, Vol 46, Supplement 1, Section: Molecular Genetics 2.66.

Public Presentations

School of Human Sciences Research Seminar Series, Spring 2012, London Metropolitan University. 'Genes involved in fetal birth weight and its concomitant sequelae'.

Graduate Student Research Symposium, London Metropolitan University, January 2010. '*HHEX-IDE* genotype is associated with birth weight in South Asians'.

Graduate Student Research Symposium, London Metropolitan University, July 2008. 'Investigation of the role of genes in adult and childhood obesity'.

Abbreviations

AC: Abdominal circumference
AGRP: agouti related protein
ALSPAC: Avon Longitudinal Study of Parents and Children
ANCOVA: analysis of covariance
AP: area postrema
APBS: Adaptive Poisson-Boltzmann Solver
ARC: arcuate
ATP: adenosine tri-phosphate
BMI: body mass index
BMR: basal metabolic rate
CART: cocaine and amphetamine related transcript
cAMP: cyclic adenosine mono-phosphate
CEU: Utah residents with Northern and Western European ancestry from CEPH collection
CDC: Centers for Disease Control
CHD: coronary heart disease
CI: confidence interval
CNV: copy number variation
CRH: cytokine receptor homology
CVD: cardiovascular disease
DEPC: Diethylpyrocarbonate
DEXA: dual energy X-ray absorptiometry
DBP: diastolic blood pressure
DMH: dorsomedial hypothalamic nucleus
DMSO: Dimethyl sulfoxide
dNTPs: deoxyribonucleotide triphosphates
DTT: dithiothreitol
DZ: dizygotic
ECG: electrocardiogram
EDTA: ethylenediaminetetraacetic acid
EGIR: European Group for the Study of Insulin Resistance
EFW: estimated fetal weight
FBS: fetal bovine serum
FL: femur length
FNIII: fibronectin type III
FOTI: fat on the outside, thin on the inside
G-CSF: granulocyte colony stimulating factor
GWAS: genome wide association study
HC: head circumference
HDL: high-density lipoprotein
HWE: Hardy-Weinberg equilibrium
IDF: International Diabetes Federation
IFG: impaired fasting glucose
Ig: immunoglobulin
IGT: impaired glucose tolerance
IOTF: International Obesity Task Force
IPTG: isopropyl β-D-1-thiogalactopyroranoside
I-TASSER: iterative threading assembly refinement server

IUGR: intrauterine growth retardation
JAK2: Janus tyrosine kinase 2
LB: Luria-Bertani
LD: linkage disequilibrium
LDL: Low-density lipoprotein
LOLIPOP: London Life Sciences Prospective Population Study
LOMETS: local meta threading sever
MAF: minor allele frequency
MALDI-TOF: matrix-assisted laser desorption/ionisation time of flight
MAPK: mitogen-activated protein kinase
MCF-7: Michigan Cancer Foundation-7
MGB: minor groove binder
M_gCl_2: Magnesium Chloride
M-MuLV-RT: moloney murine leukaemia virus reverse transcriptase
MODY: maturity onset diabetes of the young
MRI: magnetic resonance imagining
MPC: magnetic particle concentrator
MZ: monozygotic
NCBI: National Centre for Biotechnology Information
NCEP-ATP III: National Cholesterol Education Program-Third Adult Treatment Panel
NEAT: non-exercise activity Thermogenesis
NEFA: non-essential fatty acid
NFQ: non-fluorescent quencher
NHANES: National Health and Nutritional Examination Survey
NPY: neuropeptide Y
NTS: nucleus of the solitary tract
OMIM: Online Mendelian Inheritance in Man
OR: odds ratio
PAR: predictive adaptive response theory
PBS: phosphate buffered saline
PCR: polymerase chain reaction
PDB: protein data bank
Pi3K: phosphoinoitide 3 kinase
PKA: protein kinase A
POMC: pro-opiomelanocortin
PSI-BLAST: position specific iterated-basic local alignment search tool
PVN: paraventrical nucleus
PWS: Prader-Willi Syndrome
RR: relative risk
RT-PCR: reverse transcriptase polymerase chain reaction
SA: South Asian
SAP: shrimp alkaline phosphatase
SAT: subcutaneous adipose tissue
SBE: single base extension
SBP: systolic blood pressure
SDS: standard deviation score
SDS software: sequence detection system software
SGA: small for gestational age
SNPs: single nucleotide polymorphisms
SOCS3: suppressors of cytokine signalling 3

SON: supraoptic nuclei
SpD: spectral decomposition
SPSS: statistical package for the Social Sciences
STAT: signal transducer and activator of transcriptional protein
T2D: type 2 diabetes
TBE: Tris-borate-EDTA
TOFI: thin on the outside, fat on the inside
UV: ultra violet
VAT: visceral adipose tissue
VMH: ventro medial
VNTR: variable number tandem repeat
WC: waist circumference
WHO: World Health Organisation
WHR: waist to hip ratio

Abstract

Background: Global incidence of obesity and Type 2 diabetes (T2D) has increased rapidly in the last fifty years and even more rapidly in immigrant ethnic minority communities in Western countries. There is evidence of a connection between birth weight and the subsequent risk of obesity and T2D in adulthood, with several T2D susceptibility loci reported to be associated with fetal birth weight, including the leptin receptor gene (*LEPR*).

Aims: To ascertain if maternal candidate loci contribute to fetal birth weight variation in a Gujarati South Asian immigrant population in North London and to investigate three variants in the extracellular region of *LEPR* to determine their effect on LEPR protein structure and function.

Methods: Two commercially-available methods (Sequenom iPLEX and Life Technologies TaqMan) were used to genotype specific polymorphisms in the genes *HHEX-IDE*, *ADCY5* and *FTO*, while molecular cloning and computer modelling techniques were used to investigate the variants in *LEPR*.

Results: Significant associations were found with lower birth weight and rs12765131 in the *HHEX-IDE* locus (p=0.002) and with higher birth weight and markers in *FTO*, namely rs9939609 (p=0.001), rs17817449 (p=0.0273) and rs8050136 (p=0.0231). No significant associations were found with fetal birth weight and markers in the *ADCY5* gene. Although molecular cloning of three *LEPR* variants was unsuccessful, *in silico* modelling of the LEPR protein revealed that the Gln223Arg polymorphism was predicted to affect the surface electrostatic potential of the receptor and create hydrogen bonding with a neighbouring cysteine residue, potentially lowering the binding affinity with leptin.

Conclusion: These results demonstrate that maternal variants in *HHEX-IDE* and *FTO* are associated with birth weight variation in the Gujarati South Asian population. Furthermore, LEPR polymorphisms have been predicted to affect its overall structure and possibly its leptin signalling function.

Contents

Acknowledgements ii

Conference Publications iii

Public Presentations iii

Abbreviations iv

Abstract vii

List of Figures xii

List of Tables xii

1.0 Introduction 1

1.1.0 Complex disease/birth weight and obesity 1

1.2.0 Classification and Measurement of Complex Disease 3

1.2.1 Classification of Obesity in Adults 3

1.2.2 Classification and Measurement of Obesity in Children 4

1.2.3 Identification of T2D in adults 5

1.2.4 T2D identification in children 6

1.2.5 Identification of the Metabolic Syndrome 6

1.2.6 Reliability of BMI in Determining Disease Risk 9

1.2.7 Reliability of BMI in Determining Disease Risk in Children and Adolescents .14

1.2.8 BMI and Different Ethnic Minorities and Disease Risk 14

1.3.0 The importance of obesity, T2D and the metabolic syndrome in the 21st Century 16

1.3.1 Obesity in Adults and Children 16

1.3.2 T2D in Adults and Children 18

1.4.0 Obesogenic Environment 19

1.4.1 Daily Expenditure of Energy 21

1.5.0 Genetics and Obesity 22

1.5.1 Heritability of Obesity 23

1.5.1.1 Heritability of BMI 23

1.5.1.2 Heritability of Alternative Measures of Obesity 24

1.5.1.3 Heritability of Body Composition 25

1.5.1.4 Heritability of Eating Behaviour 26

1.5.2 Monogenic Forms of Obesity 27

1.5.3 Syndromic Forms of Obesity 28

1.5.4 Identification of Candidate Loci in Polygenic Obesity 28

1.5.4.1 Differences between adult and childhood obesity 35

1.6.0 Explanations into the rise of common obesity 37

1.7.0 Birth Weight 38

1.7.1 Measurement of Birth Weight 39

1.7.2 Birth weight and Increased Disease Risk in Adults and Children 40

1.7.3 Theories Explaining correlation of Birth Weight and Disease Risk 45

1.7.4 Maternal Environment 45

1.7.5 Environmental Origins of Birth Weight Variation 46

1.7.6 Effects of maternal malnutrition and body size 50

1.7.7 Placental Function 54

1.7.8 'Fetal Origin of Adult Disease' and Catch-Up Growth 55

1.7.9 Birth weight and Animal Models 57

1.7.10 Predictive Adaptive Response 59
1.7.11 Genetics and Birth weight Variation and Disease Risk 60
1.7.11.1 Epigenetics, Birth Weight and Disease Risk 60
1.7.12 Genetic Influences of Birth Weight Variation 61
1.7.12.1 Fetal Insulin Hypothesis and Disease Risk 62
1.7.13 Birth Weight Variation and Disease Risk in Ethnic Populations 70
1.8.0 Candidate Genes 71
1.9.0 LEPR and Its Ligand Leptin (*LEP*) 71
1.9.1. Identification of Leptin and Leptin Receptor using Mouse Models 72
1.9.2 Mutations within the Human forms of Leptin and Leptin Receptor 73
1.9.3 Leptin Expression and Regulation and its Role in the Adipostat 74
1.9.4 The Leptin–Melanocortin System 75
1.9.5 Leptin Resistance 78
1.9.6 Leptin Receptor 81
1.9.6.1 Leptin Receptor Extracellular Region 83
1.9.7 Leptin Receptor Signalling Pathways 89
1.9.7.1 LEPR and the JAK/STAT Pathway 89
1.9.7.2 LEPR and the Phosphoinositide 3-kinase/Phosphodiesterase 3B/cAMP Pathway 90
1.9.8 Leptin Receptor Polymorphisms 92
1.10.0 The Haematopoietically Expressed Homeobox – Insulin Degrading Enzyme (*HHEX-IDE*) region 98
1.10.1 *HHEX–IDE* and Fetal Birth Weight 98
1.10.2 *HHEX-IDE* and Childhood BMI 99
1.10.3 Location of *HHEX-IDE* 101
1.10.4 Kinesin Family Member 11 (*KIF11*) 101
1.10.5 *HHEX* Gene Function 104
1.10.6 *IDE* 104
1.11.0 *FTO* 106
1.11.1 Location of *FTO* 106
1.11.2 Function of *FTO* 106
1.11.3 Mouse studies of *FTO* 107
1.11.4 Expression of *FTO* 109
1.11.5 Associations of SNPs within Intron One of FTO with T2D and Obesity 110
1.12.0 Adenylate Cyclase Five (*ADCY5*) 111
1.12.1 *ADCY5* and Birth Weight 112
1.12.2 Location of *ADCY5* 113
1.12.3 *ADCY5* Function 114
1.13.0 South Asian Population 115
1.14.0 Aims of the investigation 117
1.14.1 Functional Studies: 117
1.14.2 Genetic Studies 117

2.0 Material and Methods 119
2.1 Reagents 119
2.2.0 Subjects 121
2.2.1 UK/Irish and South Asian Cohort 121
2.3.0 Statistical Analysis 122
2.3.1 Genotyping analysis, quantitative and qualitative 122
2.4.0 Functional Study of *LEPR* 123
2.4.1 Cell Culture 123
2.4.2 Continued Growth 124

2.4.3 Storing of Cell Pellets at -70°C 124
2.4.4 Growing Cells from Frozen 125
2.4.5 Freezing cells for long-term storage in Liquid Nitrogen 125
2.4.6 Total RNA extraction using Trizol ® Reagent 126
2.4.7 PolyA$^+$ mRNA Extraction - DynaBead Method 126
2.4.7.1 Preparation of the Dynabeads 127
2.4.7.2 Preparation of the Lysate 127
2.4.7.3 mRNA Isolation 127
2.4.8 Reverse Transcription Method using mRNA 128
2.4.9 Reverse Transcription Method using MCF-7 total RNA 131
2.4.10 Reverse Transcription Method using Placental poly A$^+$ RNA 131
2.4.11 Amplification of LEPR using PCR 132
2.4.11.1 Magnesium Titration 134
2.4.12 Agarose Gel Electrophoresis 136
2.4.13 Gel Extraction of *LEPR* PCR Product 137
2.4.14 Ligation of LEPR into pGEM T Vector 138
2.4.15 Transformation of pGEM T Vector using *E. coli* 139
2.4.16 Extraction and Purification of pGEM-T Plasmid 140
2.4.16.1 Mini Preparation of Overnight Culture 141
2.4.16.2 Maxi Preparation of Overnight Cultures 142
2.4.17 Restriction Digest of Plasmid Mini/Maxi Preparations 143
2.4.18 Ligation of the LEPR insert into Expression vector, pCMV Tag 2B 144
2.4.19 Transformation of pCMV Tag 2B expression vector using *E.coli* cells 145
2.4.20 LB Agar supplemented with Ampicillin, X-Gal and IPTG 146
2.4.21 LB agar supplemented with Kanamycin 146
2.4.22 Quantification of DNA and RNA samples 147
2.5.0 Genotyping of *HHEX-IDE* Gene locus in the South Asian Population 148
2.5.1 Sequenom iPLEX Assay Design 148
2.5.2 Quantification of DNA 148
2.5.3 MassARRAY Software 149
2.5.4 Sequenom iPLEX Reaction 150
2.5.5 Initial PCR 150
2.5.6 SAP Clean Up Method 152
2.5.7 Sequenom Single Base Extension (SBE) Reaction 153
2.5.8 MALDI-TOF Mass Spectrometry 158
2.6.0 Genotyping *FTO* variants within the South Asian Population using the Sequenom iPLEX Assay 160
2.7.0 Genotyping the variant rs11708067, close to *ADCY5* using the TaqMan Assay in South Asian Cohort. 162
2.7.1 TaqMan SNP Genotyping Assay Set-up 164
2.7.2 Allelic Discrimination Plate Read 165
2.8.0 Computer Modelling of the Leptin Receptor 167
2.8.1 UNIPROT Knowledgebase 167
2.8.2 The I-Tasser Server. 170

3.0 Results 172
3.1 *LEPR* Functional Study 172
3.1.1 Extraction of RNA and Amplification of Extracellular LEPR using PCR 174
3.1.2 Ligation of pGEM-T vector and LEPR insert and Transformation of *Escherichia coli* (JM109) 180
3.1.3 Double Restriction Digests of Mini Preparations 185

3.1.4 Ligation of pCMV–Tag 2B Expression Vector and Transformation into *E. coli* 203
3.1.5 Nested LEPR PCR 207
3.2.0 Computer Modelling of the Leptin Receptor 212
3.2.1 I-TASSER Server 213
3.2.1.1 Template Identification 213
3.2.1.2 Structure Alignment 214
3.2.1.3 Model selection and refinement 215
3.2.1.4 Structure-based functional annotation 215
3.2.2 Analysis of I-TASSER Results 216
3.2.2.1 Sequence and predicted secondary structure 217
3.2.2.2 Predicted 3D Model 222
3.2.2.3 Threading Alignments 225
3.2.2.4 Structural Analogues 227
3.2.3 Extracellular LEPR Protein Structure Prediction 229
3.2.4 CRH1 I-TASSER Results 233
3.2.4.1 Structural Alignment between CRH1Q and CRH1R Domains 235
3.2.4.2 Hydrogen Bonding in CRH1Q and CRH1R 241
3.2.4.3 Electrostatic potential in CRH1Q and CRH1R 243
3.3.5 CRH2 Domain I-TASSER Results 250
3.3.5.1 Electrostatic Potential in CRH2 253
3.3.6 Electrostatic Potential in Leptin 257

4.0 Investigating Association of Maternal T2D and Obesity Gene SNPs with Fetal Birth Weight 260
4.1.0 Genotyping of *HHEX-IDE* in Guarajti Indian Women using the Sequenom iPLEX assay 261
4.1.1 MassARRAY Software 266
4.1.2 Typer Software for Genotyping Data Analysis 266
4.1.3 *HHEX* Genotyping Success Rates 269
4.1.4. Normalisation of the *HHEX-IDE* phenotype data 273
4.1.5 Analysis of Maternal Genotype and Fetal Birth Weight 273
4.1.6 Multiple Testing Correction 275
4.1.7 Haploview Analysis to Investigate LD across *HHEX-IDE* region in SA cohort 275
4.2.0 Genotyping *FTO* in South Asian Cohort using the Sequenom iPLEX Assay 277
4.2.1 Combing *FTO* Genotyping Data 280
4.2.2 Normalisation of the *FTO* Data 281
4.2.3 Maternal Phenotype and Fetal Birth Weight 281
4.2.4 Linear regression analysis of maternal genotype at the *FTO* locus with fetal birth weight 283
4.3.0 Genotyping the *ADCY5* SNP, rs11708067, in the Maternal SA Cohort with Respect to Fetal Birth Weight 293
4.3.0 Post-read Step 293
4.3.2 Quality Control of Genotyping Data 294
4.3.3 Analysis of Maternal Genotype at rs11708067 with Fetal Birth Weight 298

5.0 Discussion 301
5.1.0 Functional analysis of the Leptin Receptor 301
5.1.1 Methodological Improvements 301
5.2.0 Computer Modelling of the Leptin Receptor 303
5.3.0 Investigation into T2D Risk Alleles with Respect to Fetal Birth Weight 320
5.3.1 Candidate Genes: *HHEX-IDE*, *ADCY5* and *FTO* 320

5.3.1.1 *HHEX-IDE* Locus 320
5.3.1.2 *FTO* Locus 324
5.3.1.3 *ADCY5* 333
5.3.2 Limitations of Candidate Gene Analysis 336
5.3.3 Environmental and Genetic Factors in Birth Weight Variation and Risk of Obesity, T2D and the Metabolic Disease in Adulthood 338
5.4 Conclusion **342**

References **344**

Appendix I: DNA Markers **419**

Appendix II: Forward, reverse and extension primer sequences for *HHEX* Sequenom iPLEX assays one and two **422**

Appendix III: Sequenom iPLEX assay calibrant containing three standard oligonucleotides of known mass: 5100, 8500 and 10,000 mass units. **428**

Appendix IV: Forward, reverse and extension primer sequences for *FTO* Sequenom iPLEX assay **429**

Appendix V: Wild type LEPR Protein FASTA Sequence **431**

Appendix VI: LEPR mRNA Sequence **433**

Appendix VII: LEPR Protein Haplotype Sequences **435**

List of Figures

Figure 1.1.1: Graphical illustration of the different biological and environmental influences that are involved in the development of obesity.. 2
Figure 1.7.1a: Graphic representation of the location of six T2D susceptibility genes and the SNPs which have been significantly associated with fetal phenotype 66
Figure 1.7.1b: Graphic representation of the location of six of the T2D susceptibility genes, and the position of the SNPs that have been significantly associated with fetal phenotype 67
Figure 1.9.1: The leptin – melanocortin pathway.. 77
Figure 1.9.2: Image depicting six of the leptin receptor isoformsand their location across the plasma membrane 82
Figure 1.9.3: A graphic representation of the extracellular domain of the leptin receptor. 85
Figure 1.9.4: Graphical representation of the hexameric leptin and leptin receptor (LR) complex.. 86
Figure 1.9.5: The complex of the CRH2 domain and mouse leptin shown as ribbon model for CRH2 and leptin molecule.. 87
Figure 1.9.6: Secondary structure ribbon model of the Ig-like domain 88
Figure 1.9.7: Activation of the JAK/STAT pathways by leptin binding to its receptor.. ...91
Figure 1.9.8: A schematic diagram of the long isoform of the leptin receptor and the correspondence of the exons with the protein domain structure.. 95
Figure 1.9.9: Graphic representation of the location of the genes *LEPR* and *LEPROT* on chromosome 1p31 and also the position of the three SNPs of interest that have been significantly associated with obesity 97
Figure 1.10.1: Graphic representation of the location of the *HHEX*-IDE region.. 103

Figure 2.4.1: Graphic representation of the long isoform of the leptin receptor, showing the three domains, along with the five domains within the extracellular region and the three SNPs of interest.....130
Figure 2.7.1: Diagrammatic representation of the TaqMan SNP genotyping assay.. 163
Figure 2.8.1: Homepage of the UniProt webpage.....168
Figure 2.8.2: Search results for leptin receptor.. 169
Figure 2.8.3: I-TASSER homepage.. 171
Figure 3.1.1: Flow chart of the LEPR functional study displaying the twelve stages to the investigation. 173
Figure 3.1.2: A Magnesium and temperature PCR optimisation of LEPR.. 176
Figure 3.1.3: 1% (w/v) agarose gel electrophoresis of leptin receptor PCR performed at the optimised conditions.. 178
Figure 3.1.4: 1% (w/v) agarose gel electrophoresis of leptin receptor PCR at the optimised conditions.. 179
Figure 3.1.5: Diagrammatic representation of the pGEM-T vector, into which the LEPR amplicon was inserted.. 182
Figure 3.1.6: 1% (w/v) agarose gel electrophoresis of ligation between LEPR and pGEM-T vector.. 184
Figure 3.1.7: Diagrammatic representation of the ligation of the LEPR insert into the pGEM-T vector. 187
Figure 3.1.8: 1% (w/v) agarose gel electrophoresis of double restriction digest with the enzymes *BamHI* and *SalI*, of seven mini preparations, with undigested mini preparations alongside.....188
Figure 3.1.9: 1% (w/v) agarose gel electrophoresis of LEPR PCR at the optimised conditions using gel extracts from the four mini preparations.....189
Figure 3.1.10: 1% (w/v) agarose gel electrophoresis of double restriction digests of mini preparations with *BamHI* and *SalI*.. 190
Figure 3.1.12: 1% (w/v) agarose gel electrophoresis of five undigested mini preparations and the corresponding double restriction digests of these five mini preparations.. ...194
Figure 3.1.13: 1% (w/v) agarose gel of double restriction digest of mini preparation numbered nineteen.. 195
Figure 3.1.14: 1% (w/v) agarose gel of the double restriction digest of mini preparation 25 µl was loaded into each lane. 196
Figure 3.1.15: 1% (w/v) agarose gel electrophoresis of LEPR PCR from gel extracts of mini preparations nineteen and twenty.. 197
Figure 3.1.16: 1% (w/v) agarose gel electrophoresis of double restriction digest of mini preparations from four different cultures.. 199
Figure 3.1.17: 1% (w/v) agarose gel electrophoresis of double restriction digest of culture 26 with the enzymes *BamHI* and *SalI*.. 201
Figure 3.1.18: 1% (w/v) agarose gel electrophoresis of the LEPR gel extract from double digest of culture 26.. 202
Figure 3.1.19: Diagrammatic representation of the pCMV-Tag 2B expression vector.. ..205
Figure 3.1.20: 1% (w/v) agarose gel electrophoresis of control and experimental ligations between pCMV-Tag 2B expression vector and LEPR insert.. 206
Figure 3.1.21: 1% (w/v) agarose gel electrophoresis of LEPR PCR using the original LEPR primers from different dilutions of the nested RT-PCR.....209
Figure 3.1.22: 1% (w/v) agarose gel electrophoresis of restriction double digests (*BamHI* and *SalI*) of six mini preparations, with undigested mini preparations loaded adjacently.. 210

Figure 3.2.1: I-TASSER results showing the submitted sequence in FASTA format and the predicted secondary sequence for the complete native protein sequence of LEPR. 219
Figure 3.2.2: A part of the predicted secondary structure from amino acid positions 138 to 227 each of the four models. A) Model A; B) Model B; C) Model C and D) Model D. 221
Figure 3.2.3: The I-TASSER predicted five models for the leptin receptor with the Lys656Asn polymorphism.. 223
Figure 3.2.4: Predicted protein structures of model 1 from four models. 224
Figure 3.2.5: A exert of the I-TASSER analysis showing ten templates located by the I-TASSER server, from the protein database.. 226
Figure 3.2.6: I-TASSER results showing the ten proteins that are structurally closest to the first model predicted by I-TASSER, using TM-align.. 228
Figure 3.2.7: Predicted secondary protein structure of the four models: A) = E, B) = F, C) = G and D) = H.. 231
Figure 3.2.8: Predicted protein structures of 'model one' for the four models E, F, G and H.. 232
Figure 3.2.9: The five best models predicted by I-TASSER of the CRH1 domain. 234
Figure 3.2.10: CHR1Q (Blue) CHR1R (Green) predicted protein structure as a ribbon model showing three beta sheet formations. 236
Figure 3.2.11: Alignment of CRH1Q (blue) and CRH1R (red) as a ball and stick model (A), improved alignment of the two larger beta sheet structures within the domain (B).. 238
Figure 3.2.12: A close up view of the alignment between the smaller beta sheet structure (amino acids 1-60) in CRH1Q (blue) and CRH1R (red), shown as a ball and stick model.. 239
Figure 3.2.13: A section of the predicted secondary structure of the CRH1 domain with the glutamine residue at 223 (A) (CRH1Q) and the arginine residue (B) (CRH1R), from amino acid positions 133 to 220 (corresponding to 112 to 159 in the complete long isoform of the leptin receptor).. 240
Figure 3.2.14: A close up of the hydrogen bond between the arginine residue at position 162 (highlighted in yellow) and the cysteine residue at position 151 (highlighted in blue), within CHR1R, as a ball and stick model. 242
Figure 3.2.15: Electrostatic potential of the CRH1Q predicted domain, mapped to the surface of the protein.. 245
Figure 3.2.16: The overall surface electrostatic potential of the CRH1R peptide.. 247
Figure 3.2.17: A close up of the surface electrostatic potential of amino acid residue 162 and surrounding residues in CRH1Q peptide, as indicated by the arrows.. 248
Figure 3.2.18: A close up of the electrostatic potential of the arginine residue at position 162 in the CHR1R peptide as indicated by the arrows.. 249
Figure 3.2.19: CRH2 domain predicted protein structure depicted as a ribbon model showing two anti-parallel beta pleated sheet structures. 251
Figure 3.2.20: The alignment of the two CRH peptide domains, CRH1 (Red) and CRH2 (blue), shown as a ball and stick model. 252
Figure 3.2.21: Electrostatic potential mapped to the surface of the CRH2 domain.. 254
Figure 3.2.22: The overall surface electrostatic potential of CRH2 domain.. 256
Figure 3.2.23: The positive surface electrostatic potential of leptin protein.. 258
Figure 4.1.1: A map of R^2 across the *HHEX* region of interest using HapMap CEU data.. 263

Figure 4.1.2: A graphic representation of the location of the *HHEX* gene and the position of the thirty-seven SNPs genotyped in the multiplex assays indicated by the circle alongside the SNP ID.. 265
Figure 4.1.4: Screen shot of assay one plate one of the Typer software used for the genotyping analysis.. 268
Figure 4.1.5: Map of R^2 using the twenty-two SNPs from this study....276
Figure 4.2.1: Graphic representation of the location of the *FTO* gene on chromosome 16q12.3 and the positon of the nine SNPs which were genotyped in the multiplex assay. 278
Figure 4.2.2: Graphic representation of the mean birth weights for the SNPs rs17817449 and rs3751812.. 286
Figure 4.2.3: Graphical representation of the mean birth weights for the SNPs rs8050136 and rs9939609.. 287
Figure 4.2.4: Graphical representation of the mean birth weights for each of the number of risk alleles for the SNPs rs9939973 and rs9940128. 288
Figure 4.2.5: Graphical representation of the mean birth weights for the number of risk alleles for the SNP rs1421085....289
Figure 4.3.1: Results of the TaqMan SNP genotyping assay for plate two, represented in plots, tables or grids using the SDS software....295
Figure 4.3.2: Allelic discrimination plot of Plate one....296
Figure 4.3.3: Allelic discrimination plot of plate two....297
Figure 5.2.1: A ball and stick model of the CRH1 domain with the Gln^{223} amino acid shown at position 162 (Blue arrow).. 307
Figure 5.2.2: Map of R^2 across the *LEPR* gene (covering 210 kb) using HapMap CEU data.. 308
Figure 5.3.1: A map of R^2 across part of the *FTO* region, 52300 kb-52385 kb using HapMap CEU data, which shows one haplotype block. 327
Figure 5.3.2: A map of R^2 across the *FTO* region, 52560kb-52600 kb using HapMap CEU data, which shows one haplotype block. 328

List of Tables

Table 1.1.1: International classification of body weight using BMI....4
Table 1.2.2: Criteria from the IDF for defining the metabolic syndrome in adults. 8
Table 1.5.1: A summary of the key genes that have been associated with T2D and common obesity in adults and children and in different populations. 32
Table 1.7.1: Associations with low birth weight at term and subsequent disease development in adulthood have been found in fifteen different population groups from around the world.. 43
Table 1.7.2: Associations with low birth weight at term and subsequent disease development in late childhood have been found in seventeen different population groups, around the world....44
Table 1.7.3: Associations with low birth weight and subsequent disease development in early childhood have been found in seven different population groups around the world. 44
Table 1.7.4: A summary of the effects on body composition of three different famines that have occurred at different times over the last sixty-nine years from three continents.49
Table 1.7.5: List of SNPs in key genes that have been significantly associated ($p<0.05$) with fetal phenotype at birth in European Caucasian populations.. 64
Table 2.4.1: Checklist for the RT-PCR 129
Table 2.4.2: Primer Sequences for *LEPR* and *G6PD* and the expected product sizes.....129

Table 2.4.3: Standard PCR conditions for three different Taq polymerases.. 133
Table 2.4.4: Volumes of 25 mM stock Magnesium Chloride concentration.. 135
Table 2.4.6: Ligation reaction of the PCR product and pCMV Tag 2B expression vector. 145
Table 2.5.1: Checklist for the initial PCR 151
Table 2.5.2: PCR cycle conditions for the initial reaction 152
Table 2.5.3: SAP Clean up method 153
Table 2.5.4: Extension reaction primer mix for assay one 155
Table 2.5.6: iPLEX reaction mix for the 20-PLEX assay (A) and for the 17-PLEX assay (B). 157
Table 2.6.1: Extension reaction primer mix for the *FTO* assay. 161
Table 2.7.1: Master mix volumes and reagents for the TaqMan genotyping assay. 165
Table 2.7.2: PCR conditions for the TaqMan SNP Genotyping Assay 165
Table 3.2.1: The eight haplotypes within the extracellular region of the *LEPR* defined by the three specific polymorphisms at Lys109Arg, Gln223Arg and Lys656Asn. 213
Table 3.2.2: Results from the I-TASSER protein prediction. C-score and TM-scores indicating the best model are given for the 'first model' for the full-length native LEPR protein sequence and for the mutations, ^{109}Arg, ^{223}Arg and ^{656}Asn. 218
Table 3.2.3: I-TASSER results for the extracellular region of the leptin receptor.. 230
Table 4.2.1: The thirty-one SNPs within the *HHEX* locus genotyped, using the Sequenom iPLEX assay.. 264
Table 4.1.2: Percentage genotyping success rates for each of the SNPs in Assay one. 270
Table 4.1.3: Percentage success rates for each SNP in assay two. 271
Table 4.1.4: Genotype success rates for the SNPs that were rerun from both assays 272
Table 4.1.5: Numbers of each genotype for each SNP and the corresponding Hardy-Weinberg values where p>0.05. 272
Table 4.1.6: ANCOVA analysis of the association of fetal birth weight with maternal SNP genotype, with the covariates parity, gestational age, fetal sex and maternal booking weight.. 274
Table 4.2.1: Genotyping combined success rates for both plates for the five runs of the FTO assay. 279
Table 4.2.2: Percentage success rates for the seven *FTO* SNPs after combining the data for the five runs. 281
Table 4.2.3: Mean fetal birth weight classified by pregnancy number. 283
Table 4.2.4: Mean fetal birth weight classified by maternal BMI. 283
Table 4.2.5: Mean birth weights for the seven *FTO* SNPs and the number of risk alleles for each of them and p values are given for those that were found to be significant ($p<0.05$).. 285
Table 4.2.6: The effect size of the nine SNPs on fetal birth weight with the covariates parity, maternal weight, gestational age and fetal sex and their p values (p=0.05) ...290
Table 4.2.7: Mean fetal birth weight grouped by genotype for each of the seven *FTO* SNPs and fetal sex 292
Table 4.3.1: Mean birth weight for the different genotypes of SNP rs11708067.. 299
Table 4.3.2: Linear regression results for the *ADCY5* SNP rs11708067 with the covariates, booking BMI, parity, gestational age and sex 300

1.0 Introduction

1.1.0 Complex disease/birth weight and obesity

Multifactoral disease is the interaction of both genetic loci and the environment, where several regions within the genome collaboratively combine with various environmental factors to increase susceptibility to a particular disease. Such a paradigm can be found in the genetics of birth weight, obesity and diabetes. For obesity, illustration of how the many different biological and environmental aspects influence each other and how they are ultimately involved in the regulation of energy is shown overleaf (Figure 1.1.1).

Birth weight is influenced by a variety of factors and has been shown to be a significant prognostic indicator for both adult obesity, Type 2 diabetes (T2D) and the metabolic syndrome, with lower birth weight, perhaps counter intuitively, predisposing to each of these (Stettler *et al.,* 2005, Corvalan *et al.,* 2007, Xiao *et al.,* 2008, Risnes *et al.,* 2009). Furthermore, obesity itself predisposes to a variety of different conditions including hyperinsulinaemia, hypertension, dyslipidaemia, T2D, cancer and cardiovascular disease with a consequent reduction in life expectancy (Must *et al.,* 1999, Kopelman and Formiguera, 1999, Wilson *et al.,* 2003). Metabolic syndrome is used to describe a collection of metabolic complications which includes insulin resistance, dyslipidaemia, hypertension and central obesity (Grundy, 2006). Thus, obesity-associated diseases are the cause of death for hundreds of thousands of individuals around the world; over 9,000 premature deaths each year in England alone are attributable to obesity and its sequelae (Department of Health, 2009). This will certainly rise as the numbers of obese individuals follows the predicted increases around the world (Yang *et al.,* 2012, Shaw *et al.,* 2010, Kelly *et al.,* 2008).

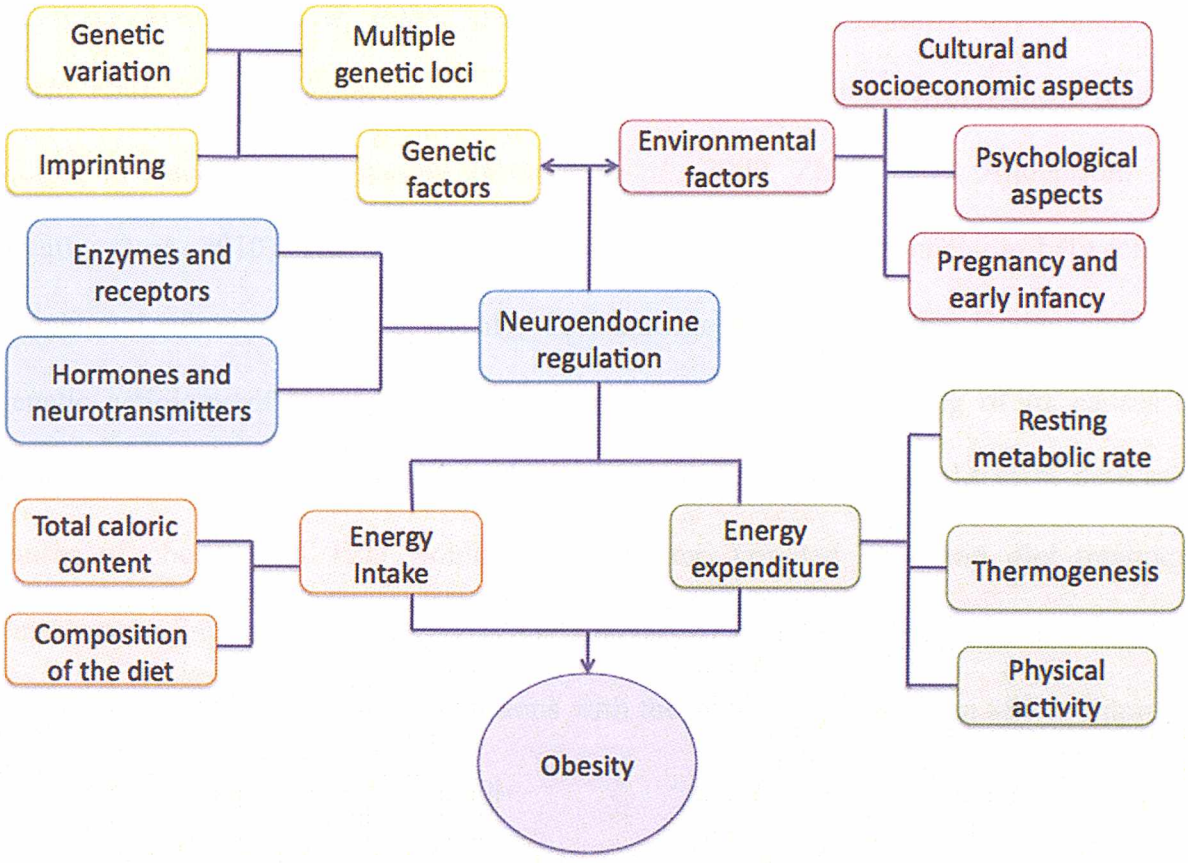


Figure 1.1.1: Graphical illustration of the different biological and environmental influences that are involved in the development of obesity. Adapted from (Balaban and Silva, 2004).

Over the last two decades, many genes have been implicated in the development of obesity and T2D (Rankinen *et al.,* 2006, Sladek *et al.,* 2007, Saxena *et al.,* 2007, Dupuis *et al.,* 2010), while recent research has correlated several T2D risk alleles with reduced fetal birth weight in Caucasian populations (Freathy *et al.,* 2009, Zhao, 2009, Andersson, 2010, Freathy *et al.,* 2010). This indicates that there is an underlying connection between birth weight and subsequent risk of obesity, T2D and metabolic syndrome. Understanding the genetic aetiology of birth weight would enable better understanding of its causes and therefore its subsequent influence in the development of disease risk. Identification of the causes of obesity and T2D would provide better, targeted drug and diet treatments. Furthermore, increased knowledge of the genetics of these complex diseases/traits would aid our understanding of their interactions with the environment, with an aim to improving our currently obesogenic environment.

1.2.0 Classification and Measurement of Complex Disease

Classification of obesity, T2D and the metabolic syndrome are important to be able to determine the worldwide prevalence rates, for both adults and children.

1.2.1 Classification of Obesity in Adults

Obesity is defined as an excess storage of adipose tissue to abnormal or excessive levels that can impair health (Spiegelman and Flier, 2001). Classification of obesity and adiposity is typically measured using body mass index (BMI), which can be used to determine the risk of excess weight on mortality and morbidity and to assess the relative body adipose tissue levels. BMI is calculated by dividing the weight in kilograms by the height in meters, squared (kg/m^2) (www.who.int). Individuals are classified into six main categories

(Table 1.1.1) ranging from underweight to super obese, two of which are sub-divided into additional categories depending on the severity of being underweight or obese. The normal, desirable BMI is between 18.50 and 24.99.

Table 1.1.1: International classification of body weight using BMI. Adapted from World Health Organisation, 2009.

Classification		BMI (kg/m^2)
Underweight	Severe thinness	<16.00
	Moderate thinness	16.00 - 16.99
	Mild thinness	17.00 - 18.49
Normal Range		18.50 - 24.99
Overweight		25.00 - 29.99
Obese	Class I	30.00 - 34.99
	Class II	35.00 - 39.99
Morbidly Obese		40.00 - 49.99
Super Obese		>50.00

1.2.2 Classification and Measurement of Obesity in Children

The measurement of overweight and obesity in children is also carried out using BMI although it is problematic as children continue to grow throughout childhood. Therefore, BMI adjusted for age and sex growth charts are used as a reference to monitor growth and to determine whether a child is overweight or obese. Various growth charts have been devised and are used. In the USA the Centers for Disease Control and Prevention produced growth charts that use data from NHANES I, II and III. These are known as the CDC 2000 growth charts. A BMI between the 85th percentile and 95th percentile is the cut-off value used for overweight and more than the 95th percentile is the cut-off value for obesity typically used in the USA (Ogden *et al.*, 2012a). In the UK the International Obesity Task

Force (IOTF) cut-off values are typically used to measure overweight and obesity in children. The cut-off values represent cut off points in which the percentiles match the adult cut-offs for an overweight BMI (25 kg/m^2) and an obese BMI (30 kg/m^2) at aged eighteen years (Cole *et al.,* 2000). It uses data published from six different countries: Brazil, Great Britain, Hong Kong, the Netherlands, Singapore and the USA, from children between six and eighteen years of age (Cole *et al.,* 2000). The WHO has devised a similar method for an up to date growth reference chart, which uses BMI for age (five to nineteen years) using percentiles that match to the adult cut-offs for overweight and obese (WHO). Overweight is > +1 standard deviation (SD) from the 50^{th} percentile, which is equal to the adult overweight cut-off. Obesity is >+2 SD from the 50^{th} percentile, which is equal to the adult obesity. However, reference groups are dependent on the population being studied, as under and over estimation of the prevalence of overweight phenotype and obesity can occur (Tuan *et al.,* 2012, Christoforidis *et al.,* 2011, Viner *et al.,* 2010, Monasta *et al.,* 2011, Khadilkar *et al.,* 2011). Therefore it is best to use more than one reference growth chart, corrected for gender, to determine the prevalence of overweight phenotype and obesity in the selected population (Hermanussen *et al.,* 2012).

1.2.3 Identification of T2D in adults

The identification of T2D is based upon certain criteria and current guidelines are given by WHO (2006). Signs and symptoms of T2D include unexplained weight loss, polyuria, polydipsia and polyphagia along with testing blood glucose levels under different conditions, after a twenty-four hour fast (fasting plasma glucose level) and two hours after a 75 g oral glucose load (two-hour plasma glucose level) and comparing these levels with normal values. A diagnosis of T2D is given with a fasting plasma glucose level ≥7.0 mmol/l and a 2-hour plasma glucose ≥11.1 mmol/l. A normal fasting glucose level is <6.0

mmol/l and a normal 2-hour plasma glucose is <7.8 mmol/l. Additional risk factors for the development of T2D are a BMI classification of >25 kg/m^2 and being more than forty years of age.

In addition a pre-diabetes state can be diagnosed, in which the blood glucose levels are higher than the normal levels but not high enough to be classified as T2D, but these can lead to T2D if they remain high. It is given by the terms impaired glucose tolerance (IGT) and impaired fasting glucose (IFG). IGT is defined as fasting plasma glucose <7.0 mmol/l and a 2-hour plasma glucose ≥7.8-11.1 mmol/l. IFG is defined as a fasting plasma glucose 6.1-6.9 mmol/l and a 2-hour plasma glucose <7.8 mmol/l.

1.2.4 T2D identification in children

The same criteria for testing T2D in adults are also used to diagnose T2D in children. The American Diabetes Association gives the following guidelines: a causal plasma level of glucose ≥11.1mmol/l or a fasting glucose level ≥7.0mmol/l or a 2-hour glucose level of ≥11.1 mmol/l. Additional criteria for testing children with T2D include a BMI >85^{th} percentile, >ten years of age and a family history of T2D.

1.2.5 Identification of the Metabolic Syndrome

There are many risk factors that contribute to the metabolic syndrome and many organisations have developed criteria for defining it, including WHO, the European Group for the study of Insulin Resistance (EGIR) and the National Cholesterol Education Program – third Adult Treatment Panel (NCEP ATP III). WHO and EGIR identify that insulin resistance is an important risk factor, but not NCEP ATP III and WHO defines

insulin resistance using IGT or diabetes whereas EGIR uses fasting insulin levels and impaired fasting glucose (Alberti *et al.,* 2006). Between them there are also differences in cut-off values for the other risk factors: blood pressure, HDL cholesterol, plasma triglycerides and the best measure of central obesity (Table 1.2.1).

Since these three groups vary on the criteria used for the metabolic syndrome the International Diabetes Federation (IDF) developed its own criteria (Alberti *et al.,* 2006). The main difference is in the measurement of central obesity, which it suggests should be measured using waist circumference and the cut-offs should be population and ethnic specific as there are differences in adiposity levels and classification of overweight and obesity. Classification of the metabolic syndrome is shown in Table 1.2.2.

Table 1.2.1: The following criteria is used for the definition of the metabolic syndrome, which has been developed by WHO, EGIR and NCEP ATP III. Adapted from Alberti *et al.*, 2006. Obesity is measured by, waist to hip ratio (WHR), waist circumference (WC) and BMI.

Risk Factor	**WHO**		**EGIR**		**NCEP ATP III**	
Insulin Resistance	Glucose Tolerance (IGT) or diabetes		Hyperinsulinaemia (Insulin fasting levels, non-diabetic)			
	Two or more of the following risk factors		*Two or more of the following risk factors*		*Three of more of the following risk factors*	
Fasting Plasma Glucose			≥6.1 mmol/l (110 mg/dl) non-diabetic		≥5.6 mmol/l (100 mg/dl)	
Blood Pressure	≥140/90 mmHg		≥140/90 mmHg or treatment			
Raised Plasma Triglycerides	≥1.7 mmol/l (150 mg/dl)		>2.0 mmol/l (178 mg/dl) or treatment		≥1.7 mmol/l (150 mg/dl)	
HDL Cholesterol	**Men**	**Women**	<1.0 mmol/dl (39 mg/dl)		**Men**	**Women**
	<0.9 mmol/l (35 mg/dl)	<1.0 mmol/l (39 mg/dl)			<1.03 mmol/l (40 mg/dl)	<1.29 mmol/l (50 mg/dl)
Obesity	**Men**	**Women**	**Men**	**Women**	**Men**	**Women**
WHR	>0.9	>0.85				
WC			≥94 cm	≥80 cm	>102 cm	>88 cm
And/or BMI	>30 kg/m^2					
Microalbuminuria	Urinary albumin excretion rate >20 µg/min or albumin: creatinine ratio ≥ 30 mg/g					

Table 1.2.2: Criteria from the IDF for defining the metabolic syndrome in adults.

Risk Factor	**Measurement and Cut-off Values**	
Central obesity	WC that is ethnic specific	
	With two of the following risk factors	
Raised Plasma Triglycerides	≥1.7 mmol/l (150 mg/dl)	
HDL Cholesterol	**Men**	**Women**
	<1.03 mmol/dl (40 mg/dl)	<1.29 mmol/dl (50 mg/dl)
Raised Blood Pressure		
Systolic or	≥130 mmHg	
Diastolic	≥85 mmHg	
	Or previous diagnosis of hypertension	
Raised Fasting Plasma Glucose	≥5.6 mmol/l (100 mg/dl)	
	Or previous diagnosis of T2D	

1.2.6 Reliability of BMI in Determining Disease Risk

BMI has been used as a measure of adiposity in many studies. High BMI has been associated with increased risk of T2D, hypertension, cancer, mortality and cardiovascular disease (Calle *et al.,* 1999, Ramachandran *et al.,* 2001a, Calle *et al.,* 2003, Berraho *et al.,* 2012, Park and Kim, 2012, Staiano *et al.,* 2012). BMI is one of the commonest and simplest methods to use and its reliability and specificity to evaluate disease risk and assess relative body fat, has been investigated over the past ten years (Dalton *et al.,* 2003, Kontogianni *et al.,* 2005, Nevill *et al.,* 2006, Shah and Braverman, 2012).

In a recent review, a total of thirty-two studies were used in a meta-analysis that looked at sensitivity (the probability that an individual who already has the condition of interest will produce a positive result), specificity (the probability that an individual who does not have the condition will produce a negative result) and likelihood ratio (the odds that the test result occurs in those who have the condition against those who do not) of BMI. It was found that although the specificity was good, the sensitivity of BMI was only 50%, indicating that individuals who are obese are not classified as such, using the current cut off values. But the sensitivity increases when BMI is >30 kg/m^2 where excess adiposity can be detectable in both sexes. This suggests that excess levels of adiposity in individuals whose BMI fall below 30 kg/m^2 are not detected. Therefore BMI is a better method at detecting excess adiposity when BMI is >30 kg/m^2 as compared to <30 kg/m^2. BMI is a good method at evaluating changes in body adiposity over a period of time as rises in body weight are probably due to increases in adipose levels (Okorodudu *et al.,* 2010).

Furthermore, the sensitivity of BMI affects athletes and body builders as they are normally categorised as overweight or obese, when they are not, as they have a higher muscle mass

compared to non-athletes (Witt, 2005, Nevill *et al.,* 2006, Ode *et al.,* 2007). As muscle mass weighs more than adipose tissue, BMI is unable to distinguish between the two as total body weight is used in the calculation. Therefore other measures have been developed to better assess the level of adiposity and disease risk. Simple methods include waist circumference, percentage body fat, fat free mass, the ratio of waist to hip circumference (WHR), waist to height ratio (WHtR) and skin fold thickness as well as more sophisticated methods including air displacement plethysmography, magnetic resonance imaging (MRI) and dual energy X–ray absorptiometry (DEXA)(Goodpaster *et al.,* 2003, Goodpaster *et al.,* 2005, Wang *et al.,* 2005, Vega *et al.,* 2006, Cheung, 2007, Pischon *et al.,* 2008). The classification of waist circumference indicates that females between 81-88 cm (32-35 ins) and males who are between 94-102 cm (37-40 ins) are overweight. Individuals with values more than those stated would be classified as obese. For WHR, men would be classified as obese with a value more than 0.9 and 0.85 for women (WHO). More recently, waist to height ratio is increasingly being used instead as it is a relatively simple non-invasive method to use as for a given height there is an adequate level of stored fat. A boundary value of 0.500 to indicate whether the amount of fat is in excess and can lead to health complications. This has been further developed into the Ashwell © Shape chart (Ashwell, 1998). Therefore it has been suggested to use other methods such as waist circumference and/or WHtR alongside BMI, as a better indication of adiposity and disease risk.

However this has only increased the debate as to which method is more accurate and better to use. Waist circumference and WHR are thought to be the better methods to use to evaluate disease risk with respect to central adiposity (Vega *et al.,* 2006; Cheung *et al.,* 2007) as they are still simple and easy to use and correlate highly to the development of obesity-related diseases (T2D, cardiovascular disease, hypertension and risk factors for the

metabolic syndrome) (Goodpaster *et al.,* 2003, Snijder, 2003, Grundy, 2004, Goodpaster *et al.,* 2005, Wang *et al.,* 2005, Meisinger *et al.,* 2006, Bobbioni-Harsch *et al.,* 2009, Ashwell *et al.,* 2012, Park and Kim, 2012). However abdominal obesity can arise from either an increase in subcutaneous adipose tissue (SAT, i.e. under the skin), visceral adipose tissue (VAT, i.e. surrounding organs in the abdominal region including the liver, stomach and intestines) or a combination of both (Freedland, 2004). Thus WC and WHR are unable to distinguish between the two locations, they only state that the fat is within a central location (Goodpaster *et al.,* 2005). More high-tech methods, such as MRI are more accurate at distinguishing between these two locations, although they are more expensive and time consuming to use and not easily available (Browning *et al.,* 2011).

The actual location of excess adipose tissue is important to the health of an individual. Research suggests that internal locations of adipose tissue, i.e. surrounding organs such as the liver, stomach and intestines is more detrimental to the health of an individual compared to a subcutaneous location (Gastaldelli *et al.,* 2009) and contribute to the development of insulin resistance and the metabolic syndrome (Pouliot *et al.,* 1992, Ross *et al.,* 2002a, Ross *et al.,* 2002b, Pickhardt *et al.,* 2012, Barreira *et al.,* 2012).

Several studies have been carried out to compare the VAT in individuals within different BMI categories and taking into consideration levels of physical activity with surprising results. In one study, fifty-four females were classified into four groups based on their BMI (Thomas *et al.,* 1998a). When MRI was used as a measure of body fat it showed that as BMI increases, so does the amount of visceral adipose tissue however, it was noted that some individuals who are within the normal BMI range had just as high a fat content as those who are obese (Thomas *et al.,* 1998). This indicates that individuals within the

normal BMI range can be overweight or even obese internally and are more at risk of developing metabolic complications associated with obesity.

This finding has led to the term 'TOFI' – thin on the outside, fat on the inside to describe these individuals. A further term 'FOTI' has also been coined to characterise those who are fat on the outside and thin on the inside (Thomas *et al.,* 2012). Research has shown that having a BMI above the normal range but being physically active can result in a lower level of visceral fat and therefore decrease the risk of T2D and CHD and mortality, when compared to those who are overweight or obese and unfit (Stevens *et al.,* 2002, Hu *et al.,* 2004b, Hu *et al.,* 2004a). A comparison of fat at different locations (total, subcutaneous, subcutaneous abdominal, internal, visceral, non-visceral), was undertaken between fifty men who were classified as slim-fit, slim-unfit, fat-fit and fat-unfit using waist girth, BMI and body mass (O'Donovan *et al.,* 2009). Results showed that although internal and visceral fat and the proportion of internal fat were higher in fat-fit and highest in fat-unfit men, there was not any significant differences between the slim-fit, slim-unfit and the fat-fit, for either total, subcutaneous, subcutaneous abdominal, internal and visceral fat after adjustment for age and height ($p<0.05$). There were also no significant differences between the slim-fit and slim-unfit for waist girth, total fat levels, internal fat and visceral fat levels (O'Donovan *et al.,* 2009).

The exact mechanisms by which visceral adiposity contributes to metabolic consequences of obesity are unclear as not all obese individuals are insulin resistant or develop T2D or cardiovascular disease (Despres and Lemieux, 2006). Visceral fat has been shown to be different from subcutaneous fat in terms of lipolytic activity, response to insulin and endocrine function, which may contribute to the increased metabolic risks. Intra-abdominal

adipocytes are enlarged and hyperlipolytic which makes them resistant to the antilipolytic effect of insulin. Non-essential fatty acid (NEFA) moves to the liver, which may affect liver metabolism, leading to increased glucose production and hepatic insulin resistance (Mauriege *et al.,* 1995, Mittelman *et al.,* 2002). A proinflammatory state is induced in an obese individual with high abdominal adiposity. Adipose tissue releases a variety of inflammatory cytokines including interleukin-6 and tumour necrosis factor-α, along with the hormone adiponectin, which is involved with regulation of glucose and fatty acid metabolism (Cote *et al.,* 2005, Greenberg and Obin, 2006, Hajer *et al.,* 2008). High levels of macrophages occur within the adipose tissue (Weisberg *et al.,* 2003), which in turn releases plasma C-reactive protein. In obese individuals with abdominal obesity, levels of interleukin-6 (IL-6) and tumour necrosis factor-α (TNF-α) are increased along with the plasma C-reactive inflammatory marker but adiponectin levels decrease (Yudkin *et al.,* 1999, Berg and Scherer, 2005). These changes could increase the altered metabolic state, however further research is needed to fully understand endocrine function within visceral adiposity (Després and Lemieux, 2006).

The development of visceral adiposity has been proposed to be due to a dysfunction of the subcutaneous adipose tissue to store excess fat. As a consequence an alternative location is used which is undesirable as it affects the functionality of the organs it surrounds, which leads to the altered metabolic state (Després and Lemieux, 2006). Improvements in the metabolic state have been demonstrated through reduction in intra-abdominal adipose tissue through diet and exercise but not through reduction in subcutaneous tissue (Uusitupa *et al.,* 2003, Langendonk *et al.,* 2006, Shojaee-Moradie *et al.,* 2007). This further supports the view that visceral adiposity results from dysfunctionality in storage of excess fat.

1.2.7 Reliability of BMI in Determining Disease Risk in Children and Adolescents

Measurement of childhood obesity in determining disease risk is problematic. As with adults, the BMI classification method has been used to measure overweight and obesity but drawbacks remain (Sabin and Shield, 2008). As with adults, the distribution of excess adipose tissue is important and the deposition can occur subcutaneously and intra-abdominally within growing children. It is becoming clear that central localisation of adipose tissue has similar health complications as with adults, although further research is needed. Previous work has shown that centrally located adipose tissue in children aged between five to seventeen years was associated with adverse concentrations of LDL and HDL cholesterol, insulin and triacylglycerol (Freedman *et al.,* 1999). Further work shows a relationship with adverse atherogenic lipoprotein profile and an increasing waist circumference (Flodmark *et al.,* 1994). Therefore, to provide a better measurement of overweight and obesity in children, waist circumference and BMI was combined and age-related waist circumference percentile curves have been developed for British children, which follows a similar method in other countries (McCarthy, 2001). As WHtR is increasingly used with adults to determine disease risk in relation to central adiposity the method is being adopted to use as a better indication with children and adolescents (McCarthy and Ashwell, 2006, McCarthy *et al.,* 2006).

1.2.8 BMI and Different Ethnic Minorities and Disease Risk

Particular ethnicities are at increased risk of obesity and the metabolic consequences that arise as a result; therefore adequate assessment to determine disease risk is also of importance within these populations. One such population is the Asian population, who develop T2D when they are classified as overweight (>25 kg/m^2) rather than obese (>30

kg/m^2) using BMI, when compared with European populations (Yoon *et al.,* 2006). A significant factor within this is that South Asians are known to have a different body composition, which is characterised by the deposition of adipose tissue in a central location and a lower muscle mass (Banerji *et al.,* 1999, Chandalia *et al.,* 2007), which has been termed the Indian 'thin-fat' phenotype. As a result, a lower threshold has been developed to classify adiposity and disease risk for BMI. A value of >23 kg/m^2 indicates a moderate to high risk and >27 kg/m^2 indicates a high to very high risk (WHO, 2004). However even though research has shown differences in the threshold for BMI and disease risk in other populations (Wang *et al.,* 1996, Deurenberg *et al.,* 1998, Lear *et al.,* 2007, Carroll *et al.,* 2008), alternative thresholds have not been agreed upon (Katzmarzyk *et al.,* 2011).

European thresholds for WC are currently used for all ethnicities, 80-88 cm for women and 94-102 cm for men however different thresholds may also need to be produced to determine disease risk. Data has shown that with BMI there are differences between ethnic groups. The value of WC that increases the likelihood to develop two or more cardiometabolic risk factors was compared in Caucasian and African-American men and women which were similar to the currently used cut-offs for African-American and Caucasian men but it was actually higher for Caucasian and African-American women, 4 cm and 9 cm respectively (Katzmarzyk *et al.,* 2011). Furthermore, Hispanic and black men and women were found to have lower WC to a corresponding overweight BMI classification compared to Caucasian men and women (Okosun *et al.,* 2000). Different thresholds for WC have been proposed to be population specific (Alberti *et al.,* 2009). More recently, a systematic review and meta-analysis has shown that WHtR is a better measure of adiposity to detect disease risk for T2D, cardiovascular disease and hypertension than BMI and WC in both sexes and for different ethnic groups (Ashwell *et*

al., 2012). Using more than one measure of adiposity could be preferential to determine disease risk than using just one method.

1.3.0 The importance of obesity, T2D and the metabolic syndrome in the 21st Century

Obesity, T2D and the metabolic syndrome are becoming a worldwide health issue. Prevalence rates for each of these diseases have increased rapidly over the last thirty years (Cripps and Ozanne, 2006, Finkelstein *et al.,* 2012, Ogden *et al.,* 2012b) in many Westernised countries including USA and United Kingdom.

1.3.1 Obesity in Adults and Children

In 2007, the Health Survey for England found that 41% of men and 32% of women were classified as overweight, with 24% of both men and women classified as obese. This makes over half the adult population either overweight or obese, (67% and 56% respectively). This has been a steady increase since 1980 when just 6% of men and 8% of women were overweight or obese. By 2050 the levels for obese men and women could be as high as 60% and 50% respectively (Foresight Report, 2007). Similar findings were found in USA as the prevalence of obesity among adults has doubled between 1980 and 2004 from 15% to 33% (Flegal *et al.,* 2002, Ogden *et al.,* 2006). It is estimated that by 2030, there will be 1.35 billion and 573 million adults overweight and obese worldwide respectively (Kelly *et al.,* 2008).

Tackling this problem will soon become a global issue as the cases of obesity in developing countries are escalating at an alarming rate, as they adopt a more Westernised lifestyle. Countries with high prevalence rates include China, India, Brazil, Algeria, Egypt

and Colombia (Yajnik, 2004, Wu, 2006, Mitchell *et al.,* 2012). Furthermore, this indicates that specific ethnic populations are at a higher propensity to develop obesity and its sequelae as they either adopt or migrate into obesogenic environment. There is a significant association between immigrants who have resided in the US for more than ten years and the increased risk of developing obesity (Goel *et al.,* 2004). African-American and Hispanic women have the greatest increase in the prevalence of obesity between the National Health and Nutrition Examination Survey (NHANES) carried out in 1988-1994 and 1999-2000 in USA (Flegal *et al.,* 2002, Ogden *et al.,* 2002). Similar results were found in adolescents where the greatest increase occurred in non-Hispanic blacks (13.4% to 23.6%) and Mexican Americans (13.8% to 23.4%) (Flegal *et al.,* 2002, Ogden *et al.,* 2002). South Asians are another ethnic population with a higher propensity to develop obesity through international migration (Landman and Cruickshank, 2001, Ramachandran *et al.,* 2001b, Barnett *et al.,* 2006, Misra and Khurana, 2009) but also via rural to urban relocation in India (Ebrahim *et al.,* 2010).

In addition, obesity is also important in young children and adolescents: with the prevalence rates increasing at a similarly alarming rate to adults (Ogden *et al.,* 2012a). Current predictions suggest that the total number of overweight and obese children will soon overtake the total number of overweight and obese adults. It is estimated that worldwide over forty-two million children under five years of age, are overweight (www.who.int). Data from the Health Survey for England 2007, found that 31% of boys and 30% of girls aged two and fifteen years were overweight and obese using the UK 1990 Growth Reference chart. These individuals are not only more likely to become obese adults (Bjorge *et al.,* 2008) but the higher levels of adipose tissue can have an impact on many of the physiological systems within the body, notably cardiovascular, endocrine, respiratory

and gastrointestinal (Han *et al.,* 2010) which can lead to hyperinsulinaemia, hypertension, and fatty liver disease in later life. Research has found that being obese or overweight between the ages of fourteen and nineteen is associated with increased mortality as an adult (Bjorge *et al.,* 2008). Further effects of childhood overweight and obesity include the beginning of atherogenic processes with the full development of coronary heart disease (CHD) occurring years later (Weiss and Caprio, 2005). Cardiovascular disease is associated with high BMI in children and adolescents. In adults, the progression from insulin resistance to T2D is between five and ten years. The occurrence of T2D at a much earlier age could be due to the accelerated development of insulin resistance, hyperinsulinemia and impaired glucose tolerance (Fagot-Campagna, 2000).

1.3.2 T2D in Adults and Children

With respect to T2D, the increased risk was higher for Asians, Hispanics and blacks compared to Caucasians in women after adjustment for age and BMI, the odds ratio was 2.26 (1.70-2.99), 1.86 (1.40-2.47) and 1.34 (1.12-1.61) respectively (Shai *et al.,* 2006). It was also found that for Asians in particular the association between increasing BMI and greater weight gain with the risk of T2D was more evident compared with the Hispanics and blacks in America (Shai *et al.,* 2006). Even though African Americans have the highest rates of coronary heart disease, the pattern of atherogenic plasma lipids is much lower which is an important risk factor in CHD. T2D as a result of obesity is highest in Mexican Americans (Crossrow, 2004).

With a wide variation in the development of obesity in different ethnic populations, further research is needed within these cohorts to understand the interactions between the environment and multiple genetic loci. However, currently, ethnic populations including

South Asians are still a relatively understudied population within this area, even though they have a higher propensity to develop obesity, T2D and cardiovascular disease in an obesogenic environment. Longitudinal studies in the UK such as the London Life Sciences Prospective Population Study (LOLIPOP) study are beginning to address this imbalance (Chambers *et al.,* 2000, Chambers *et al.,* 2001, Chambers *et al.,* 2008, Chambers *et al.,* 2009, Lindgren *et al.,* 2009, Zabaneh *et al.,* 2009).

1.4.0 Obesogenic Environment

The rapid rise in the occurrence of obesity has led to various questions about its origin, whether it is completely genetic or environmental. Clearly both are important to its development but determining how much is still to be discovered as within the obesity epidemic there are still individuals that are of normal weight and adipose levels (Maes *et al.,* 1997, Bulik and Allison, 2001, Carnell *et al.,* 2008). Thus, the contribution of the environment to the development of obesity is a significant factor as it has changed drastically in the last fifty years, compared to the human genome, which develops over thousands of years, through evolution.

Environmental influences thought to be important for the development of obesity include lifestyle, health, food and nutrition. These have changed since the end of rationing in the mid 1950's, when food started to become more abundant (Lustig, 2011). Nowadays, there is relatively easy access to food with a higher energy and fat content. In addition, there is the added problem of a more sedentary lifestyle, which has developed with respect to the work place and leisure time. These have become mainly office based and more time is now spent indoors watching TV and playing computer games respectively, along with an increase in the use of motorised transport and better control of indoor temperatures. This

environment, which has been termed obesogenic, has led to a high calorie daily intake along with a reduction in energy expenditure, resulting in the excess energy intake being stored as adipose tissue rather than utilised.

The development of the obesogenic environment has occurred in many Western countries such as USA, United Kingdom, Australia and those in Western Europe, which have high levels of overweight and obese adults and children (WHO). Correspondingly high numbers of overweight and obese individuals are now being seen in developing countries, such as those in Eastern Europe, South America and Asia, as their demographics change to become more urbanised, leading to a more obesogenic environment (WHO). China is one such country that has seen a rapid rise in obesity as it undergoes 'nutritional transition' (James, 2008). In the last fifteen years obesity rates have increased thirty-fold (Saules *et al.,* 2007). In 2007, there were 184 million people who were overweight and thirty-one million people obese (Wu, 2006, Reynolds *et al.,* 2007). This includes both adults and children who are also at risk of developing the metabolic syndrome (Ko and Chan, 2008). During this time, dramatic changes have been made to the physical environment, which has seen a reduction in the daily energy expenditure by around 300-400 kcal d^{-1}. Within the main cities, many improvements have been made, including a better motorway infrastructure, which has resulted in the numbers of people owning a car increasing by 500% within the last twenty years (Levine, 2007). This has resulted in a reduction in the use of the bicycle and contributed to a further reduction in the amount of energy used typically around 200 kcal d^{-1}. The introduction of home entertainment and computers contributes to an additional 400-600 kcal d^{-1} loss in energy expenditure (Abdulla *et al.,* 2008). Food consumption has increased substantially rather than decreasing as a result of the environmental changes and in particular the amount of fat and sugar in food has increased. In the early 1980s, 14% of

the daily intake of energy was from fat, however this increased to 27.3% and 32.8% in more urban areas by 1997 (Abdulla *et al.,* 2008). This is considerably higher than the guideline daily allowances of fat and is typically stored rather than utilised as the environment has changed.

1.4.1 Daily Expenditure of Energy

Energy is required on a daily basis for three different functions: basal metabolic rate, thermogenesis and activity thermogenesis (Levine, 2002). The energy needed for respiration and other core functions, at rest, is the basal metabolic rate (BMR). The energy required after a meal is the thermogenesis and there are two components within activity thermogenesis: exercise and non-exercise activity thermogenesis (NEAT). BMR and thermogenesis remain fairly constant and only account for 60% and 10% of an individual's daily requirement respectively, leaving 30% due to activity thermogenesis (Levine, 2002). This is quite a high percentage and can vary widely between individuals. The total daily energy expenditure between two adults of a similar size can be as high as 8368 kJ or 2000 calories, which makes it an important factor when considering the daily energy expenditure between obese and lean individuals (Levine and Kotz, 2005). NEAT is involved in the activity levels during occupation and leisure time, and can each vary widely in terms of the energy required and used. This could account for the differences in energy expenditure between obese and leaner individuals, as highly active people are known to expend more energy than less active people and to be thinner than those who are less active (Levine *et al.,* 2006). More weight or adipose is gained through over-feeding if there is no increase in NEAT and less adipose is gained if there is an increase in NEAT. Since modernisation, the demographic of the population has changed such that the majority of people spend most of their time sat throughout their day whether at work, commuting to and from home or

during their leisure time. This change, coupled with the abundance of high calorie foods, is providing a situation in which energy intake remains relatively high but energy output is declining, resulting in obesity. This situation has already occurred in many Western countries in which overweight and obesity is high within the population but as many undeveloped countries become more modern this problem will, and has, become an issue, in particular in Asian countries like India (Ramachandran *et al.,* 2001b, Ebrahim *et al.,* 2010) and China (Wu, 2006, Wang *et al.,* 2007).

Identification of these problem areas within the environment suggests routes by which they can be addressed and changed. Attempts have been made however, changing attitudes and lifestyles will take time. One of the problems is that high calorie and fatty foods are cheap and readily available, which makes them convenient, particularly in this 24-hour society, which has also changed over the decades leading to one that includes a high number of working hours per week. Those from a disadvantaged background, low-income families for example, are more likely to buy these types of foods and suffer from obesity and the related diseases it causes.

1.5.0 Genetics and Obesity

Although the environment has developed more recently into a more obesogenic one, obesity has always occurred, albeit to a lesser degree, both as a significant symptom of a number of syndromes and as a rare Mendelian single gene disorder, indicating the importance of genetics to its development. These have been studied to understand the development of common obesity and the different pathways that are involved in energy regulation.

1.5.1 Heritability of Obesity

In studies of complex disease, heritability is often calculated to help determine the genetic contribution made to the disease or trait. One method to measure heritability (h^2) is to use monozygotic (MZ) and dizygotic (DZ) twins, as they are 100% and 50% genetically identical, respectively. If a trait is in part genetic then there will be a greater similarity with MZ twins than with DZ twins. An estimate of the heritability is determined from twice the difference in correlation between MZ (rMZ) and DZ twins (rDZ): $h^2=2(rMZ-rDZ)$ (Lustig, 2011).

1.5.1.1 Heritability of BMI

Many twin, adoption and family studies have been carried out, since the late 1980s, 1990s and 2000s to estimate the genetic influence of obesity measures (Stunkard *et al.,* 1986a, Stunkard *et al.,* 1986b, Turula *et al.,* 1990, Segal *et al.,* 2009). The heritability estimate of BMI has been found to be between 50 and 90% (Maes *et al.,* 1997, Nan *et al.,* 2012). An early study by Stunkard *et al* (1986a), comprised of 1974 MZ and 2097 DZ adolescent and adult twins to investigate the heritability of BMI. The heritability estimate was 77% for adolescents and 84% for adults. Additionally there was a higher correlation for overweight with MZ twins than compared with DZ twins.

As the environment is an important factor in obesity, twin studies can be designed to distinguish between the contribution of genetics and the environment. Genetic effects are divided into additive and dominant. Environmental effects are divided into shared or common environment and unique or non-shared environment.

In a large study, two groups of MZ and DZ twins at different ages in childhood, seven years and ten years, were analysed to estimate the contributions of genetics and the environment to BMI (Haworth *et al.,* 2008b). The heritability estimates were 60% and 72% for age seven and ten years respectively. The contribution due to shared environment was much lower, 12-22%. Genetic and environmental influences on the mean difference between obese and normal weight children were also investigated using DeFries-Fulker extremes analysis (DeFries and Fulker, 1988). It was found that the genetic and environmental influences on obesity were also similar across the range of BMI, both quantitatively and qualitatively. This indicates that the same genes and environmental factors that contribute to the range of BMI are the same for obesity and suggests that obesity is simply an extreme of normal variation.

The influence of BMI through childhood into adulthood seems to be genetic. The heritability of BMI from the age of four to aged eleven increased, from 48% to 78% respectively (Haworth *et al.,* 2008a). The heritability estimate for BMI was 56-86% from early adulthood into middle age (Silventoinen *et al.,* 2007). In addition this study found that between the ages of one and eighteen years, a high amount of influence on weight itself was found to be genetic. The correlation of BMI at age eighteen continued to increase from the BMI at aged one until seventeen years, r=0.32 to r=0.91 respectively. Between 81-95% of this correlation is due to additive genetic factors, with a small amount due to unique environmental factors.

1.5.1.2 Heritability of Alternative Measures of Obesity

BMI is one measure of obesity and determining disease risk. The genetic influences of other measures of obesity, such as WC, WHR, skinfold thickness and weight, have also

been investigated using twin studies. These also show a high degree of heritability even when accounting for environmental influences (Bodurtha *et al.,* 1990, Allison *et al.,* 1994, Rose *et al.,* 1998, Nelson *et al.,* 1999, Schousboe *et al.,* 2004, Wardle *et al.,* 2008). WC and BMI showed a high heritability in children aged between eight and eleven years, 77% for each, despite the occurrence of the obesogenic environment (Wardle *et al.,* 2008). Furthermore the results also showed that there was a genetic influence on WC, which was independent of BMI (40%). 60% of the genetic influence on WC was also common to BMI. In an adult twin study similar results were also found, the genetic influences on WHR were independent of BMI (28% and 48%) and WC (46% and 66%) in males and females, respectively (Nelson *et al.,* 1999).

1.5.1.3 Heritability of Body Composition

Body composition has been previously shown to be an important risk factor in the development of obesity, T2D and the metabolic syndrome. Twin studies investigating the distribution of body fat have shown it to be under genetic control, when body fat distribution was measured by DEXA in young and older adults (Cardon *et al.,* 1994, Carey *et al.,* 1996, Malis *et al.,* 2005, Skidmore *et al.,* 2009) and additionally when body composition was split into different classifications based upon relative fitness and adiposity, termed somatotype (Song *et al.,* 1994, Peeters *et al.,* 2003, Reis *et al.,* 2007, Peeters *et al.,* 2007). There are three different somatotypes: endomorph (high fat deposits with a large waist), mesomorph (muscular with low adiposity, a small waist and large shoulders) and ectomorph (low adiposity with thin limbs and are slim). Each of these show a high degree of heritability in adolescence and adults and are also sex specific (Peeters *et al.,* 2003; Peeters *et al.,* 2007). Significant differences between MZ and DZ twins, aged

seven to nineteen years, have been found between the three somatotypes and for height but not for BMI and weight (Reis *et al.*, 2007)

Overall these different twin studies into the heritability of measures of obesity indicate that genetic influences account for a high proportion of the variation in BMI, body composition and body fat even in an obesogenic environment. However, further investigation is needed within twin studies to fully understand gene-environment interactions and the total genetic variance due to candidate genes (Silventoinen *et al.*, 2010).

1.5.1.4 Heritability of Eating Behaviour

Another aspect to consider in the heritability of obesity is eating behaviour as it is a significant modifier of the environmental influences and many candidate genes that have been identified are expressed in the central nervous system. As a result of this it has been proposed that obesity is a neurobehavioural disorder (O'Rahilly and Farooqi, 2008).

Many twin studies have been conducted to determine the heritability of eating behaviour, which includes eating styles, the responsiveness to food cues and satiety. Even though measurement of this phenotype is difficult as it is conducted using self-reported questionnaires, which makes them unreliable due to under-reporting, they do show that there are genetic influences to these aspects (Keski-Rahkonen *et al.*, 2007, Carnell and Wardle, 2008, Keskitalo *et al.*, 2008, Sung *et al.*, 2009).

One study into satiety and responsiveness to food cues in two groups of twins, three and five years of age and eight and ten years of age, found that children with a high BMI had a lower satiety responsiveness ($r=-0.19$ and $r=-0.22$, respectively) and higher food cue

responsiveness (r=0.18 in both groups). This was also the same for WC, r=-0.23 for satiety responsiveness and r=0.20 for higher food cue responsiveness (Carnell and Wardle, 2008). For eating rate, in another study it was found to be highly heritable (62%) and a faster eating rate was correlated with obesity in ten and twelve year old twins (Llewellyn *et al.*, 2008).

1.5.2 Monogenic Forms of Obesity

Within rare Mendelian single gene disorders, over 200 cases have been identified with mutations in one of eleven genes (Bell *et al.*, 2005a, Rankinen *et al.*, 2006, Mazen *et al.*, 2011, Buchbinder *et al.*, 2011, Hung *et al.*, 2012, Rouskas *et al.*, 2012, Saeed *et al.*, 2012). Of those that have been identified the genes for leptin (*LEP*), the leptin receptor (*LEPR*), pro-opiomelanocortin (*POMC*) and the melanocortin 4 receptor (*MC4R*), have been found to be involved in appetite regulation, via the leptin-melanocortin pathway (Bell *et al.*, 2005). This pathway identifies the hypothalamus as the hub of energy control regulating energy intake and expenditure. These Mendelian disorders are characterised by severe early onset, typically due to hyperphagia. Mutations in the *MC4R* gene are the most frequent autosomal dominant forms of monogenic obesity (Vaisse *et al.*, 1998, Yeo *et al.*, 1998), representing 1-6% of obesity in children and adults (Farooqi *et al.*, 2003). In cases of mutations within the leptin and leptin receptor genes, administration of recombinant forms of the gene product can be given and a reduction in the obese phenotype is seen (Farooqi *et al.*, 1999).

1.5.3 Syndromic Forms of Obesity

Of the twenty syndromes in which obesity is a common phenotype there are only four that present with severe hyperphagia, so unlike the Mendelian disorders, the cause of the increase in adiposity can arise from increased food consumption, decreased thermogenesis (in which brown adipose tissue is involved and higher levels in rodents have been shown to be resistant to obesity (Almind *et al.,* 2007, Guerra *et al.,* 1998, Vitali *et al.,* 2012)) and/or environmental factors (Bell *et al.,* 2005b). One such syndrome is Prader-Willi Syndrome (PWS). It is caused by the absence of the paternally expressed critical region on chromosome 15q11-13 (Sode-Carlsen *et al.,* 2009). It is a complex disorder with a variety of clinical features including muscular hypotonia, hypogonadism, short stature and developmental delay. Individuals with PWS gain weight within the first six years of life leading to morbid obesity. The body composition of PWS individuals FOTI, i.e. the excess adipose tissue is deposited predominantly in a subcutaneous location rather than viscerally when compared to an obese individual without PWS. There is also a higher percentage body fat content, with less lean tissue at a given BMI but with a lower ratio between visceral and subcutaneous fat (Theodoro *et al.,* 2006). The reduced visceral fat depots are thought to contribute to the higher insulin sensitivity and a better lipid profile even though individuals with PWS often die prematurely due to the complications that obesity causes, notably arterial hypertension, respiratory insufficiency and cardiovascular disease (Einfeld *et al.,* 2006), although the etiology remains unknown (Brambilla *et al.,* 2011).

1.5.4 Identification of Candidate Loci in Polygenic Obesity

Even though the heritability estimates for obesity related traits are high and key genes involved in the energy regulation pathway have been found through monogenic and

syndromic forms of obesity, the full contribution of genetics to the development of common obesity has yet to be elucidated.

The most comprehensive review in the literature, The Human Obesity Gene Map: the 2005 update (Rankinen *et al.,* 2006), identifies 135 different candidate genes either associated or linked with obesity and obesity-related phenotypes (body weight, BMI, WHR, WC, body composition and body fat distribution), in humans, 127 of which have been identified from association studies. The 127 candidate genes include *MC4R*, neuropeptide Y (*NPY)*, peroxisome proliferation activated receptor (*PPARG)*, peptide YY (*PYY)*, neuropeptide Y receptor Y2 (*NPY2R)*, adiponectin (*ADIPOQ)*, which are involved in the melanocortin pathway along with leptin (*LEP*) and the leptin receptor (*LEPR*). These have been associated with overweight and obesity related phenotypes; body weight, BMI, body composition and fat distribution (Rankinen *et al.,* 2006). Furthermore twenty-two genes associated with obesity-related phenotypes are replicated in at least five different studies, of which twelve genes associated have been replicated in >ten studies. These include *LEP, LEPR*, *PPARG, ADIPOQ, ADRB3* and *ADRB2*.

Since the comprehensive review in 2005, further studies have shown significant associations with measures of obesity, mainly through the use of genome wide association studies (GWAS) (Day and Loos, 2011), which has radically changed the study of complex disease. In this type of study single nucleotide polymorphisms (SNPs) are genotyped in both cases and controls to detect disease risk variants, for a particular trait, throughout the whole genome. As there are >ten million SNPs in the human genome it makes them a good genetic marker. However this is a vast number to genotype therefore the development of the HapMap project (http://hapmap.ncbi.nlm.nih.gov) and genotyping technology has

made it easier. In a genomic region, SNPs that are close together are more likely to be inherited together, than is expected by chance, in a block type pattern, known as a haplotype. This is due to linkage disequilibrium (LD) and the 'correlation' between two SNPs is measured using r^2 and D'. An r^2 value of one indicates complete LD, whereas zero indicates linkage equilibrium, or no linkage. Allele frequency and recombination between the two SNPs determines the value of r^2. Within a haplotype block of high LD, where $r^2 >$ 0.8, a selective number of SNPs, known as Tag-SNPs, can be chosen and act as proxy markers for the others in high LD in order to reduce the number of SNPs that need to be genotyped. GWAS has increased the number of novel genetic loci associated with a particular disease, as it is a non-candidate gene approach i.e. it is hypothesis free. Associations with a particular SNP indicate a region in the genome, which influences the risk of a disease.

In 2006 and 2007, with the use of this technology two genes were significantly associated with BMI: *FTO* and *INSIG2* (Herbert *et al.,* 2006, Dina *et al.,* 2007, Hinney *et al.,* 2007, Frayling *et al.,* 2007a, Scuteri *et al.,* 2007). Further GWAS have replicated the significant result with these two genes (Rosskopf *et al.,* 2007, Liu *et al.,* 2008, Bauer *et al.,* 2009, Heard-Costa *et al.,* 2009, Meyre *et al.,* 2009, Thorleifsson *et al.,* 2009). However it is thought obesity development from variants within *INSIG2* may occur in those already predisposed to moderate obesity through additional genetic variants and environmental factors (Korner *et al.,* 2008), as a large study involving Caucasian, Afro-Caribbean and Indian subjects found a lack of association (Smith *et al.,* 2007).

Furthermore initial GWAS identified common variants near *MC4R* to be associated with common obesity along with *FTO* and *INSIG2* (Loos *et al.,* 2008, Chambers *et al.,* 2008,

Cauchi *et al.,* 2009). Variants with *MC4R* had previously been associated with severe forms of obesity in children and adults in Caucasian populations (Vaisse *et al.,* 1998, Mergen *et al.,* 2001, Rosmond *et al.,* 2001) as well as Japanese (Kobayashi *et al.,* 2002), Chinese (Wang *et al.,* 2006) and Pima Indians (Ma *et al.,* 2004). This further supports a role for *MC4R* in the development of common obesity where it was initially thought to only be involved with rare monogenic forms. It is thought that a decrease in *MC4R* activity may result from mutations within the gene affecting its function in weight regulation pathway (Lubrano-Berthelier *et al.,* 2003).

Fifty novel genetic loci have additionally been associated with obesity related phenotypes along with further replication of the initial associations in GWAS. These include *TMEM18, KCTD15, GNPDA2, SH2B1, MTCH2, NEGR1, NRXN3, MSRA, LYPLAL1, CTNNBL1* and *FAP2B* (Liu *et al.,* 2008, Heard-Costa *et al.,* 2009, Lindgren, 2009, Renstrom *et al.,* 2009, Thorleifsson *et al.,* 2009, Willer *et al.,* 2009, Zhao *et al.,* 2009, Heid *et al.,* 2010, Beckers *et al.,* 2011, Hong and Oh, 2011, Fox *et al.,* 2012, Wang *et al.,* 2012). In one study alone, eighteen new loci were associated with BMI, at the level of genome wide significance of $p<5x10^{-8}$, along with confirming associations in fourteen previously identified loci (Speliotes *et al.,* 2010). Significant associations with WHR were found with thirteen novel loci, ($p<1.9x10^{-9}$-$1.8x10^{-40}$), including *VEGFA, NFE2L3, TBX15-WARS2, LY86, HOXC13* and *ADAMT59* (Heid *et al.,* 2010). Significant associations with these fifty novel loci have recently been replicated in different ethnic populations, including Chinese (Cheung *et al.,* 2010) and African-Americans (Ng *et al.,* 2012).

A limitation of GWAS is that due to the high numbers of SNPs being genotyped, this increases the type one errors from the multiple-hypothesis problem that arises (Seng and

Seng, 2008, Shuldiner, 2008). Therefore to identify a true result and reduce the number of false-positives, the P-value has to be stringent, with a significance level of $<10^{-7}$ and this means that a large number of participants need to be genotyped.

Table 1.5.1: A summary of the key genes that have been associated with T2D and common obesity in adults and children and in different populations. N.B Diabetes is rare disease seen in childhood with a prevalence rate of 0.35 per 1000 children (Lipman *et al.*, 2013) therefore there is a lack of research supporting involvement of genes for T2D in this group.

Gene	Phenotype	Population Studied	Adult Studies	Child Studies
FTO	BMI, HC, weight, WC, subcutaneous adipose tissue, WHR	European Caucasian	Frayling *et al.*, 2007; Dina *et al.*, 2007; Speliotes *et al.*, 2010	Meyre *et al.*, 2009; Dina *et al.*, 2007; Hinney *et al.*, 2006; Cauchi *et al.*, 2009
		Korean	Hong and Oh, 2011	
		Chinese	Cheung *et al.*, 2010	Wang *et al.*, 2012
		American Caucasian	Scuteri *et al.*, 2007	Zhao *et al.*, 2009
		Asian Indian	Saw *et al.*, 2012	Dwivedi *et al.*, 2012
		East Asian	Li *et al.*, 2012; Wen *et al.*, 2012	
MC4R	BMI, WC, weight	European Caucasian	Meyre *et al.*, 2009; Cauchi *et al.*, 2009; Loos *et al.*, 2008	Meyre *et al.*, 2009; Cauchi *et al.*, 2009; Loos *et al.*, 2008
		Indian Asians	Chambers *et al.*, 2008	Dwivedi *et al.*, 2012
		Korean	Hong and Oh, 2011	
		Chinese	Cheung *et al.*, 2010	Wu *et al.*, 2010
		American Caucasian	Thorleifsson *et al.*, 2009	Zhao *et al.*, 2009
BDNF	BMI, WC, weight, overall skin fold thickness	European Caucasian	Bauer *et al.*, 2009; Thorlesfsson *et al.*, 2009; Speliotes *et al.*, 2010	den Hoed *et al.*, 2010
		Korean	Hong and Oh, 2011	
		Chinese		Wu *et al.*, 2010
		American Caucasian	Thorlesfsson *et al.*, 2009	Zhao *et al.*, 2009
MTCH2	BMI, WC, weight	European Caucasian	Bauer *et al.*, 2009; Willer *et al.*, 2009;	

			Speliotes *et al.*, 2010; Renstrom *et al.*, 2009	
		Korean	Hong and Oh, 2011	
TMEM18	BMI, subcutaneous adipose tissue, WC, weight, overall skin fold thickness	European Caucasian	Willer *et al.*, 2009; Speliotes *et al.*, 2010; Hong and Oh, 2011	den Hoed *et al.*, 2010; Rask-Andersen *et al.*, 2012
		Chinese		Wang *et al.*, 2012
		American Caucasian		Zhao *et al.*, 2009
NEGR1	BMI, WC, weight, subcutaneous adipose tissue	European Caucasian	Bauer *et al.*, 2009; Willer *et al.*, 2009; Speliotes *et al.*, 2010; Renstrom *et al.*, 2009; Fox *et al.*, 2012	den Hoed *et al.*, 2010
		Chinese	Cheung *et al.*, 2010	
		American Caucasian		Zhao *et al.*, 2009
GNPDA2	BMI, WHtR, percentage fat mass, overall skinfold thickness	European Caucasian	Willer *et al.*, 2009; Speliotes *et al.*, 2010; Renstrom *et al.*, 2009	den Hoed *et al.*, 2010
		Chinese	Cheung *et al.*, 2010	Wu *et al.*, 2010
		American Caucasian		Zhao *et al.*, 2009
LEPR	BMI, fat mass, impaired leptin signalling,	European Caucasian	Chagnon *et al.*, 2000; Quinton *et al.*, 2001;	
		Tunisian	Ben Ali *et al.*, 2009	
SEC16B	BMI	European Caucasian	Speliotes *et al.*, 2010	den Hoed *et al.*, 2010
		Korean	Hong and Oh, 2011	
		Chinese	Cheung *et al.*, 2010	
PPARG	BMI, WHR, fat mass, T2D, HC, subscapular skinfold, total skinfold	European Caucasian	Meirhaeghe *et al.*, 2005; Scott *et al.*, 2007; Ghoussaini *et al.*, 2005	Lagou *et al.*, 2008
		Korean	Kim *et al.*, 2004	
		Indian	Chauhan *et al.*, 2010	
HHEX	T2D	European Caucasian	Sladek *et al.*, 2007; Scott *et al.*, 2007; Saxena *et al.*, 2007	

		Japanese	Omori *et al.*, 2008	
		Korean	Lee *et al.*, 2008	
		Indian	Chauhan *et al.*, 2010	
		Chinese	Wu *et al.*, 2008; Han *et al.*, 2010	
		Asians	Ng *et al.*, 2008	
CDKAL 1	T2D	Japanese	Omori *et al.*, 2008	
		Korean	Lee *et al.*, 2008	
		European Caucasian	Scott *et al.*, 2007; Saxena *et al.*, 2007	
		Indian	Chauhan *et al.*, 2010	
		Chinese	Wu *et al.*, 2008	
		Asian	Ng *et al.*, 2008	
CDKN2 A/2B	T2D	Japanese	Omori *et al.*, 2008	
		Korean	Lee *et al.*, 2008	
		Indian	Chauhan *et al.*, 2010	
		Chinese	Wu *et al.*, 2008	
		Asian	Ng *et al.*, 2008	
		European Caucasian	Scott *et al.*, 2007; Saxena *et al.*, 2007	
ADCY5	T2D, fasting glucose	South Asian	Rees *et al.*, 2011	
		European Caucasian	Dupuis *et al.*, 2010;	Windholz *et al.*, 2011; Barker *et al.*, 2011
		Asian Indian	Vasan *et al.*, 2011	

Due to the use of Tag-SNPs, even though the chromosomal region has been identified as having a significant association with risk of a disease, resequencing and fine mapping of the loci are needed to determine the true disease-causing variant (Seng and Seng, 2008). However the majority of these novel genes seem to be expressed in the hypothalamus indicating a role in energy regulation (Bauer *et al.,* 2009), so further investigation of these novel loci which would lead to further insights in body weight regulation, with a view to determine the gene function and the pathways that they are involved in (Day and Loos, 2012).

1.5.4.1 Differences between adult and childhood obesity

The prevalence rates of obesity are increasing for both adults and children therefore determining the genes that contribute to the development of this disease and the pathways that they are involved with are important. Overweight or obese children are also more likely to be overweight or obese as adults (Whitaker *et al.,* 1997, Freedman *et al.,* 2001). As a result genes with a role in childhood obesity are thought to represent a different etiology compared to adult onset obesity (Blakemore and Froguel, 2008). In childhood obesity variations in genes involved in energy and metabolism have a more direct effect (Meyre *et al.,* 2004), possibly determined by hyperphagia. Several loci associated with BMI in both adult and childhood obesity have a more pronounced effect in childhood (den Hoed *et al.,* 2010, Elks *et al.,* 2010). There have been examples of gene-hyperphagia interaction effects. There were up to three times more variance in weight and fat mass gain between pair differences when twelve pairs of monozygotic twins were overfed by more than 1000 calories for six days a week, for a total of eighty-four days in a 100 day overfeeding period (Deriaz *et al.,* 1993), suggesting particular individuals are more prone to gain weight than others, indicating genotypic effects (Bouchard, 2008). Whereas in adult

obesity the development of the phenotype is likely to occur over time resulting from subtle variations in genes involved in energy expenditure and metabolism. In addition with adult onset obesity there is a longer exposure to the environment compared to childhood obesity therefore the effects of genes are likely to contribute much stronger in the development of obesity.

Research investigating the three known variants within the extracellular region of the leptin receptor have only found associations with adult obesity (Ukkola *et al.,* 2000, Quinton *et al.,* 2001, Mattevi *et al.,* 2002, Salopuro *et al.,* 2005, Fairbrother *et al.,* 2007, Furusawa *et al.,* 2010, Sun *et al.,* 2010, Yang *et al.,* 2011) and not with children who are obese (Echwald *et al.,* 1997, Endo *et al.,* 2000, Pyrzak *et al.,* 2009, Komsu-Ornek *et al.,* 2012, Angel-Chavez *et al.,* 2012), indicating a difference in etiology. However, these studies were rather small in size, (<1000 subjects) which indicates that they were not powerful enough to detect an effect, but larger studies are easier to undertake in adults.

Novel loci identified through numerous GWAS have also been investigated in childhood obesity. In one study nine loci have been significantly associated with BMI in children: *INSIG2, FTO, MC4R, TMEM18, GNPDA2, NEGR1, BDNF* and *KCTD15* (Zhao *et al.,* 2009). Moreover the effect size of BMI was more pronounced with *TMEM18, SEC10B* and *KCTD15* in children and adolescence compared to adults (den Hoed *et al.,* 2010). In a recent study, loci previously associated individually with childhood BMI were investigated together on weight, height and BMI from birth to age eleven years in the Avon Longitudinal Study of Parents and Children (ALSPAC) (Elks *et al.,* 2010). These eight loci (*FTO, MC4R, TMEM18, GNPDA2, KCTD15, NEGR1, BDNF* and *ETV5*) were used to produce an 'obesity risk allele score' where the total number of risk alleles were

determined, ranging between two and fifteen. Measurements of weight, height and BMI were repeatedly taken from birth to eleven years of age and converted to standard deviation score (SDS). No significant associations were found with a high obesity risk allele score and birth weight but who are within the 5% who 'fail to thrive' despite sufficient nutrition in the early postnatal environment (p=0.009). At nine years of age each additional risk allele was associated with greater weight ($p=1.2x10^{-17}$), greater BMI ($p=1.4x10^{-19}$), greater height (p=0.0009). The differences between those with ≤4 and ≥13 risk alleles equalled 3.5 kg in body weight, 1.4 kg/m^2 in BMI and 2 cm in height. There was an increased risk of overweight (OR=1.4 per risk allele, 95% CI 1.10-1.19, $p=6.3x10^{-11}$) and obesity (OR=1.17 per risk allele, 95% CI 1.07-1.26, p=0.0002). In addition, there was a positive association in weight gain between birth and six weeks (0.119 SDS/allele/year, 95% CI 0.023-0.216, p=0.001) but not from six weeks to eleven years (0.004 SDS/allele/year 95% CI 0.004-0.005) and a positive correlation with early infancy length gain (0.158 SDS/allele/year, 95% CI 0.032-0.284). These results suggest that these loci influence greater weight gain in childhood and are involved in a pathway that leads to increase adult obesity risk (Elks *et al.,* 2010).

1.6.0 Explanations into the rise of common obesity

Knowledge that genetic factors make up a large contribution to the development of obesity, has lead to eight different hypotheses being proposed to explain the development of common obesity, in the current obesogenic environment (Walley *et al.,* 2009). These are the thrifty gene hypothesis, the fetal programming hypothesis, the sedentary lifestyle hypothesis, the ethnic shift hypothesis, the predictive release hypothesis, the increased reproductive fitness hypothesis and the assortative mating hypothesis.

1.7.0 Birth Weight

One theory into the origins of obesity and its sequelae is the fetal programming hypothesis. This suggests that exposure to adverse conditions such as under or overnutrition *in utero* causes metabolic changes in the offspring in preparation for its postnatal conditions. However these adaptations can be detrimental to the health of the offspring when the postnatal conditions do not match to the ideal.

Birth weight has long been known to have a profound influence on the health of an individual in adulthood (Rich-Edwards *et al.,* 1999, Syddall *et al.,* 2005, Finken *et al.,* 2006, Barker *et al.,* 2009) and investigations into factors that influence fetal growth have been carried out since the mid 20th century (Illingworth, 1950, Karn and Penrose, 1952, Banik *et al.,* 1967, Simpkiss, 1968). At this time, birth order, gestational age and maternal age were first considered and found to have an effect on birth weight. Since then, the number of intrinsic and extrinsic factors affecting birth weight includes smoking, parity, maternal weight and weight gain, gestational diabetes, maternal nutrition, environmental toxins and functioning of the placenta (Magnus *et al.,* 1985, MacLaughlin *et al.,* 2005, Chen *et al.,* 2009a, Bae *et al.,* 2011a, Mamun *et al.,* 2011, Roseboom *et al.,* 2011, Laitinen *et al.,* 2012).

Furthermore, research suggests birth weight to be 10-45% heritable (Clausson *et al.,* 2000, Baird *et al.,* 2001, van Dommelen *et al.,* 2004, Hur *et al.,* 2005, Svensson *et al.,* 2006, Beardsall *et al.,* 2009), indicating a strong contribution of genetics but estimates are inconsistent despite the use of MZ and DZ twins. The variability in estimates have been attributed to the unequal sharing of nutrients and placental blood supply (Machin, 1996) but Vlietinck *et al.,* in 1989 (Vlietinck *et al.,* 1989) identified gestational age to explain

42% of the variance of birth weight, with 23% due to random environment, 22.5% due to additive genetic factors and 0.7% to maternal age. In twins total variance is split into genetic (A) contribution, shared or common environment (C) and unique environment (E) components, where reducing the common and unique environmental factors will increase the heritability. However covariates also influence each of these components, therefore the design of the study is also important in estimating heritability. A recent study by Gielen *et al.,* (Gielen *et al.,* 2008) used a model to determine heritability of birth weight across gestational age, as it is likely to change during gestation and different covariates are likely to explain variance over different stages of pregnancy. The covariates included were split into different categories: maternal factors (gestational age, parity and maternal age), twin specific factors (zygosity, sex and birth order) and placental factors (chorionicity, fusion of placentas and insertion of umbilical cord). The heritability was found to decrease during gestation from 38% at twenty-five weeks to 15% at forty-two weeks and the covariates explained more of the common environmental variance than genetic variance. Unique environmental variance is dependent on covariates and heritability was highest for those first born of women with more than two previous pregnancies and two separate placentas; 52% at twenty-five weeks to 30% at forty-two weeks. Therefore, birth weight is a complex trait, but the precise molecular mechanisms underpinning birth weight are still relatively unknown.

1.7.1 Measurement of Birth Weight

At term, the range of normal birth weight is 2500 g to 4000 g. Low birth weight is classified as <2500 g, and high birth weight as >4000 g. Low birth weight includes newborns termed small for gestational age (SGA) i.e. less than the 10^{th} percentile for either birth weight or birth length, and those who have intrauterine growth retardation (IUGR).

Low birth weight is connected to many congenital conditions and syndromes but is considered here as an isolated phenotype. Low birth weight as part of a syndrome (e.g. Silver-Russell syndrome (OMIM 180860) and Three M syndrome (OMIM 273750)) is reflected by poor fetal growth and is accompanied by additional phenotypes and likely to have variable etiology and differing impacts on adult health.

1.7.2 Birth weight and Increased Disease Risk in Adults and Children

One interesting aspect of this isolated phenotype is the correlation of low birth weight at term with disease risk in later life (Weiss and Caprio, 2005, Sabin and Shield, 2008). Birth weight and birth length, normalised for gestational age, are markers for fetal growth, therefore understanding the connection between birth weight and its sequelae may lead to more effective clinical intervention.

Low and high birth weight for gestational age (at term) predisposes to the onset of obesity, diabetes and the metabolic syndrome (Table 1.7.1). In one study significant associations were found with lower birth weight ($p<0.001$), head circumference ($p<0.001$), smaller ponderal index (a measure of leanness using birth weight (kg)/length (m^3) ($p<0.001$)) and shorter length ($p=0.004$) and T2D and impaired glucose regulation compared with those with normal glucose tolerance in 2019 Chinese subjects (Xiao *et al.,* 2008). There is a close link between these conditions because they are all connected with energy regulation.

Numerous studies have found associations between birth weight and these pathological sequelae as adults, either singularly or collectively (Table 1.7.1). The majority of these studies have been carried out in European or American populations. However, similar findings have been replicated in different ethnicities including Chinese, Japanese and

Indian (Stein *et al.,* 1996, Schroeder *et al.,* 1999, Levitt *et al.,* 2000, Anazawa *et al.,* 2003, Tian *et al.,* 2006, Corvalan *et al.,* 2007, Xiao *et al.,* 2008). This suggests that the lasting effect of low birth weight at term on the health of the individual is not restricted to one ethnicity, indicating the mechanism(s) involved may be the same or similar.

The increased risk of developing CVD, T2D, obesity, impaired glucose tolerance or higher fasting glucose was found to be one to two times higher for those of lower birth weight, compared to those of normal weight at term (Hales *et al.,* 1991, Carlsson *et al.,* 1999, Leong *et al.,* 2003, Rich-Edwards *et al.,* 2005, Wadsworth *et al.,* 2005, Tian *et al.,* 2006, Xiao *et al.,* 2008). In some cases the risk has been found to be up to four times higher for fasting glucose (Tian *et al.,* 2006) or six times higher (Hales *et al.,* 1991). However, it is possible that the estimated risk of developing adverse health problems are not as high as those stated. It is dependent on the variables that are controlled for within the analysis and not all studies will have been conducted in the same manner, with the same variables measured (Carlsson *et al.,* 1999).

In addition, high birth weight and high ponderal index at term are also associated with adult disease risk of obesity, T2D and cardiovascular disease (Sorensen *et al.,* 1997, Eriksson *et al.,* 2001a, Parsons *et al.,* 2001, Wei *et al.,* 2003, Tian *et al.,* 2006). This indicates that both high and low birth weights contribute to the adult disease phenotype. A U-shaped relationship is normally observed, with birth weight within the normal range having no effect or association with adult disease risk. However it seems that low birth weight has the greatest risk on adult disease. This is contributed to the fact that individuals with low birth weight at term undergo rapid catch up to normal weight during early infancy and this weight gain continues during childhood. High BMI and abdominal fatness occurs

as a result. Obese or overweight children are more likely to be obese or overweight as adults (Sabin and Shield, 2008).

The number of children and adolescents that are overweight (a BMI between >85th percentile and 95th percentile adjusted for age) and obese (a BMI >95th percentile adjusted for age) (Kuczmarski *et al.,* 2002) and the number of young people with T2D (defined as having a random plasma glucose concentration (tested without regard to timing of food intake) ≥200 mg/dl, fasting plasma glucose ≥126 mg/dl and a 2-hour plasma glucose ≥200 mg/dl) has also increased over the last thirty years, in line with the rise in adults (Ogden *et al.,* 2012a; Jefferies *et al.,* 2012). As a consequence, the age of onset of T2D has decreased duringthis time frame. This is of concern as it was previously, predominately, a disease more commonly found in those over forty years of age. Associations between low birth weight at term and aspects of the adult disease risk phenotype have been found between the ages of twelve and eighteen years (Table 1.7.2), in addition, to the under ten years age group (Table 1.7.3) (Crowther *et al.,* 1998, Ong and Dunger, 2004, Ibanez *et al.,* 2008). The study by Dabelea *et al.,* (Dabelea *et al.,* 1999) investigated the incidence of T2D in 3061 Pima Indian children and young adults. Higher prevalence of T2D was found with low birth weight at term in those aged between ten and fourteen and fifteen and nineteen years of age ($p<0.001$) and mean 2-hour plasma glucose concentration was significantly higher in those aged between ten and fourteen years and fifteen and nineteen years with low birth weight ($p<0.01$).

Table 1.7.1: Associations with low birth weight at term and subsequent disease development in adulthood have been found in fifteen different population groups from around the world. The three measures of disease are all part of the spectrum of the Metabolic Syndrome. British and American populations are of European Caucasian descent.

Phenotype	Population	Author
Type 2 Diabetes, impaired glucose tolerance, impaired fasting glucose, impaired fasting insulin, insulin resistance	British	(Hypponen *et al.*, 2003)
	Swedish	(Lithell *et al.*, 1996)
	Chinese	(Tian *et al.*, 2006)
	British	(Wadsworth *et al.*, 2005)
	Chinese	(Xiao *et al.*, 2008)
	British	(Phillips *et al.*, 2005)
	British	(Hales *et al.*, 1991)
	Swedish	(McKeigue *et al.*, 1998)
	Swedish	(Carlsson *et al.*, 1999)
	Japanese	(Anazawa *et al.*, 2003)
	American	(Burke *et al.*, 2004)
Cardiovascular disease, stroke, hypertension, blood pressure, coronary heart disease	British	(Reynolds *et al.*, 2001)
	Danish	(Schack-Nielsen *et al.*, 2002)
	American	(Rich-Edwards *et al.*, 1999)
	British	(Lawlor *et al.*, 2004)
	Finnish	(Eriksson *et al.*, 2001b)
	Dutch	(Roseboom *et al.*, 2000)
	Finnish	(Eriksson *et al.*, 2000)
	British	(Lawlor and Leon, 2005)
	British	(Barker *et al.*, 1990)
	South Indian	(Stein *et al.*, 1996)
	British	(Law *et al.*, 2002)
	Finnish	(Eriksson *et al.*, 2000)
	Finnish	(Forsen *et al.*, 1999)
	British	(Martyn *et al.*, 1996)
	Norweign	(Risnes *et al.*, 2009)
	Swedish	(Leon *et al.*, 1996)
	American	(Rich-Edwards *et al.*, 2005)
	Italian	(Banci *et al.*, 2009)
	Danish/Finnish	(Andersen *et al.*, 2010)
	American	(Curhan *et al.*, 1996)
	South African	(Levitt *et al.*, 2000)
Measures of obesity (WHR, BMI, HC, WC, Skinfold thickness, percentage fat)	Finnish	(Yliharsila *et al.*, 2007)
	Guatemalan	(Corvalan *et al.*, 2007)
	American	(Stettler *et al.*, 2005)
	Belgian	(Loos *et al.*, 2001)
	Belgian	(Loos *et al.*, 2002)
	British	(Kuh *et al.*, 2002)
	British	(Law *et al.*, 2002)
	Guatemalan	(Schroeder *et al.*, 1999)
	Finnish	(Eriksson *et al.*, 2001a)
	Swedish	(Byberg *et al.*, 2000)
	Chinese	(Tian *et al.*, 2006)
	American	(Leong *et al.*, 2003)

Table 1.7.2: Associations with low birth weight at term and subsequent disease development in late childhood have been found in seventeen different population groups, around the world. The three measures of disease are all part of the spectrum of Metabolic Syndrome.

Phenotype	Populations	Author
T2D, impaired glucose tolerance, impaired fasting glucose, impaired insulin resistance, fasting insulin	Japanese	(Sugihara *et al.*, 2008)
	South African	(Crowther *et al.*, 1998)
	Pima Indians	(Dabelea *et al.*, 1999)
	Finnish	(Eriksson *et al.*, 2003)
	South Indian	(Bavdekar *et al.*, 1999)
	Chilean	(Soto *et al.*, 2003)
	British	(Ong *et al.*, 2004)
	Chilean	(Mericq *et al.*, 2005)
	Taiwanese	(Wei *et al.*, 2003)
	British	(Singhal *et al.*, 2003)
	Korean	(Kim *et al.*, 2006)
	South Indian	(Raghupathy *et al.*, 2010)
Cardiovascular disease, stroke, hypertension, blood pressure, coronary heart disease	Philippine	(Adair and Cole, 2003)
	Jamaican	(Forrester *et al.*, 1996)
	British	(Barker *et al.*, 1990)
	Finnish	(Osmond *et al.*, 2007)
Measures of obesity (BMI, WC, AC, WHR, skinfold thickness, fat mass, fat free mass	British	(Ong *et al.*, 2000)
	British	(Barker *et al.*, 1997)
	British	(Sayer *et al.*, 2004)
	Jamaican	(Walker *et al.*, 2001)
	British	(Parsons *et al.*, 2001)
	Brazilian	(Gonzalez *et al.*, 2010)
	Danish	(Sorensen *et al.*, 1997)
	Israelian	(Friedlander *et al.*, 2002)
	Danish	(Barker *et al.*, 1997)

Table 1.7.3: Associations with low birth weight and subsequent disease development in early childhood have been found in seven different population groups around the world.

Phenotype	Population	Author
T2D, impaired fasting insulin, impaired glucose tolerance	Pima Indian	(Dabelea *et al.*, 1999)
	British	(Crowther *et al.*, 1998)
	Indian	(Bavdekar *et al.*, 1999)
	British	(Ong *et al.*, 2004)
	Chilean	(Mericq *et al.*, 2005)
	Spanish	(Ibanez *et al.*, 2008)
Blood pressure, Total cholesterol	Jamaican	(Forrester *et al.*, 1996)
	Indian	(Bavdekar *et al.*, 1999)
Measures of obesity (BMI, weight, WC, percentage fat,	British	(Crowther *et al.*, 1998)
	British	(Ong *et al.*, 2000)
	Indian	Bavdekar *et al.*, 1999
	British	Ong *et al.*, 2004
	Chilean	Mericq *et al.*, 2005
	Spanish	Ibanez *et al.*, 2008
	German	Karaolis-Parekert *et al.*, 2006

1.7.3 Theories Explaining correlation of Birth Weight and Disease Risk

There are two major theories proposed to explain the correlation between birth weight and obesity, T2D and the metabolic syndrome, though the precise mechanisms involved remain unclear. What is certain is that both *in utero* conditions and subsequent lifestyle have a role in the development of obesity and the metabolic consequences.

1.7.4 Maternal Environment

All pregnancies have some degree of restriction in fetal growth, as it is important for the survival of the mother and subsequent offspring. Although the growth trajectory of the fetus is influenced by the fetal genome, it is predominately determined maternally, through maternal body size, a process termed maternal constraint. Animal cross breeding studies (Walton and Hammond, 1938) have helped to demonstrate the importance of the maternal environment along with human half sibling studies (Morton, 1955) and studies involving the transfer of embryos in humans (Brooks *et al.,* 1995). In humans, birth weight was similar with half siblings of the same mother (R = 0.58) however those with the same father had different birth weights (R = 0.1). Furthermore, for human embryos that were transferred into a recipient mother, the recipient mother had more influence on the growth of the fetus than the donor mother, suggesting that genetics plays a lesser role at this time. Fetuses tend to have a higher birth weight when they are transferred into a large uterus. Additional contributing factors are maternal age, presence of more than one fetus, placental size (affecting the nutrition and oxygen supply) and whether or not the mother has previously given birth (parity) (Karn and Penrose, 1952, Banik *et al.,* 1967, Godfrey, 2002). Maternal age and parity increases birth weight whilst presence of more than one fetus and a smaller placenta decreases birth weight.

1.7.5 Environmental Origins of Birth Weight Variation

Hales and Barker (1992) (Hales and Barker, 1992) proposed the *'fetal origin of adult disease'* hypothesis, or the *'thrifty phenotype hypothesis'*, which suggests that the onset of adult disease is the result of an adaptive response to adverse conditions during fetal growth that are meant to promote survival in the post-natal environment.

Optimum maternal nutrition is essential for fetal growth and development during pregnancy. Throughout this time the fetus is sensitive to many environmental cues, particularly pertaining to the post-natal surroundings. Sensing unfavourable conditions, such as low nutrition, (maternally through the placenta), the fetus undergoes many adaptations, which has been termed 'developmental plasticity' (Hales and Barker, 1992). Fetal adaptations include changes in metabolism, hormone production and the growth and function of various organs, while neurodevelopment is maintained, called 'programming'. The changes that occur within the fetal organs include reductions in nephron number in the kidney, pancreatic β-cell number, skeletal muscle mass and fibre composition, decreased expression of components of the insulin signal transduction pathway in skeletal muscle and proliferation of hepatocytes (Garofano *et al.,* 1997, Gruppuso *et al.,* 2005, Ozanne *et al.,* 2005, Zandi-Nejad *et al.,* 2006, Calkins and Devaskar, 2011). Environmental cues impacting on uterine environment, such as maternal malnutrition, placental dysfunction and smoking are likely to be indicative of the postnatal surroundings therefore induced fetal changes are more likely to be beneficial in adulthood. Problems arise when the environmental conditions, experienced by the fetus differ to those the fully developed adult experiences. If they are much more favourable than expected, e.g. such as the obesogenic environment that has developed in the Western world, this leads to the increased risk of obesity, T2D and the metabolic syndrome (Hales and Barker, 1992).

1.7.5.1 Supportive Evidence for the 'Thrifty Phenotype Hypothesis'

Various sets of data support the 'Thrifty Phenotype' theory. These include detailed birth records from England and Finland (Hales *et al.,* 1991, Barker *et al.,* 1993, Eriksson *et al.,* 1999) and famine data from Holland (Painter *et al.,* 2006, Stein *et al.,* 2006), Russia (Stanner *et al.,* 1997), China (Yang *et al.,* 2008) and Nigeria (Hult *et al.,* 2010) (Table 1.7.4). Birth records have demonstrated that those with a lower birth weight at term had an increased risk of impaired glucose tolerance (OR 6.6, 95% CI 1.5-28, $p<0.001$) (Hales *et al.,* 1991) and a lower ponderal index had an increased risk of developing coronary artery disease (Hazard Ratio 2.07 $p<0.0001$) as adults (Erkisson *et al.,* 1999). Subsequent research has replicated these findings (Table 1.7.1). Significant associations with low and high birth weight at term have been identified with impaired glucose tolerance, obesity, stroke, hypertension and bone mass and strength in European, American, Chinese and Japanese populations (Osmond *et al.,* 1993, Curhan *et al.,* 1996, Eriksson *et al.,* 2001b, Roseboom *et al.,* 2001b, Anazawa *et al.,* 2003, Kensara *et al.,* 2005, Phillips *et al.,* 2005, Tian *et al.,* 2006, Oliver *et al.,* 2007, Osmond *et al.,* 2007).

Famine data has been utilised to assess the link between food restriction in pregnancy and susceptibility to adult disease, with famine defined as the total number of calories consumed maternally being less than 1000 kcal each day. Offspring born during a famine were found to be 200 g lighter at birth (Roseboom *et al.,* 2001a). As adults, survivors were found to have raised blood glucose levels (Ravelli *et al.,* 1998), impaired insulin secretion (de Rooij *et al.,* 2006a) and had higher WC and hip circumference compared to those not exposed to famine *in utero* (Ravelli *et al.,* 1999, Stein *et al.,* 2007). Hypertension was increased (an elevated systolic and diastolic blood pressure) (Stein *et al.,* 2007), as was the

occurrence of coronary heart disease (CHD) (defined by presence of angina pectoris, Q waves on ECG and history of coronary revascularisation) (Roseboom *et al.*, 2000) with earlier onset (Painter *et al.*, 2006). These findings were replicated in Chinese (Yang *et al.*, 2008) and Nigerian populations (Hult *et al.*, 2010).

China suffered from natural disasters between 1959 and 1961 that resulted in famine in three regions of the country. Current BMI status was analysed with respect to birth year. Those born in 1964 were included as controls. Women who were born between 1959 and 1961 had a significantly higher BMI than those who were born in 1964 ($p<0.02$). No significant difference was found with men who were born in the same time frame. This suggests that the risk of being overweight or obese was much higher for those women who were born during the famine. This result replicates previous findings, which found higher WC, BMI and HC in women exposed to famine *in utero* compared to men (Ravelli *et al.*, 1999; Stein *et al.*, 2007), suggesting there is a sex specific effect although both men and women are likely to store adipose tissue in an intra-abdominal location (Ravelli *et al.*, 1999).

The Biafran famine in Nigeria lasted for nearly two years, between June 1967 and January 1970. The study by Hult *et al.*, (2010) found that those who were exposed to the famine during fetal life and early infancy had a higher systolic blood pressure ($p<0.0001$), diastolic blood pressure ($p<0.0001$), waist circumference ($p=0.0011$), BMI ($p=0.016$) and plasma glucose level ($p=0.04$) compared to those born after the famine. This indicates that the risk of severe hypertension, T2D and obesity is higher in this group than those who were not exposed to the adverse conditions during fetal or in childhood.

Table 1.7.4: A summary of the effects on body composition of three different famines that have occurred at different times over the last sixty-nine years from three continents.

Famine	Phenotype	Author
Dutch (1944-1945)	Earlier onset CAD	Painter *et al.*, 2006
	Higher weight (women), higher WC, higher HC, higher BMI, higher WHR, higher mid thigh circumference, higher subscapular skinfold thickness	Stein *et al.*, 2007
	Higher BMI (women), higher WC, higher HC	Ravelli *et al.*, 1999
	Higher total cholesterol (women), higher triglycerides levels, higher LDL cholesterol levels, higher BMI, higher WC	Lumey *et al.*, 2009
	Higher LDL:HDL cholesterol ratio, lower plasma HDL cholesterol, higher LDL cholesterol, higher total cholesterol	Roseboom *et al.*, 2000
	Higher glucose levels	de Rooij *et al.*, 2006; Ravelli *et al.*, 1998
	Higher prevalence of CHD	Roseboom *et al.*, 2000
China (1959-1961)	Higher BMI (women)	Yang *et al.*, 2008
Biafran (1967-1970)	Higher systolic blood pressure, higher diastolic blood pressure, higher WC, higher BMI	Hult *et al.*, 2010

However, with both these studies even though significant associations have been found there are limitations to each study. No birth weight data from any of the subjects were analysed with respect to the phenotype data. Therefore no direct association with low birth weight can be made and the subsequent risk of developing the adult phenotype but these studies fit with the idea that poor fetal nutrition reflected by low birth weight is associated with increased risk of obesity and T2D in later life.

1.7.6 Effects of maternal malnutrition and body size

One of the important influences on maternal constraint and the growth and birth weight of the fetus is the health, nutritional status, and body size of the mother (Godfrey and Barker, 2000, Belkacemi *et al.,* 2010). This is of importance not only before pregnancy, (including preconception and preimplantation) but also throughout pregnancy, even during early gestation (Godfrey, 2002, McMillen *et al.,* 2008). The exact mechanisms involved in how maternal nutrition affects the offspring both *in utero* and postnatally, are still not fully understood but it is thought that the timing in which poor fetal environment occurs is of significance (McMillen *et al.,* 2008). Data from the Dutch famine shows that for women who were in early gestation, at the time of the famine, their offspring had normal birth weights while those who were in either middle or late stages of pregnancy had lower birth weights at term and were smaller in size (Calkins and Devaskar, 2011). This suggests that fetuses in the early stages of development could adjust to the lack of nutrition much better than those who were accustomed to a certain level of nutrition and calorie intake, which restricted their growth in the final stages of pregnancy. In contrast, other research has shown that those experiencing suboptimal conditions during early to mid-gestation were proportionally smaller in weight and length at birth, known as symmetric growth restriction. Whereas during late gestation the offspring was disproportionately small, known as asymmetric growth restriction (Kramer *et al.,* 1990). Maternal smoking, maternal size, pregnancy related hypertension and sexual dismorphism also affect fetal growth, which suggest additional factors are also important throughout pregnancy along with nutrition (Hindmarsh *et al.,* 2002).

Furthermore the growth trajectory of the fetus is established during early gestation as well as the development of the placenta, which undergoes its maximal growth during this time

(Osgerby *et al.,* 2004). Therefore changes in maternal weight at this time can have a detrimental effect. Substantial research suggests that poor maternal nutrition during periconception and at the time of implantation is crucial to fetal growth and future health. Offspring who experienced poor maternal nutrition during early gestation, or as embryos, were found to be more at risk of a high body mass index, coronary heart disease and glucose intolerance as adults (de Rooij *et al.,* 2006b, Painter *et al.,* 2006).

Additional evidence comes from animal models, in which sheep model are regarded as a good comparison to humans as they typically give birth to one or two offspring at a time, compared to rodent models as they typically have larger numbers of offspring per birth (Vuguin, 2007). Maternal malnutrition during preconception has been associated with poor fetal renal vascular nephron development and renal function in adult sheep (Lloyd *et al.,* 2012), as well as poor fetal cerebral development in sheep (Ranade *et al.,* 2012). Early and mid-gestation maternal malnutrition was found to increase adipocyte size in offspring of beef cattle, compared to those fed a normal diet (Long *et al.,* 2012). A direct relationship has been found between maternal weight gain during periconception and implantation (forty-five days before conception and seven days after conception) and the weight of the fetus and placenta at fifty-five days gestation in sheep. However this relationship is affected during maternal undernutrition, in which the weight of the fetus and placenta decreases (Edwards and McMillen, 2002b, Edwards and McMillen, 2002a, MacLaughlin *et al.,* 2005). When maternal malnutrition was continued more severely and up to thirty days after conception similar results were found with enhanced activation of the fetal pituitary-adrenal axis (Bloomfield *et al.,* 2003). Peri-implantation undernutrition in sheep did not affect offspring birth weight or postnatal growth but an altered cardiovascular function was identified in the adult offspring (Gardner *et al.,* 2004). Angiotensin II activity occurs

during this time so poor maternal nutrition could increase the activity and affect cardiovascular function in the adult offspring, which can lead to hypertension.

In addition, the type of malnutrition seems to be important in fetal growth. In rats that were given a low protein diet 4.25 days after conception, there was a decrease in the number of cells within the blastocyst, which could influence the growth trajectory as the offspring had a low birth weight and increased blood pressure in later life (Kwong *et al.,* 2000). A restricted protein maternal diet has also been found to be associated with hypertension in offspring in rats (Augustyniak *et al.,* 2010), brain development in rat offspring (Torres *et al.,* 2010) and changes in the fetal renin-angiotensin system in the mouse (Goyal *et al.,* 2009). However, diets high in sugar, fat and energy (junk food) have also been found to have an adverse effect on fetal growth and future health. Female pregnant rats fed on a diet consisting of energy-rich chow produced offspring that had a higher risk of developing non-alcoholic fatty liver disease and had a higher liver mass than those fed on a normal diet (Bayol *et al.,* 2010). Maternal overweight and obesity at conception, has been shown to have a detrimental effect on fetal growth and has future health implications for the offspring. Babies born to those who are either overweight or obese are more likely to be larger in size and have a higher percentage body fat compared to those born to mothers who are of normal weight before pregnancy (Catalano *et al.,* 2009).

Moreover maternal body size and composition can influence the growth and development of the fetus. Women who have a high weight or body mass index (BMI) >25 kg/m^2 prior to pregnancy are more likely to have a larger infant at birth (Mandal *et al.,* 2011, Heude *et al.,* 2012) whereas smaller sized women, i.e. those who are smaller in weight and height (BMI <25 kg/m^2) tend to have a smaller fetus at birth (Voigt *et al.,* 2011). High BMI before

pregnancy is a risk factor for developing insulin resistance and gestational diabetes, which can also affect the fetus. It can lead to fetal overnutrition and overgrowth resulting in high birth weight (McMillen *et al.,* 2008). Complications can arise as a result including preeclampsia and the requirement of a caesarean section. There is an increased the risk of death in men through coronary heart disease who are born to obese mothers (Forsen *et al.,* 1997). The numbers of women who are overweight or obese has increased over the last thirty years, so have the numbers who carry excess weight at the start of pregnancy, rising from 25% to 35% between 1991 and 2001 (LaCoursiere *et al.,* 2005).

Maternal underweight (<18 kg/m^2) before pregnancy similarly influences fetal growth and placental developments with maternal overweight/obesity. Preterm birth (<thirty-seven weeks) is the main cause of neonatal mortality and can lead to low birth weight (Branum and Schoendorf, 2002). It has been debated whether maternal underweight increases, decreases or has no effect on preterm birth and low birth weight. However, a systematic meta-analysis of the literature has found that maternal underweight increases the risk of preterm birth and low birth weight compared to normal maternal weight (Han *et al.,* 2011). The meta-analysis included seventy-eight studies (cohort and case-control), including a total of over a million women from North America, Western Europe and Asia. Preterm birth was grouped into either <thirty-seven weeks, thirty-two and thirty-six weeks and <thirty-two weeks and low birth weight was grouped as moderate low birth weight (1500-2500 g), very low birth weight (<1500 g) and extremely low birth weight (<1000 g). In both pooled case-control and cohort studies maternal underweight had an increased risk of preterm birth at <thirty-seven weeks, Relative Risk (RR) 1.21 (95% CI 1.14-1.28) and OR 1.25 (95% CI 1.35-1.77) respectively. There were also similar findings for thirty-two and thirty-six weeks for the pooled cohort studies (RR 1.25 95% CI 1.09-1.43). A higher risk

was also found with low birth weight. In both pooled cohort studies maternal underweight had an increased risk of low birth weight, RR 1.50 (95% CI 1.34-1.94), with moderately low birth weight (1500-2500 g), RR 2.10 (95% CI 1.59-2.76) and very low birth weight RR 1.54 (95% CI 1.22-1.94). The Odds Ratio for case-control studies was 1.81 (95% CI 1.16-2.84).

In addition, maternal body composition has similar effects on offspring health. Lower maternal weight gain during gestation (fifteen to thirty-five weeks) was associated with raised blood pressure in seventy-seven offspring at eleven years of age in (Godfrey *et al.,* 1994). SBP increased by 10.7 mmHg (95% CI 5.7-15.6, p=0.001) each log mm decrease in mothers triceps skinfold thickness and increased 0.6 mmHg (95% CI 0.1-1.0, p=0.02) with a 1 kg decrease in weight gain. A further study of 296 children found that mothers with a tricep skinfold thickness lower than the mean (15 mm) at eighteen weeks gestation and a reduced pregnancy weight gain was associated with higher blood pressure in offspring (Clark *et al.,* 1998). SBP increased by 11.3 mmHg (95% CI 2.2-20.4) and DBP increased by 10.1 mmHg (95% CI 3.2-17.1) with a 1 kg decrease per week in pregnancy weight gain.

1.7.7 Placental Function

The supply of nutrients from the mother to the fetus and the removal of waste occur via the placenta and is important for normal fetal growth and development (Belkacemi *et al.,* 2010). Trophablast cells are key in the establishment of the placenta, which develops during implantation and early gestation (John and Hemberger, 2012). Therefore changes to the structure and size of the placenta would have a detrimental effect on its function, affecting fetal growth and development (Jansson *et al.,* 2006, Fowden *et al.,* 2008). In addition, maternal diet and body composition (underweight and overweight) directly

affects the supply of nutrients to the fetus via the placenta and indirectly through placental structure and function (Godfrey *et al.,* 1997, Belkacemi *et al.,* 2010, Higgins and Mc Auliffe, 2010).

As the placenta plays such an important role in fetal growth research it is not surprising that placental weight has been associated with size at birth, where low placental weight is a risk factor for low birth weight at term and still birth (Nohr *et al.,* 2005, Husslein *et al.,* 2012, Hutcheon *et al.,* 2012, Roland *et al.,* 2012). Furthermore high and low placental weight has been associated with increased blood pressure, impaired glucose tolerance, T2D, stroke and coronary artery disease in adults (Martyn *et al.,* 1996, Leon *et al.,* 1998, Moore *et al.,* 1999, Forsen *et al.,* 2000, Thame *et al.,* 2000).

This suggests the placenta is able to make short-term changes to fetal nutrient supply based on maternal nutritional status to optimise fetal growth *in utero* but also adapt fetal growth to maximise its survival in the postnatal environment (Lewis *et al.,* 2012).

1.7.8 'Fetal Origin of Adult Disease' and Catch-Up Growth

The 'fetal origins of adult disease' theory (Hales and Barker, 1992) suggests that problems in adulthood only arise when there is a mismatch between the intra-uterine environment and postnatal environments. If postnatal conditions are better nutritionally than the uterine environment, catch-up growth occurs, in which the offspring increases its weight rapidly to within the normal range and in some instances above this, as a natural response to nutritional deprivation (Calkins and Devaskar, 2011). It is this rapid increase in weight that seems to cause adverse effects in later life (Ong *et al.,* 2000, Leong *et al.,* 2003, Dennison

et al., 2006, Gonzalez *et al.,* 2010, Yu *et al.,* 2011). Therefore, the early postnatal environment is an additional critical period in the development of the offspring.

The main risk factor in the development of T2D and the metabolic syndrome is adiposity. Adipose tissue localises to either a visceral or subcutaneous location but research has shown that visceral adipose tissue (i.e. surrounding organs such as the liver, intestines and stomach) is more detrimental to the health of an individual compared to a subcutaneous location (Pouliot *et al.,* 1992, Ross *et al.,* 2002a, Ross *et al.,* 2002b, Gastaldelli *et al.,* 2009, Barreira *et al.,* 2012, Pickhardt *et al.,* 2012). However the exact mechanisms remain unknown: whether visceral adipose tissue contributes to features of insulin resistance (through altered free fatty acid metabolism and adipokine release), leading to the metabolic syndrome or a marker of dysfunctional adipose tissue (the lipid overflow-ectopic fat model) in an environment of excess energy (Després and Lemieux, 2006).

Additionally the timing of weight gain, i.e. how many weeks after birth and the type of weight gain, is of importance. In humans, for weight gain in the early postnatal period, the impact on health can be detected up to ten years afterwards (Singhal *et al.,* 2003, Soto *et al.,* 2003, Stettler *et al.,* 2005, Gonzalez *et al.,* 2010). Children at the ages of two and four and a half years, born SGA had more total adipose tissue than lean tissue compared with those of the same age, who were within the normal weight range at birth, regardless of their weight gain and BMI (Dulloo *et al.,* 2006). In addition, catch up growth in preschool years can be just as detrimental to the later health of an individual, as catch up growth in the first year of life (Parsons *et al.,* 2001, Eriksson *et al.,* 2003, Ong and Dunger, 2004, Ezzahir *et al.,* 2005, Ibanez *et al.,* 2008). Thus, undergoing catch-up growth, irrespective of timing and current weight status in childhood can be a risk factor for overweight and

other conditions later on (Walker *et al.,* 2001, Karaolis-Danckert *et al.,* 2006, Gonzalez *et al.,* 2010).

Analysis of the famine data mentioned earlier, showed that as the Leningrad famine spanned 800 days, those who were malnourished during fetal growth endured the same conditions during early infancy (Stanner *et al.,* 1997). However, the Dutch famine was much shorter which meant that there was more food available in the postnatal environment, and children exposed to poor nutrition *in utero* underwent accelerated weight gain (Roseboom *et al.,* 2000; Eriksson *et al.,* 2001). As a result, the incidence of T2D, cardiovascular disease and obesity in later life was much higher in those from the Dutch famine (Ravelli *et al.,* 1999; Roseboom *et al.,* 2000; Painter *et al.,* 2006; Stein *et al.,* 2007) when compared to individuals from the Leningrad famine, as they had lower cases of adult disease (Stanner *et al.,* 1997, Koupil *et al.,* 2007).

1.7.9 Birth weight and Animal Models

Animal models have been used to determine the effects of low birth weight through maternal undernutrition and the problems that arise within the offspring in the postnatal environment with respect to growth and development. These are important as they can aid in our understanding of the same/similar mechanisms that occur in humans.

Animal models show that the initial period after birth is important. In both rats and mice, those that were growth restricted *in utero* and exhibited post-natal catch up growth, developed obesity and diabetes (Jimenez-Chillaron and Patti, 2007, Shahkhalili *et al.,* 2010) and renal dysfunction (Boubred *et al.,* 2009) as mature adults. Mice show that catch-up growth within the first three weeks after birth impacts on adiposity and glucose

intolerance by six months of age. However, if catch-up growth occurs after three weeks, during weaning, no increased health risks are observed (Jimenez-Chillaron *et al.,* 2006). Nutritional programming has been suggested to affect normal hypothalamic neural organisation (Velkoska *et al.,* 2008). Maternal protein restriction during fetal and postnatal growth leads to an immaturity of the hypothalamus (Coupe *et al.,* 2011). Furthermore there were gene expression differences of neurodevelopmental process (cell differentiation and cytoskeleton organisation) in the hypothalamus of IUGR rats between birth and twelve days but early catch up growth up to five days after birth improved hypothalamic neurodevelopment (Coupé *et al.,* 2011). In addition, overfeeding in the postnatal development lead to persisted elevations of adipose tissue and plasma leptin levels leading to cardiac fibrosis in adulthood (Velkoska *et al.,* 2008).

Life expectancy is affected by catch up growth during the suckling period in rodents (Jennings *et al.,* 1999, Ozanne and Hales, 2004, Langley-Evans and Sculley, 2006). Changes in metabolic pathways are thought to be involved with the reduction in life span of rats born from a mother on a protein-reduced diet (Chen *et al.,* 2009b). Telomeres, repetitive nucleotide regions at the end of each chromosome to protect against deterioration, have been found to shorten in the pancreas and in the kidney in rats exposed to fetal maternal protein restriction and undergo catch up growth in early postnatal life compared to controls in two separate studies (Jennings *et al.,* 1999, Tarry-Adkins *et al.,* 2009). There were fewer longer telomeres and more, shorter telomeres compared to controls ($p<0.01$) (Tarry-Adkins *et al.,* 2009). In addition, those who were growth restricted *in utero* and underwent postnatal catch up growth had a shorter life span ($p=0.01$), while those experiencing restriction in postnatal life had longer kidney telomeres

($p<0.05$) and increased longevity ($p<0.01$). Whether this is an animal specific mechanism, further research is needed.

1.7.10 Predictive Adaptive Response

The 'predictive adaptive response' (PAR) is a theory, proposed by Gluckman *et al.,* (2005) (Gluckman *et al.,* 2005) is an extension of the 'fetal origins of adult disease'. It focuses on plasticity and the critical time points within fetal development that environmental cues can influence. The adaptations the offspring undergoes are of benefit in the future environment, not the immediate surroundings, to confer a survival advantage. More favourable conditions would be a disadvantage. Furthermore, it indicates maternal constraint evolved to promote survival and can influence the adaptive response if it restricts fetal growth more than normal. Our ancestors would have benefitted from the adaptations to the fetal metabolic physiology arising as a result of poor fetal environment, in a nutritional environment where food was scarce and availability was unpredictable. However, as the environment has improved substantially these adaptations now have adverse consequences. Additional factors include our increased life span, which used to be much shorter compared with our ancestors. It has increased from ~45 years of age in the 1840s to ~80 years of age at present (Westendorp, 2006). Along with this, there has been an increase in the age in which women bear offspring, with many more women having children over the age of thirty-five. Between 1990 to 2001 the number of women having children increased by 30%, 47% and 190% for the maternal ages thirty-five to thirty-nine years, forty to forty-four years and forty-five to forty-nine years respectively (Blickstein, 2003).

However this theory has been criticised. The PAR theory suggests that the adaptations the fetus undergoes is advantageous for survival in adulthood and increase reproductive

success but in humans reaching reproductive age is dependent on surviving infancy, therefore it has been suggested that the fetal adaptations are important for surviving infancy rather than adulthood (Rickard and Lummaa, 2007). The reliability of the prediction will decrease over time, as there is more chance for environmental conditions to change, in addition in previous generations without medical care 60-75% of born children survived into infancy. Furthermore humans have a much longer lifespan and the examples supporting this theory are in animals, which have a shorter life span reducing the chance for a difference in predicted and actual environments. Thus it is difficult to extrapolate these findings to humans (Rickard and Lummaa, 2007).

1.7.11 Genetics and Birth weight Variation and Disease Risk

Humans are genetically diverse: the human genome contains a range of common variants, including large numbers of copy number variations, insertions, deletions and SNPs (Shastry, 2009, Alkan *et al.,* 2011). Although there are Mendelian causes of diabetes and obesity, due to genetic mutations (Hinney *et al.,* 2010, Blakemore and Froguel, 2010, Steck and Winter, 2011), there is no evidence that a single genotype interacts with different environmental stressors to give an increased propensity to develop common forms of T2D and obesity in later life.

1.7.11.1 Epigenetics, Birth Weight and Disease Risk

Heritable changes that affects gene expression without modification in the DNA sequence is termed epigenetics (Bird, 2007) and the expression of alleles according to maternal and paternal origin (genetic imprinting) is mediated by DNA methylation and histone modification. Fetal growth factors are known to be important to fetal growth and a number

are imprinted (John and Surani, 2000, Fowden *et al.,* 2011), including insulin-like growth factor (IGF2), which is paternally expressed and insulin-like growth factor receptor (IGF2R), which is maternally expressed and are involved in the insulin-like growth factor system (Figure 1.7.4). Additional imprinted genes are the maternally expressed pleckstrin homology-like domain family A member 2 (*PHLDA2*) (Frank *et al.,* 1999, McMinn *et al.,* 2006) and pleiomorphic adenoma gene-like 1 (*PLAGL1*) (Frost and Moore, 2010) and the maternally expressed, non-protein coding, transcript H19. Expression of these imprinted genes has been associated with birth weight (Johnston *et al.,* 2003, Salas *et al.,* 2004, Apostolidou *et al.,* 2007, Diplas *et al.,* 2009, Adkins *et al.,* 2010, Koutsaki *et al.,* 2011, Petry *et al.,* 2011).

1.7.12 Genetic Influences of Birth Weight Variation

Genetic studies have shown associations between fetal birth weight and the VNTR within the insulin (*INS*) gene (Dunger *et al.,* 1998) and the maternal genotype of the glucokinase (*GCK*) gene, which is involved in the first rate-limiting step in glucose metabolism (Hattersley *et al.,* 1998, Weedon *et al.,* 2006). Furthermore the insulin VNTR has been associated with measures of obesity, BMI (Heude *et al.,* 2006), adiposity level (Heude *et al.,* 2004), and the metabolic syndrome and obesity in children (Santoro *et al.,* 2006). Mutations in the genes *INS*, *HNF4A, TCF2, PAX4* and *GCK* can cause a subset of diabetes known as maturity onset diabetes of the young (MODY) (Ellard and Colclough, 2006, Froguel and Velho, 1993, Motzkau *et al.,* 2012, Yamagata *et al.,* 1996). However, these mutations are relatively rare, as MODY accounts for 1-2% of all diabetes cases, of which 32% are caused by mutations in *GCK* and <1% by mutations in *INS* (Gardner and Tai, 2012). Therefore they do not contribute to common causes of low birth weight. Nonetheless it demonstrates that genetics is influential in fetal growth.

1.7.12.1 Fetal Insulin Hypothesis and Disease Risk

The fetal insulin hypothesis is an alternative explanation to the 'fetal origins of adult disease theory', suggesting that genetics is more influential in birth weight than the environment (Hattersley and Tooke, 1999). It proposes that variants in either fetal or maternal genes, involved with insulin secretion and/or sensitivity may affect fetal growth, leading to lower birth weight at term and predisposition to T2D, obesity and coronary artery disease in adulthood (Freathy *et al.,* 2009, Hattersley and Tooke, 1999). Insulin is a key fetal growth hormone and is involved in the insulin growth factor axis and is involved in the development of T2D in later life (Elijah *et al.,* 2011). This connection had stimulated the investigation of various T2D susceptibility genes/loci for an association with birth weight. A summary of the number of studies providing significant associations between several fetal T2D risk alleles and fetal phenotype is shown in Table 1.7.5, while their location in the genome is shown in Figures 1.7.1a and 1.7.1b. These associations have been found with both higher and lower birth weight and interestingly fetal growth characteristics *in utero*. This suggests that T2D susceptibility loci are also involved in fetal growth but the mechanism involved may be different as variants in different genes have been associated with either an increase or decrease in fetal phenotypes.

Higher birth weight at term additionally increases the risk of obesity, T2D and the metabolic syndrome and has been associated with higher BMI and WC in adults (Rillamas-Sun *et al.,* 2012, Zhao *et al.,* 2012) as well as lower birth weight. The risk of adult disease increasingly progresses during childhood and into adulthood (Dietz, 1998). Several loci have been associated with measures of obesity in both adults and children, including *FTO, MC4R, BDNF* and *MTCH2* (Frayling *et al.,* 2007b, Thorleifsson *et al.,* 2009, Willer *et al.,* 2009). However there has been a lack of association between these loci and birth weight,

even though those recently reported were large studies, indicating these loci have small or no effects on birth weight (Andersson *et al.,* 2010a, Kilpelainen *et al.,* 2011a). There was an initial small effect with *FTO* and high birth weight (p=0.013) and *MTCH2* (p=0.012) with low birth weight but this did not survive correction for multiple testing (Kilpelainen *et al.,* 2012). Additional evidence has shown that *FTO* is expressed in the placenta and is related to fetal growth (Bassols *et al.,* 2010, Sebert *et al.,* 2010).

Table 1.7.5: List of SNPs in key genes that have been significantly associated ($p<0.05$) with fetal phenotype at birth in European Caucasian populations. Risk genotype and p values are given for each of the studies. AC = abdominal circumference, FL = femur length, EFW = estimated fetal weight, HC = head circumference, [1] = 3rd trimester, [2] = 2nd trimester, PI = ponderal index.

Gene	SNP	Risk Allele	Fetal Phenotype	P Value	Author
ADCY5	rs9883204	CC	Lower birth weight	7×10^{-15}	Freathy *et al.*, 2010
			Smaller AC[1]	0.002	Mook-Kanamori *et al.*, 2011
			Smaller FL[1]	0.05	
			Smaller EFW[1]	2.5×10^{-4}	
	rs11708067	GG	Lower birth weight	0.004	Andersson *et al.*, 2010
CDKAL 1	rs4712523	GG	Lower birth weight	0.01	Andersson *et al.*, 2010
				0.01	Zhao *et al.*, 2009
	rs10946398	CC		2×10^{-5}	Freathy *et al.*, 2009
	rs7756992	GG		0.04	Andersson *et al.*, 2010
				8×10^{-5}	Zhao *et al.*, 2009
CCNL1/ LEKR1	rs900400	CC	Lower birth weight	2×10^{-35}	Freathy *et al.*, 2010
			Smaller HC[2]	0.05	Mook-Kanamori *et al.*, 2011
			Smaller FL[2]	0.02	
			Smaller HC[1]	0.02	
			Smaller AC[1]	8.85×10^{-4}	
			Smaller FL[1]	4.26×10^{-4}	
			Smaller EFW[1]	4.19×10^{-5}	
HHEX-IDE	rs1111875	TT	Lower birth weight	0.004	Freathy *et al.*, 2009
				8×10^{-5}	Andersson *et al.*, 2010
CDKN2 A/2B	rs10811661	TT	Higher birth weight	0.02	Andersson *et al.*, 2010
				0.03	Pulizzi *et al.*, 2009
JAZF1	rs864745	GG	Higher birth weight	0.02	Pulizzi *et al.*, 2009

TCF7L2	rs7903146	CC	Higher birth weight	0.001	Freathy *et al.*, 2007
MTNR1B	rs1387153	TT	Higher birth weight	0.02	Zhao *et al.*, 2009
GCK	rs1799884	AA	Higher birth weight	0.02	Weedon *et al.*, 2006
FTO	rs1421085	CC	Higher BMI	0.02	Cauchi *et al.*, 2009
			Higher PI	0.005	

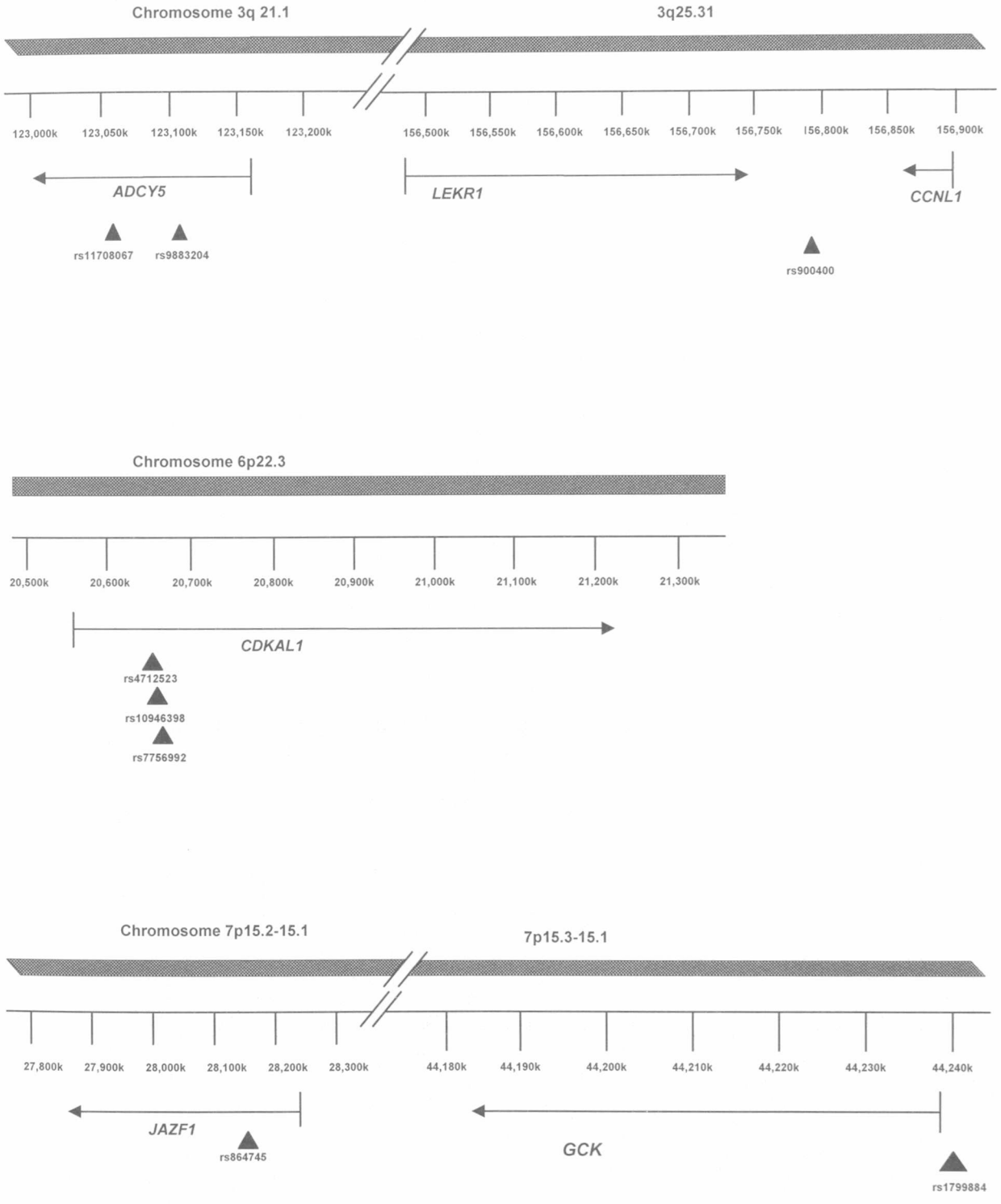


Figure 1.7.1a: Graphic representation of the location of six T2D susceptibility genes, *ADCY5, LEKR1, CCNL1, CDKAL1, JAZF1* and *GCK* and the SNPs that have been significantly associated with fetal phenotype (Table 1.7.11)

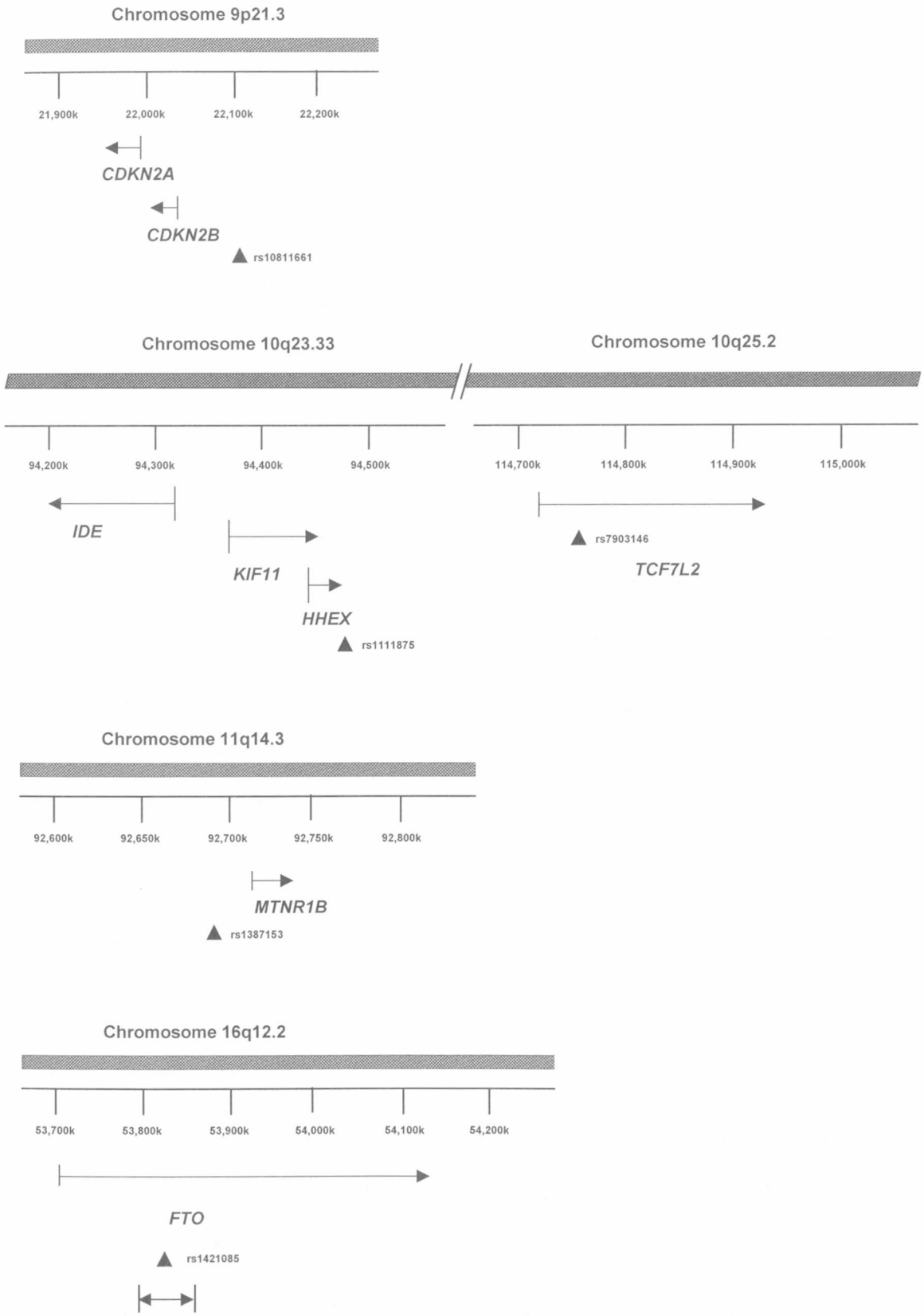


Figure 1.7.1b: Graphic representation of the location of six of the T2D susceptibility genes, *CDKN2A/2B, HHEX-IDE, TCF7L2, MTNR1B* and *FTO* and position of the variants that have been significantly associated with fetal phenotype (Table 1.7.11). The region highlighted by the arrow across part of the *FTO* gene represents the region that will be investigated with respect to fetal birth weight.

The mechanism by which sequence variants might be directly responsible for lower birth weight is yet to be determined, however one possibility is by influencing insulin secretion, as insulin is an important intrauterine growth factor. SNPs within *HHEX-IDE* and *CDKAL1* have previously been associated with decreased beta cell function (Grarup *et al.,* 2007, Pascoe *et al.,* 2007) and research into mouse embryos have shown that *HHEX* is expressed during development of the pancreas and *HHEX* null mice lack the ventral pancreas (Bort *et al.,* 2004). Pancreatic islets and beta cells express *ADCY5* (Dupuis *et al.,* 2010, Leech *et al.,* 1999), which encodes adenylate cyclase 5 and catalyses the generation of ATP, which is converted to cyclic AMP and pyrophosphate (Hanoune *et al.,* 1997). Associations have been found with *ADCY5* with increased fasting glucose level and increased risk of T2D in adults and children (Barker *et al.,* 2011, Dupuis *et al.,* 2010, Rees *et al.,* 2011, Vasan *et al.,* 2011, Windholz *et al.,* 2011). Variants in *HHEX-IDE*, *CDKAL1* and *ADCY5,* may influence fetal insulin secretion *in utero* resulting in low birth weight and the development of T2D in later life. Further investigation is needed to determine the exact role as the function of *FTO* is yet to be elucidated. Possible roles include an involvement in DNA repair, nucleic acid modification (Gerken *et al.,* 2007, Sanchez-Pulido and Andrade-Navarro, 2007) and regulation of gene expression (Wu *et al.,* 2010).

In addition, further investigation is needed to determine if the true variants have been identified. *ADCY5* is located within a region of high LD therefore it is possible that the true birth weight lowering risk allele has not been identified. The *HHEX-IDE* locus covers a region of over 200 kb, and the SNP rs900400 lies between *CCNL1* and *LERK1*. The function of the latter is currently unknown whereas *HHEX* is a regulator of RNA polymerase 2 transcription and *IDE* is an enzyme involved in insulin degradation. The SNP

is also 800 bp downstream from an uncharacterised transcript region that is highly conserved (Ryckman *et al.,* 2012).

It is possible that the genes involved with BMI are different to those involved with fetal growth as there is a lack of association between them. Furthermore, it is also possible that the timing of the effect of these genes on weight/birth weight is important. Both *HHEX-IDE* and *CDKAL1* have been associated with BMI in childhood in two separate studies (Winkler *et al.,* 2010, Zhao *et al.,* 2010), and they have previously been associated with birth weight (Freathy *et al.,* 2009; Zhao *et al.,* 2009; Andersson *et al.,* 2010). The opposite association of *FTO* and *MTCH2* on birth weight suggests a possible regulatory mechanism/role to control fetal growth within normal maternal constraints.

Additional variants in T2D susceptibility genes have not been found to be significantly associated with birth weight, which makes the loci associated with lower birth weight (Table 1.7.11), interesting candidates to be investigated further, including *HHEX-IDE* and *ADCY5.* The majority of studies between the fetal loci with birth weight have been conducted in European Caucasians, in which strong significant associations have been found with fetal birth weight therefore the role of the fetal and maternal loci in fetal growth in ethnic minorities in an obesogenic environment, is of particular interest. While research into *FTO* and birth weight has produced inconclusive results including from different ethnicities therefore further investigation is needed to clarify its role in birth weight. As a result, this makes *HHEX-IDE*, *ADCY5* and *FTO* interesting candidates to be investigated in the maternal South Asian population. Variants in these genes have previously been associated with T2D and obesity and the risk of developing these diseases is higher in this ethnic population in an obesogenic environment, compared to European Caucasians,

therefore these variants may also contribute to fetal growth. Research has found no association with variants in *LEPR* and fetal birth weight (Rand *et al.,* 2001) so it is not thought to contribute to fetal growth therefore it was not included in genotyping with the other three genes.

1.7.13 Birth Weight Variation and Disease Risk in Ethnic Populations

One aspect that needs to be investigated further is that the majority of studies into the relationship between birth weight and risk of adult disease have been conducted in European Caucasian populations. Many ethnic minorities are at increased risk of adult disease and South Asians are one such group. Development of T2D occurs within a much younger age group and this occurs at a lower abdominal obesity classification compared to European Caucasians.

One possibility is that the Indian 'thin-fat' (thin on the outside, fat on the inside) phenotype develops *in utero* indicating a different aetiology between the two populations. Studies have found that South Asian (SA) babies are smaller and lighter at birth compared to white Caucasians (Modi *et al.,* 2009, Yajnik, 2002) and they have a significantly higher volume of visceral (0.012 to 0.023 L, $p=0.001$), and subcutaneous adipose tissue (deep (0.003 to 0.017 L, $p=0.006$), superficial (0.006 to 0.043 L, $p=0.011$)) (Modi *et al.,* 2009). They have relatively less muscle and viscera and significant differences have been found in mid arm circumference ($p<0.001$) and tricep skinfold ($p<0.001$) (Yajnik *et al.,* 2002). The thin-fat phenotype has been found to persist from birth to four years of age with subscapular skinfold thickness significantly larger in Mysore infants compared with white Caucasians ($p<0.001$) indicating higher truncal adiposity. Other measurements including height,

weight, head circumference and body mass index were all found to be smaller in SAs ($p<0.001$) compared to four year old Caucasians (Krishnaveni, 2005).

Additionally immigrant Indians who have adopted a Westernised lifestyle have a higher birth weight than native Indian babies (Harding *et al.,* 2004), suggesting environmental cues may have a stronger effect than genetics. In any case, this further supports the idea that additional research is needed to identify the mechanisms involved in birth weight irrespective of ethnicity.

1.8.0 Candidate Genes

Many candidate genes have been identified for a role in the development of obesity, with those having a functional role in energy regulation being of particular interest, notably leptin and the leptin receptor. In addition, birth weight is a risk factor for obesity, T2D and the metabolic syndrome and genes involved in obesity and T2D have been associated with size at birth, indicating there maybe a complex interaction of genes involved with birth weight and weight gain through postnatal life into adulthood. Leptin is also known to be involved in the reproductive system, fertility and expressed during pregnancy. As a result, polymorphisms within three candidate genes will be investigated with respect to birth weight, *HHEX-IDE*, *FTO* and *ADCY5*, while functional implications of polymorphisms in *LEPR* will be investigated.

1.9.0 LEPR and Its Ligand Leptin (LEP)

LEPR and its ligand leptin are candidates for obesity through their well-established role in monogenic obesity (Clement *et al.,* 1998, Montague *et al.,* 1997, Strobel *et al.,* 1998). However, as these mutations are rare within the general population they cannot explain the

rapid rise of polygenic obesity but it indicates that these genes may still be involved to a lesser degree (Rankinen *et al.,* 2006). Significant associations in large GWAS have not replicated the associations of variants in these two genes (Ben Ali *et al.,* 2009, Yang *et al.,* 2011, Yiannakouris *et al.,* 2001). Further research is needed to fully understand the mechanisms involved and their implications not only in energy regulation but also in reproduction and the immune system, as leptin is also important in these systems.

1.9.1. Identification of Leptin and Leptin Receptor using Mouse Models

The identification of leptin and subsequently its receptor marked a turning point into the genetic causes of obesity. It had been hypothesised for years that there was a circulatory 'factor' produced to regulate energy intake, since the identification of the spontaneously obese mouse (*ob/ob)* and the later discovery of the *db/db* mouse, in the late 1940's/early 1950's, which had been used as animal models into human obesity. The *ob/ob* mouse fails to make the circulating factor and the *db/db* mouse is unable to respond to it. However the cause of the obese phenotype still remained elusive even though extensive biochemical, physiological and nutritional studies were performed. The breakthrough occurred in 1994 when leptin was identified as the circulating factor through positional cloning (Zhang *et al.,* 1994), which was a new technique, first used in the late 1980s, to identify candidate genes (Collins, 1992). The factor was termed leptin from the Greek 'leptos' meaning thin. A year later the leptin receptor was also cloned (Tartaglia *et al.,* 1995).

Other rodent models that develop the obese phenotype due to single gene mutations include the fat mouse, (*fa/fa*), the tubby mouse (*tub/tub*), the yellow agouti mouse (*agouti*) and the Zucker rat (*fa/fa*); of these the fat mouse is of importance as it occurs due to mutations in the leptin receptor (Trayhurn *et al.,* 1999).

1.9.2 Mutations within the Human forms of Leptin and Leptin Receptor

These findings within the mouse models led to the identification of human forms of leptin and the leptin receptor and mutations within these human orthologues were subsequently detected. Mutations within leptin and leptin receptor genes are mainly identified in those with early onset obesity. Subjects have a normal birth weight but soon develop obesity within three to four months, through constantly eating due to hyperphagia and have a high percentage of body fat.

Two different mutations within leptin have been described in two, separate, consanguineous families. In one of the families, a C to T base substitution at codon 105 results in a non–synonymous arginine to tryptophan change within the leptin protein. Three individuals homozygous for this mutation have been found where their BMI was >30 kg/m^2, circulating leptin levels were low and they were hyperphagic (Strobel *et al.,* 1998). Other members of the family were either heterozygous or homozygous for the wild type with normal weight and circulating leptin levels.

Within a consanguineous Pakistani family, a frame-shift mutation was identified in three siblings, in codon 133, resulting from the deletion of the G nucleotide, altering the reading frame and leading to fourteen subsequent different amino acids and the introduction of a premature stop codon (Montague *et al.,* 1997). The resulting protein is unable to be secreted and remains within the cell due to misfolding and is consequently degraded by the proteasome (Rau *et al.,* 1999). These subjects were found to have a high BMI and body fat with low levels of circulating leptin. Family members without the obese phenotype were either heterozygous or homozygous for the wild type undeleted allele.

Only one mutation in the leptin receptor has been identified so far, that gives rise to monogenic obesity, which has been identified within three siblings in a consanguineous family of Kalibian descent. Direct nucleotide sequencing showed a G to A base change in the +1 position of intron sixteen (one base after the 3′ end of exon sixteen) and expression analysis showed an abnormal LEPR, which lacks exon sixteen. This results in the LEPR protein lacking the transmembrane and intracellular domains (Clement *et al.,* 1998). The truncated leptin receptor produced is similar to the short isoform and binds leptin promoting its half-life, contributing to the high levels of circulating leptin. Individuals with this mutation have a high weight and circulating leptin levels.

Individuals with early onset obesity resulting from mutations in leptin were also found to have either hypogonadotropic hypogonadism with delayed puberty or primary amenorrhoea (Strobel *et al.,* 1998). Upon administration of recombinant leptin, fertility is restored within *ob/ob* mice (Chehab *et al.,* 1996) and puberty has been shown to begin in young women with defects in either leptin or leptin receptor, also upon administration of recombinant leptin (Clement *et al.,* 1998, Farooqi *et al.,* 1999).

These studies on leptin and the leptin receptor provide evidence that leptin is involved within energy balance but it was soon discovered that these mutations were rare within the general population, indicating that other genetic polymorphisms within energy regulation are involved in the development of common obesity.

1.9.3 Leptin Expression and Regulation and its Role in the Adipostat

The main site of expression of leptin is within the adipose tissue where it acts as a satiety factor by relaying the level of stored adipose to the hypothalamus to regulate food intake

and energy expenditure via the leptin-melanocortin system. Leptin receptor expression has also been found in these locations, indicating that leptin acts not only centrally but also peripherally, by increasing glucose and lipid metabolism independently. Other sites of expression identified include stomach, skeletal muscle, osteoblasts and the placenta (Fruhbeck, 2006, Trayhurn and Bing, 2006).

Leptin secretion appears to be highly regulated over twenty-four hours. Minimum levels of leptin occur in the morning between 8 am and 12 pm, then rise throughout the rest of the day peaking between midnight and 4 am. High leptin levels reduce the desire to eat and increase energy expenditure whereas low leptin levels increase food intake and reduce energy expenditure (Sinha *et al.,* 1996). These levels are indicative of requiring food during the early and middle part of the day, particularly after fasting during sleep and less so during the later part of the day.

With regards to measuring leptin levels in women, adjustments need to be made, as levels are much higher in females compared to males and vary with respect to their menstrual cycle (Hardie *et al.,* 1997, Quinton *et al.,* 1999). Leptin levels rise slightly during the follicular phase (days one to fourteen) then rise further with the surge in levels of the luteinising hormone and during the luteal phase (days fifteen to twenty-eight). Many studies use postmenopausal women to minimise these hormonal effects.

1.9.4 The Leptin–Melanocortin System

Leptin is released from the adipocytes into the blood where it is transported to the brain. The signal is received by the leptin receptor on the arcuate nucleus, which is one of the

nuclei within the hypothalamus, the ventro-medial and paraventricular being the other two, and it is the most important component regarding energy balance. Within the arcuate nucleus there are two sets of neurons: agouti–related protein (AGRP) and neuropeptide Y (NPY) plus pro–opiomelanocortin (POMC) and cocaine and amphetamine related transcript (CART). AGRP and NPY produce orexigenic signals to promote food intake and reduce energy expenditure via downstream effector neurons. POMC and CART neurons produce anorexigenic signals to reduce food intake and increase energy expenditure (Bell *et al.,* 2005b). Other hormones and neural signals involved in short to long-term appetite regulation include peptide YY_{3-36}, ghrelin and insulin (Figure 1.9.1).

Other single gene mutations have been described within this system resulting in the obese phenotype, the affected genes being *POMC, MC4R, NPY* and *NPYR2*. This indicates that this is an important pathway involved in energy regulation and although mutations within these genes are rare, less harmful gene variants could contribute to common obesity.

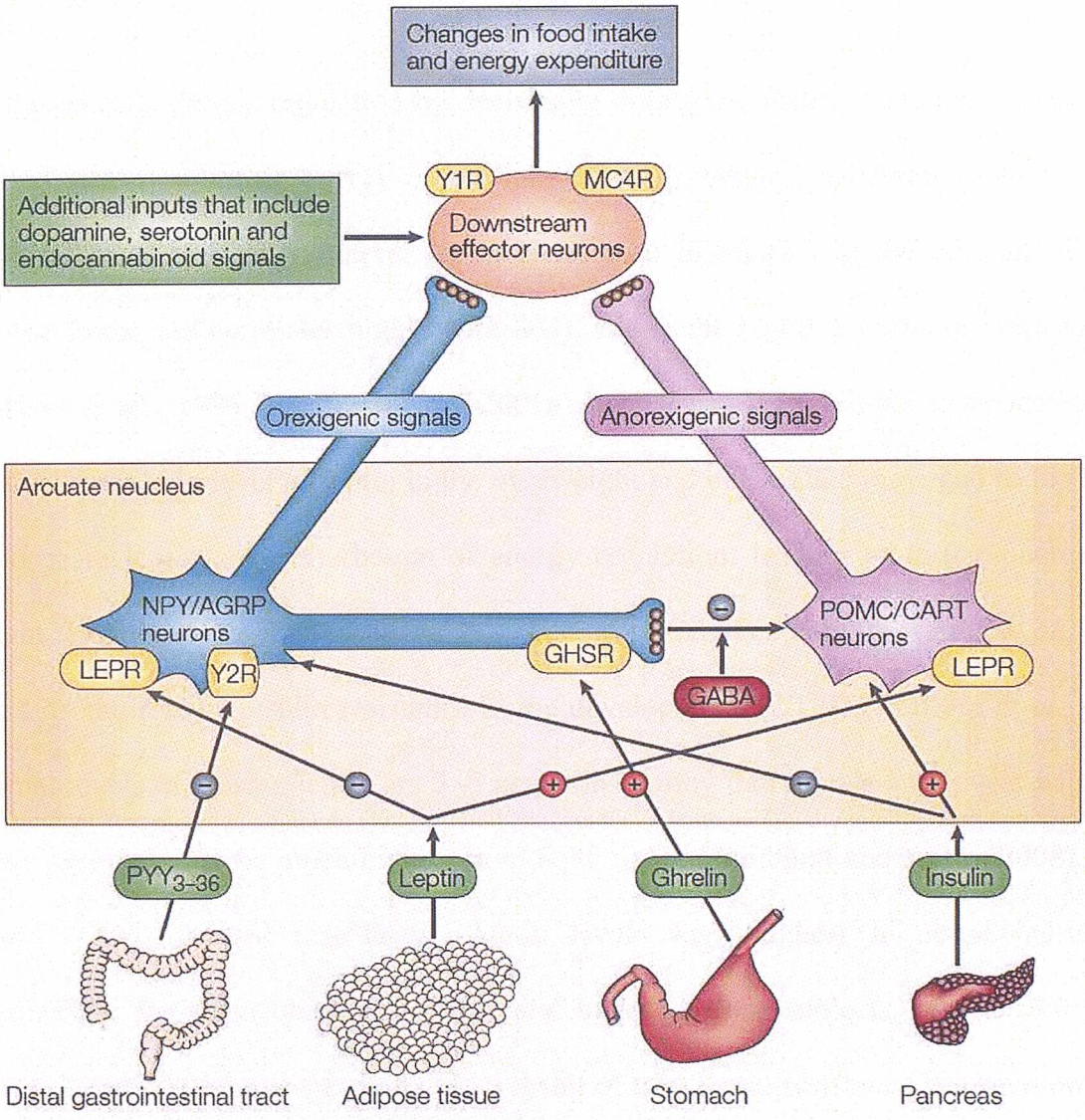


Figure 1.9.1: The leptin – melanocortin pathway. The pathway indicates the hormones that are involved in activating the two sets of neurons involved in reducing and increasing food intake and energy expenditure. Used with permission from one of the authors from Bell *et al.*, 2005.

1.9.5 Leptin Resistance

Leptin controls energy regulation by decreasing during the fasting state and increasing in the fed state, promoting energy expenditure. During weight maintenance, when energy intake is equal to energy output, leptin is released in relation to the amount of stored adipose tissue and correlates highly with BMI, due to the higher volume of adipose tissue (Haffner *et al.,* 1996, Wauters *et al.,* 2001a), which has been termed hyperleptineamia. This constant high level of leptin in the overweight and obese state may lead to the loss of response to leptin and perturbation of energy regulation, leading to further increases in weight. The phrase 'leptin resistance' has been proposed to explain this (Arch *et al.,* 1998), based on the role of insulin resistance in the development of T2D (Belfiore *et al.,* 1979). Normal levels of plasma leptin are 3–5 ng/ml in healthy individuals and levels within the obese were found to be much higher, from 8–90 ng/ml (Anubhuti and Arora, 2008). Sinha *et al.,* (1996) showed that fasting leptin levels were highest in obese individuals, intermediate for those obese with T2D, and lowest in lean subjects, at 41.7±9.0 ng/ml, 30.8±6.7 and 12.0±4.4 respectively. As a result of this, leptin resistance has been proposed to explain the loss of response of leptin to regulate energy homeostasis (Anubhuti and Arora, 2008). Though the mechanism is unknown it is thought it is a combination of resistance at the receptor, (the level of signalling) and a decreased ability of the blood brain barrier to transport leptin (Anubhuti and Arora, 2008). Support for both of these mechanisms have been found. Animal studies have shown that obese rats with high levels of leptin only respond to the hormone when it is given centrally not peripherally (Halaas *et al.,* 1997, Ramsey *et al.,* 1998, Van Heek *et al.,* 1997). A further study has shown a failure of leptin to activate the STAT3 signal pathway in the hypothalamus (Munzberg *et al.,* 2004).

Despite leptin resistance and hyperleptinaemia being considered key contributing factors in the development of the obese state, the importance and use of the term 'leptin resistance' has recently been questioned (Myers *et al.,* 2012). There are several reasons for this. Firstly, leptin resistance has been suggested to be selective (Mark *et al.,* 2002; Anubhuti and Arora, 2008; Singh *et al.,* 2010), present only in the peripheral circulation rather than centrally, since administration of leptin centrally can still lead to signal activation (Ramsey *et al.,* 1998, Van Heek *et al.,* 1997) and leptin is involved in increasing renal sympathetic nerve activity which is part of the autonomic nervous system control (Singh *et al.,* 2010). High levels of leptin are detrimental to the health of the obese individual and are associated with additional health complications, such as cardiovascular disease, T2D and high levels of cytokines, which lead to what has been termed a proinflammatory state. The leptin resistance argument does beg the question: if high levels of leptin lead to leptin resistance why does that occur only in the hypothalamus bearing in mind that the receptor is expressed in many peripheral tissues (spleen, kidneys, pancreas and adipocytes) and the sympathetic nervous system is still activated? The likely explanation is that there is a dual aspect to leptin resistance: a high level of the cytokine leads to loss of the central energy regulation function with an opposite effect peripherally. High levels lead to the pro-inflammatory state, increased activation of the sympathetic nervous system and to the development of cardiovascular disease. Thus, there could be more than one underlying mechanism in the development of leptin resistance, leading to the obese phenotype explaining the contradictory evidence between polymorphisms in the leptin receptor and obesity. This dual effect could be mediated by the six isoforms of LEPR (see section 1.9.9) as the soluble isoform transports leptin across the blood-brain barrier and the long isoform is involved in cell signalling.

Leptin is a pleiotrophic cytokine, as it is not only involved in energy regulation but reproduction, the cardiovascular system and in immunity. Therefore, any definition of leptin resistance needs to take this into account. High levels of circulating leptin are involved in the metabolic complications of obesity and its sequelae: T2D, cardiovascular disease, hypertension, cancer and the pro-inflammatory state that arises as a consequence of higher levels of adipose tissue, through cell size and number. Leptin is suggested as an important link between obesity and the development of cardiovascular disease (Singh *et al.,* 2010) as it affects blood pressure, platelet aggregation, formation of arterial thrombosis and inflammatory vascular response (Beltowski *et al.,* 2002, Bodary *et al.,* 2002, Chaldakov *et al.,* 2001, Cooke and Oka, 2002, Konstantinides *et al.,* 2001). Furthermore, the pro-inflammatory state is likely to have a significant rolesince leptin induces the inflammatory cytokines TNF-α and IL-6 (Bastard *et al.,* 2006) and is associated with high levels of adipose tissue found in obese individuals. Adipose tissue also secretes many more adipokines, which have been linked to the modulation of immune function and inflammation either a directl and indirect way (Hersoug, 2007). This suggests that higher levels of adipose tissue not only increases the production of leptin but also additional adipokines (IL-6, IL-8, C-reactive protein and plasminogen activator inhibitor-1) and these are also implicated in both in the pro-inflammatory state (Esposito *et al.,* 2002, Festa *et al.,* 2001) and with insulin resistance (Cinti, 2012) that develop as a result of obesity. Therefore adipose tissue is not just confined to leptin resistance. Indeed, adipose tissue can be infiltrated by macrophages, which produce the pro-inflammatory cytokines tumour necrosis factor-α (TNFα), interleukin-6 (IL-6) and interleukin-1α (IL-1α). These cytokines can interfere with insulin receptor signalling (Gregor and Hotamisligil, 2011), which can impact on insulin resistance. So does leptin resistance lead to obesity or does obesity lead to leptin resistance? It is difficult to distinguish between mechanisms, which

predispose to weight gain and those that occur as a result of it (Myers *et al.,* 2010). In addition to other mechanisms discussed, cellular pathways may be mediators in the development of leptin resistance through attenuation of leptin receptor signalling.

1.9.6 Leptin Receptor

After the identification of leptin and its similarity to the cytokine family, the leptin receptorwas found to belong to the class I cytokine receptor family, which includes the receptors for interleukin-6, granulocyte colony stimulating factor and leukaemia inhibitory factor (Tartaglia *et al.,* 1995). It is located on chromosome 1p31 across 220.995 kb and contains twenty exons. Six isoforms of the leptin receptor are known to be produced and can be classified into short, long or secreted classes depending on their structure. The isoforms all contain the same extracellular and intramembrane domains with the main difference being the intracellular domain, which differs in size. LEPRa, LEPRc, LEPRd and LEPRf contain short structures. LEPRe is the secreted form that lacks the intracellular domain and its main function is to transport circulating leptin across the blood brain barrier. LEPRb is the longest isoform, which mediates the effects of leptin (Figure 1.9.2). The *LEPR* is expressed ubiquitously in a variety of different tissues ranging from the hypothalamus to the kidneys, pancreatic ß–cells, adipose tissue, ovaries and placenta (Trayhurn *et al.,* 1999), which indicates the pleiotropic function of leptin.

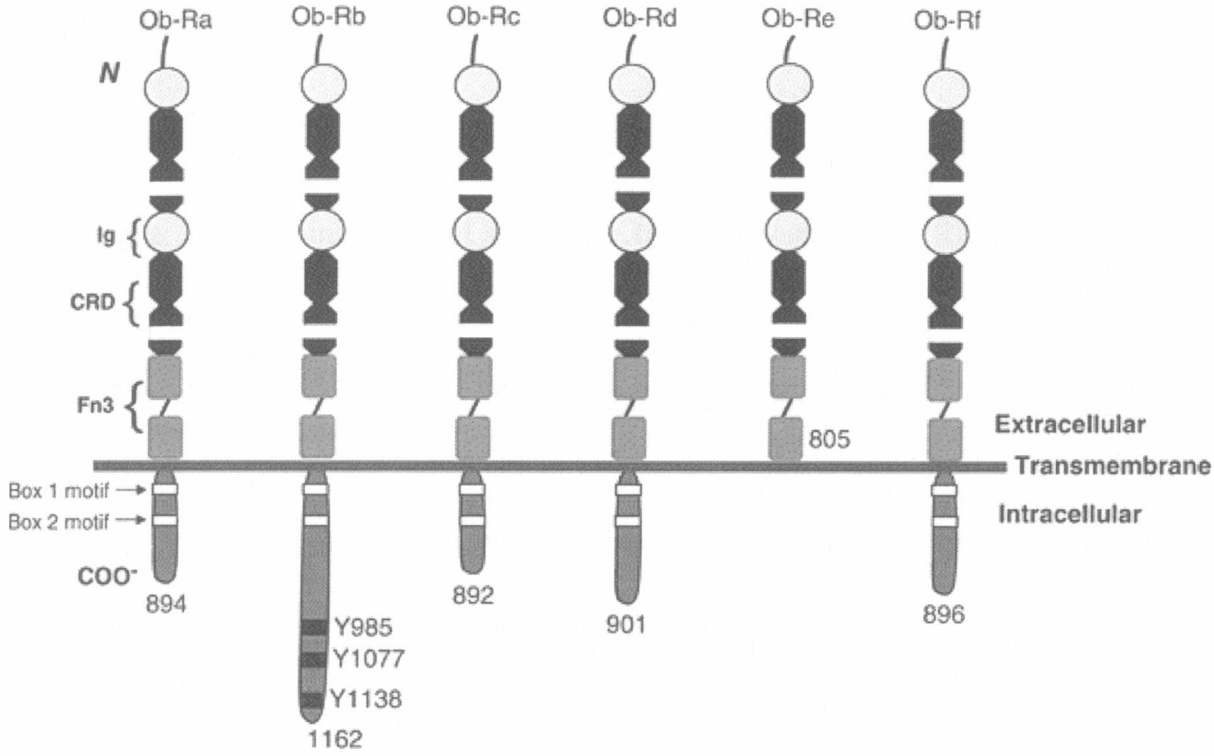


Figure 1.9.2: Image depicting six of the leptin receptor isoformsand their location across the plasma membrane. The three domains, Ig-like, CRH2 and FNIII, are also shown. LEPRe (Ob-Re) is the soluble form of the receptor, which lacks the transmembrane domain. LEPRb (Ob-Rb) is the long receptor with LEPRa (Ob-Ra) and LEPRc (Ob-Rc) slightly smaller isoforms. Box 1 situated within exon seventeen is required for Jak binding. Within exon eighteen Box two and STAT3 are required for STAT signalling. Located on the Rb isoform are the tyrosine residues important for signalling. Taken from Ceddla, 2005 (Ceddia, 2005).

1.9.6.1 Leptin Receptor Extracellular Region

The extracellular domain binds its ligand, leptin and can be further divided into subdomains: at the amino terminal a cytokine receptor homology (CRH) domain (CRH1), followed by immunoglobulin (Ig)-like domain, a second CRH domain (CRH2) with two fibronectin type III (FNIII) folds (Figure 1.9.3). Interestingly, the CRH1 and CRH2 sub-domains have also been shown to contain two separate fibronectin type III folds (Haniu *et al.*, 1998), which are found in a variety of extracellular proteins within the cytokine 1 superfamily and are involved in cell surface signalling.

The CRH2 domain is important for the high affinity binding of leptin (Fong *et al.,* 1998, Sandowski *et al.,* 2002, Zabeau *et al.,* 2004) whereas the Ig-like and FNIII domains are critical in LEPR activation (Zabeau *et al.,* 2004; Fong *et al.,* 1998). The function of the CRH1 domain remains elusive as it is less conserved compared to the other domains and upon deletion there is no detrimental effect on activation (Peelman *et al.,* 2006). Leptin, like many other cytokines that belong to the gp130 superfamily, interact with their receptors by forming a complex through three different binding sites: I, II and III (Peelman *et al.,* 2004). Binding site II occurs between the CRH2 domain of the receptor and the A and C helices of leptin whereas binding site III occurs between the Ig-like domain of the receptor and the N-terminus of helix D of leptin. Leptin shares similarities in structure with the cytokines IL-6 and the granulocyte colony-stimulating factor (G-CSF), which have receptors that belong to the same type I cytokine receptor family as LEPR. The G-CSF has been found to form a 2:2 tetrameric complex with its receptor, based on the three binding sites (Layton *et al.,* 2001) while IL-6 forms a hexameric complex with two IL-6 molecules, two gp130 chains and two IL-6 α-receptors (Peelman *et al.,* 2006). Due to the stoichiometry of the signalling complex it is thought that leptin forms a similar hexameric

complex with its receptor (Figure 1.9.4) (Zabeau *et al.,* 2003). Building a hexameric model of the leptin and LEPR complex, the conserved amino acids Lys^{13}, Lys^{86}, Lys^{89} and Phe^{92}, which are situated within the A-B helix loop of leptin, fit within the CRH2 domain of the LEPR (Figure 1.9.5) and form the centre of binding site I (Iserentant *et al.,* 2005). Furthermore, a cluster of amino acids within the Ig-like domain, Leu^{370}, Ala^{407}, Tyr^{409}, His^{417} and His^{418} (Figure 1.9.6), forms the main centre of binding within binding site III, and Ile^{501}, Phe^{502}, Leu^{503}, Leu^{504}, Ser^{505}, Asp^{615} in CRH2 (Figure 3.3.2) are located at the centre of binding site II. Mutations within this region affect the affinity for leptin to bind (Iserentant *et al.,* 2005; Peelman *et al.,* 2006).

As leptin is a pleiotropic hormone that is involved with control of energy regulation but also with immunity, inflammation, cardiovascular system and cancer, development of agonists and antagonists are important, however little is known about the structure of LEPR and the complex with leptin (Carpenter *et al.,* 2012, Iserentant *et al.,* 2005). Investigation of how leptin binds to its receptor has been performed through the use of computer modelling using similarities in structure between leptin and other cytokines, notably gp130 and G-CSF, as well as LEPR and other members of the class 1 cytokine receptor superfamily.

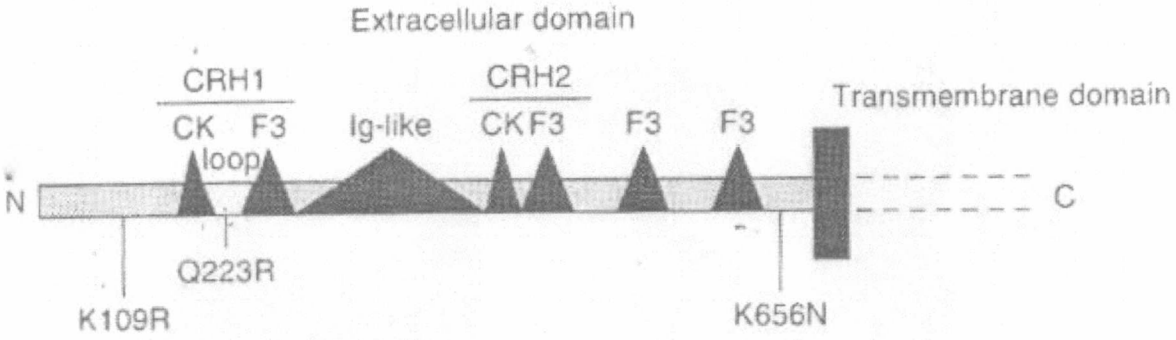


Figure 1.9.3: A graphic representation of the extracellular domain of the leptin receptor. Also located are the sub-domains and the three polymorphisms Lys109Arg (K109R), Gln223Arg (Q223R) and Lys656Asn (K656N). Taken from Stratigopoulos *et al.,* (2009) (Stratigopoulos *et al.,* 2009)

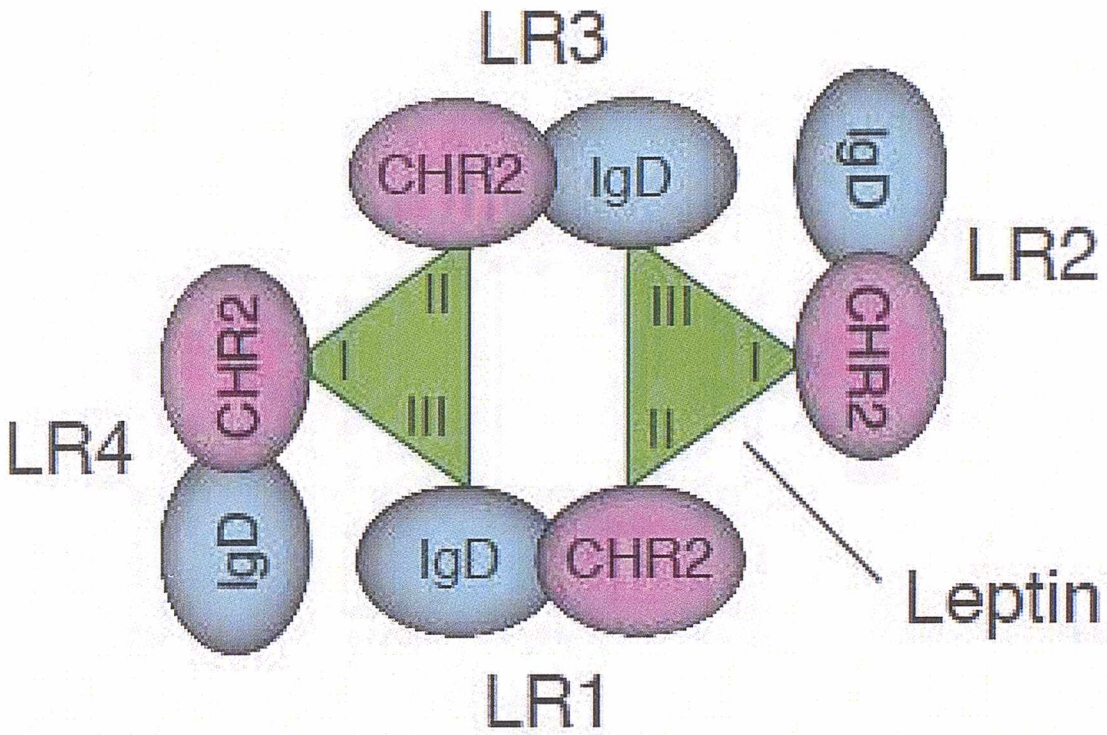


Figure 1.9.4: Graphical representation of the hexameric leptin and leptin receptor (LR) complex. There are four leptin receptor molecules (two as dimers) with two molecules of leptin. The three binding sites are shown as green triangles, which show where two binding sites are in relation to leptin and leptin receptor molecules. Only the Ig-like domain (IgD) and CRH2 domain of the extracellular region of the leptin receptor is shown. Taken from Peelman *et al.*, 2006 (Peelman *et al.*, 2006a).

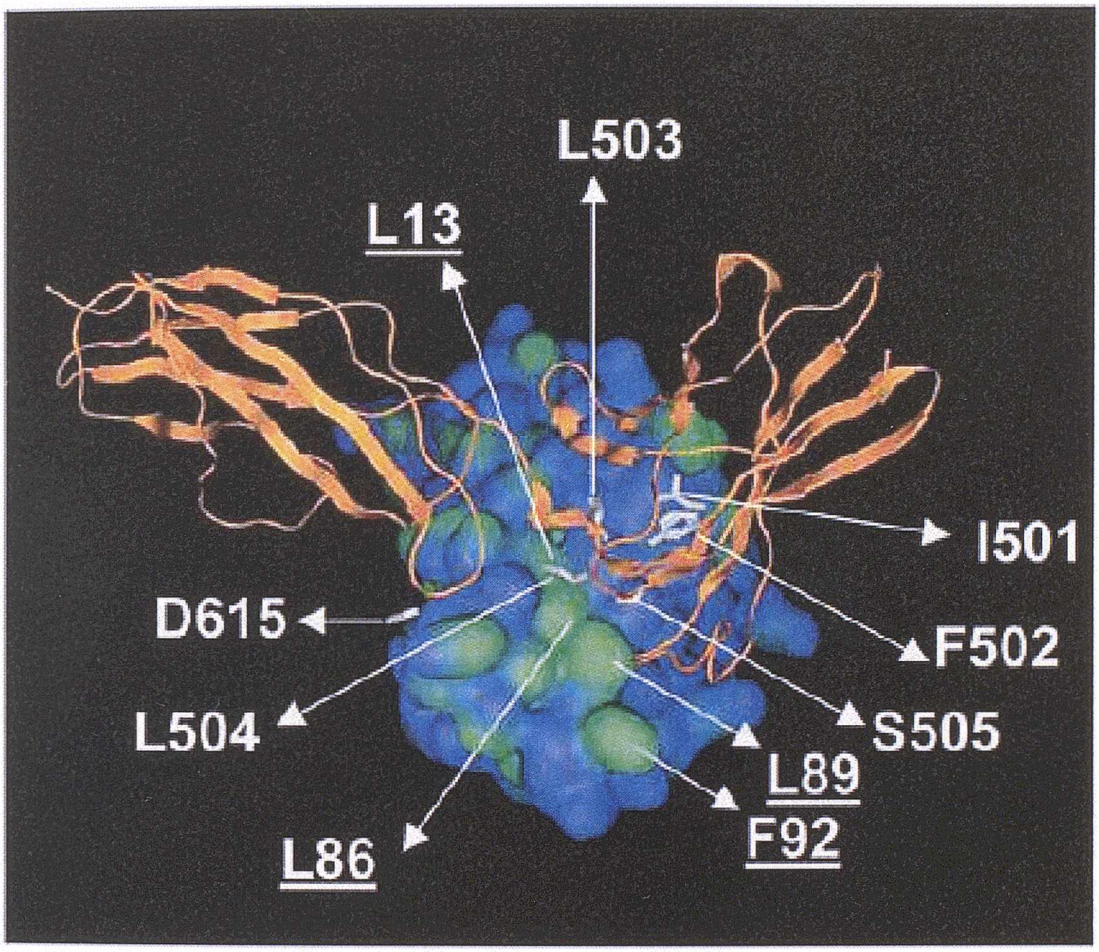


Figure 1.9.5: The complex of the CRH2 domain and mouse leptin shown as ribbon model for CRH2 and leptin molecule. The surface leptin molecule displays hydrophobicity (blue indicates hydrophilic areas while green indicates hydrophobic areas). Shown in white are the amino acid residues that are important in binding between the two molecules, with those underlined located in leptin molecule and the white sticks indicate side chain residues in the CRH2 domain. Taken from Iserentant *et al.*, 2005.

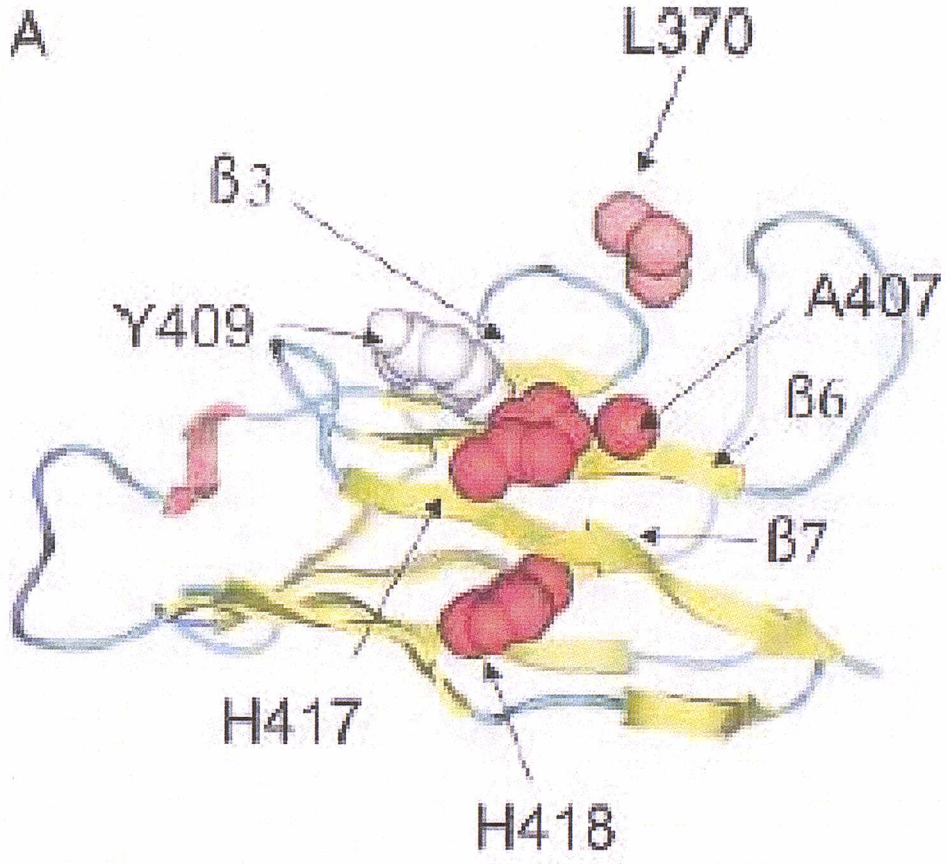


Figure 1.9.6: Secondary structure ribbon model of the Ig-like domain. It displays amino acid residues that affect leptin receptor activity when they are mutated. Taken from Peelman *et al.*, 2006b.

1.9.7 Leptin Receptor Signalling Pathways

Of all the isoforms of the leptin receptor, LEPRb has been studied the most since it mediates the physiological functions of leptin notably metabolic rate, reproduction and growth (Louis and Myers, 2007). Through this isoform leptin has been shown to signal using multiple pathways. These include the Janus tyrosine kinase 2 (JAK2)/cytosolic signal transducer and activator of transcription protein 3 (STAT3) pathway, the mitogen–activated protein kinase (MAPK) cascade, the phosphoinositide 3-kinase/phosphodiesterase 3B/cAMP pathway and 5'-AMP-activated protein kinase pathway (Figure 1.9.7).

1.9.7.1 LEPR and the JAK/STAT Pathway

Activation of this pathway is initiated by the binding of leptin to LEPRb, four tyrosine residues located on the intracellular domain of the receptor are phosphorylated following the activation of constitutively associated JAK2, which undergoes autophosphorylation motifs located on intracellular domain (Banks *et al.,* 2000, Kloek *et al.,* 2002). Once activated, the tyrosine residue at position 1138 recruits the STAT3 transcription factor, which is subsequently translocated to the nucleus where it binds to specific DNA sequences that activate various genes involved with energy regulation (Vaisse *et al.*, 1996). These include POMC and the signalling inhibitor, suppressors of cytokine signalling three (SOCS3). SOCS3 is involved with the negative regulation of leptin. The translated protein of SOCS3 binds to the 985 phosphorylated tyrosine residue, on the intracellular domain of the leptin receptor, which inhibits further signalling.

1.9.7.2 LEPR and the Phosphoinositide 3-kinase/Phosphodiesterase 3B/cAMP Pathway

This pathway is involved in insulin function and glycogen synthesis and is mainly activated via the binding of insulin to its receptor. Activation is initiated by the phosphorylation of an insulin receptor substrate leading to the activation of phosphoinositide 3-kinase (PI3K). Research has shown that leptin can also activate the pathway, suggesting there is cross talk between this pathway and the leptin signalling pathway (Niswender and Schwartz, 2003), as leptin is now considered to be involved in glucose homeostasis and controlling insulin sensitivity in perpherial tissue (Morton and Schwartz, 2011). Phosphorylation of JAK2 can phosphorylate the insulin receptor substrates one and two.

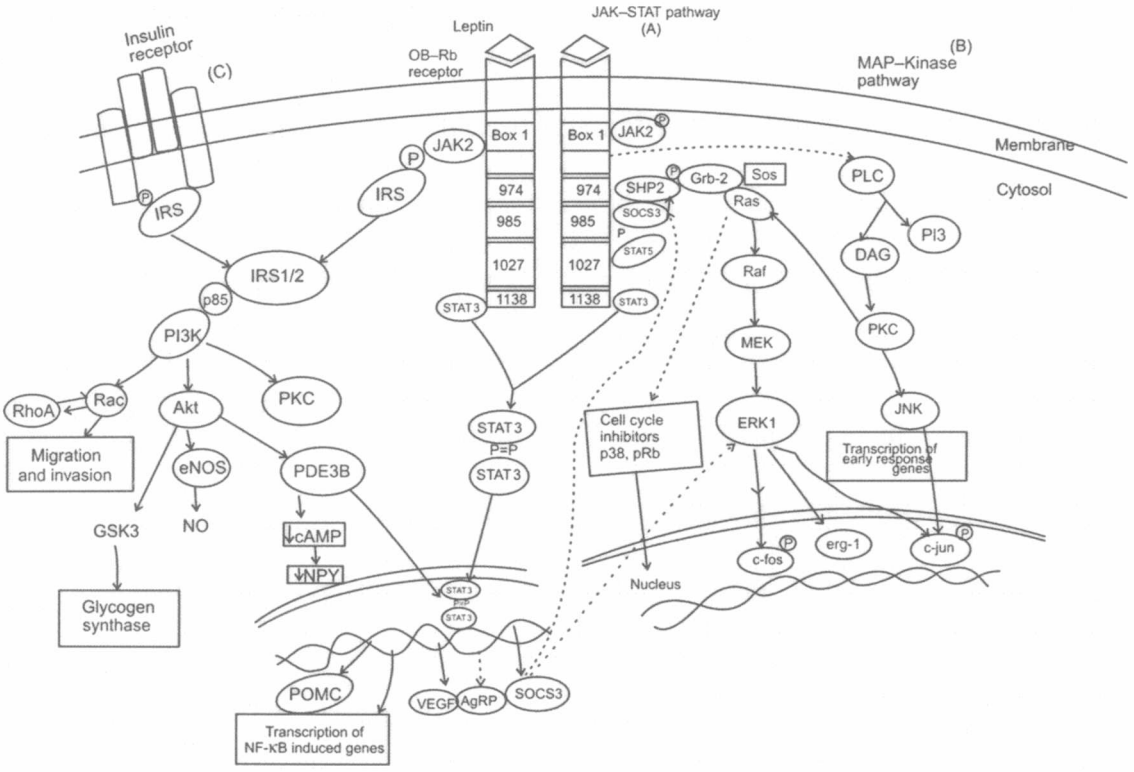


Figure 1.9.7: Activation of the JAK/STAT pathways by leptin binding to its receptor. Taken from Anubhuti and Arora, 2008.

1.9.8 Leptin Receptor Polymorphisms

Since the identification of the *db/db* mouse (Tartaglia *et al.,* 1995), only one polymorphism within the leptin receptor is known to result in monogenic obesity in humans (Dubern and Clement, 2012), as along with mutations within leptin, these are rare within the general population. Despite this, several studies have found associations between common polygenic obesity and mutations within the *LEPR* gene (Ben Ali *et al.,* 2009, Chagnon *et al.,* 2000, Mattevi *et al.,* 2002, Quinton *et al.,* 2001, Wauters *et al.,* 2001b, Yiannakouris *et al.,* 2001), in particular three mutations in the coding regions: Gln223Arg, Lys109Arg and Lys656Asn.

The mutation at codon position 223 results in a glutamine to arginine change (CAG to CGG) in exon six, (Gln223Arg) and a mutation at codon 656 results from a lysine to asparagine codon change (AAG to AAC) in exon fourteen (Lys656Asn). The Lys109Arg mutation results from a change at codon position 109, (AAG to AGG) in exon four. The former two mutations are non-conservative whereas the Lys109Arg mutation is conservative (Paracchini *et al.,* 2005). These mutations have been associated with impaired leptin signalling (De Luis *et al.,* 2006, Quinton *et al.,* 2001), higher levels of leptin (van Rossum *et al.,* 2003, Yiannakouris *et al.,* 2001), and higher levels of fasting glucose and insulin (Chiu *et al.,* 2004, de Luis *et al.,* 2008, Lakka *et al.,* 2004, Takahashi-Yasuno *et al.,* 2004, Wauters *et al.,* 2001b).

On the other hand, several studies have not found association between any of the *LEPR* polymorphisms and obesity (Dias *et al.,* 2012, Gotoda *et al.,* 1997, Komsu-Ornek *et al.,* 2012, Pyrzak *et al.,* 2009, Silver *et al.,* 1997). As the associations with obesity are not very robust, it has been proposed that they may be population specific or they influence

intermediate traits or phenotypes (Salopuro *et al.,* 2005) due to the associations with higher levels of fasting glucose and insulin. Wauters *et al.,* (2001) found associations with glucose and insulin levels in women with the polymorphisms Lys109Arg and Gln223Arg, with the presence of the mutant allele, suggesting that these mutations could interact with factors associated with impaired glucose tolerance including hyperinsulinaemia or insulin resistance. In another study 6-7% of the variability of insulin sensitivity was attributed to the Arg^{223} allele, which was additionally associated with insulin resistance (Chiu *et al.,* 2004), supporting the view that this polymorphism could contribute to the initiation of factors leading to insulin resistance in a subset of individuals. However being homozygous for both major alleles at Gln^{223} and Lys^{109} were found to be beneficial to individuals changing to a much healthier lifestyle compared to those who are heterozygous and homozygous for the mutant allele, indicating a modulator effect to prevent conversion from IGT to T2D (Salopuro *et al.,* 2005).

As these SNPs occur in the extracellular region of the receptor, where its ligand leptin binds, any polymorphisms in this region could have functional implications (Figure 1.9.8). The majority of the associations between *LEPR* and obesity are with the Gln223Arg mutation, however evidence also suggests a role for both Lys109Arg and Lys656Asn. Therefore it is possible that particular haplotypes within *LEPR* are important in the development of obesity and not just one locus. Analysis of the mutations Lys109Arg, Lys656Asn and another polymorphism in *LEPR*, Pro1019Pro with obesity was studied. The eight different possible haplotypes were studied with respect to the development of obesity. Individuals carrying the minor alleles at the three loci (GCA) were found to be 4.41% and 3.50% higher in lean mass and fat mass respectively than those without these mutations (Liu *et al.,* 2004). In both multi-allele and allele-wise tests, significant

associations were found between this minor allele haplotype and with lean mass (p=0.006 and p=0.005 respectively) and fat mass (p=0.020 and p=0.012 respectively). No significant associations were found with the other haplotypes (Liu *et al.,* 2004). In another study the same three polymorphisms were analysed with respect to T2D risk in a Chinese population. A different haplotype was found to confer a higher risk of T2D, AGC, with minor alleles for Lys109Arg and Pro1019Pro and the major allele for Lys656Asn (Odds Ratio 1.69). Additionally the heterozygous diplotype (GGT/AGC) was found to be higher in T2D cases than in controls, at 17.0% and 8.2% respectively (Qu *et al.,* 2008). Preliminary data has suggested that particular *LEPR* haplotypes decrease the binding avidity with leptin therefore having a functional effect in the development of obesity (Mechan and Blakemore, *Pers Com*).

There are several possible mechanisms for the effect of these polymorphisms in the development of obesity: low expression of the leptin receptor, reduced ligand binding or signalling of the long isoform of the leptin receptor, differences in receptor trafficking, increased expression of non-signalling receptor isoforms (including the soluble isoform) or decreased ability to transport leptin across the blood-brain barrier (Anubhuti and Arora, 2008, Myers *et al.,* 2012).

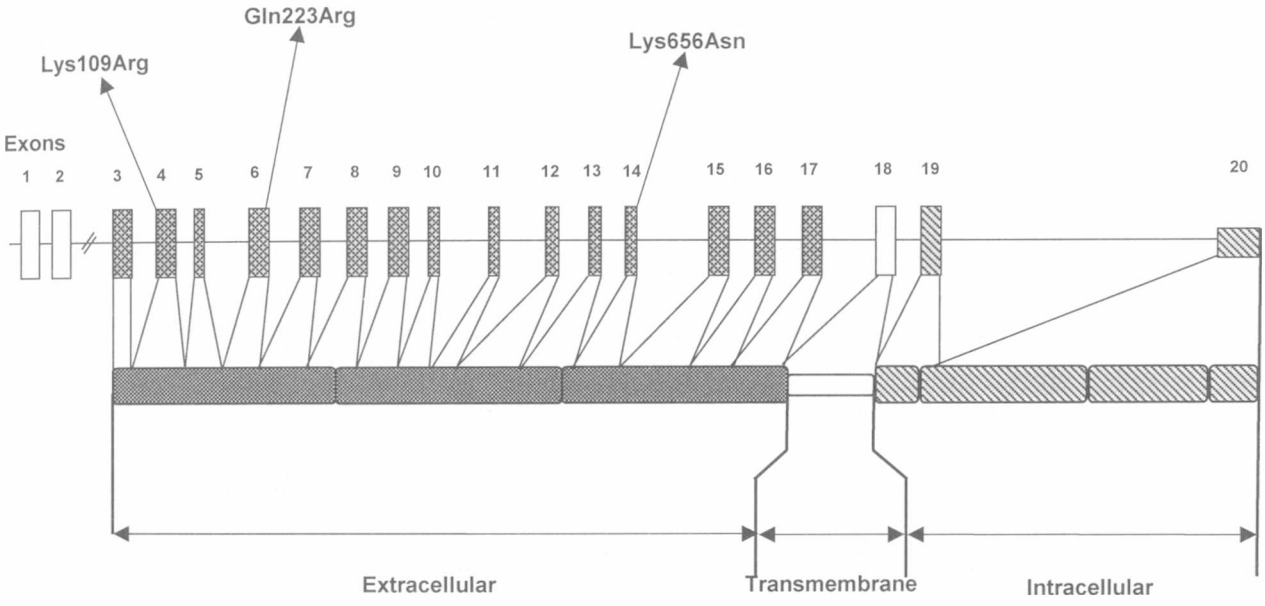


Figure 1.9.4: A schematic diagram of the long isoform of the leptin receptor and the correspondence of the exons with the protein domain structure. Also shown are the three polymorphisms that occur within the gene sequence, which are of importance. Adapted from Matsuoka *et al.*, (1997).

1.9.9 LEPROT

Located adjacent to the *LEPR* gene is the leptin receptor overlapping transcript (*LEPROT*). LEPROT belongs to a family of small proteins, which form part of functional protein complexes with other proteins involved in protein trafficking to lysosomes (Touvier *et al.*, 2009). Research has shown LEPROT to be involved in transporting internalised LEPRb to the lysosomes and has been found to regulate the distribution of LEPRb between the plasma membrane and intracellular compartments, though the pathway involved is unknown (Couturier *et al.*, 2007). Furthermore, *LEPROT* has been suggested to be involved in *LEPR* gene expression, as they are both transcribed by the same promoter (Couturier *et al.*, 2007). In addition a short CNV lies across the first two exons of both the *LEPR* and *LEPROT* loci (Kidd *et al.*, 2008, Matsuzaki *et al.*, 2009, Park *et al.*, 2010, Korbel *et al.*, 2007) (Figure 1.9.5) and recent research has found that copy number variation at this locus is associated with glucose homeostasis and energy metabolism (Jeon *et al.*, 2010). Further genetic and functional investigations of the *LEPROT* locus would be of interest to ascertain whether SNPs within this region affect LEPROT protein structure, and thus its function in transporting and influencing cell surface expression of LEPRb. In addition, it is clear that it would be wise to investigate the effect of the CNV that is located across the *LEPROT/LEPR* locus and whether they are both ultimately involved in leptin resistance and the obese phenotype.

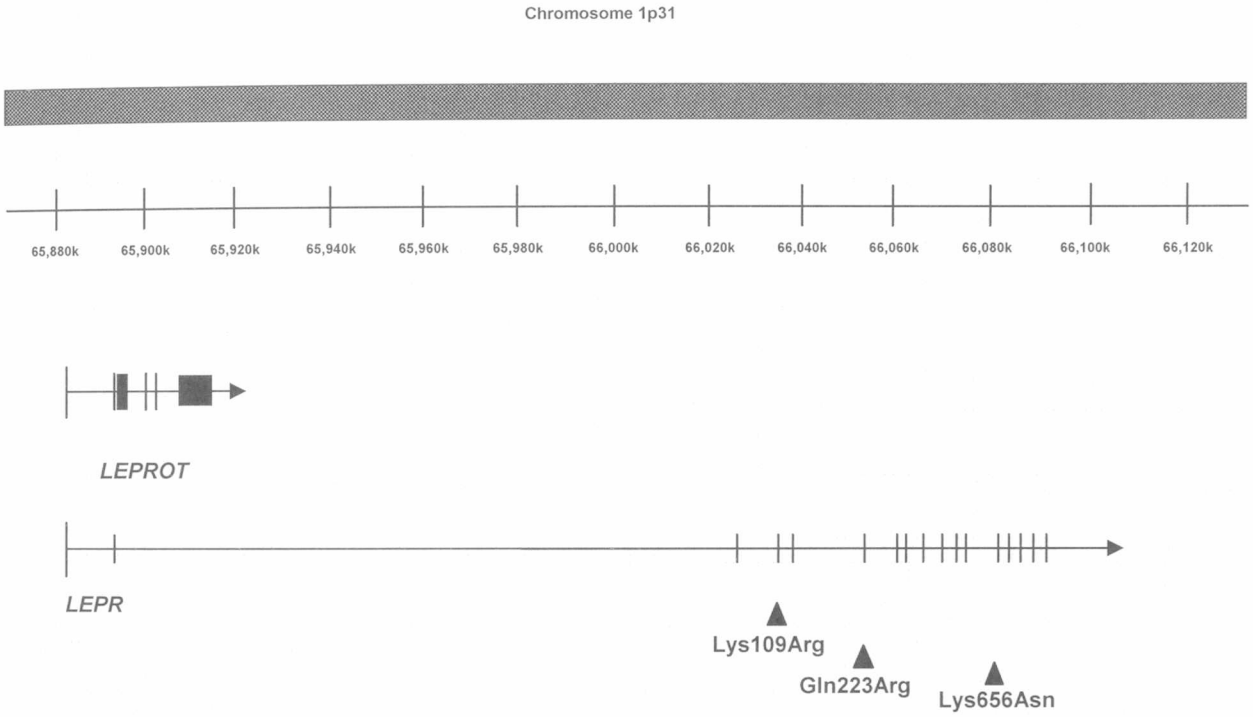


Figure 1.9.5: Graphic representation of the location of the genes *LEPR* and *LEPROT* on chromosome 1p31 and also the position of the three polymorphisms of interest, that have been significantly associated with obesity.

1.10.0 The Haematopoietically Expressed Homeobox – Insulin Degrading Enzyme (*HHEX-IDE*) region

Genome wide association studies have shown that markers within *HHEX-IDE* show association with T2D (Grarup *et al.,* 2007, Saxena *et al.,* 2007, Scott *et al.,* 2007, Sladek *et al.,* 2007, Zeggini *et al.,* 2007). Previous studies have also shown significant association of variants in this region with the development of T2D in a range of different ethnicities, including Europeans, (Cauchi *et al.,* 2008, Staiger *et al.,* 2008, van Vilet-Ostaptchouk, 2008), Japanese (Furukawa *et al.,* 2008, Horikawa *et al.,* 2008, Horikoshi *et al.,* 2007, Omori *et al.,* 2008), Chinese (Wu *et al.,* 2008) Asians (Ng *et al.,* 2008) and Indians (Chauhan *et al.,* 2010).

1.10.1 *HHEX–IDE* and Fetal Birth Weight

Subsequent research has found that the SNP rs1111875 within the fetal *HHEX–IDE* gene region shows strong evidence of influencing fetal birth weight (Freathy *et al.,* 2009; Pulizzi *et al.,* 2009; Andersson *et al.,* 2010).

A total of 19,200 offspring from four different cohorts were investigated for an association between five T2D susceptibility loci including rs1111875 at *HHEX-IDE*, with birth weight (Freathy *et al.,* 2009). In the meta-analysis each risk allele of this SNP was associated with decreased birth weight by 14 g (95% CI 4-23, p=0.004). In addition rs10946398 at *CDKAL1* loci was also significantly associated with lower fetal birth weight, with each risk allele reducing birth weight by 21 g (95% CI 11-31, $p=2x10^{-5}$) but no association was found with the other three loci. The risk alleles for rs1111875 and rs10946398 were combined and each additional risk allele was associated with a decrease birth weight by 17 g (95% CI 10-24, $p=5x10^{-7}$). There were 4% of offspring who had four risk alleles and

where found to be 80 g lighter compared to the 8% who had no risk alleles. As maternal and fetal genotypes are 50% correlated, where paired fetal and maternal genotype data was available, (7799 in number) analysis was performed adjusting for fetal and maternal genotype accordingly in addition to fetal sex and gestational age. Fetal genotype remained significant after adjusting for maternal genotype for rs10946398 (p=0.0005) and for rs1111875 (p=0.003) but there was more of an effect on birth weight, a reduction of 36 g (95%CI 56-16) and 29 g (95% CI 48-10) respectively, indicating the maternal locus has an opposite effect on birth weight. This was confirmed when maternal genotype was adjusted for fetal genotype as the effect size was 21 g (95% CI 1-42) and 7 g (95% CI 12-26) respectively although they were not significant (p>0.05).

In a similarly large study of 25,164 offspring of European descent, each risk allele of rs1111875 was found to reduce birth weight by 16 g (95%CI 24-8, $8x10^{-5}$) (Andersson *et al.,* 2010).

1.10.2 *HHEX-IDE* and Childhood BMI

In addition, SNPs within the *HHEX-IDE* have also been significantly associated with pediatric BMI (Winkler *et al.,* 2010, Zhao *et al.,* 2010). In one study twenty SNPs at eighteen loci previously associated with T2D were investigated in 3592 children (discovery cohort) and in an additional 3592 children (replication cohort) aged between two and eighteen years of age, with respect of childhood BMI (Zhao *et al.,* 2009). Two SNPs within *HHEX-IDE* were genotyped, rs1111875 and rs7923837 with the latter being significantly associated with higher BMI in childhood in the discovery (p=0.0013) and replication cohorts (p=0.023) and upon combining these data sets ($p=1.01x10^{-4}$). There was a significant association with rs3751812 in *FTO* ($p=3.8x10^{-5}$) and higher BMI but no

association with the other SNPs. In both genders, being homozygous for the minor allele was associated with higher BMI and the effect was strongest in the two to six year age range (p=0.000877) compared to those between the ages of seven and eleven (p=0.07446) and twelve and sixteen (p=0.1225).

However, in another study two different SNPs within the *HHEX-IDE* locus rs5015480 and rs10882102 were investigated with respect to BMI in 646 children at aged eight (Winkler *et al.,* 2010) and each risk allele of both SNPs were significantly associated with lower BMI at aged eight, ($p<0.05$) compared to higher BMI with the SNP rs3751812, in the study by Zhao *et al.,* (2009). In addition, there was an interactive effect between birth weight, genotype and lower BMI at aged eight (p=0.01). Upon further analysis birth weight was stratified for gestational age and fetal sex, being born large for gestational age (>90^{th} percentile) had the strongest association with lower BMI at aged eight (p=0.003) compared to those born small for gestational age (<10^{th} percentile) (p=0.83) and within the normal range, (10^{th} to 90^{th} percentile) (p=0.30).

These two studies provide contradictory evidence for SNPs within *HHEX-IDE* gene region to influence childhood BMI. It could suggest that additional factors are also involved as Zhao *et al.,* (2010) found *FTO* to also be significantly associated which has previously associated with BMI in children and adults. The study by Winkler *et al.,* (2010) was performed in a much smaller cohort, which may have resulted in the different result. Different SNPs within *HHEX-IDE* were also investigated between these two studies, which may indicate they have opposite effects on BMI and they could be in LD with the causal variant. In addition the study by Winkler *et al.,* (2010) analysed BMI at one time point during childhood whereas Zhao *et al.,* (2010) investigated the associated between the ages

of two and eighteen which found that the rs7923837 was more significantly associated in early childhood (two and six years of age) indicating that the effect of these SNPs is related to age.

1.10.3 Location of *HHEX-IDE*

The *HHEX* gene is situated on chromosome 10, 200 kb downstream of the *IDE* gene, position 10q23-25.

HHEX spans across 7.4 kb and contains four exons which produce four transcripts but only three are protein coding, comprising of one 270 amino acids in length and two 98 amino acids in length.

IDE spans across 122.39 kb and contains twenty-five exons which produces eight transcripts but only three are protein coding. These differ in length with transcript one comprising of 1019 amino acids, transcript two of 257 amino acids and transcript three of 464 amino acids.

1.10.4 Kinesin Family Member 11 (*KIF11*)

Located between *HHEX* and *IDE* is Kinesin Family Member 11 gene (*KIF11*), spanning 62.3 kb and contains twenty-two exons all of which are coding to produce one transcript, which is 1056 amino acids in length. The gene encodes a motor protein and belongs to the kinesin-like protein family (Le Guellec *et al.,* 1991). This family of proteins are involved in creating a bipolar spindle during cell mitosis, chromosome positioning and centrosome separation, and moving vesicles and organelles in the cell by acting on microtubules

(Kapitein *et al.,* 2005, Sawin *et al.,* 1992, Walczak *et al.,* 1998). They convert chemical energy (ATP) into mechanical energy. The protein has a catalytic or motor domain and a tail domain, which is important to interact with the target protein. The tail domain is less conserved across the family, which reflects the wide range of functions of kinesins within the cell, while the motor domain is highly conserved within the family of kinesin proteins (Hirokawa, 1998, Kim and Endow, 2000, Vale and Fletterick, 1997). Regulation of kinesin proteins is not that well defined but one mechanism is thought to be through control of binding to microtubules. This requires an additional protein cyclin dependent kinase 1 (CDK1), which phosphorylates a single, threonine residue within the tail domain (Blangy *et al.,* 1995, Sawin and Mitchison, 1995).

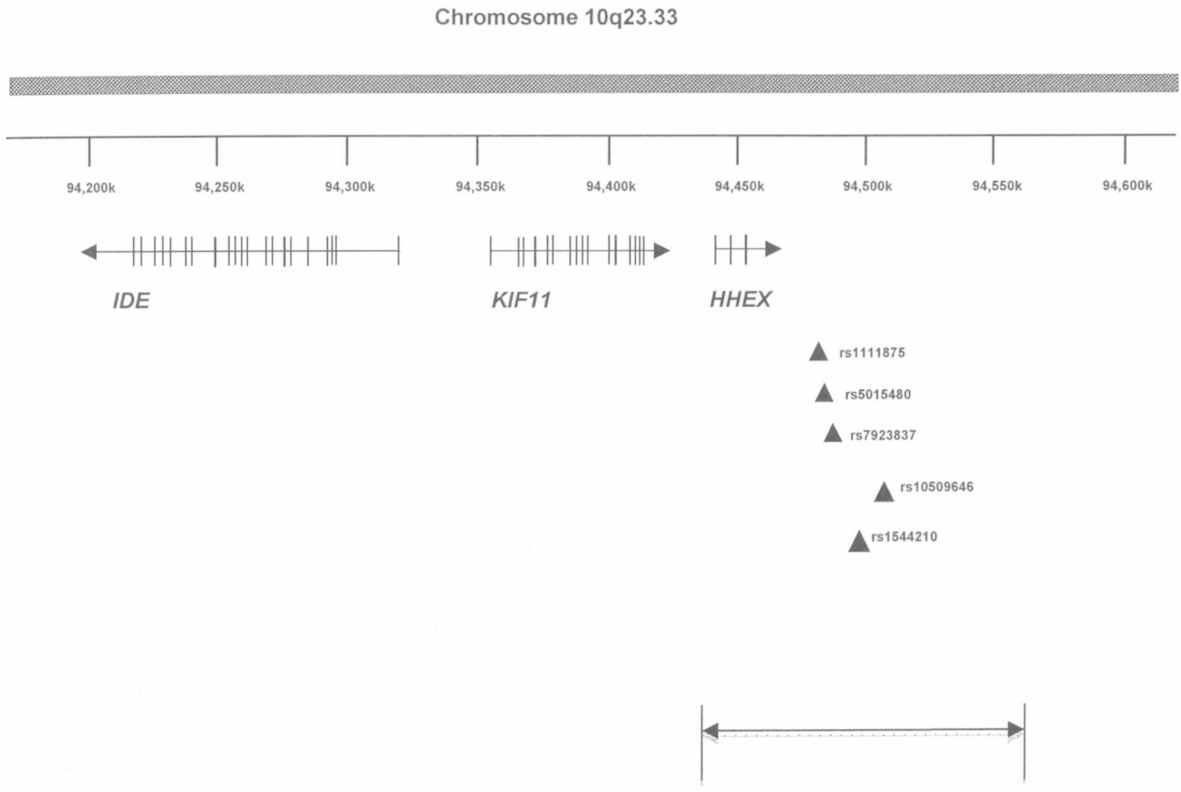


Figure 1.10.1: Graphic representation of the location of the *HHEX-IDE* region. It spans across 200 kb and is located on chromosome 10q23-25. Also depicted are the three SNPs that have been associated with T2D and fetal birth weight. The arrows indicate the region of this locus that will be investigated through genotyping in this research.

1.10.5 *HHEX* Gene Function

The *HHEX* gene encodes hematopoietically expressed homeobox, which is a member of homeobox transcription factor family. Mouse studies have found *Hhex* to be expressed in early development, within the different cell layers of the blastocyst (mesoderm and endoderm) (Keng *et al.,* 1998, Keng *et al.,* 2000, Thomas *et al.,* 1998b) and within the tissues liver, thyroid and lung during the embryonic stage (Bogue *et al.,* 2000, Keng *et al.,* 1998, Martinez-Barbera *et al.,* 2000). These studies suggest a role for *HHEX* in differentiation and/or organogenesis, regulation of cell proliferation and haematopoiesis (Jayaraman *et al.,* 2000, Keng *et al.,* 1998, Keng *et al.,* 2000, Newman *et al.,* 1997). Furthermore additional research showed *Hhex* to be expressed in gallbladder and pancreas in early development as well as replicating expression in the liver, lung and thyroid, further supporting a role for *Hhex* in the development of organs from the foregut endoderm of the blastocyst. Expression of *Hhex* in the thyroid, lung and liver was also found in the mature adult suggesting a role to maintain function of these organs (Bogue *et al*, 2000). Bort *et al.,* (2004) showed that *Hhex* was important in the proliferation of cells, which develop into the pancreas.

1.10.6 *IDE*

The insulin-degrading enzyme gene is ubiquitously expressed in a variety of tissues; liver, kidney, skeletal muscle, pancreas and adipocytes (Cotsapas *et al.,* 2010, Duckworth *et al.,* 1990, Yfanti *et al.,* 2008) and codes for a zinc metallopeptidase, which degrades insulin resulting in the termination of its activity (Affholter *et al.,* 1990, Rawlings *et al.,* 2008, Seta and Roth, 1997). It has been found to localise to the cytoplasm but it is also found within the extracelluar space, cell membrane, peroxisome and mitochondrion in some cell

types (Leissring *et al.,* 2004, Morita *et al.,* 2000, Seta and Roth, 1997). Distinct iosforms of the enzyme result from alternative splicing of the transcript. The active site of IDE mainly cleaves basic residues (arginine and lysine) and hydrophobic residues, (leucine and tyrosine) indicating that the recognition of the substrate is led by the tertiary structure of the protein rather than the amino acid sequence (Kurochkin, 1998).

Insulin is not the sole substrate of this enzyme; it has been shown to degrade a wide variety of other substrates including glucagon, amylin, bradykinin, calcitonin, kallidin and ubiquitin (Bennett *et al.,* 2000, Farris *et al.,* 2003, Ralat *et al.,* 2011, Saric *et al.,* 2003). It also has regulatory roles in cell growth and development, peroxisomal fatty acid oxidation and proteasome activity (Fawcett *et al.,* 2007, Hamel *et al.,* 2001, Kuo *et al.,* 1993). *IDE* regulates translocation of insulin from the cytoplasm to the nucleus (Harada *et al.,* 1993).

As *IDE* is involved in the breakdown of insulin it is thought to have a role in the development of T2D. Animal models have shown that a reduced *IDE* enzyme activity can induce the diabetic phenotype. In the Goto-Kakizaki rat a missense mutation in *IDE* gene resulted in a reduction in insulin degradation (Fakhrai-Rad *et al.,* 2000) and *IDE* knockout mice had higher levels of insulin and were glucose intolerant (Farris *et al.,* 2003). Furthermore significant associations have been found with variants in *IDE* and T2D in humans, rs2209772, rs1887922 and rs2251101 (Gu *et al.,* 2004, Karamohamed *et al.,* 2003, Kwak *et al.,* 2008). One study genotyped fourteen SNPs in LD across the *IDE* gene region. Haplotype genotypes were constructed based on common haplotypes across the region and associations were found with fasting insulin ($p=0.0009$), 2-hour insulin level ($p=0.0027$), homeostasis model assessment of insulin resistance ($p=0.001$) and BMI ($p=0.0067$) (Gu *et al.,* 2004). However, the exact mechanism in which these variants

contribute to impaired glucose and insulin homeostasis is unknown and there has been contradictory evidence as one study has found no association with ten variants across the *IDE* gene and T2D (Groves *et al.,* 2003).

1.11.0 *FTO*

The *FTO* gene was initially identified as a candidate through two large GWA studies in 2007 (Hinney *et al.,* 2007, Scuteri *et al.,* 2007) making it an exciting candidate gene. So far several studies have identified a cluster of SNPs, within intron one, that are highly associated with obesity in both adults and children (Dina *et al.,* 2007, Frayling *et al.,* 2007b, Hotta *et al.,* 2008, Peeters, 2007).

1.11.1 Location of *FTO*

The *FTO* gene is situated on chromosome 16q12.2 across 417 kb and comprises of nine exons. The transcribed protein is 58282 Da in size consisting of 505 amino acids.

1.11.2 Function of *FTO*

As yet the function of the gene remains elusive however structural and functional studies have shown that *FTO* is involved with DNA repair, nucleic acid modification and metabolism. *FTO* codes for an Fe (II) and 2–oxoglutarate (2-OG)–dependent dioxygenase (Sanchez–Pulido and Andrade–Navarro, 2007; Gerken *et al.,* 2007) and is part of the non-haem dioxygenase superfamily. Within this superfamily is the AlkB homologue protein family and the AlkB protein is involved with DNA repair (Sundheim *et al.,* 2006, Yu *et al.,* 2006).

Fe (II) and 2–oxoglutarate dependent dioxygenases are commonly found not only in eukaryotes but also bacteria and catalyse many reactions that involve the oxidation of an organic substrate using a dioxygen molecule (Aravind and Koonin, 2001). Both Fe(II) and 2OG are required for the normal functioning of the FTO protein, acting as a cofactor and a co-substrate respectively, with oxygen as another co–substrate (Gerken *et al.,* 2007) and succinate and carbon dioxide are by products. The reactions that 2OG dioxygenases are involved with include DNA repair, fatty acid metabolism and nucleic acid modifications.

Comparison of three members of the AlkB homologue protein family, ABH2, ABH3, and AlkB, have shown that FTO can be split into four structural and functional regions within the NH_2-terminal domain (Gerken *et al.,* 2007). Within the catalytic core is a double stranded β-helix (DSBH) in a 'jellyroll fold' (two anti-parallel beta sheets), which is conserved in each ABH member. Two of the beta strands form a hairpin that covers the substrate recognition site. The structure of the active site is phylogenetically conserved between species within each paralog but is slightly different structurally between each of the members. *FTO* contains a putative nuclear localisation signal, which no other member of this family contains, along with two long insertions that vary in length in different species (Sundheim *et al.,* 2006).

1.11.3 Mouse studies of *FTO*

As with most human genes, animal models have been used to determine the protein function. *FTO* was initially identified in a mouse (Peters *et al.,* 1999) with a 1.6 Mb deletion created by insertional mutagenesis on chromosome 8. Within this deletion are the genes *Ftm*, *Fto*, *Ft1* and the *Iroquois B* cluster containing *Irx3, 5* and *6*. The authors also identified a large gene of unknown function that they called *Fatso* (*FTO*). The resulting

phenotype from the deletion was fused toes and hyperplasia of the thymus and the mouse was referred to as the *Fto* mouse, for fused toes. Homozygous mutations for the 1.6 Mb deletion, were embryonically lethal, lacking neural tube formation, left–right asymmetry and polydactyly. Heterozygous mutations resulted in the fusion of the forelimb digits and thymic hyperplasia (van der Hoeven *et al.,* 1994). When the human gene linked to obesity was identified as the homologue of the *Fatso* mouse gene, the gene was re-named as the fat mass and obesity associated gene but retained the FTO symbol (Fischer *et al.,* 2008).

Mouse knockouts of *FTO* have also been investigated. A homozygous deletion within exons two and three, which is replaced by a neomycin STOP cassette, are viable initially but die after four weeks. Before death they show postnatal growth retardation, a reduction of adipose tissue and lean mass but have increased energy expenditure, sympathetic nervous system activity, relative hyperphagia and reduced spontaneous locomotor activity (Fischer *et al.,* 2009).

Research by Church *et al.,* (Church *et al.,* 2009) has shown that an A/T point mutation leading to a substitution of phenylalanine to isoleucine at position 367 (Fto^{I367F}), results in a lean phenotype in male mice. This is accompanied by a reduction in weight and fat mass, even when physical activity and food intake remained the same, indicating an increase in metabolism or in a failure to store fat. The reduction in body weight is thought to be because of the decrease in fat mass where a reduction of 16 to 18% was found between the Fto^{I367F} and wild type mice at twenty-four weeks, using DEXA. There was also a gender difference as only the males had the lean phenotype. It is thought that within females the sex hormones are protective. The Fto^{I367F} mutation was found to be expressed at lower

levels within the brain and liver, when compared to the wild type. This data gives support for a functional role of *FTO* in energy balance and the development of obesity.

1.11.4 Expression of *FTO*

FTO is expressed in many human tissues including adipocytes, skeletal muscle, kidney, heart and adrenal glands (Dina *et al.,* 2007, Frayling *et al.,* 2007b, Gerken *et al.,* 2007, Kloting *et al.,* 2008), however the main site, with the highest levels of expression were within the hypothalamus. Specific regions, which express the protein of the hypothalamus and brainstem have been identified: arcuate (ARC), supraoptic nuclei (SON), nucleus of the solitary tract (NTS), area postrema (AP), dorsomedial hypothalamic nucleus (DMH), paraventrical nucleus (PVN) and ventromedial (VMH). Within these identified areas the distribution was found to be throughout the ARC, whereas distinct areas were found within the SON, notably within the dorsal *vs*. ventral region (Fredriksson *et al.,* 2008).

The expression analysis indicates that *FTO* has a possible role in energy regulation and relaying information from central and peripheral locations. However a specific role of *FTO* in feeding, whether through preference, hunger or satiety, is still unclear. The use of animal models has been important for this. These studies have shown that *FTO* expression is controlled by the feeding and fasting state. *FTO* is down regulated, by around 60% in the arcuate nucleus of the hypothalamus in the mouse during forty-eight hours of food deprivation (Gerken *et al.,* 2007). Another study looking at different obese mice (Lep^{ob}, $Lepr^{db}$, *tub*, Cpe^{fat} and A^{y}) showed that *Fto* expression in the hypothalamus did not differ compared with the fed control (+/+) mouse, but they were decreased in the fasted Lep^{ob} mouse compared with the fed Lep^{ob} mouse (20%, $p<0.03$). Further analysis showed that within the +/+ mouse, fasting was associated with a twofold decrease ($p<0.01$) in *Fto*

expression within the mesenteric adipose tissue. On exposure to 4°C for thirty minutes, expression of *Fto* decreased by 30% ($p<0.01$) in the hypothalamus and twofold ($p<0.001$) in the mesenteric fat of the +/+ mouse (Stratigopoulos *et al.,* 2008).

Expression analysis of *FTO* within adipose tissue has revealed varying results and implications of *FTO* in the development of obesity are still unclear. *FTO* expression analysis within visceral and subcutaneous adipose tissue has been carried out in a group of twenty-nine women and twenty-six men. Results showed that levels were significantly higher in subcutaneous than in visceral adipose tissue but there were no differences between the sexes (Klöting *et al.,* 2008). Levels within the subcutaneous adipose tissue were also correlated with BMI, percentage body fat, sex and age but not with visceral fat. Another study investigating mRNA levels of *FTO* in skeletal muscle and adipose tissue found no significant difference between the two tissues (Grunnet *et al.,* 2009).

1.11.5 Associations of SNPs within Intron One of FTO with T2D and Obesity

FTO was initially identified as a candidate gene in the development of T2D and obesity through Genome Wide Association Studies (GWAS) (Dina *et al.,* 2007; Frayling *et al.,* 2007; Scuteri *et al.,* 2007). The SNPs within the *FTO* gene were found to be strongly associated with T2D but this was removed when the BMI was adjusted for, indicating that the association of *FTO* with T2D is mediated through BMI. Subsequent studies have since found *FTO* to be associated with obesity in children and adults in different populations: Japanese (Hotta *et al.,* 2008), South Asian and Chinese (Al-Attar, 2008), along with Europeans (Jacobsson *et al.,* 2008, Peeters, 2007, Cauchi *et al.,* 2009). Within intron one, a cluster of roughly forty SNPs are in strong linkage disequilibrium ($r^2> 0.80$ in CEU from HapMap) in Caucasian populations. One study found the cluster of SNPs to be strongly

associated with early–onset and severe obesity in both adults and children of European descent (P value of 1.67 x 10^{-26}) (Dina *et al.,* 2007). Another study found BMI, hip circumference and weight to be associated with rs9930506 among the cluster of SNPs within intron 1($p=8.6x10^{-7}$, $p=3.4x10^{-8}$ and $p=9.1x10^{-7}$ respectively).

Further investigation of the SNP rs9939609, which is within this cluster, has revealed that those who are homozygous for the A allele and heterozygous weigh 3 kg and 1.5 kg respectively more than those who are homozygous for the T allele (Frayling *et al.,* 2007). Children aged between ten and eleven years of age who carry the minor allele at rs9939609 consume more food, in particular total energy and fat, than those who do not (Timpson *et al.,* 2008). Another study has shown that this SNP is associated with impaired satiety responsiveness in children aged between eight and eleven years old, even after controlling for BMI (Wardle *et al.,* 2009). The exact mechanism is unknown but it may be involved with appetite response. One study found that rs9939609 is associated with impaired satiety responsiveness, with homozygous AA for the minor allele being significantly lower (Wardle *et al.,* 2009).

It is possible that the associated variant is in linkage disequilibrium with an un-typed causal variant (Jowett *et al.,* 2010). Other possibilities include the associated SNPs could influence the expression, splicing or affect the function of FTO (Meyre *et al.,* 2010).

1.12.0 Adenylate Cyclase Five (*ADCY5*)

A genome wide association study has shown that the SNP rs11708067 within ADCY5 is association with fasting glucose level ($p=7.1x10^{-22}$) and indices of β-cell function ($2.5x10^{-12}$) in 118,475 European Caucasians (Dupuis *et al.,* 2010). A further study found a different

SNP, within *ADCY5*, rs2877716, significantly associated with 2-hour glucose level ($p=4.2x10^{-16}$) and T2D OR 1.12 (95%CI 1.09-1.15, $p=4.8x10^{-18}$) (Saxena *et al.,* 2010). The associated SNP from this study is in high LD with rs11708067. These results both identify *ADCY5* as a novel susceptibility locus for T2D. The significant associations between fasting glucose, 2-hour glucose and 2-hour insulin levels and the SNP rs9883204 have also been replicated in an Asian Indian population (Rees *et al.,* 2011; Vasan *et al.,* 2011).

However, there is contradictory evidence as two studies have shown a lack of association with fasting glucose levels in 6822 Chinese men and women aged between fifty and sixty years of age (Hu *et al.,* 2010) and glucose stimulated insulin levels in 6694 Danish men and women aged between forty and fifty years of age (Boesgaard *et al.,* 2010). This suggests that the effect on glucose and insulin homeostasis maybe related to age, as rs11708067 has been significantly associated with increase in fasting glucose levels in 16,000 healthy children aged between nine and sixteen years of age (Barker *et al.,* 2011).

1.12.1 *ADCY5* and Birth Weight

Subsequent research has identified variants within *ADCY5* with lower birth weight and ponderal index (Freathy *et al.,* 2010; Andersson *et al.,* 2010).

The SNP rs9883204 was initially investigated in 10,623 European Caucasians and replicated in an additional 27,591 European Caucasians with respect to birth weight, birth length, birth head-circumference and ponderal index, (Freathy *et al.,* 2010). Each additional risk allele (C) was significantly associated with a 0.063-s.d (95% CI 0.079-0.047) lower birth weight ($p=7x10^{-15}$), a 0.044-s.d (95% CI 0.066-0.022) lower birth length

($p=4x10^{-5}$), a 0.025-s.d (95% CI 0.048-0.004) smaller head circumference (p=0.030) and a 0.032 (95% CI 0.055-0.009) lower ponderal index (p=0.006).

In another study significant associations were found with a different SNP, rs11708067, and lower birth weight in 4,213 European Caucasians, p=0.004, with each additional risk allele decreasing birth weight by 33 g (95% CI 55-10) (Andersson *et al.,* 2010). When this SNP was combined with rs1111875 in *HHEX-IDE* and rs7756992 in *CDKAL1* (which has previously been significantly associated with lower birth weight; Freathy *et al.,* 2009; Zhao *et al.,* 2009), they were significantly associated with lower birth weight, each additional risk allele decreased birth weight by 22 g (95% CI 34-10, p=0.0003). Those with five to six risk alleles were found to weigh 110 g (95% CI 42-179) lighter at birth compared to those with zero to one risk allele.

In addition, the C risk allele of rs9883204 has been significantly associated with fetal growth characteristics abdominal circumference, femur length and estimated fetal weight in the third trimester, with a combined effect of -16.9 g ($p=4.2x10^{-5}$) (Mook-Kanamori *et al.,* 2010). This suggests that the effect of this locus could be on fetal growth itself.

1.12.2 Location of *ADCY5*

The *ADCY5* gene is located on chromosome 3q13.2-q21 across 167.46 kb and contains twenty-one exons. Ten transcripts are produced but only six are protein coding, with the longest protein containing 1261 amino acids.

1.12.3 *ADCY5* Function

ADCY5 is a member of the adenylyl cyclase (AC) family, which encodes an enzyme which catalyses the generation of cAMP, from ATP. AC is activated through activation of adenylyl cyclase stimulatory G protein coupled receptors. cAMP is a second messenger which activates protein kinase A (PKA) that is involved in signal transduction and can regulate ion channels and affect gene expression by binding to DNA (de Rooij *et al.,* 1998). An increase in the level of cAMP increases the level of PKA. ADCY5 is one of nine membrane bound isoforms (AC1-9), with one soluble form (sAC) (Buck *et al.,* 1999, Paterson *et al.,* 1995). Each isoform, although encoded by different genes still have the same protein structure; a large glycoprotein which has a cytoplasmic NH_2 terminus followed by two repeats of a transmembrane domain and a cytoplasmic region (Krupinski *et al.,* 1989). The transmembrane domain (M1 and M2) is comprised of six helices, which loop across the plasma membrane, while the cytoplasmic regions (C1 and C2) are sub-organised further into a and b regions. The C1a and C2b regions together form the head to tail motif and form the catalytic region (Tang and Gilman, 1995). The cytoplasmic regions are conserved across the isoforms while transmembrane domains are less so, suggesting this region is important for membrane localisation and maybe involved in regulation of AC activity (Seebacher *et al.,* 2001).

ADCY5 is expressed in the kidney, heart, pancreas, brain and liver (Ludwig and Seuwen, 2002) but there is an uneven tissue expression of the nine isoforms as AC1 is mainly expressed in the brain while AC2, AC3 and AC8 are not expressed in the liver and kidney (Defer *et al.*, 2000). This reflects the distinct properties each isoform has and the downstream effects of activation are dependant on the cell type; AC within the liver responds to glycagon and muscle cells respond to adrenalin (Ludwig and Seuwen, 2002).

1.13.0 South Asian Population

The southern region of the Asian continent is called South Asia and contains many countries including India, Pakistan, Sri Lanka, Bangladesh, Korea, China and Japan (Ramaraj and Chellappa, 2008). Therefore the term South Asian population can be used to refer to individuals from each of these different countries. But all individuals, even from different countries from this region are known to have a different body composition compared to European Caucasians. They have a higher adipose tissue content and less lean tissue, which can affect the diagnosis of T2D and classification of obesity as they can develop T2D at a lower BMI category when compared to European Caucasians. They would be classified as overweight rather than obese. Therefore a different system of classification has been developed to account for this.

In a traditional diet the majority of energy density consumed comes from carbohydrates with a low amount from fat, due to the high amount of pulses and grains ingested with a variable amount of fruit, vegetables, meat, fish and diary (Abeywardena *et al.,* 2003; Misra *et al.,* 2009). Migration to Western countries results in a change in their dietary habits and this is of concern as within these countries there is an obseogenic environment, with vast quantity of processed food that is high in energy density, sugar and fat that is easily available. Thus there is an increase in the total energy density intake, with a decrease in the percentage energy from carbohydrates and an increase in energy from fat (Simmons and William, 1997). The changes in dietary habits are thought to contribute to the increased risk of developing obesity and its sequelae in an obesogenic environment along with genetic, epigenetic and gene-environment interactions (Misra and Khurana, 2011), particularly in first generation and second generation migrants (Popkin and Udry, 1998; Gordon-Larsen *et al.,* 2003). In addition, the term 'thin-fat' phenotype is used to describe

the body composition of the South Asian population, as they have a higher amount of adipose tissue in a central location with a lower muscle mass (Banerji *et al.,* 1999; Chandalia *et al.,* 2007).

1.14.0 Aims of the investigation

It is clear that the candidate genes *HHEX-IDE*, *FTO*, *ADCY5* and *LEPR* have a role in birth weight obesity and T2D from the vast amount of research conducted to date. LEPR However, further investigation is needed to determine if variants in the maternal genes *HHEX-IDE*, *FTO* and *ADCY5* are associated with birth weight in additional populations, apart from European Caucasians. Previous research has found a lack of association between variants in *LEPR* and birth weight therefore these will not be investigated. In addition, if known polymorphisms or particular haplotypes in *LEPR* affect the binding affinity of the recptor for leptin. Therefore a genetic and functional approach will be taken to investigate the candidate genes mentioned.

1.14.1 Functional Studies:

This is to analyse the functional implications of common coding variants in the extracellular region of the *LEPR* protein product.

a. Using site directed mutagenesis, naturally occurring haplotypes, some of which have been associated with obesity will be constructed, focussing on those known or predicted to be of interest.

b. Protein products will be expressed using the pCMV-Tag 2B expression vector.

1.14.2 Genetic Studies

a. Genotype T2D-associated alleles in the *HHEX-IDE* region, in maternal South Asian Cohort to investigate association with fetal birth weight.

b. Genotype *FTO* in maternal South Asian cohort and investigate association with respect to fetal birth weight

c. Genotype novel (discovered by the genome scan currently being undertaken at Imperial) candidate genes associated with obesity and inflammation in the adult South Asian cohort.

2.0 Material and Methods

2.1 Reagents

Ampicillin

A stock solution of 50 mg/ml Ampicillin was made by dissolving 0.5 g of ampicillin sodium salt (Sigma-Aldrich) in 10 ml of deionised water and then it was passed through a sterile filter unit with a 0.2 μM pore size for sterilisation. 1 ml aliquots of this solution were pipetted into 1.5 ml Eppendorf tubes and stored at -20°C.

Diethylpyrocarbonate (DEPC)

A 0.1% (v/v) stock solution of DEPC (Sigma Aldrich) was made by adding 1ml of DEPC solution to 999 ml of deionised water, in a fume hood. Deactivation was carried out by autoclaving at 120°C for twenty minutes (Boxer Autoclave).

Dimethyl Sulfoxide (DMSO) and Fetal Bovine Serum (FBS)

1ml of DMSO (Sigma-Aldrich) was added to 9 ml of FBS (Fisher Scientific), to make the media for freezing cells. Five of these solutions were made and stored at –20°C for future use. The remainder DMSO was made into 10 ml aliquots and also stored at –20°C.

0.5M Ethylenediaminetetraacetic acid (EDTA) pH 8

20 g of sodium hydroxide pellets (Fisher Scientific) were dissolved in 400 ml of deionised water, to which 186 g of EDTA (Fisher Scientific) was added, with the volume being made up to 1 L with deionised water and autoclaved for twenty minutes at 120°C.

Luria-Bertani (LB) Agar

The components of LB agar (Sigma Aldrich) were in tablet form and four tablets were dissolved in deionised water to a total volume of 200 ml. The mixture was shaken and left for the tablets to dissolve after which time it was autoclaved for twenty minutes at 120°C. The final volume contains 36.56 g/L of tryptone, 18.28 g/L yeast extract, 18.28 g/L sodium chloride, 54.88 g/L agar and 6.4 g/L inert tableting aids.

LB Broth

Four tablets (1.1 g each; Sigma-Aldrich) were dissolved in deionised water to a total volume of 200 ml. The mixture was shaken by hand and left to dissolve before autoclaving at 120°C for twenty minutes. The final volume contains the components enzymatic digest casein at 40 g/L, yeast extract (low sodium) at 20 g/L, sodium chloride at 20 g/L and 8 g/L of inert agents.

Phosphate Buffered Saline (PBS), pH 7.4

The components of PBS (0.01 M phosphate buffer, 0.0027 M potassium chloride and 0.137 M sodium chloride; Sigma-Aldrich) were in tablet form. Five tablets were dissolved by hand, in a total volume of 1 L deionised water and autoclaved for twenty minutes at 120°C. After which time, in a fume cupboard, 50 ml aliquots were made.

Supplemented RPMI

The cell culture media was supplemented with three additional reagents for cell growth. 50 ml of FBS was added to give a final concentration of 10%, 1 ml of insulin to give a final concentration of 0.01 mg/ml and a combination antibiotic of penicillin/streptomycin to give a final concentration of 10 U/10 μg/ml in the final volume of 500 ml of the medium.

Tris-Borate-EDTA buffer (TBE)

1 L of 5X TBE was made by adding 20 ml 0.5 M EDTA (pH 8), 54 g of Tris base (Fisher Scientific) and 27.5 g Boric acid (Fisher Scientific) first then the volume was made up to 1 L with deionised water.

X-Gal

A stock solution of 50 mg/ml was made by dissolving 0.1 g of X–Gal powder (Fisher Scientific) in 2 ml of N,N-dimethylformamide. 500 µl aliquots were made in 1.5 ml Eppendorf tubes and stored at -20°C. The Eppendorf tubes were wrapped in coloured tape so that the solution does not degrade, as it is sensitive to light.

2.2.0 Subjects

2.2.1 UK/Irish and South Asian Cohort

In total, 1811 DNA samples were included in the genetic studies, of which 896 are UK/Irish and 915 are of South Asian decent. The samples were collected from women who were recruited at Northwick Park Hospital (Northwest London Hospitals) between 1997 and 1999 whilst they were attending an 18-week antenatal scan. Genomic DNA samples were extracted using a Gentra Puregene Buccal Cell kit following the manufacturer's instructions (Qiagen). A questionnaire was used to ascertain details about family origin along with different family diseases including asthma, stroke and heart disease. Information regarding the child was also collected including birth weight, sex and gestation age.

The South Asian population studied here were originally from North West India, from the state of Gujarat, and are 2nd generation migrants, who have settled in North West London, which has a large South Asian community.

2.3.0 Statistical Analysis

All statistical analysis was carried out using the program SPSS version 18. Golden Helix SVS, data analysis software was used to analyse *FTO* genotypes with respect to birth weight using linear regression.

2.3.1 Genotyping analysis, quantitative and qualitative

Case Control Analysis: Data sets were divided into phenotype groups and the genotypes analysed for association with phenotypes. Genotyping data was initially tested for Hardy Weinberg equilibrium (HWE) by comparing observed allele frequency with the expected allele frequency ($p^2 + 2pq + q^2 = 1$). Statistically significant distortions of this equilibrium in the whole data set can indicate problems with detection of specific alleles. Pearson's c^2 analysis enables us to determine whether expected genotype frequencies differ significantly from observed genotype frequencies ($p=<0.05$). Data that conformed to HWE was then tested further for association of alleles with specified phenotypes, again utilising Pearson's Chi^2 analysis with the null hypothesis that there is no significant difference in allele frequencies when cases and controls are compared.

Population based studies: To analyse the South Asian population for quantitative traits using the general linear univariate analysis of covariance (ANCOVA) method was used. Data was tested for normality, using histograms and box plots; if it did not conform it was

transformed appropriately before analysis. Traits explored were those associated with maternal adiposity and neonatal size. Covariates that were considered included gestational age, parity and sex of neonate.

2.4.0 Functional Study of *LEPR*

2.4.1 Cell Culture

Michigan Cancer Foundation-7 (MCF-7) cells, are a breast cancer cell line, that are known to express the leptin receptor (Garofalo *et al.,* 2004), were grown in supplemented RPMI in 25 cm^2 Nunclon cell culture flasks (poly-D-lysine coated, vent closed cap) and placed in a duel gas cell culture incubator (Heraeus, Hera cell) at 37°C and 5% carbon dioxide. MCF-7 cells are an adhesive cell line so they grow attached to the bottom of the flask.

Cells were observed under the microscope daily for growth and contamination. Once the cells had grown and covered most of the surface area of the flask, (75-100% confluent) they were then either washed with PBS and split into two to three flasks to continue to grow, if more cells were needed. Or they were centrifuged in a bench-top centrifuge (Eppendorf 5810R) at 160 g for five minutes and stored at -70°C as a cell pellet for RNA extraction.

Cell numbers where checked by using a haemocytometer. Twenty microlitres of cell solution was added to the main chamber of the haemocytometer. Cells were counted using the middle chamber containing twenty-five small squares under an inverted microscope, viewed at a total magnification of x100 then x400. Calculations were performed to estimate the number of cells within the ~7 ml total volume (1-2 ml trypsin and 5 ml of

RPMI). This was carried out when the cells became 75-100% confluent to determine the numbers of flasks to use for continued growth.

2.4.2 Continued Growth

Under sterile conditions the RPMI was removed and the cells were washed with 5 ml of sterile PBS to remove any RPMI that may have been left behind. After the cells had been washed for a second time 1-2 ml of 0.25% trypsin-EDTA (Sigma-Aldrich) solution was added, so that a thin layer was formed over the cells and placed in the incubator for three to four minutes. Once the cells had been incubated they were checked under the microscope, at x400 total magnification. If most of the cells were in suspension then 15 ml of pre-warmed supplemented RPMI was added, to split into three flasks (5 ml per $25cm^3$ flask). If some cells remained adhered, the trypsin solution was swirled around the flask to help loosen the remaining cells. However caution was taken so that the incubation time was no more than five minutes. Once the cell suspension was aliquotted into the three flasks they were then placed back in the incubator and left to grow.

2.4.3 Storing of Cell Pellets at -70°C

The method is the same as in 2.5.2 however once the RPMI was added all the contents of the flask were removed and placed in a 15 ml Falcon tube and centrifuged using a bench top centrifuge at 160 g for five minutes. The supernatant was removed and discarded and the cell pellet re-suspended in any remaining RPMI. The cell pellet was washed two more times with non-supplemented RPMI, the supernatant being removed each time and then flash frozen in liquid nitrogen and stored at -70°C for DNA extraction at a later stage.

2.4.4 Growing Cells from Frozen

Three 5 ml aliquots of RPMI supplemented with FBS were pre-warmed by adding it to a water bath (Grant) set to 37°C. MCF-7 cells contained in 2 ml cyrovials were taken out of liquid nitrogen storage, whilst ensuring that personal safety was maintained by wearing a lab coat, gloves and insulated gloves. The cells were then placed in the 37°C water bath for a maximum of five minutes, during which time they were checked and gently flicked. Once they had thawed they were subsequently added to 5 ml of pre-warmed RPMI and gently flicked again to ensure complete mixing, after which time they were centrifuged at 160 g for five minutes using a bench-top centrifuge. The liquid was removed and another 5 ml of pre-warmed RPMI was added to the cells, gently flicked then and centrifuged again. A further 5 ml of pre-warmed supplemented RPMI was added to the cell pellet where it was re-suspended a second time then under sterile conditions, the cells were transferred to a 25 cm^3 cell culture flask and placed in the incubator.

2.4.5 Freezing cells for long-term storage in Liquid Nitrogen

The method is the same as 2.5.2 and 2.5.3 however, once the cells had been washed twice in RPMI, the cell pellet were re-suspended in 1ml of DMSO and FBS and transferred to a 2 ml cyrovial. The cyrovial was flash frozen in liquid nitrogen then stored at -70°C overnight before being stored in an appropriate storage vessel containing liquid nitrogen.

2.4.6 Total RNA extraction using Trizol ® Reagent

Two frozen MCF-7 cell pellets were used for the extraction of total RNA. The pellets were defrosted on ice before use. DEPC treated tips and Eppendorfs were also used to prevent RNA degradation. 750 µl of Trizol reagent (Invitrogen) was added to the cell pellet and the cells were then lysed by repetitive pipetting after which time the homogenised samples were incubated for five minutes at room temperature. 200 µl of chloroform was subsequently added and the Eppendorf tube was shaken vigorously by hand for fifteen seconds before further incubating at room temperature for ten minutes. The samples were then centrifuged at 12,000 g, at room temperature for thirty minutes, instead of the recommended fifteen minutes at 4°C, due to lack of equipment. After centrifugation the upper aqueous layer was removed, into a clean tube as this contained the total RNA. DNA and protein were within the lower red, phenol-chloroform phase. 500 µl of isopropyl alcohol was added to the aqueous phase to precipitate the RNA. Samples were incubated for ten minutes at room temperature and then centrifuged for thirty minutes at 12,000 g at room temperature. The supernatant was removed and the RNA washed once with 1ml of 75% ethanol. The DNA was then left to air dry for five to ten minutes before dissolving the RNA in 30 µl of DEPC treated water and storing at -70°C, if the cDNA step was not performed.

2.4.7 PolyA$^+$ mRNA Extraction - DynaBead Method

PolyA$^+$ mRNA was extracted from total RNA using magnetic Oligo(dT)$_{25}$ Dynabeads® that binds to the polyA$^+$ mRNA. 50 µl of lysis buffer containing 2.5 x 10^5 MCF-7 cells was used to extract between 3-250 ng of mRNA for six separate reactions.

2.4.7.1 Preparation of the Dynabeads

The Dynabeads® were prepared by resuspending them within the Storage buffer and washing them in 30 µl of Lysis/Binding buffer using the magnetic particle concentrator (MPC). After the Dynabeads had been resuspended 30 µl was removed and placed in six separately labelled RNase-free Eppendorf tubes for each cDNA reaction to be performed later. They were then placed on the Dynal MPC. When the suspension was clear the supernatant was removed, the tubes were then also removed from the MPC and placed on ice. 30 µl of Lysis/Binding buffer (supplied as 100 mM Tris-HCl (pH 7.5), 500 mM LiCl, 10 mM EDTA (pH 8), 1% LiDS and 5 mM dithiothreitol (DTT)) was added to the beads and the tubes were replaced on the MPC and the clear supernatant removed again. The beads were resuspended in the Lysis/Binding buffer to 30 µl and left on ice.

2.4.7.2 Preparation of the Lysate

This method was carried out using frozen cell pellets. The cell pellet was removed from -70°C and 1 ml of Lysis/Binding buffer was added as soon as it defrosted. Lysis of the cell pellet was carried out by pipetting several times. 50 µl of the lysate was added to the Eppendorf tubes containing the Dynabeads. The remainder of the lysate was placed back at -70°C for later use.

2.4.7.3 mRNA Isolation

All the reagents supplied with the Dynabead kit were brought up to room temperature, from storage at 4°C, except 10 mM Tris-HCl, which was left on ice, whilst preparing the Dynabeads. After 50 µl of the lysate was added to the six separate Eppendorf tubes

containing the Dynabeads, they were mixed by pipetting and placed on a sample mixer for five minutes at room temperature to enable the beads to anneal to the mRNA. After the time had elapsed the sample tubes were placed on the MPC and the supernatant was discarded. The tubes were removed from the MPC to resuspend the Dynabead-mRNA complex in 100 µl of Washing Buffer A (supplied as 10 mM Tris-HCl (pH 7.5), 0.15 M LiCl, 1 mM EDTA and 0.1% LiDS) and mixed by gently flicking the tubes. The sample tubes were again placed on the MPC to remove the supernatant. Washing the mRNA was carried out once more with the Wash Buffer A, then the complex was resuspended in 100 µl of Washing Buffer B (supplied as 10 mM Tris-HCl (pH 7.5), 0.15 M LiCl and 1 mM EDTA) and transferred to new, labelled tubes. The new tubes were placed on the MPC and the clear supernatant discarded. 100 µl of Wash Buffer B was again added to the tubes and mixed then placed back on the MPC and the clear supernatant discarded the supernatant. 100 µl of ice cold 10 mM Tris-HCl (supplied as 10 mM Tris-HCl, pH 7.5) was added to each of the tubes that had been removed from the MPC. Prior to the addition of the RT-PCR reagents the Eppendorfs were placed back on to the MPC to allow the removal of the supernatant. Eppendorf tubes were then placed in a water bath at 42°C for sixty minutes. The reaction was terminated by placing the tubes at 70°C for ten minutes then stored at –20°C.

2.4.8 Reverse Transcription Method using mRNA

Before the mRNA isolation method was performed the reverse transcription mix was prepared (Table 2.5.1) to a total volume of 50 µl and kept on ice. The reverse transcription method is a two-step PCR and was prepared according to the manufacturer's instructions (Fermentas). Six RT-PCRs were carried out in separate RNase free Eppendorf tubes consisting of *LEPR* forward primer, *LEPR* reverse primer, Glucose-6-phosphate

dehydrogenase (*G6PD*) forward primer, *G6PD* reverse primer, Random Hexamer primers and a water blank reaction containing *G6PD* forward primer.

Table 2.4.1: Checklist for the RT-PCR

REAGENT	VOLUME FOR ONE REACTION (µL)
dNTPs	1.60
DEPC water	34.00
Primers (100 pm/l)	2.00
5XBuffer	8.00
RiboBlock	1.00
RT Enzyme	1.50

Primer sequences can be found in Table 2.4.2. Primers were supplied at 100 µM/µl and were diluted 1:2 for a final concentration of 15-20 pmoles in 50 µl (Eurogentec). RiboLock RNase inhibitor (Fermentas) was used in the reaction, which prevents the degradation of mRNA by any RNases.

Table 2.4.2: Primer Sequences for *LEPR* and *G6PD* and the expected product sizes.

PRIMER NAME	PRIMER SEQUENCE	RT-PCR RNA BAND SIZE
LEPRF	5' - GGATCCAACTTGTCATATCCAATTACTCCTTG - 3'	2457bp
LEPRR	5' - ACCTGCATCACTCTGGTGTTT - 3'	
G6PDF	5' - CCAGATAGGCTGGAACCGCATCAT - 3'	205bp
G6PDR	5' - TGAGGATAACGCAGGCGATGTTGT - 3'	

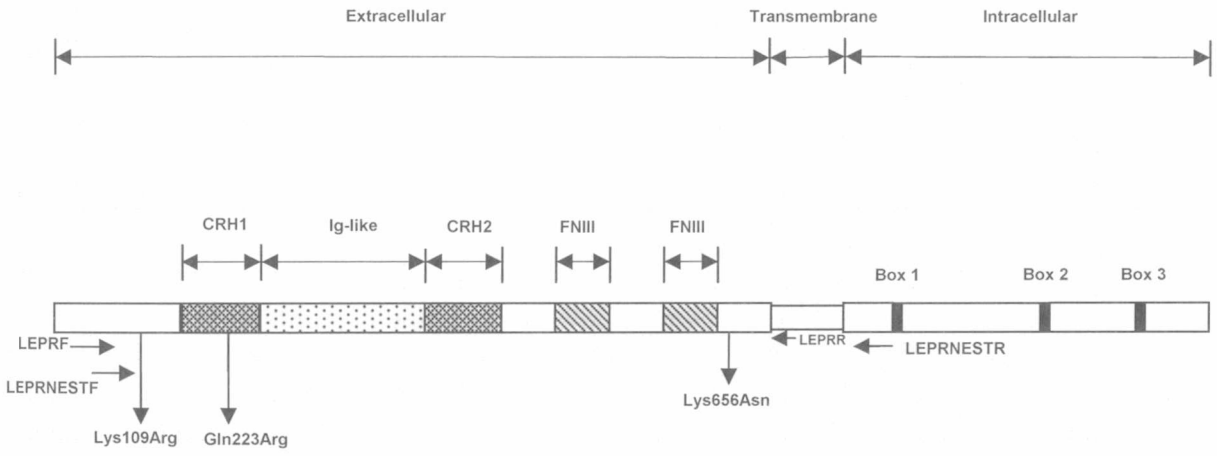


Figure 2.4.1: Graphic representation of the long isoform of the leptin receptor, showing the three domains (extracellular, transmembrane and intracellular) along with the five domains within the extracellular region and the three SNPs of interest. Also depicted is the location of the *LEPR* forward and reverse primers and the nested primers.

2.4.9 Reverse Transcription Method using MCF-7 total RNA

The RT-PCR method was performed as a two-step reaction and the first strand cDNA synthesis step was carried out using Moloney Murine Leukaemia Virus Reverse Transcriptase (M-MuLV RT) according to the manufacturer's instructions (Fermentas). Total RNA extracted from MCF-7 cells using Trizol ® was used. All reagents were defrosted on ice and remained there whilst the reaction mixture was prepared. Two reactions were carried out using the MCF-7 RNA and a third was set up as a negative control. DEPC treated tips and Eppendorf tubes were used. 27.8 µl of DEPC treated water and 5 µl of RNA was added to the bottom of the Eppendorf tubes, followed by 8µl of reaction buffer, 1.2 µl of RiboLock RNase inhibitor, 4 µl of dNTPs, 3.5 µl of reverse primer to the side of the tube before pulse spin centrifugation at 12,000 rpm in a micro-centrifuge (Eppendorf 5415C). 0.5 µl of reverse transcriptase enzyme was added to the mixture before centrifuging the tubes for a second time, after which time they were placed in a water bath at 45°C for one hour. The reaction was subsequently stopped, by placing the tubes on ice for ten minutes.

2.4.10 Reverse Transcription Method using Placental poly A^+ RNA

The RT-PCR method was performed as a two-step reaction and the first strand cDNA synthesis step was prepared using SuperScript® according to the manufacturer's instructions (Invitrogen). Human placental poly A^+ RNA was used as the starting material, which contains mRNA from three different sources (Agilent Technologies). All reagents were defrosted on ice where they remained while the reaction was carried out. Initially 1 µl of reverse primer, 1 µl 10mM dNTPs, 1 µl of mRNA and 11 µl DEPC treated water were

pipetted into a DNase and RNase free Eppendorf tube and heated to 65°C for five minutes and it was then incubated on ice for one minute. Contents were centrifuged briefly, after which 4 µl 5X First Strand buffer (supplied with the enzyme), 1 µl 0.1 M DTT (supplied with the RT enzyme), 1 µl RiboLock RNase inhibitor and 1 µl of SuperScript III RT were added. Contents were mixed by pipetting then incubated at 55°C for sixty minutes, which is the maximum reaction temperature and time suggested due to the final length of the cDNA and with using a gene specific primer. The final step in the reaction was to inactivate the enzyme by heating the contents to 70°C for fifteen minutes. First strand cDNA synthesis was also carried out using the reverse primer from the Glucose-6-phosphate dehydrogenase (*G6PD*) gene, as this is constitutively expressed in all tissues.

2.4.11 Amplification of LEPR using PCR

The extracellular region of the LEPR was PCR-amplified using previously made cDNA. Standard PCR conditions were used for the three different Taq DNA Polymerases that were used as provided by their respective supplier (Table 2.4.3).

Table 2.4.3: Standard PCR conditions for three different Taq polymerases. New England Biolabs (NEB), Applied Biosystems and KAPA Biosystems. Deionised water was used to make the total PCR volume up to 25µl.

PCR Reagent	Supplier		
	NEB	Applied Biosystems	KAPA Biosystems
PCR Buffer	2.5 µl (10X stock)	2.5 µl (10X stock)	5 µl (5X stock)
Magnesium Chloride (25 mM)	2.5 µl	2.5 µl	2.0 µl
Forward Primer (10mM)	0.5 µl	0.5 µl	0.75 µl
Reverse Primer (10 mM)	0.5 µl	0.5 µl	0.75 µl
dNTPs (10 mM)	0.5 µl	0.5 µl	0.5 µl
Taq Polymerase	0.125 µl	0.25 µl	0.1 µl
DNA	50 ng	50 ng	50 ng
Total PCR volume	25 µl	25 µl	25 µl

Standard PCR conditions were also used for each polymerase and were as follows:

Initial denaturation: 95°C for 30 seconds

Then 30 cycles of

Denaturation: 30 seconds at 95°C

Annealing: 30 seconds at 56°C

Extension: 30 seconds at 72°C

Then a final single step:

Final extension: 72°C for 5 minutes

With the exception that the final extension was held at 72°C for ten minutes for AmpliTaq Gold polymerase (Applied Biosystems), while for KAPA Taq polymerase (KAPA Biosystems) the conditions varied in length at each step. The initial denaturation lasted for two minutes, denaturation and annealing steps were held for fifteen seconds at the same respective temperatures as above and the extension step was held for forty-five seconds. The final extension lasted for one and a half minutes.

2.4.11.1 Magnesium Titration

The magnesium titration was performed to give final concentrations of 1mM to 4mM at 0.5 mM increments. The following formulae were used to calculate the volume of stock magnesium chloride to be used in each of the reactions for the standard PCR conditions provided by each supplier. To maintain the total PCR volume at 25 µl the volume of water added was reduced. The example given is for 1 mM.

Final Concentration (mM) = Dilution Factor x Stock Solution

Volume of $MgCl_2$ to be added (µl) = Total PCR volume (µl) x Dilution factor

Dilution factor = final concentration / stock solution

= 1 mM/25 mM

= 0.04

Volume of $MgCl_2$ to be added = total PCR volume x dilution factor

$= 25\ \mu l \times 0.02$

$= 1\ \mu l$

Table 2.4.4: Volumes of 25 mM stock Magnesium Chloride concentration. These were used to give final concentrations between 1 mM and 4 mM in 0.5 increments in the PCR.

Final Concentration	**Volume in 25 µl**
1.0 mM	1.0 µl
1.5 mM	1.5 µl
2.0 mM	2.0 µl
2.5 mM	2.5 µl
3.0 mM	3.0 µl
3.5 mM	3.5 µl
4.0 mM	4.0 µl

The distilled water and $MgCl_2$ were added to the appropriately labelled tubes then the remaining reagents, bar the cDNA, were combined into a master mix, by multiplying the volume for one reaction by the total number of reactions needed. The master mix was subsequently added to all the tubes followed by the cDNA. A 1:10 dilution of all the primers was used. Primer sequences can be found in Table 2.4.2. All the reagents were kept on ice whilst they defrosted and throughout the reaction set up. The reagents were gently flicked so that they are all mixed then placed on the thermal cycler (G-Storm, GS4).

2.4.12 Agarose Gel Electrophoresis

Once the PCR had been completed, the products were subjected to electrophoresis on an agarose gel. For a 1% (w/v) mini gel the following method was used.

600 ml of 1X TBE was made in total from a 5X stock solution. One gram of agarose was then added to 100 ml of 1X TBE buffer in a conical flask (250 ml). The mixture was heated at thirty-second intervals in a microwave until all the agarose had dissolved and the mixture was clear and colourless. Distilled water was added to the mixture if the total volume was lower than the 100 ml marked on the flask. The flask was cooled under the cold tap until it was hand hot at which time 3 µl of ethidium bromide (10 mg/ml) was added and mixed in to the molten agarose solution. The gel apparatus was set up, including combs within the appropriate slots, and the agarose solution poured into it and left to set for around thirty minutes. Fifteen microlitres of ethidium bromide was added to the rest of the 1 X TBE buffer and added to the gel tank, up to the maximum level, once the gel had set.

In some cases a 1.2% (w/v) or a 2% (w/v) mini gel was used in which 1.2 g and 2 g was added to 100 ml of 1X TBE respectively. Where a 1% (w/v) midi gel was used, for that 2 g of agarose was added to 200 ml of 1X TBE. 600 ml was used for the gel tank.

Eight microlitres of the PCR product was loaded with 2 µl of 5X loading dye (Bioline) and loaded into each well, using GeneRuler (Fermentas), Hyperladder I (Bioline) or Easy Ladder II (Bioline) as the DNA marker (Appendix I).

The gel was electrophoresed at 100 volts for thirty minutes or until the dye had moved half the distance between the well and the bottom of the gel. The DNA bands were then visualised and was photographed using a High Performance UV Transilluminator and camera (GelDoc-it Imagining System, UVP).

2.4.13 Gel Extraction of *LEPR* PCR Product

As the optimisation PCR of *LEPR* was non-specific, the band correlating to the extracellular region of the *LEPR*, which was 2457 bp in size, was extracted from an agarose gel using a gel extraction kit after another set of PCRs was performed, to increase the yield. The kit uses spin columns which binds the DNA until the final step in which it is eluted during the addition of nuclease free water and centrifuged. The agarose gel was made as described earlier however three combs were taped together to create bigger wells so that a larger volume of the PCR product could be electrophoresed in one well. After the gel had been electrophoresed for thirty minutes the gel was placed on a portable UV transilluminator to excise the bands, taking care to remove as much extraneous agarose as possible and ensuring that correct personal protective equipment was used, in this case a lab coat, disposable gloves and a UV-blocking face shield. Once the bands had been excised they were placed in a 1.5 ml Eppendorf tube and weighed. An empty Eppendorf tube was also weighed. The difference in weight was calculated after which three volumes of Buffer QG was added. The tubes were then placed in a water bath at 55°C for ten minutes, or until the gel had dissolved, inverting two to three times to ensure full mixing. The colour of the mixture was checked so that it did not differ from the original colour of the buffer, which indicates problems with the TBE buffer. If a change from yellow to orange or violet was observed, then 10 µl of 3 M sodium acetate, pH 5 was added and

mixed. Once the gel had been dissolved, one volume of isopropanol was added to the solution, mixed and subsequently added to the spin column, and centrifuged for one minute at 12,000 rpm. The spin column has a maximum volume of 800 µl therefore any remainder of the gel mixture was added to the spin column, after the initial spin and the flow through discarded. Once all the gel mixture had been added to the spin column, 750 µl of the wash buffer PE (after the addition of 100% ethanol) was added and centrifuged for one minute at 12,000 rpm. It was recommended to incubate for two to five minutes for subsequent salt sensitive reactions. The flow through was discarded and the column was spun, empty, for an additional one minute to remove any buffer that remained. After this the spin column was placed in a 1.5 ml Eppendorf tube and 30 µl of EB buffer was added directly to the column. These were left for one minute to increase the final concentration of the DNA, before being centrifuged for one minute and the eluted DNA collected in the Eppendorf tube.

The LEPR PCR amplicon was also extracted from an agarose gel from smaller wells, in which only 25 µl was added. This meant that for each LEPR PCR (25 µl total), all of the product was loaded into one well. The weights of each of the gel extracts were measured.

2.4.14 Ligation of LEPR into pGEM T Vector

Before the ligation reaction was conducted the purified LEPR products were quantified to calculate the volume to be used for a 1:4 ratio with the vector. This was calculated using the following equation:

ng of vector X kb size of insert X 4

kb size of vector 1

50 X 2.4 X 4 = 160 ng
3.0 1

The reagents were left to defrost on ice after, which time the pGEM T vector and control insert DNA were briefly centrifuged. Three ligation reactions were set up in separate 0.5 ml Eppendorf tubes, one at a time using the reagents and volumes in Table 2.4.5 and mixed by pipetting. The ligation buffer was shaken before use each time. The tubes were wrapped in cling film and left overnight at room temperature for the maximum number of transformants. The ligations were set up as duplicates so that one set could be electrophoresed to check that it had been successful.

Table 2.4.5: Reagents and volumes used in the ligation of the *LEPR* insert into the pGEMT vector.

Reagent	LEPR Insert (µl)	Positive Control (µl)	Negative Control (µl)
2X Buffer	5	5	5
pGEMT	1	1	1
PCR Product	3	0	0
Control insert	0	2	0
T4 Ligase	1	1	1
dH_2O	0	1	3
Total	10	10	10

2.4.15 Transformation of pGEM T Vector using *E. coli*

After the ligation reaction had been completed overnight the tubes were spun briefly and placed on ice. 4 µl of each ligation was added to three labelled 15 ml Falcon tubes. A fourth tube was set up using 0.1 ng of uncut plasmid (supplied with the JM109 competent *E. coli* cells) to calculate transformation efficiency of the competent cells. 200 µl of JM109

competent *E. coli* cells were taken out of -70°C to defrost on ice. After which time 50 µl, of these cells, was added to each of the falcon tubes, gently flicked and placed on ice for twenty minutes. The pipette tips used to transfer the cells were cut to avoid shearing the competent cells. Once the incubation was complete, the tubes were then placed at 42°C for fifty seconds, then placed on ice for a further two minutes. 950µl of the transformants was added to each of the four falcon tubes and placed in a shaking incubator at 37°C for 1.5 hours at 320 rpm. After which time, 100 µl were placed on duplicate LB agar plates supplemented with ampicillin, X-Gal and IPTG and left to grow overnight at 37°C (see section 2.4.20). However, for the transformants containing pGEM-T and LEPR all of the 1ml was spread onto three plates for the maximum number of white colonies.

2.4.16 Extraction and Purification of pGEM-T Plasmid

Successful colonies were picked and inoculated in 10 ml of LB broth for growth overnight. Glycerol stocks were made from each bacterial culture and stored at –20°C; 500 µl of culture to 500 µl of 70% sterile glycerol.

To identify successful transformants i.e. those that had the LEPR insert in the correct orientation, mini preparations were initially carried out from the overnight cultures, followed by restriction digests. Those that produced 2.5 kb and 3 kb bands on an agarose gel were then selected for a maxi preparation to increase the yield and concentration of the plasmid for the restriction digestion and the subsequent ligation of the LEPR insert into the pCMV Tag 2B expression vector.

2.4.16.1 Mini Preparation of Overnight Culture

A GeneJET mini preparation kit (Fermentas) was used to extract and purify plasmid DNA. 5 ml of bacterial culture were added to a 15 ml Falcon tube and centrifuged at 4000 rpm for five minutes. Supernatant was then removed and the pellet was resuspended using 250 µl of resuspension solution. RNase A (supplied with the kit) was added to the solution prior to first use. The solution was then transferred to a 1.5 ml Eppendorf tube to which 250 µl of lysis buffer was added and mixed by inverting six times. This was followed by 350 µl of neutralising solution. Once the contents had been mixed, by inverting the tube six times it was then centrifuged for five minutes at 12000 rpm, using a microcentrifuge. The supernatant was removed without disturbing the pellet and transferred to a spin column provided with the kit, where it was centrifuged again for one minute at 12000 rpm. The flow through was discarded and 500 µl of wash buffer was added. 170 ml of 100% ethanol was added to the wash buffer prior to first use. Tubes were centrifuged for a fourth time and the flow through discarded. Columns were washed with another 500 µl of wash buffer and centrifuged for one minute at 12000 rpm. Flow through was discarded and the spin columns were centrifuged empty to remove any residual wash buffer. Fifty microlitres of elution buffer was then added to the centre of the spin column, left to incubate at room temperature for two minutes then centrifuged for a final time at 12000 rpm to collect the elute which contains the plasmid DNA. Once the plasmid DNA had been collected 5 µl was mixed with 2 µl of 5X loading dye, which was loaded onto a 1.0% (w/v) agarose midi gel and electrophoresed to check that the mini prep had been successful.

2.4.16.2 Maxi Preparation of Overnight Cultures

Extraction and purification of the plasmid was carried out using a PureLink™ HiPure plasmid filter purification kit (Invitrogen) which uses a filter column containing an anion-exchange resin, to allow the negatively charged phosphate backbone of the DNA to bind to the positively charged resin of the column and the impurities to be washed away. DNA is eluted in the high salt elution buffer and desalted by ethanol precipitation for use in subsequent ligation/restriction digest reactions.

Conical flasks were fitted underneath the filter columns, to collect the flow through and were prepared before use by adding 30 ml of equilibrium buffer and left to drain by gravity flow. During which time the 100 ml of overnight culture, which was grown overnight at 37°C in two 50 ml Falcon tubes, was centrifuged at 4000 g to pellet the cells, in a bench top centrifuge. All medium was removed taking care not to dislodge the pellet. 10 ml of Resuspension Buffer, into which RNase A had previously been added, was added to each one of the pellets until homogeneous, after which time it was transferred to the second pellet for resuspension. 10 ml of Lysis Buffer was subsequently added and mixed by inverting five times and incubated at room temperature for five minutes. 10 ml of Precipitation Buffer was added and mixed immediately by inverting and centrifuged at 15,000 g for ten minutes at room temperature. The supernatant was carefully removed and loaded into the equilibrated column and left to drain by gravity flow. 60 ml of Wash Buffer was added to the column and again left to drain by gravity flow. Once the Wash Buffer had completely drained the conical flask was removed and a sterile 50 ml Falcon tube was placed underneath. 15 ml of Elution Buffer was added to the column and allowed to drain by gravity flow to elute the DNA. 10.5 ml of isopropanol alcohol was added to the eluted

DNA and mixed well to precipitate the DNA before centrifugation at 15,000 g for thirty minutes at 4°C. The supernatant was discarded and 5 ml of 70% ethanol was added to the pellet, followed by centrifugation at 15,000 g for five minutes at 4°C. The pellet was then allowed to air dry for ten minutes after which time the plasmid DNA was resuspended in 500 µl of TE buffer and stored at -20°C.

2.4.17 Restriction Digest of Plasmid Mini/Maxi Preparations

After the plasmid DNA was extracted by mini prep or maxi prep methods they were then subjected to a double restriction digest to determine which of the plasmids contain the LEPR insert in the correct orientation. As there are two restriction enzymes, the buffer supplied with the *BamHI* enzyme was used but as this does not provide optimum conditions for *SalI*, the volume of this enzyme was doubled. The restriction digest was set up in 0.5 ml Eppendorf tubes by adding 5 µl of enzyme buffer, 1 µl of *BamHI* enzyme (10 Weiss units/µl), 2 µl of *SalI* enzyme (10 Weiss units/µl), 2 µl of plasmid DNA with the remaining total being taken up to 50 µl with deionised water. The tubes were then wrapped in cling film and placed at 37°C overnight.

Once the products had been digested overnight, 10 µl was mixed with 3 µl 5X loading dye and subsequently subjected to electrophoresis on a 1% (w/v) agarose midi gel. To determine whether the plasmid had been digested successfully and contained the LEPR insert in the correct orientation, 7 µl of each undigested mini or maxi prep, mixed with 2 µl was loaded alongside each of the corresponding double digests.

Plasmids that had been successfully digested and produced bands of 3 kb and 2.5 kb were thought to contain the *LEPR* insert in the correct orientation. Those that did were used for ligation into the expression vector pCMV Tag 2B.

As the digested product of the mini prep contains two bands, 3 kb and 2.5 kb, the 2.5 kb band had to extracted from the gel using the method previously described (2.4.8). Therefore more double digests of the mini preparation had to be set up to produce a large enough quantity to load into the large wells within the agarose gel, to be used in the ligation into pCMV Tag 2B Expression Vector.

2.4.18 Ligation of the LEPR insert into Expression vector, pCMV Tag 2B

The digested LEPR that was extracted from the gel was used to ligate into the expression vector. Five ligation reactions were set up, three as controls and two experimental. The ideal insert to vector molar ratio of DNA was calculated using the following calculation to get a 1:1 ratio:

$$X\ \mu g \text{ of insert} = \frac{(\text{base pairs of insert})\ (0.1\ \mu g \text{ of pCMV Tag 2B})}{4.3 \text{ kb of pCMV Tag 2B}}$$

$$\frac{2500 \times 0.1}{4.3} = 58.12\ \mu g$$

Before the ligations were set up the pCMV Tag 2B vector was digested using the restriction enzymes *SalI* and *BamHI*, to enable the ligation to occur. Ligations were set up as in Table 2.5.7, in duplicate so that one set could be electrophoresed, to check that the ligation had been successful. Three controls were carried out to determine the effectiveness of the double restriction digest of the vector (control one), to determine whether any uncut plasmid remained (control two) and to verify if any of the original plasmid remained (control three). Two experimental ligations were performed to determine the best insert to vector ratio.

Table 2.4.6: Ligation reaction of the PCR product and pCMV Tag 2B expression vector.

Ligation Components	**Control Reactions**			**Experimental Reactions**	
	1	**2**	**3**	**4**	**5**
pCMVTag 2B (µl)	1	1	-	1	1
PCR insert (µl)	-	-	2	2	4
10X Buffer (µl)	1	1	1	1	1
T4 Ligase (µl)	1	-	1	1	1
dH_2O (µl)	7	8	6	5	3
Total Volume (µl)	10	10	10	10	10

2.4.19 Transformation of pCMV Tag 2B expression vector using *E.coli* cells

After the ligation reaction had been completed overnight the tubes were spun briefly and placed on ice. 2 µl of β-mercaptoethanol was added to five separately labelled 15 ml Falcon tubes, which already contained 50 µl of competent *E. coli* cells and incubated on ice for ten minutes. After which time 2 µl of the control reactions and 4 µl of the

experimental reactions were added to the five 15 ml Falcon tubes. These were then left to incubate for a further thirty minutes on ice before being added to a water bath set to 42°C for thirty seconds before incubating on ice for an additional two minutes. 950 µl of NY broth was added to each of the five Flacon tubes, which prior to use had been preheated to 42°C. The tubes were then incubated at 37°C in a shaking incubator at 320 rpm, for one hour. After which time, 100 µl were plated on duplicate LB agar plates supplemented with kanamycin and left to grow overnight at 37°C. However for the experimental transformants all of the 1 ml was spread onto four plates for the maximum number of colonies.

2.4.20 LB Agar supplemented with Ampicillin, X-Gal and IPTG

200 ml of LB agar was prepared as previously described. After it was autoclaved and had cooled so that it was hand hot, ampicillin, X-Gal and IPTG (Fermentas) were added to give a final concentration of 50 µg/ml for the ampicillin and 100 µg/mleach for X-Gal and IPTG. 10 ml of this mixture was subsequently poured into twenty 90 mm petri dishes and left to set before storing at 4°C until needed.

2.4.21 LB agar supplemented with Kanamycin

200 ml of LB agar was made and after autoclaving left to cool to hand hot after which time 1 ml of 10 mg/ml stock solution of kanamycin was added and mixed well before pouring 10 ml into twenty 90 mm petri dishes and left to set before being stored at 4°C until needed.

2.4.22 Quantification of DNA and RNA samples

The concentration of DNA samples was determined by using an UV spectrophotometer (Cecil) initially by using a 1 in 1000 dilution of the DNA sample. If this failed to produce a positive reading the dilution was changed to a 1 in 500. Prior to use the spectrophotometer was set to 260 nm and a water blank was used as the negative control (blank). The concentration of DNA samples was calculated by using the equation below. An OD reading of 1 at 260 nm equals 50 μg/ml.

DNA concentration (μg/ml) = 50 μg/ml x OD260 X dilution factor

For the determination of RNA samples the same method was used except the spectrophotometer was set to 260 nm and the following calculation was used. An OD reading of 1 at 260 nm equals 40 μg/ml.

RNA concentration (μg/ml) = 40 μg/ml x OD260 x dilution factor.

2.5.0 Genotyping of *HHEX-IDE* Gene locus in the South Asian Population

2.5.1 Sequenom iPLEX Assay Design

The Sequenom iPLEX assay was used to genotype thirty-seven selected SNPs surrounding the *HHEX-IDE* gene locus in the South Asian (SA) cohort. In brief the SNPs of interest were selected using the National Centre for Biotechnology Information (NCBI) website and the HapMap database. The MassARRAY assay design software was then used to design a set of three primers for each SNP in a multiplex assay.

2.5.2 Quantification of DNA

To determine DNA concentration, twenty samples were selected at random from the SA cohort (see section 2.4.20).

Once the concentration of the DNA was calculated this was then extrapolated to the rest of the South Asian samples. As there was a wide range of concentrations from the samples tested a 1/100 dilution was done for all the samples so that the final working concentration would be close to 2.5 ng/µl, in a total volume of 20 µl, which is within the concentration range of the iPLEX assay. This would be sufficient for ten iPLEX reactions. All the South Asian DNA samples were placed in numerical order in eight 96-well plates and stored at –20°C.

2.5.3 MassARRAY Software

This software was used in designing the primers for the thirty-seven SNPs so that a maximum number of reactions can be carried out in one assay without any of the primers or products interfering with one another or a reduction in performance in any of the SNPs when the Matrix-Assisted Laser Desoprtion/Ionisation Time-Of-Flight (MALDI-TOF) stage is performed. Three sets of primers are designed; forward and reverse primers are designed for the multiplex PCR amplification with a third primer (extension primer) for the iPLEX extension reaction. The extension primer hybridises adjacent to the SNP, which is then extended by a single terminal base. The product produced varies in mass depending on the base added and are easily differentiated using MALDI-TOF mass spectrometry.

The Ensembl database was used to acquire the sequence surrounding each SNP for uploading into the MassARRAY software. The SNP itself was placed in square brackets with a forward slash separating the two alleles. The SNP information was copied and pasted into the SNP Group and the Assay Group name was changed, for the input file for the Assay design. There were four different aspects of the software to consider. Firstly the assay design was set up to be single base extension (SBE) Mass Extend. The presets tab was selected and the settings were changed to high multiplexing (thirty-six) with regard to the SBE presets (iPLEX). Once that was changed the maximum multiplex level was set to forty. SNP capture details were set to an amplicon length of between 80 bp and 160 bp, with the optimum being 100 bp. The extend primer design was set so that the length of the extended primer was between 15 bp and 30 bp and the MS multiplexing remained unchanged. The lower limit mass of the extension primer was left at 4500 Da and the upper limit was left at 8500 Da. The minimum distance between the expected mass peak of two

assay analytes was kept at fifty. The minimum distance between two MassEXTEND primers was kept at ten. Once the program was completed the final results were displayed indicating how many assays the SNPs could be incorporated into, the full report saved as an excel file, including information regarding the assays and the primers. Primers were labelled as the SNP ID, with either forward (F), reverse (R) or extension (E) used as an extra identifier. The F and R primers were ordered at a different concentration to the extension primers, 25 nmoles and 200 nmoles respectively.

2.5.4 Sequenom iPLEX Reaction

The primers were centrifuged then rehydrated using deionised water. The forward and reverse primers were made up to a final concentration of 100 pmoles/μl and the extension primers were made up to a final concentration of 500 pmoles/μl and stored at –20°C. Primers were labelled with the last four digits of their SNP ID along with either F, R or E. Primer sequences are given in Appendix II.

2.5.5 Initial PCR

Two microlitres of the South Asian DNA samples were transferred from the eight 96 well plates into two 384 well plates, with water placed randomly as blanks. Details of the sample positions were manually recorded. This was carried out twice as two assays were performed on each plate. The plates were labelled accordingly; Assay one plate one, Assay one plate two, Assay two plate one and Assay two, plate two. Each assay was carried out separately to reduce possible contamination and error.

Table 2.5.1: Checklist for the initial PCR

Reagent	Volume for one reaction in 384 plate (µl)	Volume for 500 reactions in 384 plate (µl)
dH_2O	0.850	425.0
PCR buffer	0.625	312.5
$MgCl_2$ (25 mM)	0.325	162.5
dNTPs (25 mM)	0.100	50.00
Primer Mix (500 nM)	1.000	500.0
HotStart Taq (5 U/µl)	0.100	50.00
Total (µl)	3.000	1500

The PCR mix was made for a total of 500 reactions to provide sufficient quantities for the whole plate (Table 2.5.1). The primer mix was made first, adding 2.5 µl of the forward and reverse stock primers (100 pmoles/µl) with the rest of the final 500 µl total being made up with dH_2O. The rest of the reagents were added into the 2 ml Eppendorf tube with the HotStart Taq being added last. The 3 µl of PCR mix was dispensed into each well, using a 16-channel pipette. The plates were then centrifuged so that all the reagents were mixed well together and placed on a Peltier thermal cycler (PTC-225) (Table 2.5.2).

Once the initial PCR was completed the Sequenom iPLEX extension reaction was performed and the resulting mixture of products were ran on the Sequenom MassARRAY system. The different stages of the Sequenom iPLEX extension reactions were all performed either by myself, Ms Marlene Attard or Dr Petros Takousis at some stage during the project. The final stage of the reaction were performed by, either Ms Marlene Attard or Dr Petros Takousis.

Table 2.5.2: PCR cycle conditions for the initial reaction

PCR Step	Temperature	Time	Number of cycles
Initial denature	95°C	15 minutes	
Denaturation	95°C	20 seconds	
Annealing	65°C	30 seconds	
Extension	72°C	1 minute	X 4
Denaturation	95°C	20 seconds	
Annealing	58°C	30 seconds	
Extension	72°C	1 minute	X 4
Denaturation	95°C	20 seconds	
Annealing	53°C	30 seconds	
Extension	72°C	1 minute	X 38
Final extension	72°C	3 minutes	

2.5.6 SAP Clean Up Method

Once the initial PCR was completed the products were then treated with shrimp alkaline phosphatase (SAP), which is a cleanup step to degrade any remaining dNTPs and primers to prevent them from being carried over into subsequent reaction. The reagents were added to a 1.5 ml Eppendorf tube in the order given in Table 2.5.3. The tube was vortexed, centrifuged and allowed to reach room temperature to ensure that the enzyme mixes with the other reagents before adding 2 µl to each well using a 16-channel pipette. Once the enzyme mix had been added the plate was sealed with an adhesive lid, centrifuged then placed on to the thermal cycler for incubation under the following conditions: 37°C for twenty minutes, 80°C for five minutes and 4°C for five minutes. The temperature was increased to 80°C to deactivate the SAP enzyme. Once the clean up step was completed the extension was carried out.

Table 2.5.3: SAP Clean up method

Reagent	Volume for 1 reaction (µl)	Volume for 384 reactions (µl)
Water (HPLC Grade)	1.53	810.9
SAP Buffer	0.17	90.1
SAP Enzyme	0.30	159.0
Total (µl)	2.00	1060.0

2.5.7 Sequenom Single Base Extension (SBE) Reaction

The volume of each extension primer used in the reaction varied depending on the mass. Primers with a high mass were added at a higher concentration in the primer mix as they are harder to ionise compared to the lower mass primers. This was determined by using the Regression method for dilution of extension primers as recommended by Sequenom, detailed in Appendix A of the iPLEX Gold Application Guide. In summary, the concentrations of the extension primers were adjusted on an individual basis to ensure strong primer peaks well above background, accounting for each primer's molecular mass and the total number of primers in the multiplex. The volume of the primer mix was made up to 498.2 µl with dH_2O (Tables 2.5.4 and 2.5.5).

Equation 1:

LN (Mass of the primer) – 7.82 = Final reaction concentration (µM)

Equation 2:

Concentration in Primer mix (µM) = Final reaction concentration (µM) x 11.194

Equation 3:

Volume in 1 plate (µl)	= Concentration in Primer mix (µM)	x Total volume in 1 plate	/ Extension primer stock concentration (µM)

The worked example is for the SNP rs882136E from Table 2.6.9.

Equation 1:

LN (4481.9) – 7.82 =

8.408 – 7.82 = 0.588 µM

Equation 2:

0.588 x 11.194 = 6.58 µM

Equation 3:

6.58 x 498.2 / 500 = 6.56 µl

Table 2.5.4: Extension reaction primer mix for assay one

Extension Primer Name	Primer Mass	Final Reaction Concentration (μM)	Concentration in Primer Mix (μM)	Volume in 1 384 Plate (μl)
rs882136E	4481.9	0.588	6.58	6.56
rs17851141E	4584.0	0.610	6.83	6.81
rs2497349E	4752.1	0.646	7.24	7.21
rs12262390E	4857.2	0.668	7.48	7.45
rs10509646E	5169.4	0.731	8.18	8.15
rs1544210E	5252.4	0.746	8.36	8.33
rs34266926E	5386.5	0.772	8.64	8.61
rs2488068E	5600.6	0.811	9.07	9.04
rs2497311E	5741.8	0.836	9.35	9.32
rs17107841E	6011.9	0.881	9.87	9.83
rs2497350E	6195.0	0.911	10.20	10.17
rs9420592E	6372.2	0.940	10.52	10.48
rs2488067E	6396.2	0.943	10.56	10.52
rs11187146E	6497.2	0.959	10.74	10.70
rs7081351E	6678.4	0.987	11.04	11.00
rs4933236E	6831.5	1.009	11.30	11.26
rs17374868E	7027.6	1.038	11.61	11.57
rs12765131E	7056.6	1.042	11.66	11.62
rs10882102E	7325.8	1.079	12.08	12.04
rs2096177E	7672.0	1.125	12.60	12.55
			Water	304.99
			Total	498.2

Table 2.5.5: Extension reaction primer mix for those SNPs within assay two

Extension Primer Name	Primer Mass	Final Reaction Concentration (μM)	Concentration in Primer Mix (μM)	Volume of Primers in 1 384 Plate
rs2229328E	4488.9	0.589	6.60	6.57
rs2497304E	4866.2	0.670	7.50	7.47
rs2488087E	4995.3	0.696	7.79	7.77
rs2488073E	5139.4	0.725	8.11	8.08
rs1418387E	5228.4	0.742	8.30	8.27
rs2497309E	5445.6	0.783	8.76	8.73
rs11187144E	5554.6	0.802	8.98	8.95
rs1539330E	5828.8	0.851	9.52	9.49
rs2275729E	5902.9	0.863	9.66	9.63
rs11187161E	5934.9	0.869	9.72	9.69
rs10437472E	6072.0	0.891	9.98	9.94
rs11187158E	6246.1	0.920	10.30	10.26
rs11187152E	6745.4	0.997	11.16	11.12
rs11597458E	6799.5	1.005	11.25	11.21
rs947591E	7062.6	1.043	11.67	11.63
rs12780253E	7148.7	1.055	11.81	11.76
rs11187157E	7419.9	1.092	12.22	12.18
			Water	335.46
			Total	498.2

The reagents and volumes for the extension step are given in Table 2.5.6. They vary for the two assays due to the numbers of SNPs in each, twenty and seventeen named high PLEX and low PLEX respectively.

Once the plate set up was complete, it was centrifuged then placed on a thermal cycler (Table 2.5.7).

Table 2.5.6: iPLEX reaction mix for the 20-PLEX assay (A) and for the 17-PLEX assay (B).

A)

Reagent	Concentration in 9 µl	Volume for 1 reaction (µl)	Volume for 384 reactions (µl)
Water (HPLC)	N/A	0.619	328.07
iPLEX Buffer	0.222X	0.200	106.00
iPLEX Termination Mix	1.000X	0.200	106.00
iPLEX Enzyme	1.000X	0.041	498.20
Total		2.000	1038.27

B)

Reagent	Concentration in 9 µl	Volume for 1 reaction (µl)	Volume for 384 reactions (µl)
Water (HPLC)	N/A	0.7395	391.97
iPLEX Buffer	0.222X	0.2000	106.00
iPLEX Termination Mix	0.500X	0.1000	53.00
iPLEX Enzyme	0.500X	0.0205	498.12
Total		2.0000	1060.00

Table 2.5.7: PCR cycle times and temperatures for the extension reaction

PCR Step	Temperature	Time	Number of cycles
Initial Denature	94°C	30 seconds	
Denaturation	94°C	5 seconds	
Annealing	52°C	5 seconds	
Extension	80°C	5 seconds	X 4
Denaturation	94°C	5 seconds	
Annealing	52°C	5 seconds	
Extension	80°C	5 seconds	X 39
Final Extension	72°C	3 minutes	

2.5.8 MALDI-TOF Mass Spectrometry

Once the extension primer reaction has been performed the products were treated with an anion exchange resin to remove adduct-forming ions that can complicate with the MALDI-TOF mass spectrum. The MassExtend clean resin is initially spread onto a solid block containing 384-wells. A small amount of the resin is placed on the left hand side of the block and spread across the block so that it fills all of the wells. Any remaining resin is removed and the block is left to air dry for thirty minutes. After which time the plates were briefly spun and the adhesive film removed. The plate was then placed upside down over the block and carefully aligned then both the block and plate were tipped upside down so that the resin falls into the wells. 16 µl of dH_2O was then pipetted into each of the wells. Adhesive lids were then placed on the plate and the plate was regularly rotated for ten minutes so that the resin was suspended within the solution. The plate was then centrifuged in a bench top centrifuge for five minutes at 3000 rpm.

15-25 nl of the extended primer products were robotically (MassARRAY NanoDispenser) dispensed onto SecrtroCHIP array containing matrix spots (3-hydroxylpicolinic acid) spots in a 384 well format. Calibrant was also spotted onto the chip ten times. This contains three standard oligonucleotides of known mass that is used to compare with the samples: 5100, 8500 and 10,000 mass units (Appendix III).

Data files, including plate layouts and the expected mass for each of the three genotypes expected (Appendix IV), were uploaded onto the Sequenom server so that each sample can be identified.

2.6.0 Genotyping *FTO* variants within the South Asian Population using the Sequenom iPLEX Assay

To ascertain whether variants in the *FTO* gene influence fetal birth weight in the South Asian population, selected variants that have been previously published and found to be associated with obesity and type II diabetes were genotyped using the Sequenom iPLEX assay. The *FTO* SNPs include rs9939609, rs17817449, rs1421085, rs3751812, rs8050136, rs1121980, rs9940128, rs7193144 and rs9939973. The parameters were the same as for the design of the *HHEX-IDE* assay (Section 2.5).

2.6.1 Sequenom iPLEX Reaction

The protocol for the initial PCR, clean up method and extension reaction is the same as in sections 2.5.7, 2.5.8 and 2.5.9 the extension reaction primer mix used is given in Table 2.6.1. The high PLEX protocol was performed, as there are twenty-seven SNPs within the assay. Two 384-well plates were used to genotype all the samples, which also included several water blanks and replicates. *FTO* primer names (forward, reverse and extension) and sequences are given in Appendix IV.

Table 2.6.1: Extension reaction primer mix for the *FTO* assay.

Extension Primer Name	Primer Mass	Final Reaction Concentration (uM)	Concentration Primer Mix (uM)	Volume in 1 384 plate (ul)
rs3751812E	4432.9	0.577	6.46	6.43
rs11781222E	4528.9	0.598	6.70	6.67
rs4872145E	4534.9	0.600	6.71	6.69
rs13439692E	4890.2	0.675	7.56	7.53
rs10098490E	5078.3	0.713	7.98	7.95
rs7833754E	5200.4	0.736	8.24	8.21
rs2928684E	5241.4	0.744	8.33	8.30
rs17698981E	5413.5	0.777	8.69	8.66
rs9939973E	5613.7	0.813	9.10	9.07
rs9940128E	5621.7	0.814	9.12	9.08
rs8050136E	5780.8	0.842	9.43	9.39
rs10421768E	6004.9	0.880	9.85	9.82
rs17089358E	6037.9	0.886	9.92	9.88
rs1421085E	6207.0	0.913	10.22	10.19
rs721183E	6407.2	0.945	10.58	10.54
rs34968988E	6495.2	0.959	10.73	10.69
rs6986233E	6590.3	0.973	10.90	10.86
rs9939609E	6688.4	0.988	11.06	11.02
rs7834536E	6828.5	1.009	11.29	11.25
rs2978477E	7050.6	1.041	11.65	11.61
rs17817449E	7082.6	1.045	11.70	11.66
rs17089331E	7210.7	1.063	11.90	11.86
rs4872153E	7219.7	1.065	11.92	11.87
rs7193144E	7624.0	1.119	12.53	12.48
rs1121980E	7632.0	1.120	12.54	12.49
rs34146184E	7837.1	1.147	12.84	12.79
rs10104250E	7997.2	1.167	13.06	13.01
			Water	228.17
			Total	498.2

2.7.0 Genotyping the variant rs11708067, close to *ADCY5* using the TaqMan Assay in South Asian Cohort.

To investigate whether there was an association between the maternal *ADCY5* locus and birth weight, South Asian samples were genotyped using the TaqMan SNP Genotyping Assay for the variant rs11708067. The SNP rs9883204 has been previously associated with low birth weight and this variant is in high LD with rs11708067. Real time PCR was chosen to easily genotype the SNP rs11708067 as it uses two primers to amplify the region of interest (forward and reverse) and uses two different fluorescent minor groove binder (MGB) probes to discriminate between the two alleles. At the 3′ end of the probes the non-fluorescent quencher (NFQ) is attached, which prevents the dye from emitting a signal when the two are in close proximity to each other. During the PCR cycles, the primers bind to their complementary sequences and the same occurs with the two probes, which bind to the region surrounding the SNP. DNA Taq polymerase extends the new DNA strand from the 3′ end of the primer to the 5′ end of the probe and cleaves the dye in the process. When this occurs the dye increases its fluorescent signal, as it is no longer close to the quencher and indicates which allele is present in the sample. If the VIC dye is emitted only then samples are homozygous for the one allele and if the FAM dye is only present then the samples are homozygous for the alternative allele. If both dyes are emitted then the samples are heterozygous, so have both alleles (Figure 2.7.1).

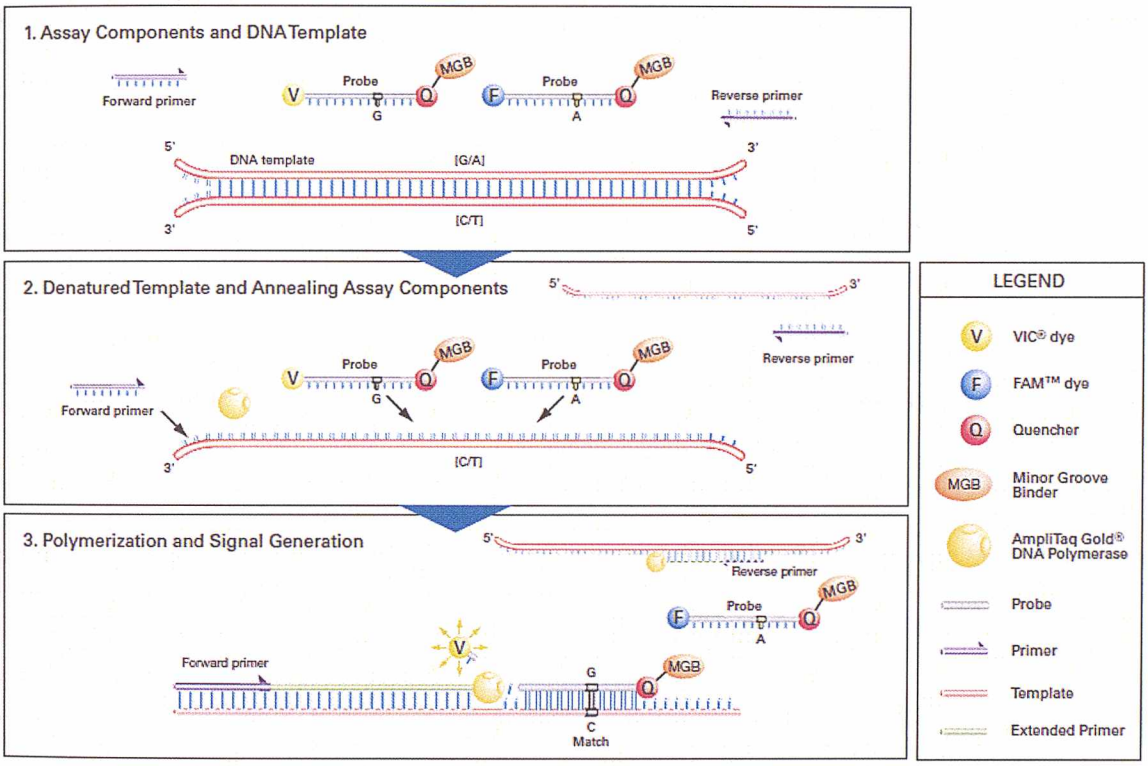


Figure 2.7.1: Diagrammatic representation of the TaqMan SNP genotyping assay. Three stages of the reaction process is shown; 1) all the components of the assay 2) during the denaturation step of the PCR, where the forward primer binds to the template DNA and depending on the allele at the SNP of interest one of the two probes will bind 3) extension step of the PCR, where the DNA polymerase is extending the forward primer based on the template strand, and cleaving the dye from the probe as it extends towards it.

Once the PCR step was performed the plates were analysed during the post-read step where the genotypes were called depending upon the fluorescence of the samples.

2.7.1 TaqMan SNP Genotyping Assay Set-up

DNA samples were diluted 1 in 25 for a final concentration of around 20 ng, in 2.45 µl and set up on eight 96-well plates, arranged according to their sample number. Within these plates, duplicates and water blanks were also placed randomly. Once the 96-well plates had been set up the DNA was transferred to two 384-well plates.

As there were 768 samples, the genotyping assay, which contains the primers and the two probes, was ordered as small scale, which was enough for 1500 reactions and the genotyping master mix was ordered to provide enough mixture for 400 reactions.

Once the DNA was added to the 384-well plates the master mix was prepared. The genotyping mixture was stored in restricted light, in an ice bucket with a lid, to prevent repeated exposure to light, as this would degrade the fluorescent dyes contained within it. The master mix was made to contain enough 2 X PCR master mix and genotyping assay mix for 384 reactions, with sufficient additional mix to allow for any pipetting variability, so the volume for one reaction was multiplied by 500 (Table 2.7.1).

Table 2.7.1: Master mix volumes and reagents for the TaqMan genotyping assay.

Reagent	1 X PCR using 40X mix	1 X 384 Plate
Taqman universal PCR Master Mix (2)	2.5 µl	1250 µl
Assay Mix	0.125 µl	62.5 µl
DNA	2.4 µl each	2.4 µl each
Total	5 µl	5 µl each

Once the master mix had been added to the 384 well plates they were centrifuged to aid in mixing, then placed on to the Real Time PCR machine (ABI 7900HT). SDS software, provided by the supplier Applied Biosystems, was used to set up the PCR program under the specified conditions (Table 2.7.2) and to select the two fluorescent dyes that were being used. The two alleles, A and G were given a reference to distinguish between the three genotypes. The A allele was given the reference LT-1 and the G allele was given the reference LT-2.

Table 2.7.2: PCR conditions for the TaqMan SNP Genotyping Assay

PCR Conditions	Duration and Temperature
Hold Step	10 min at 95°C
Melting step for 40 cycles	15 secs at 92°C 1 min at 60°C

2.7.2 Allelic Discrimination Plate Read

After the PCR had been completed an endpoint plate read was carried out to determine the genotype the DNA samples. This was carried out using the RT-PCR machine. The SDS

software uses the fluorescent signals measured during the RT-PCR reaction for each allele to determine the genotype for each sample. These are displayed as an allelic discrimination plot, which were initially checked as a quality control step. This is to ensure that the success rate was >90%, the water controls had been given no genotype and three distinct genotypes were viewed.

2.8.0 Computer Modelling of the Leptin Receptor

2.8.1 UNIPROT Knowledgebase

The protein sequence of the Human leptin receptor was downloaded from the protein knowledgebase, UNIPROT (www.uniprot.org). Leptin receptor was searched within the database (Figure 2.8.1), which revealed several entries (Figure 2.8.2). The Human leptin receptor was selected (P48357). The results for this entry were searched to identify the sequence information amongst the other details displayed on the page including properties, function and location. As there are several isoforms of the leptin receptor, care was taken to select the correct one, isoform B (OBRb) under the Sequence heading. The FASTA format link was selected and the leptin receptor sequence (Appendix V) was copied from the resulting page that was displayed. It was then saved into a word document to be used to upload onto the I-TASSER server.

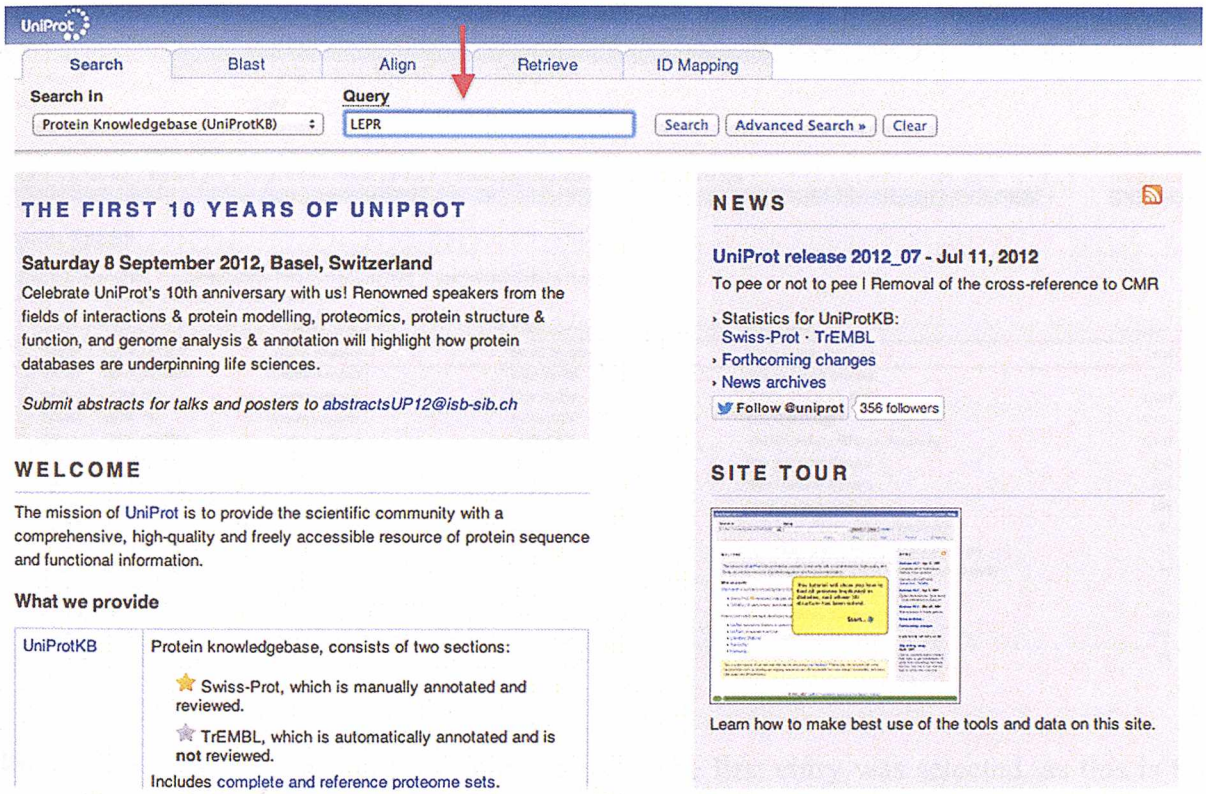
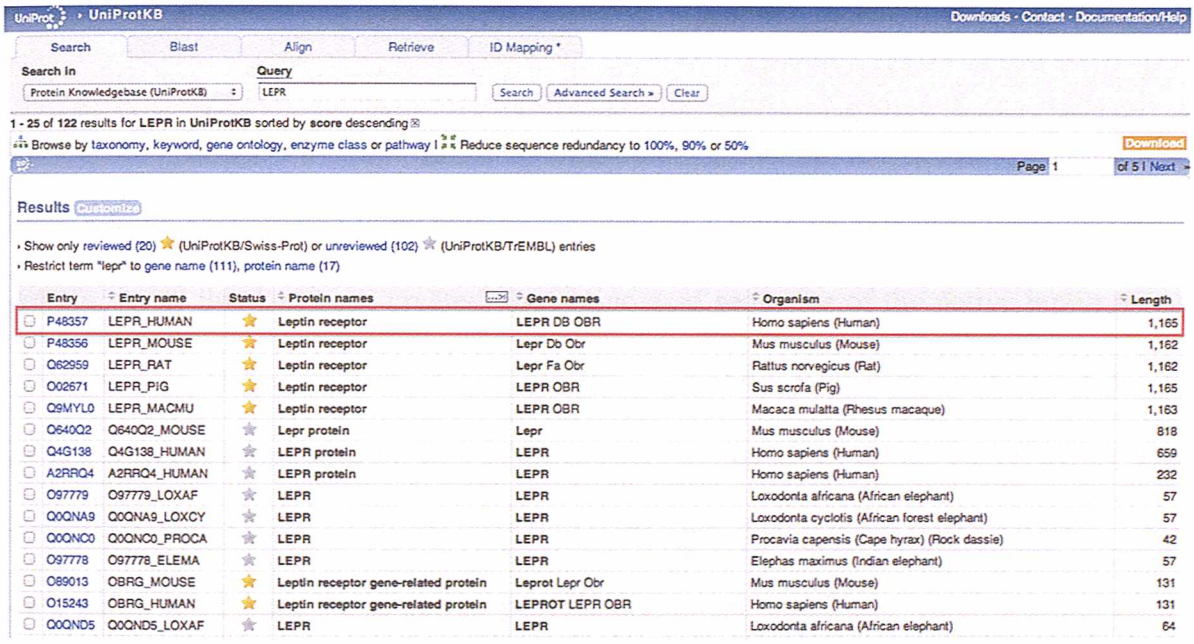


Figure 2.8.1: Homepage of the UniProt webpage. This was used to search for the leptin receptor sequence, under the query option (highlighted in red).



Entry	Entry name	Status	Protein names	Gene names	Organism	Length
P48357	LEPR_HUMAN	★	Leptin receptor	LEPR DB OBR	Homo sapiens (Human)	1,165
P48356	LEPR_MOUSE	★	Leptin receptor	Lepr Db Obr	Mus musculus (Mouse)	1,162
Q62959	LEPR_RAT	★	Leptin receptor	Lepr Fa Obr	Rattus norvegicus (Rat)	1,162
O02671	LEPR_PIG	★	Leptin receptor	LEPR OBR	Sus scrofa (Pig)	1,165
Q9MYL0	LEPR_MACMU	★	Leptin receptor	LEPR OBR	Macaca mulatta (Rhesus macaque)	1,163
Q640Q2	Q640Q2_MOUSE	☆	Lepr protein	Lepr	Mus musculus (Mouse)	818
Q4G138	Q4G138_HUMAN	☆	LEPR protein	LEPR	Homo sapiens (Human)	659
A2RRQ4	A2RRQ4_HUMAN	☆	LEPR protein	LEPR	Homo sapiens (Human)	232
O97779	O97779_LOXAF	☆	LEPR	LEPR	Loxodonta africana (African elephant)	57
Q0QNA9	Q0QNA9_LOXCY	☆	LEPR	LEPR	Loxodonta cyclotis (African forest elephant)	57
Q0QNC0	Q0QNC0_PROCA	☆	LEPR	LEPR	Procavia capensis (Cape hyrax) (Rock dassie)	42
O97778	O97778_ELEMA	☆	LEPR	LEPR	Elephas maximus (Indian elephant)	57
O89013	OBRG_MOUSE	★	Leptin receptor gene-related protein	Leprot Lepr Obr	Mus musculus (Mouse)	131
O15243	OBRG_HUMAN	★	Leptin receptor gene-related protein	LEPROT LEPR OBR	Homo sapiens (Human)	131
Q0QND5	Q0QND5_LOXAF	☆	LEPR	LEPR	Loxodonta africana (African elephant)	64

Figure 2.8.2: Search results for leptin receptor. The first entry was selected, as this is the human form of the protein (highlighted in red).

2.8.2 The I-Tasser Server.

Before the I-TASSER server (http://zhanglab.ccmb.med.umich.edu/I-TASSER/) could be used, a login was created to be able to use the site. This login was used to upload a query protein. After this had been created, the amino acid sequence of the native leptin receptor was uploaded to the relevant section on the I-TASSER web page (Figure 2.8.3). Each protein sequence uploaded on to I-TASSER server was given an ID for reference use, to distinguish between the many proteins that were uploaded. Further down the page, options were given to restrain the search or provide additional templates, which was not used. A notification to the email address provided when assigning a login was sent when the structure of the query protein had been determined. Only one query protein could be uploaded at a time therefore once the structure of one query protein had been determined, another was subsequently uploaded. The same method was used to upload each additional file. The native LEPR was uploaded first and given the reference 'LEPRnative'.

The mutations at the three different positions within the leptin receptor (Lys109Arg, Gln223Arg, and Lys656Asn) were created as separate files, from the native leptin receptor sequence. Only one mutation was created from the native leptin receptor sequence at one time and uploaded on to the I-TASSER server. They were given the ID references LEPR109, LEPR223 and LEPR656, respectively.

Once the query protein structure had been predicted the results were viewed as a separate webpage.

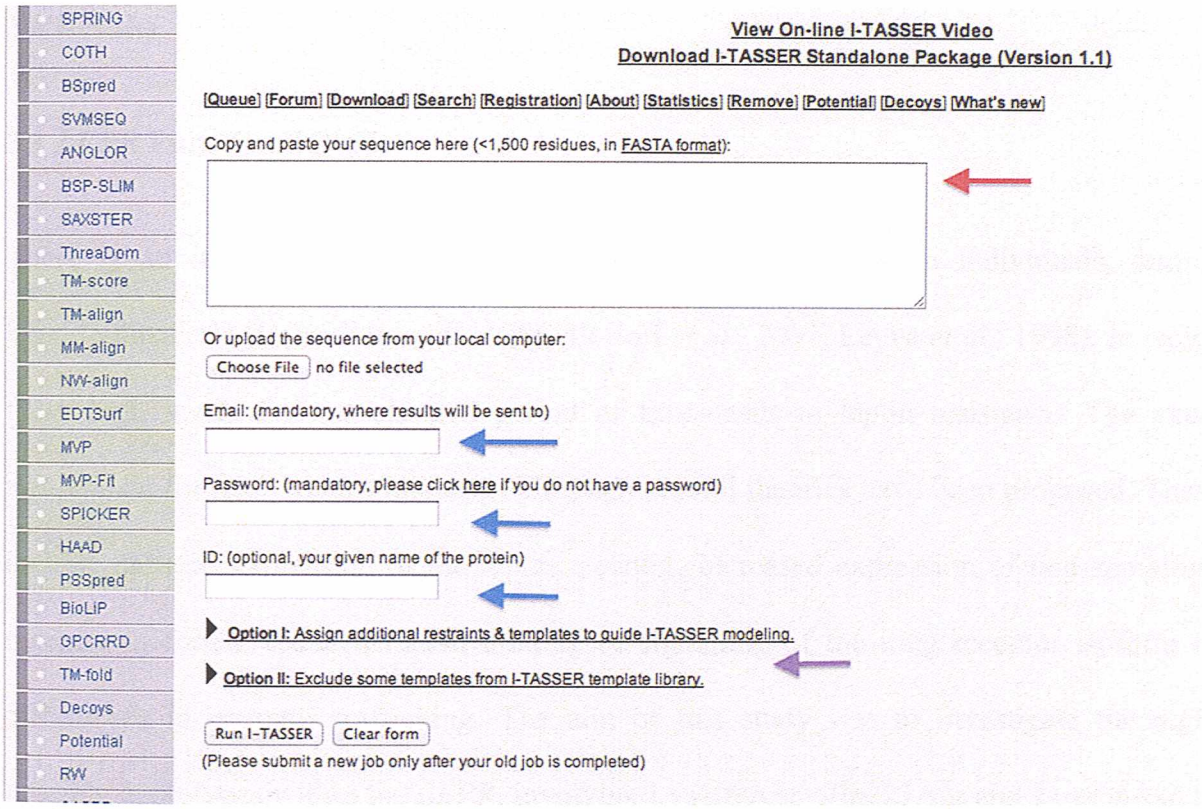


Figure 2.8.3: I-TASSER homepage. The leptin receptor sequence was submitted in FASTA format (highlighted in red) and login details were provided along with an ID for the protein (highlighted in blue). The options given were left unchanged (highlighted in purple).

3.0 Results

3.1 *LEPR* Functional Study

There is an increase in the circulating levels of leptin in obese individuals, termed hyperleptinaemia (Considine *et al.,* 1996, Enriori *et al.,* 2007, Leyva *et al.,* 1998), in which these high levels for a prolonged period of time leads to leptin resistance. The exact mechanism for this remains unknown although several theories have been proposed. These include reduced expression of the leptin receptor, increased expression of non-signalling receptor isoforms, reduced ligand binding or signalling of the long receptor isoform or differences in receptor trafficking. The aim of this study was to investigate the eight possible haplotypes within the *LEPR*, involving Lys109Arg, Gln223Arg and Lys656Asn to determine if there is a change in binding avidity between leptin with particular haplotypes in the *LEPR,* as a reduction in binding avidity could result in a modulation decrease in intracellular signalling as a result. This study involved PCR amplification of the extracellular region of *LEPR* from Michigan Cancer Foundation 7 (MCF-7) cells, sub-cloning the *LEPR* PCR product into an expression vector where the eight naturally occurring haplotypes could be generated via site-directed mutagenesis, and then transfected into a mammalian cell line for analysis of the binding avidity with the ligand leptin using a Quartz crystal microbalance (Figure 3.1.1).

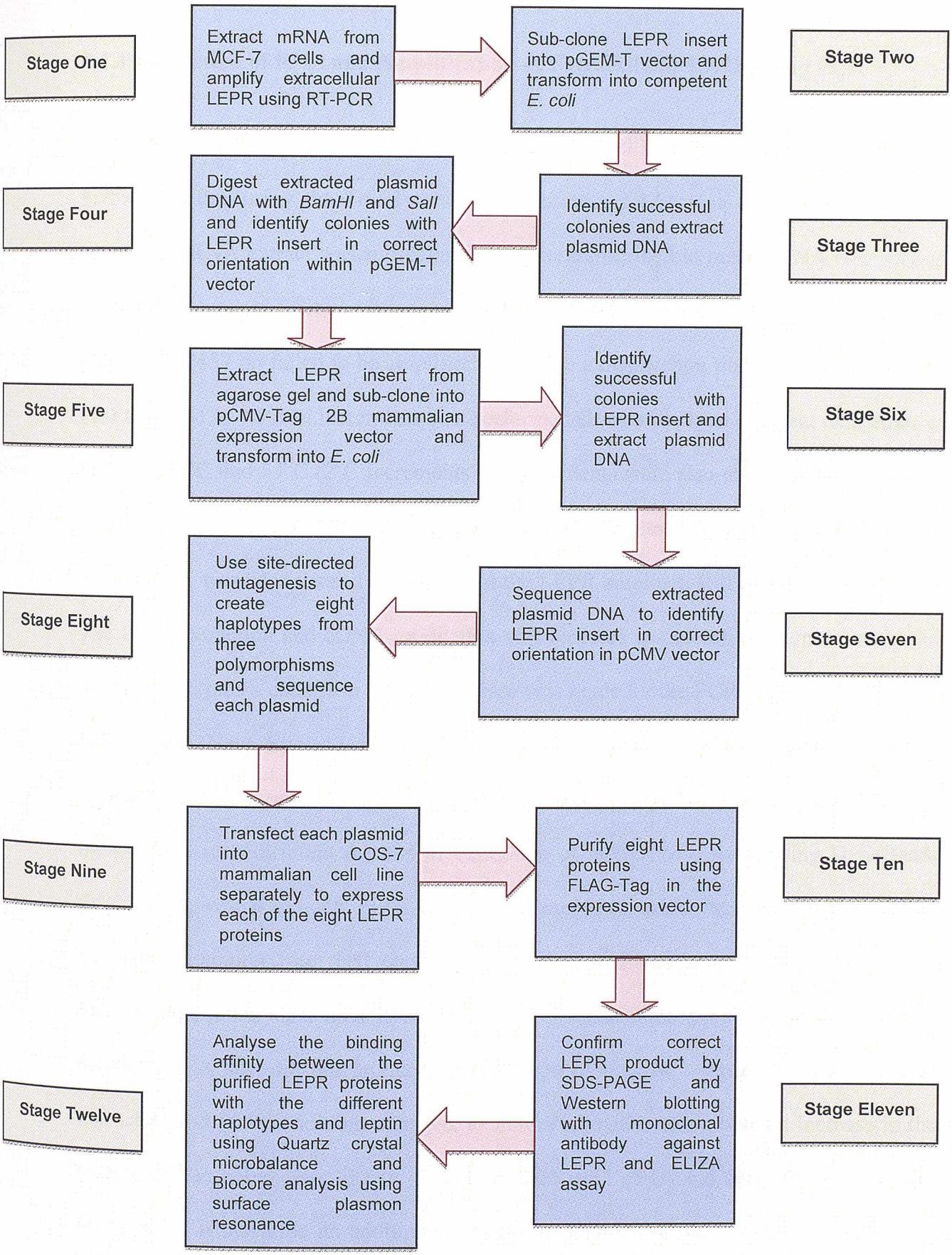


Figure 3.1.1: Flow chart of the *LEPR* functional study displaying the twelve stages to the investigation.

3.1.1 Extraction of RNA and Amplification of Extracellular LEPR using PCR

Total RNA was successfully extracted from MCF-7 cells, (from a pellet that contained $7.75x10^6$ cells), using the Trizol method (see section 2.4.6), to a concentration of 5.16 µg/µl. The RNA extract was subsequently reverse transcribed to make cDNA (see section 2.4.8). The standard PCR conditions were used (see section 2.4.11) but the expected LEPR product of 2457 bp failed to be amplified, therefore a magnesium titration, ranging from 1 mM to 4 mM in 0.5 mM increments was performed along with three different temperatures between 50°C and 54°C (2°C increments). Amendments were also made to the PCR cycle conditions; the extension step was increased to three minutes from thirty seconds and 100 ng of cDNA was used instead of 50 ng. But the LEPR amplicon still failed tobe amplified. The RNA extraction and cDNA production was repeated in case of any contamination or human error, plus a change in Taq polymerase was made for the PCR step, from Taq DNA polymerase (New England Biolabs) to AmpliTaq Gold (Applied Biosystems) using standard PCR conditions provided by the supplier (see section 2.4.11.) This did not resolve the problem so a different method of extraction of mRNA was used, using DynaBeads (Invitrogen), which extracts mRNA, by using magnetic beads that bind to the mRNA allowing separation from cell contents and buffer solutions using a magnetic particle concentrator, rather than total RNA from MCF-7 cells. Subsequent use of this mRNA-derived cDNA still did not produce the amplicon of interest, even after protocols were rechecked and PCR conditions changed, as described earlier along with an increase in the amount of Taq polymerase to 1 µl (5 U) from 0.25 µl (1.25 U) and thirty-five cycles from thirty. After which time an additional DNA Taq polymerase was used, from AmpliTaq Gold to KAPA 2G Fast HotStart (Kapa Biosystems) in the PCR. This was specifically designed to amplify large amplicons, up to 5 kb at a much faster rate, thus reducing total reaction time compared to wild-type Taq polymerases. Mutations may be created as a

result of using an enzyme without proof reading capability, which has an error rate for base incorporation of approximately 1 every 9000 nuceleotides (Tindall and Kunkel, 1988)(Tindall and Kunkel, 1998). Any such errors would be identified at stage seven of the investigation, when the correct colonies containing the pCMV vector with the LEPR insert would be sequenced to identify the insert in the correct orientation in the pCMV vector. Standard PCR and cycle conditions were used as provided by the manufacturer (see section 2.4.11.). The expected product was amplified at 56°C and at 2.0 mM magnesium chloride however the band was rather faint and the reaction was non-specific therefore the reaction conditions were changed. A magnesium titration (1.5 mM to 3.5 mM in 0.5 mM increments) was subsequently performed at two annealing temperatures, 57°C and 58°C and the number of cycles was increased to forty. The LEPR amplicon was not successfully amplified and the reaction was still non-specific therefore the PCR conditions were repeated with a magnesium titration (1.5 mM to 3.5 mM in 0.5 mM increments) at 56°C, 57°C and 58°C but the number of cycles were decreased back down to thirty-five and the volume of Taq in the reaction was increased from 0.1 µl (0.5 U) to 0.2 µl (1 U) per 25 µl reaction. The expected product was amplified under several conditions: 3 mM and 3.5 mM magnesium at 56°C and 2.5 mM at 58°C (Figure 3.1.2). However these reactions were still non-specific and produced a smear effect when the results were electrophoresed using a 1% (w/v) agarose gel (see section 2.4.12) therefore the band at the predicted size was extracted from the gel by cutting it out of the agarose and extracting it as below. This enriched for the specific band and enabled it to be ligated into the pGEM-T vector (Figure 3.1.5). The control PCR using G6PD specific primers, was performed alongside the LEPR reaction. It was also amplified with the expected product of 395 bp (Figure 3.1.2).

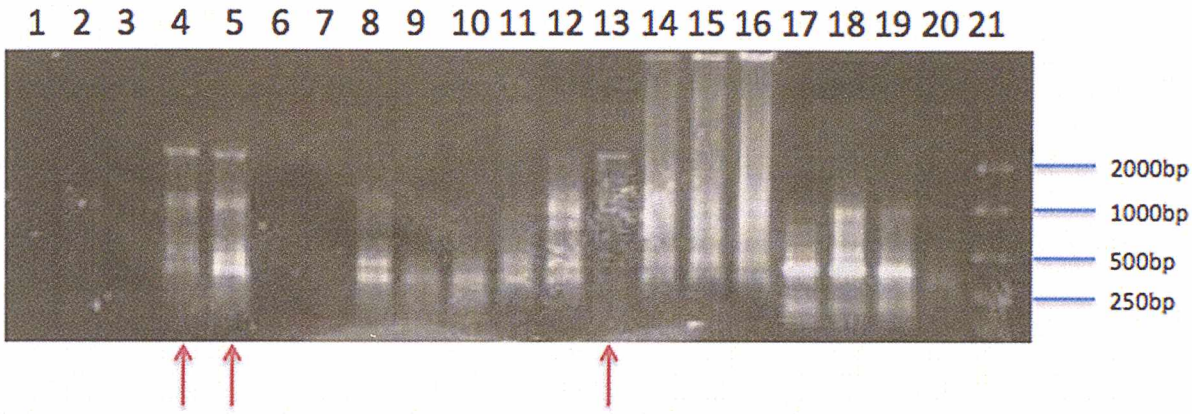


Figure 3.1.2: A Magnesium and temperature PCR optimisation of LEPR. DNA ladder (Easy Ladder) is shown in the last well (21), magnesium titration is performed (1.5 mM, 2 mM, 2.5 mM, 3 mM and 3.5 mM) at three annealing temperatures, 56°C (Lanes 1 to 5), 57°C (Lanes 7 to 11) and 58°C (Lanes 12 to 16). G6PD PCR is performed in triplicate (Lanes 17 to 20). Lanes 4, 5 and 13 (as indicated by the arrows) contain a visible band of the right size of the LEPR product. 10 µl of each PCR was loaded into the corresponding lanes.

For the agarose gel extraction process (see section 2.4.13) the PCR was performed at the optimised conditions of 2 mM magnesium chloride and the annealing temperature of 58°C, to increase the yield of LEPR so that a high quantity could be ligated in to the pGEM-T vector. However, after several attempts at this, during electrophoresis, the PCR product was not separating particularly well with such a large quantity (Figure 3.1.3) therefore only 25 µl was loaded instead into each well (Figure 3.1.4). Each time the LEPR band was extracted from the agarose gel it was weighed to determine the correct volume of extraction buffer to add. The weight of each of the gel extracts was low each time the LEPR band was extracted but they varied between 0.044 and 0.074 g.

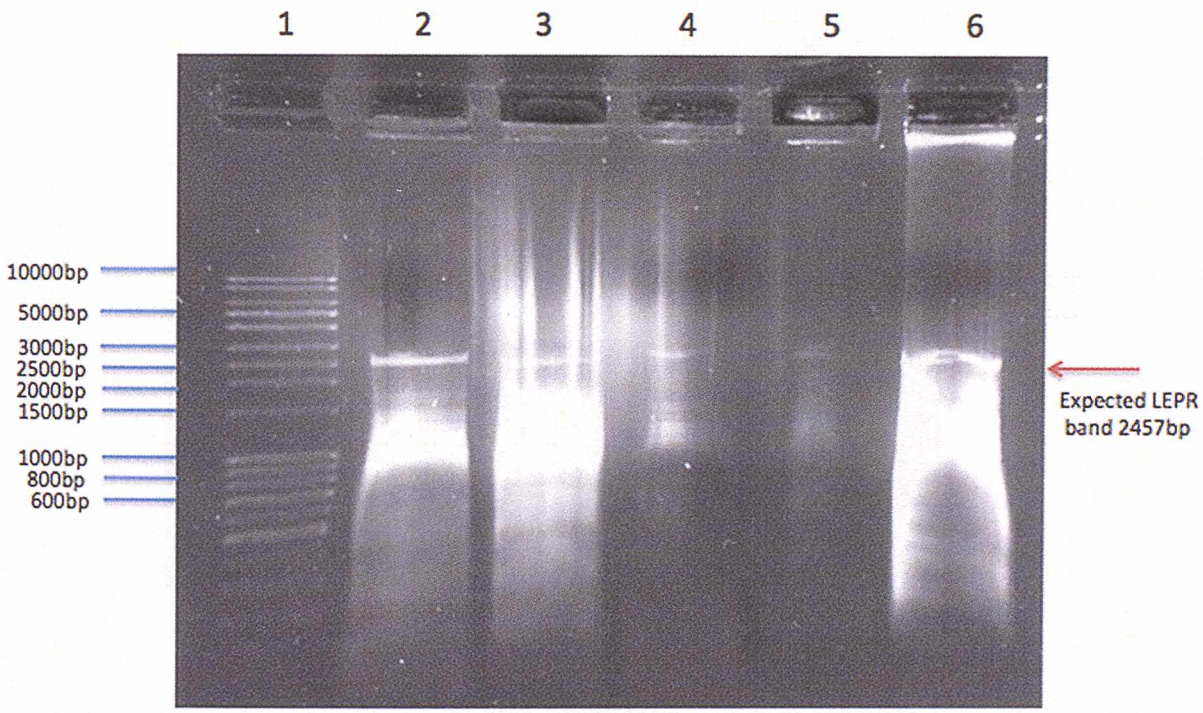


Figure 3.1.3: 1% (w/v) agarose gel electrophoresis of leptin receptor PCR performed at the optimised conditions. Twenty reactions of the PCR were performed and loaded into the five wells, which could hold 100 µl, Lanes 2 to 6. Lane One contains DNA ladder (GeneRuler).

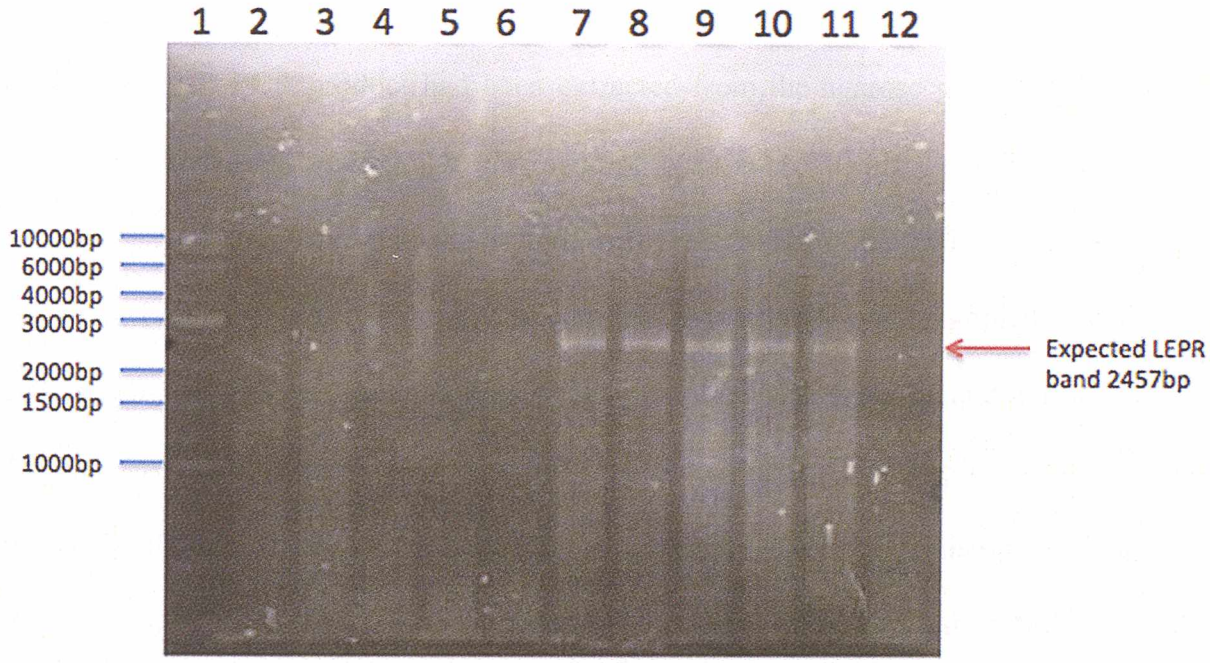


Figure 3.1.4: 1% (w/v) agarose gel electrophoresis of leptin receptor PCR at the optimised conditions. Lane one contains DNA Ladder (GeneRuler). Lanes 2 to 12 contain 20 µl of each of the ten PCRs that was performed, with lanes 7 to 11 containing a band of the expected size of LEPR.

3.1.2 Ligation of pGEM-T vector and LEPR insert and Transformation of *Escherichia coli* (JM109)

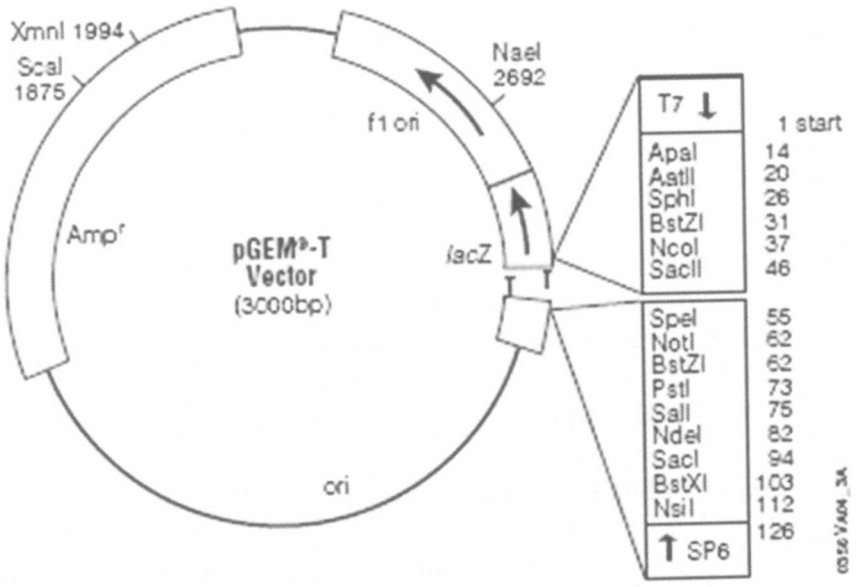
Once the LEPR PCR product had been extracted from the agarose gel the concentration was determined (see section 2.4.22) before it was subsequently ligated into pGEM-T vector (see section 2.4.14). The diagrammatic representation of the pGEM-T vector map is shown in Figure 3.1.5a, which contains an origin of replication an ampicillin resistance gene and the *lacZ* gene. T overhangs are located within the *lacZ* gene and during PCR the Taq polymerase adds an additional A base on the 3′ end of the amplicon, thus the two products will complementary bind together. Within the *lacZ* there are also multiple enzyme restriction sites (Figure 3.1.5b) to enable the PCR insert to be cleaved from the vector. The LEPR insert was initially ligated with the pGEM-T vector in a 1:3 ratio (160 ng of LEPR insert) and transformed into competent *E.coli* (see section 2.4.15). Each time LEPR was extracted from the agarose gel the DNA concentration was determined to be between 162.5 ng/µl and 275 ng/µl. The pGEM-T vector is exactly 3 kb in size and the LEPR insert is 2457 bp therefore once the two products are ligated together it produced a recombinant plasmid with a size of 5457 bp.

Detection of the successful colonies was indicated by the white selection by using *lacZ* gene, (see section 2.4.15). Colonies that had the LEPR insert incorporated within the vector produced a white colour indicating that the *lacZ* gene had been disrupted. Those that had a blue colour did not have the LEPR insert, as the *lacZ* gene had not been disrupted.

Initial ligations between LEPR PCR product and the pGEM-T vector, were followed by transformations into competent JM109 *E.coli* cells which resulted in only blue colonies being produced. The preparation of the LB agar plates was checked so that the correct

solutions of ampicillin, X-Gal and IPTG were added to produce a final concentration of 50 μg/ml, 100 μg/ml and 100 μg/ml respectively (see section 2.4.20). The three ligations (experimental, positive and negative controls) were left overnight at room temperature rather than being placed in the fridge to help the ligation process as given by the standard protocol (see section 2.4.14). A second set of each of the three ligations was performed so that a small amount could be electrophoresed to determine if the ligation was a success before transformation into *E.coli*.

A)



B)

2.A. Multiple Cloning Sequences

pGEM®-T Vector

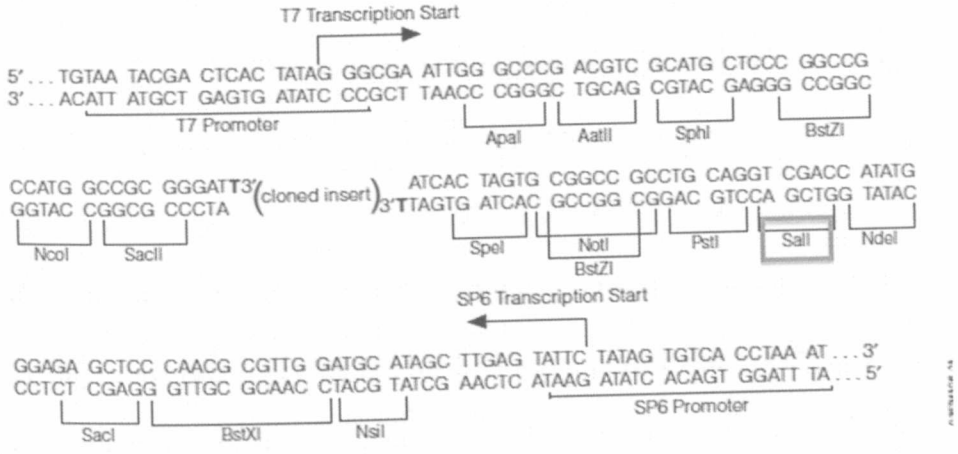


Figure 3.1.5: Diagrammatic representation of the pGEM-T vector, into which the LEPR amplicon was inserted. A) The T overhangs are located within the *lacZ* gene of the vector so that the inserted DNA of interest can disrupt the gene. B) Multiple restriction enzyme sites are also situated next to the two T overhangs so that the inserted DNA can be easily removed. The restriction enzyme *SalI* is highlighted. Accessed from pGEM-T vector system technical manual (Promega).

These results showed that the concentration of the extracted product was too low, as the PCR product was faint, despite the concentration of the extracted PCR product being determined before ligation was performed. The quantification of the extracted LEPR PCR product showed that for a 1:3 ratio, 3 µl of extracted LEPR contained 375 ng, which was more than the 165 ng calculated. One method to improve the concentration of LEPR extract in the ligation reaction was to increase the ratio of insert to vector from 1:3, to 1:4, 1:6 and 1:8 but not so that the total ligation reaction was over 10µl. This still did not seem to improve the number of white colonies with the correct insert therefore when several LEPR bands were extracted from the agarose gel, after successful PCR amplification, two gel extracts were placed into one Eppendorf tube and extracted together. In addition, all the extracts that were carried out were pooled together at the last step. Quantification of these extracted PCR products showed an increase in concentration (675 ng/µl) and this subsequently resulted in successful ligations between the LEPR insert and pGEM-T vector (Figure 3.1.6). Adjustments were also made to the transformation protocol; instead of the recommended 2 µl, 4 µl of the ligation reaction was added to 50 µl of competent *E.coli* cells to increase the number of possible successful transformations that would contain the pGEM-T vector with LEPR insert. All of the SOC media containing the transformants (1 ml) was spread out on to five agar plates to increase the chance of white colonies growing. Many ligations and transformations had to be performed but the number of white colonies was high, ranging from eleven to thirty colonies compared to over fifty blue colonies. Once the transformants were incubated overnight on LB agar plates they were placed in the fridge so that the blue colour could develop to avoid error.

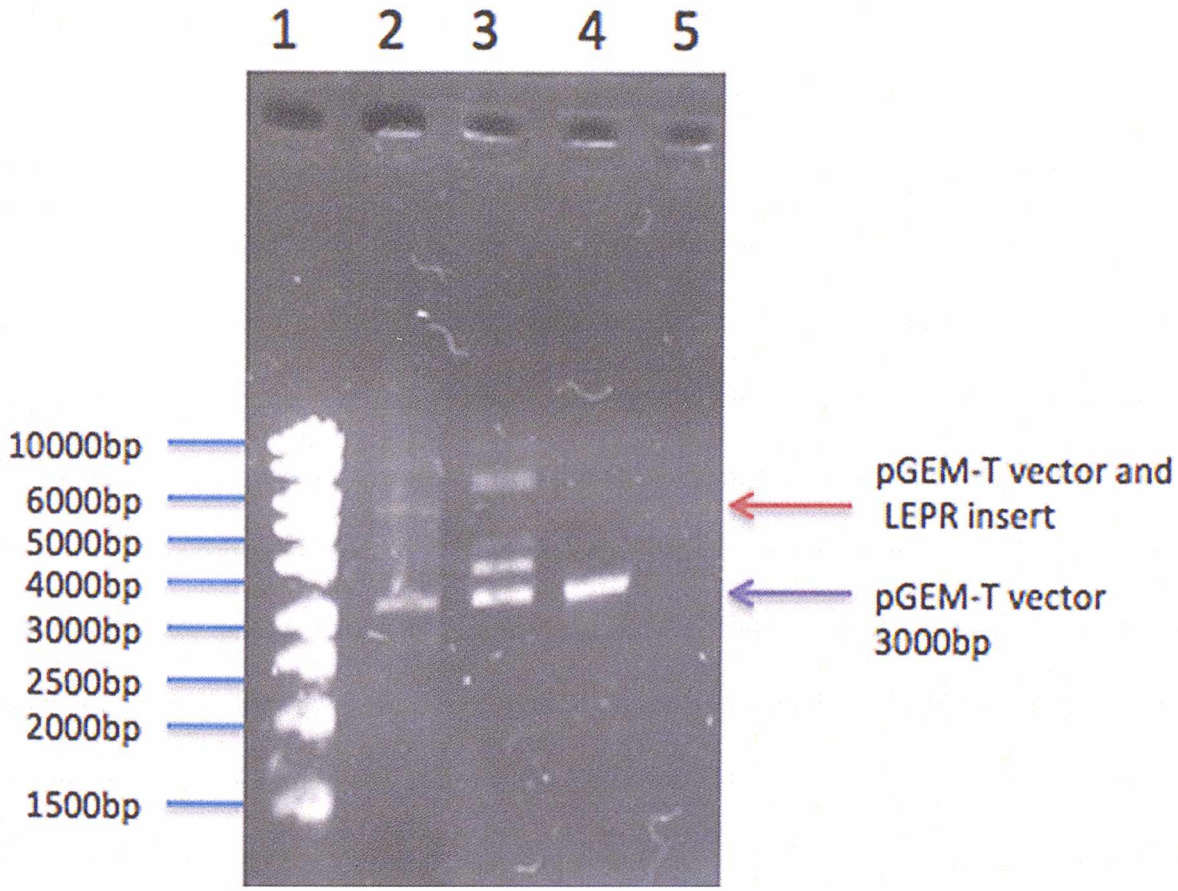


Figure 3.1.6: 1% (w/v) agarose gel electrophoresis of ligation between LEPR and pGEM-T vector. Lane 1: DNA ladder (HyperLadder I), Lane 2; ligation with LEPR gel extract, Lane 3: ligation with positive control, Lane 4: ligation with negative control. 4 µl of each ligation was loaded into the corresponding wells.

3.1.3 Double Restriction Digests of Mini Preparations

The white colonies that were identified on the LB agar plates after overnight incubation at 37°C were selected and grown as overnight cultures (see section 2.4.16) for mini preparation (see section 2.4.16.1) and subsequently set up as overnight double digests with the restriction enzymes *BamHI* and *SalI* (see section 2.5.16). The restriction site for *BamHI* (5′-G˅GATCC-3′) was incorporated into the forward primer of the LEPR (Table 2.4.2), as this enzyme wasnot within the pGEM-T vector but the *SalI* restriction site (5′-G˅TCGAC-3′) was (Figure 3.1.5). This was designed for ease of cloning so that the LEPR insert could be subsequently cleaved from the pGEM-T vector and sub-cloned into the pCMV-Tag 2B mammalian expression vector (Figure 3.1.20).

Digested mini preparations were analysed via agarose gel electrophoresis (see section 2.4.12) to identify cultures in which the LEPR insert had been ligated into the pGEM-T vector. Uncut mini preparations were electrophoresed next to the corresponding cut mini preparations to detect those that had been digested. Mini preparations that were found to have two bands upon visualisation of the double digest, corresponding to 2457 bp and 3000 bp in size indicated that the LEPR was successfully ligated into the vector, in the correct orientation. The *BamHI* enzyme restriction site was incorporated into the 5′ end of the forward LEPR primer, thus upon digestion using *BamHI* and *SalI* the LEPR insert would be digested out of the vector and two bands would be visible. If the LEPR insert is in the wrong orientation, the *BamHI* restriction site would be located within 30 bp of the *SalI* restriction site therefore upon electrophoresis this would be resolved at 5427 bp which is similar in size to the insert and vector ligated together (Figure 3.1.7; Figure 3.1.5B). Once the conditions for ligation and transformation had been optimised, twenty-eight

overnight cultures of white colonies, mini preparations and restriction digests were set up to try and identify the correct culture, that contained the LEPR insert in the correct orientation.

Before the restriction digests were set up, the concentration of the plasmid DNA extracted from the overnight cultures was determined (see section 2.4.22), so that there was ~500-1000 ng in the reaction. The concentrations of the plasmid DNA from the mini preparations ranged from 637 ng to 1120 ng, therefore 1 µl was used in the digests. After these twenty-eight experiments were performed, four cultures were identified that produced two bands upon digestion with the two enzymes *BamHI* and *SalI* at 3 kb and 2.5 kb (Figure 3.1.8). An additional band was seen at 2kb therefore this band and the one at 2.5 kb were extracted. However, after careful analysis this appeared to be due to human error (Figure 3.1.9). Many of the other cultures did not seem to be producing the desired results, producing one band at 3000 bp indicating the presence of the pGEM-T vector without the insert, or a band at 5.5 kb indicating the LEPR was in the wrong orientation (Figure 3.1.10). This was confirmed when selected cultures were digested with each enzyme individually singularly (Figure 3.1.11).

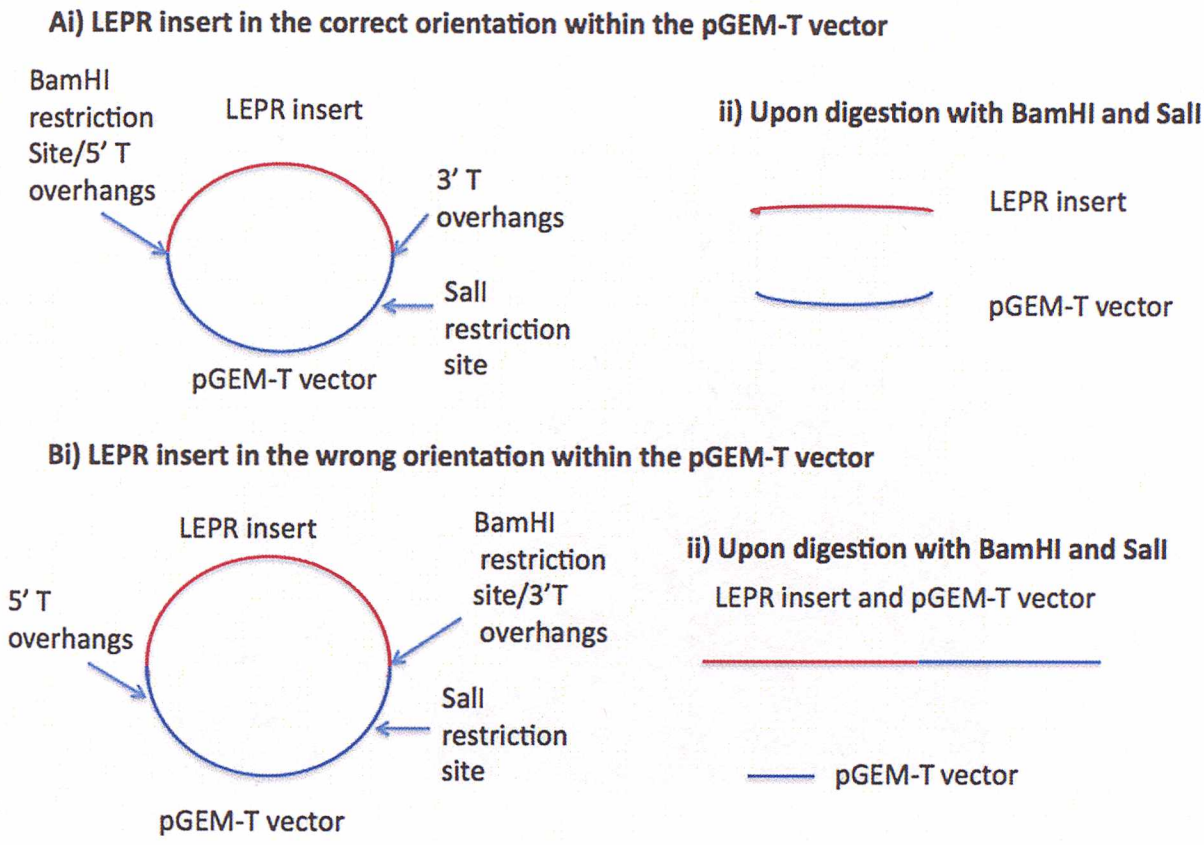


Figure 3.1.7: Diagrammatic representation of the ligation of the LEPR insert into the pGEM-T vector. The orientation of the LEPR is important for subsequent ligation of the LEPR into the pCMV Tag 2B expression vector. Ai) The correct orientation of the LEPR insert into the vector and Aii) the cleavage of the LEPR insert from the vector upon digestion with *BamHI* and *SalI* restriction enzymes. Bi) The wrong orientation of the LEPR insert in the pGEM-T vector and Bii) the failure of the LEPR to be cleaved from the pGEM-T vector upon restriction digestion

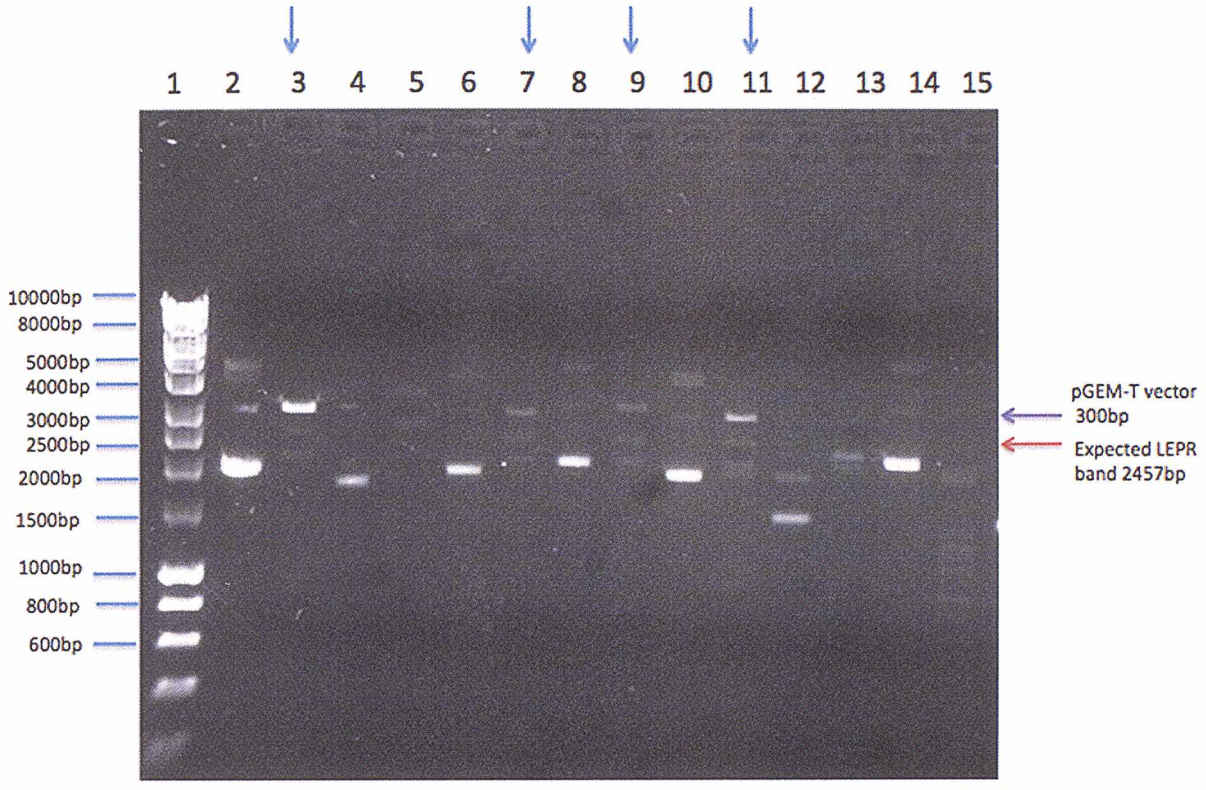


Figure 3.1.8: 1% (w/v) agarose gel electrophoresis of double restriction digest with the enzymes *BamHI* and *SalI*, of seven mini preparations, with undigested mini preparations alongside. Lanes 3, 7, 9 and 11 show bands at both 3 kb and 2.5 kb indicating the LEPR insert has been ligated out of pGEM-T vector. Lane 1: DNA ladder (Hyperladder I), Lane 2: undigested mini preparation 1, Lane 3: digested mini preparation 1, Lane 4: undigested mini preparation 2, Lane 5: digested mini preparation 2, Lane 6: undigested mini preparation 3, Lane 7: digested mini preparation 3, Lane 8: undigested mini preparation 4, Lane 9: digested mini preparation 4, Lane 10: undigested mini preparation 5, Lane 11: digested mini preparation 5, Lane 12: digested mini preparation 6, Lane 13: undigested mini preparation 6, Lane 14: undigested mini preparation 7, Lane 15: digested mini preparation 7. 5 µl of undigested mini preparation and 10 µl of digested mini preparation were loaded into the corresponding lanes.

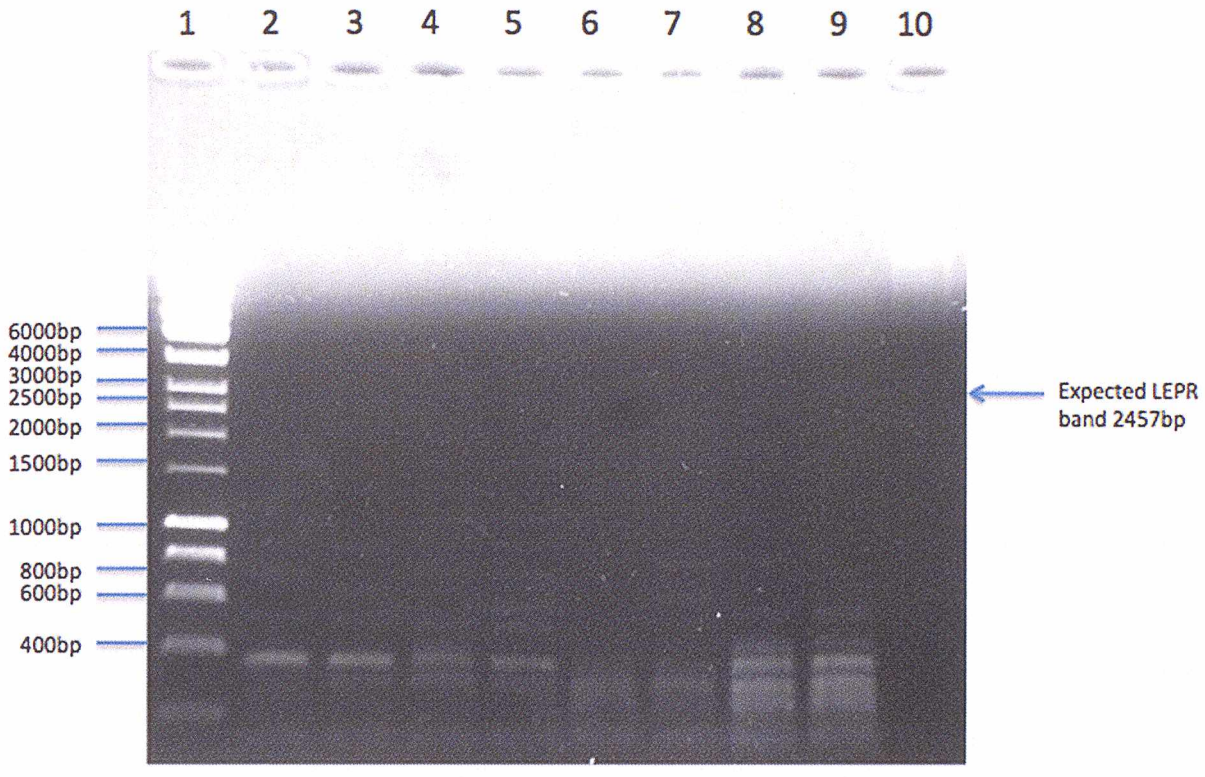


Figure 3.1.9: 1% (w/v) agarose gel electrophoresis of LEPR PCR at the optimised conditions using gel extracts from the four mini preparations (one, three, four and five) that produced two bands upon double digestion at 3 kb and 2.5 kb. Lane 1: DNA ladder (Hyperladder I), Lane 2: mini preparation one 2 kb gel extract, Lane 3: mini preparation one 2.5 kb gel extract, Lane 4: mini preparation three 2 kb gel extract, Lane 5: mini preparation three 2.5 kb gel extract, Lane 6: mini preparation four 2 kb gel extract, Lane 7: mini preparation four 2.5 kb gel extract, Lane 8: mini preparation five 2 kb gel extract, Lane 9: mini preparation five 2.5 kb gel extract, Lane 10: PCR water blank. 10 µl of each PCR was loaded into the corresponding lanes.

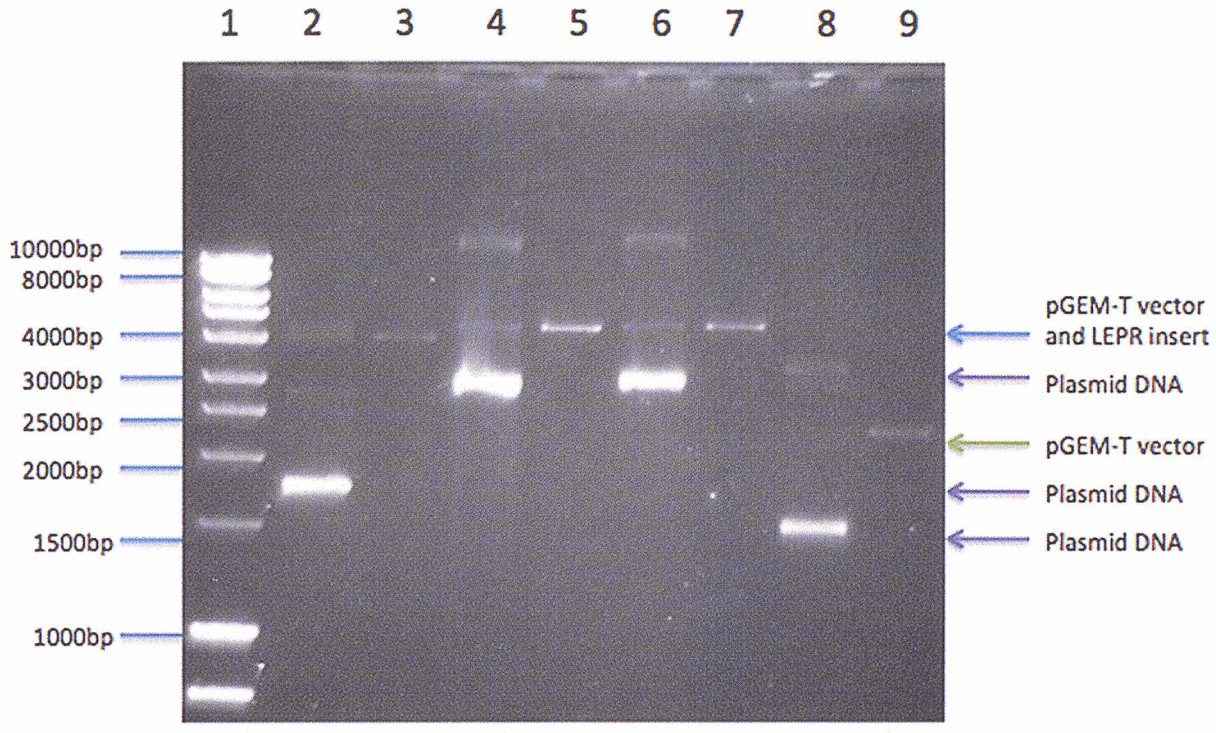


Figure 3.1.10: 1% (w/v) agarose gel electrophoresis of double restriction digests of mini preparations with *BamHI* and *SalI*. Digested mini preparations were loaded alongside the corresponding undigested mini preparation. Lane 1: DNA ladder (GeneRuler), Lane 2: undigested mini preparation 16, Lane 3: digested mini preparation 16, Lane 4: undigested mini preparation 17, Lane 5: digested mini preparation 17, Lane 6: undigested mini preparation 18, Lane 7: digested mini preparation 18, Lane 8: undigested mini preparation 19, Lane 9: digested mini preparation 19. The main band from all the mini preparations seem to be between 1.5 kb and 3.0 kb, while the size of the double digested products vary. None of which produce the two expected bands of 2.5 kb and 3 kb. Lanes 3 and 9 show mini preparations that contain empty vector while Lane 5 shows mini preparation, which contains LEPR insert in the wrong orientation. 5 µl of undigested mini preparations and 10 µl of digested mini preparations were loaded into the corresponding lanes.

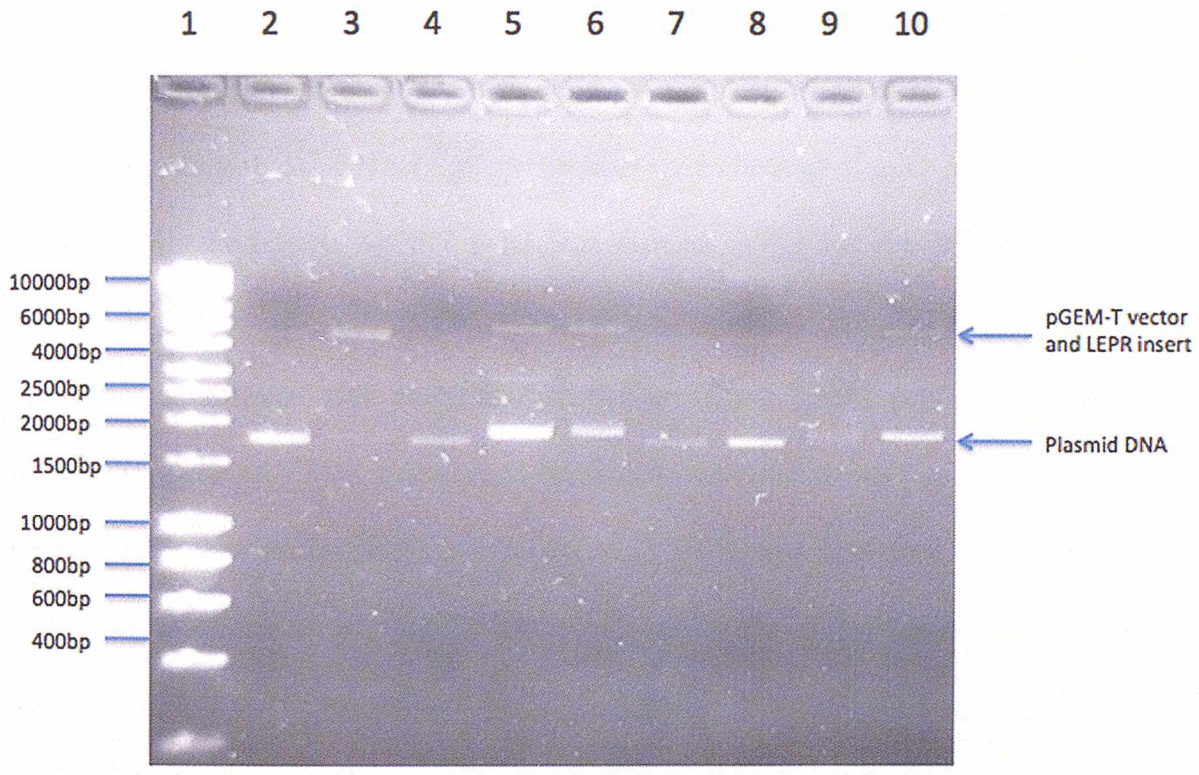


Figure 3.1.11: 1% (w/v) agarose gel electrophoresis of single restriction digests with the restriction enzymes *BamHI* and *SalI* for three mini preparations. Digested mini preparations are loaded next to the single digests. Lane 1: DNA ladder (HyperLadder I), Lane 2: undigested mini preparation 2, Lane 3: digested mini preparation 2 with *SalI*, Lane 4: digested mini preparation 2 with *BamHI*, Lane 5: undigested mini preparation 12, Lane 6: digested mini preparation 12 with *SalI*. Lane 7: digested mini preparation 15 with *BamHI*, Lane 8: undigested mini preparation 15, Lane 9: digested mini preparation with *SalI*, Lane 10: digested mini preparation 12 with *BamHI*. 5 µl of undigested mini preparations and 10 µl of digested mini preparations were loaded into the corresponding lanes.

Due to this inability to identify the correct restriction enzyme fragments from the mini-prepped recombinant plasmids, the experiment was started all over again. This included using placental poly A^+ RNA (Agilent Technologies) instead of mRNA extracted using the Dynabead method from MCF-7 cells (see section 2.4.7) for the RT-PCR (see section 2.4.11) and PCR amplifying the extracellular region of the LEPR at the previously optimised conditions. Gel extraction of the amplified LEPR product and subsequent ligation into the pGEM-T vector and transformation into competent JM109 *E.coli* cells produced fifty-six white colonies. The plasmid DNA was extracted from these colonies and set up as double restriction digests, but the restriction digest assay was changed from the standard protocol (see section 2.4.17) to try and maximise the digestion reaction. The total volume was changed from 20 µl to 50 µl, with the amount of distilled water increased up to this volume. The volumes of the digestion buffer and plasmid DNA were increased by two and a half times but the volume of the enzymes were kept the same. Two cultures were identified that produced two bands upon digestion with *BamHI* and *SalI* of which two revealed to have the LEPR insert in the correct orientation (Figure 3.1.12). Further identical double restriction digests were carried out for the two mini preparations to extract the LEPR insert so that there would be a high concentration of the insert for subsequent ligation into the pCMV-Tag 2B mammalian expression vector. When all of the digested products (250 µl) for mini preparation number nineteen were loaded into one large well across the width of the mini agarose gel by placing masking tape over the 'teeth' of the comb, the LEPR band also appeared much fainter than the pGEM-T vector (Figure 3.1.13), therefore a maxi preparation was carried out to increase the yield of the culture and plasmid DNA extracted, however this did not seem to increase the quantity of the LEPR insert after restriction digestion (Figure 3.1.14). As there were problems with the restriction digestion and low yield, a PCR was performed using the extracted product to

see if the LEPR would be re-amplified. A band was produced but it was smaller than the 2.5 kb band from the marker suggesting it might not be the LEPR insert (Figure 3.1.15).

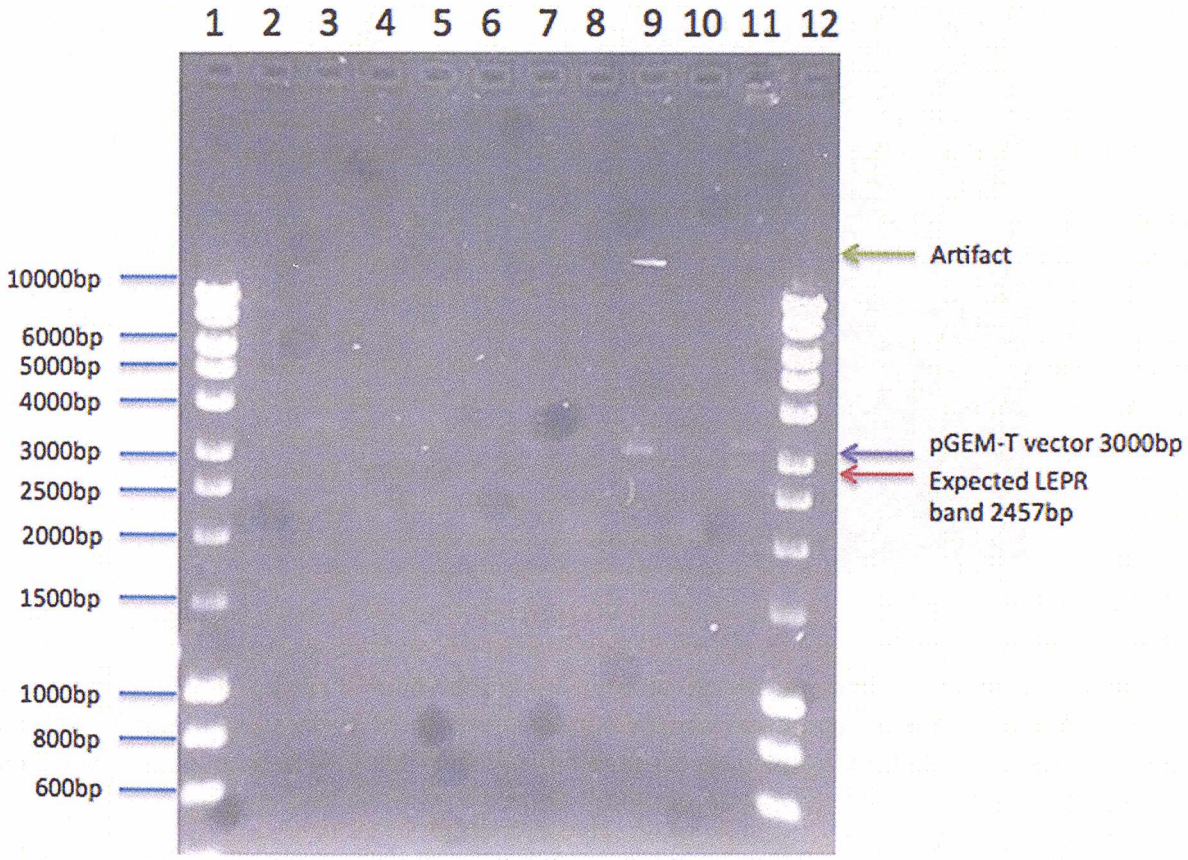


Figure 3.1.12: 1% (w/v) agarose gel electrophoresis of five undigested mini preparations and the corresponding double restriction digests of these five mini preparations. Faint bands at 3 kb and 2.5 kb can be seen in lanes 9 and 11 for mini preparations 19 and 20 respectively. Lanes 1 and 12: DNA ladder (HyperLadder I), Lane 2: undigested mini preparation 9 (5 µl), Lane 3: digested mini preparation 9, Lane 4: undigested mini preparation 10, Lane 5: digested mini preparation 10, Lane 6: undigested mini preparation 16, Lane 7: digested mini preparation 16, Lane 8: undigested mini preparation 19, Lane 9: digested mini preparation 19, Lane 10: undigested mini preparation 20, Lane 11: digested mini preparation 20. 5 µl of undigested mini preparations and 10 µl of digested mini preparations were loaded into the corresponding lanes.

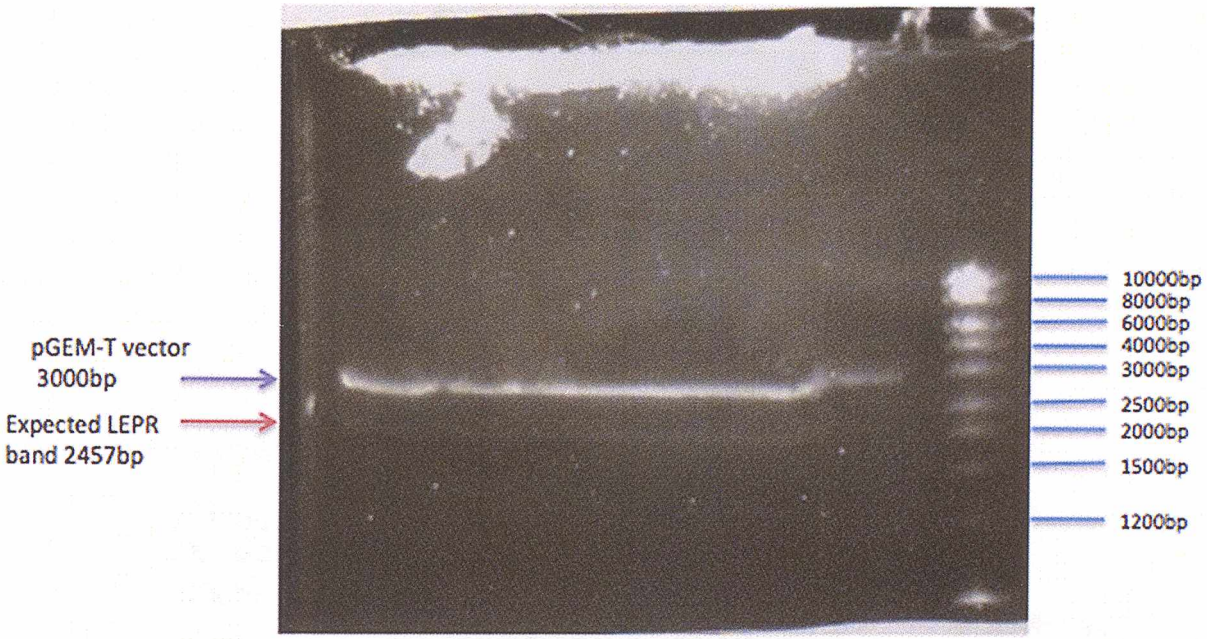


Figure 3.1.13: 1% (w/v) agarose gel of double restriction digest of mini preparation numbered nineteen. 250 µl from five separate restriction digests (50 µl total) was loaded into one well located across the width of the agarose gel, using DNA ladder (GeneRuler).

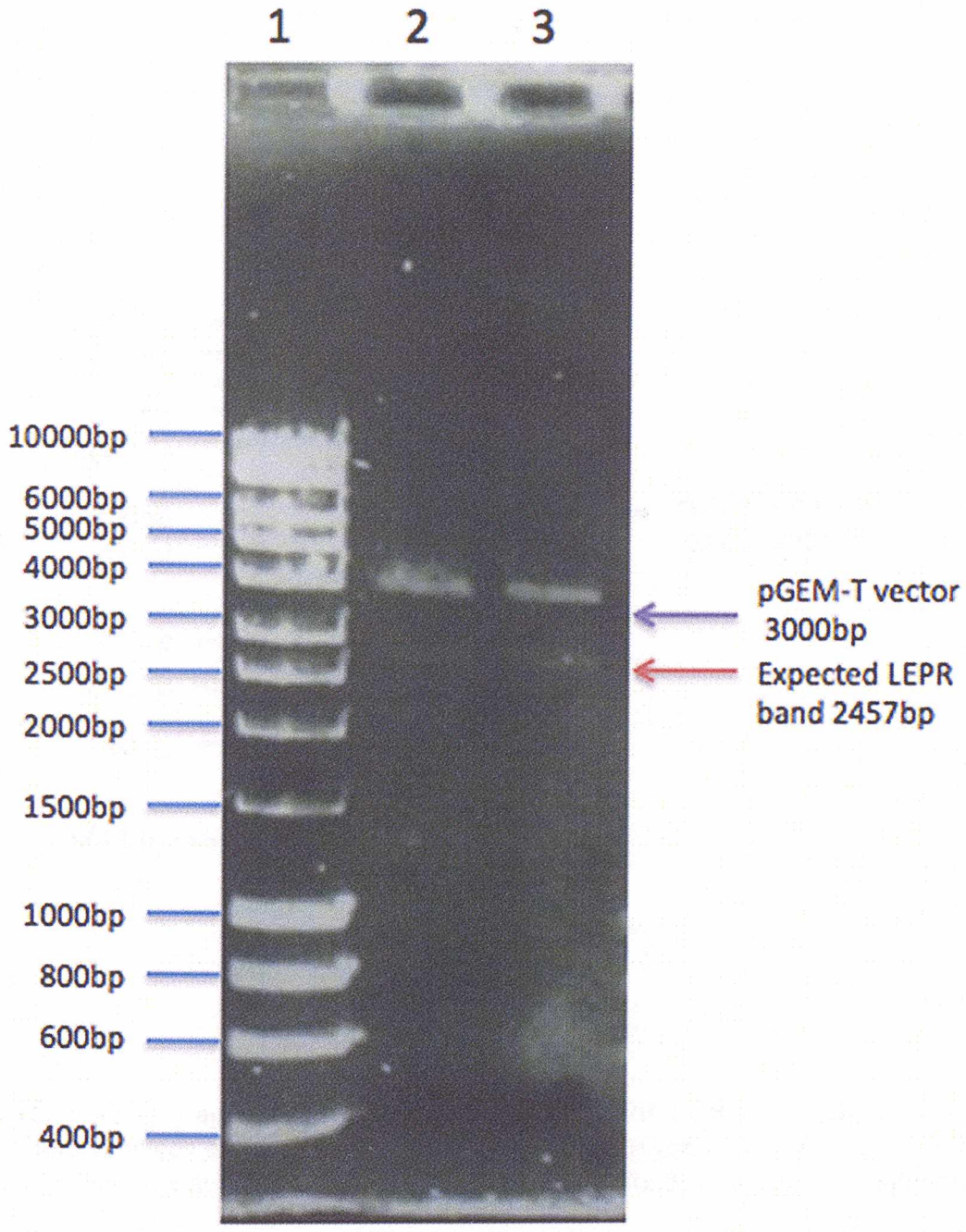


Figure 3.1.14: 1% (w/v) agarose gel of the double restriction digest of mini preparation 25 µl was loaded into each lane. Lane 1: DNA ladder (Hyperladder I)

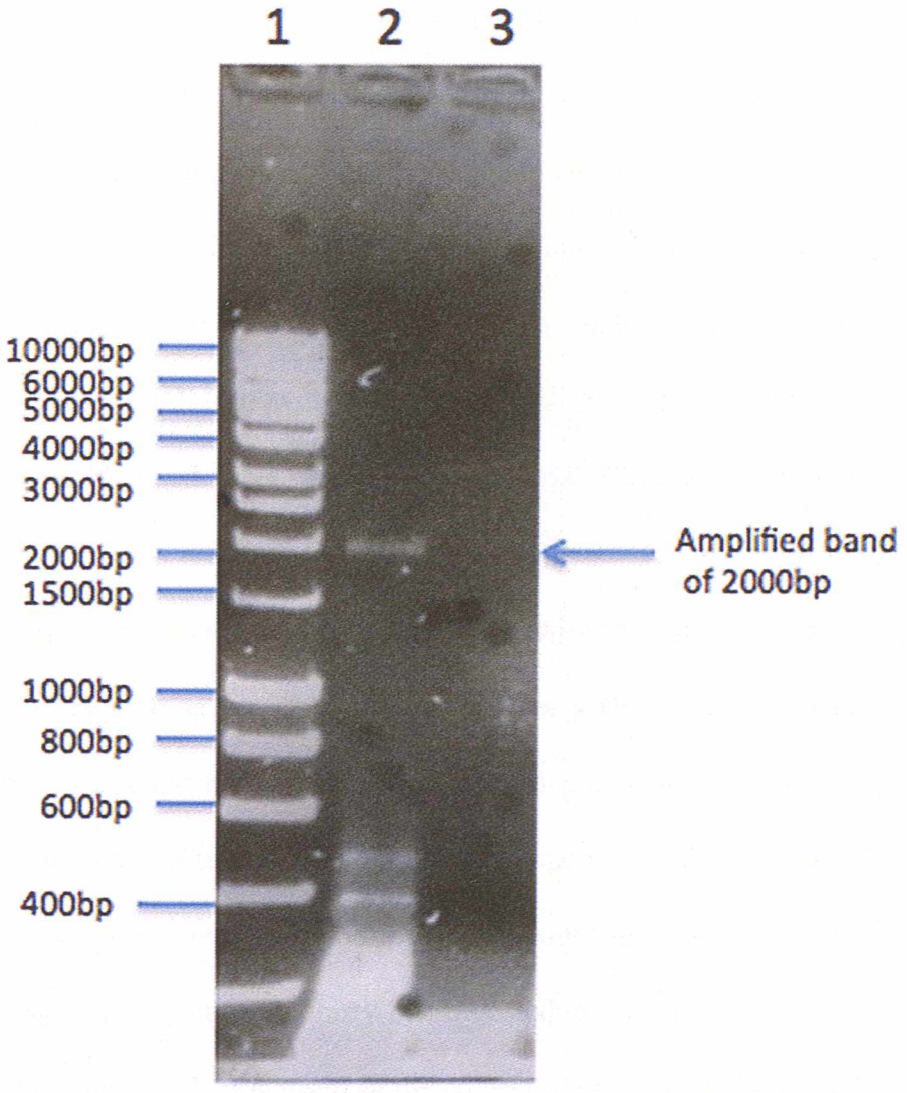


Figure 3.1.15: 1% (w/v) agarose gel electrophoresis of LEPR PCR from gel extracts of mini preparations nineteen and twenty. Lane 1: DNA ladder (HyperLadder I), Lane 2: mini preparation 19, Lane 3: water blank. 10 µl of each PCR was loaded into the corresponding lanes.

As the two cultures, from the previous experiments failed to produce the LEPR insert upon restriction digestion of the plasmid DNA, the study was started from the beginning again, at the PCR stage, for a second time. Thirty-one white colonies were observed on the LB agar plates supplemented with X-Gal, Ampicillin and IPTG following ligation with pGEM-T vector and transformation into *E. coli.* These were left to grow overnight for mini preparation and restriction digests with *BamHI* and *SalI*, the following day. The restriction digestion assay was changed for a second time to improve conditions further. The same enzymes were used from a different supplier (New England BioLabs from Fermentas), as both enzymes were compatible and had 100% activity with the same restriction buffer. The same conditions were used as the standard protocol (see section 2.5.16) but the volume of the *SalI* enzyme was reduced to 1 µl from 2 µl. In previous restriction digests using enzymes from Fermentas, a two-fold excess of *SalI* was required as it has a 50 to 100% enzymatic activity with the restriction buffer as compared to *BamHI* enzyme, which had 100% activity. One culture out of the thirty-one digested to produce two clear bands, which indicated the culture had the LEPR insert in the correct orientation (Figure 3.1.16).

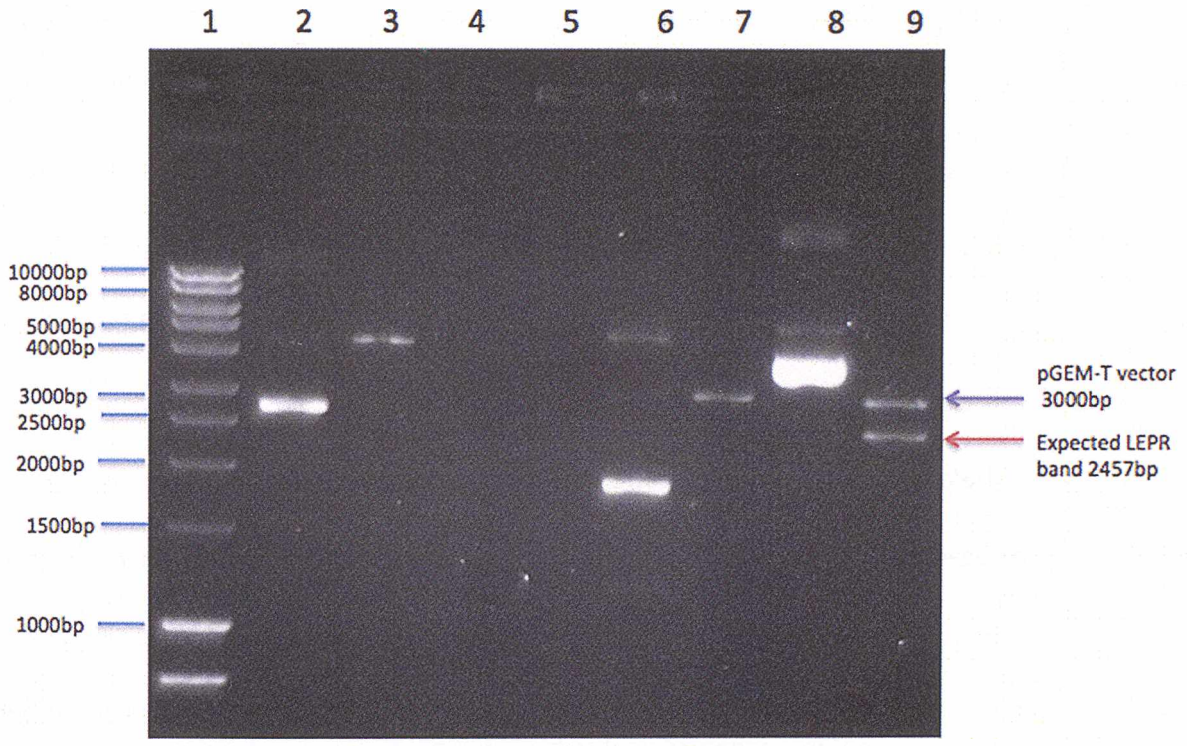


Figure 3.1.16: 1% (w/v) agarose gel electrophoresis of double restriction digest of mini preparations from four different cultures. The two separate bands for LEPR insert and pGEM-T vector can be seen in lane 9. Lane 1: DNA Ladder (GeneRuler) Lane 2: undigested mini preparation of culture 23; Lane 3: digested mini preparation of culture 23; Lane 4: undigested mini preparation of culture 24; Lane 5: digested mini preparation of culture 24; Lane 6: undigested mini preparation of culture 25; Lane 7: digested mini preparation of culture 25; Lane 8: undigested mini preparation of culture 26; Lane 9: digested mini preparation of culture 26. 5 µl of undigested mini preparations and 10 µl of digested mini preparation were loaded into the corresponding lanes.

Once the mini preparation had been located that contained the LEPR PCR product (Culture 26), five identical restriction digests were set up overnight so that the LEPR product could be extracted from the agarose gel, to increase the yield for it to be ligated into the pCMV-Tag 2B expression vector. However upon visualisation of these five restriction digests, two bands could be seen but they did not appear to have separated very well although the same conditions were used (Figure 3.1.17). The LEPR band at 2457 bp was still extracted and extracts were pooled together to increase the concentration of the extracted LEPR product. Once the LEPR product was extracted, a 4 µl aliquot was electrophoresed to ensure there was a good quantity produced (Figure 3.1.18). The culture was also set up as another overnight culture so that a maxi preparation could be carried out to further increase the yield of the plasmid DNA for further restriction digests.

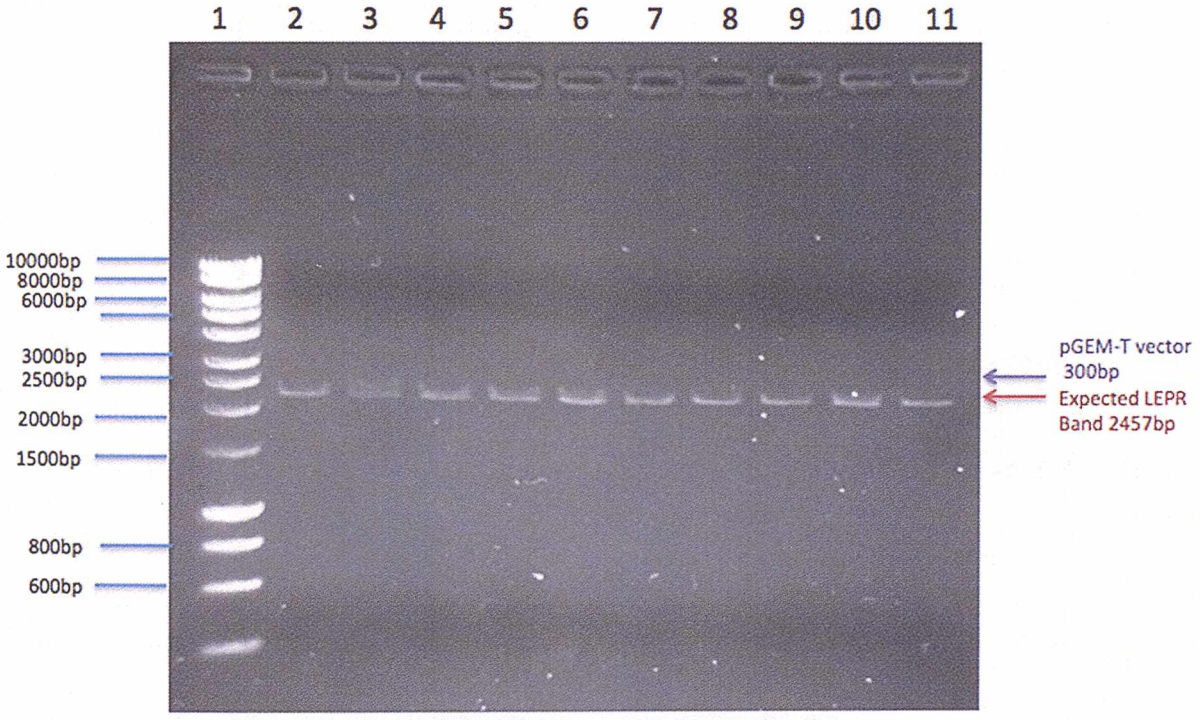


Figure 3.1.17: 1% (w/v) agarose gel electrophoresis of double restriction digest of culture 26 with the enzymes *BamHI* and *SalI*. Lane 1: DNA ladder (GeneRuler); Lanes 2-11: Culture 26 double digested mini preparation. 20 µl of digested mini preparation was loaded into each of the corresponding lanes.

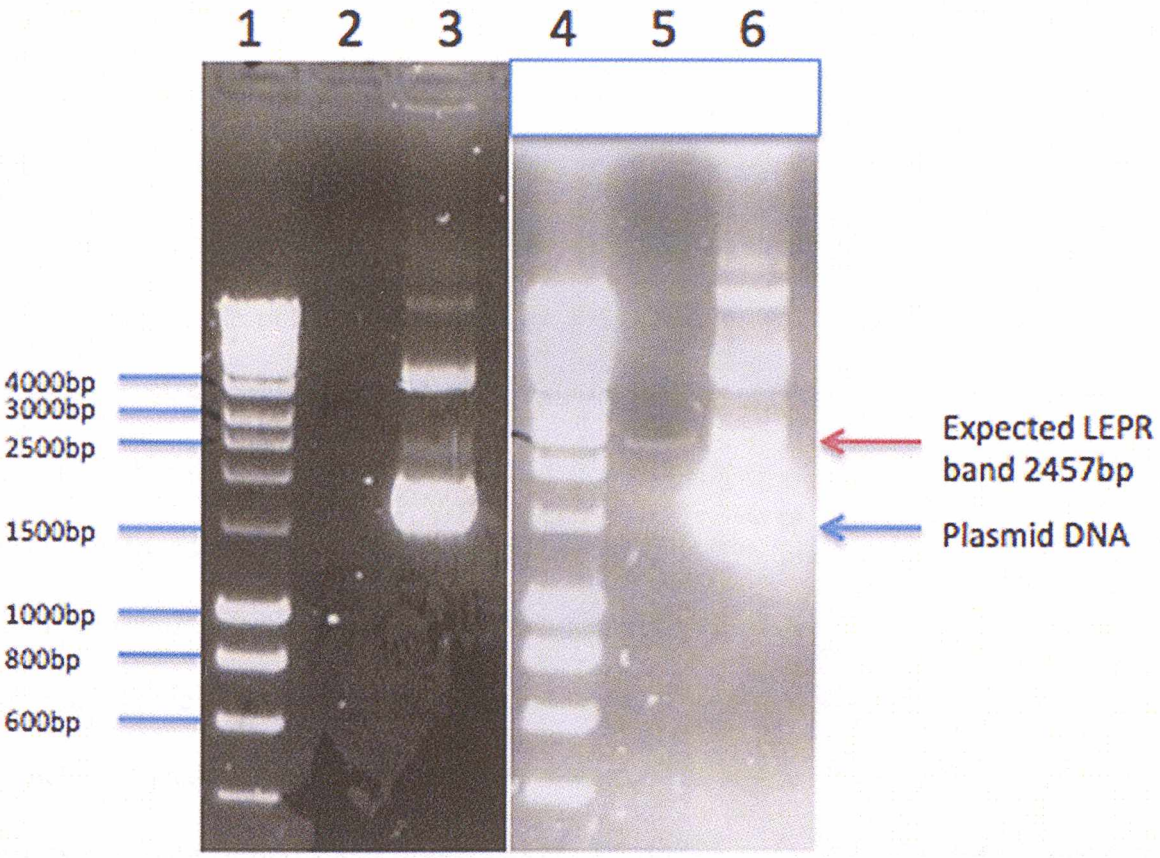


Figure 3.1.18: 1% (w/v) agarose gel electrophoresis of the LEPR gel extract from double digest of culture 26. Lane 1 and 4: DNA ladder (Hyperladder I), Lanes 2 and 5: LEPR gel extract, Lane 3 and 6: Maxi preparation of culture 26. 5 µl of both the gel extract and maxi preparation was loaded into the corresponding lanes.

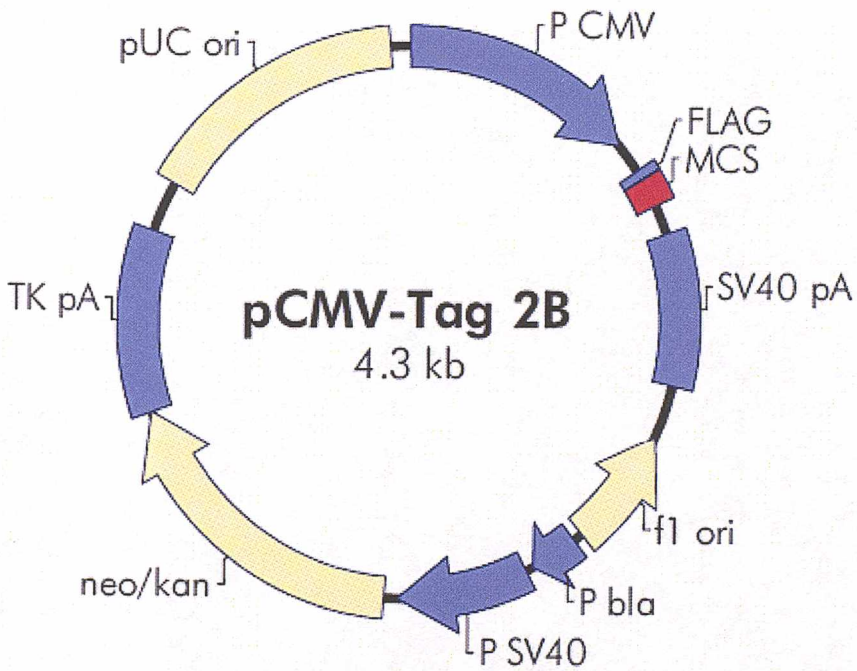
3.1.4 Ligation of pCMV–Tag 2B Expression Vector and Transformation into *E. coli*

Once the LEPR band had been extracted from the agarose gel after the double digest, it was then ligated into the pCMV-Tag 2B expression vector. The diagrammatic representation of the pCMV-Tag 2B vector map is shown in Figure 3.1.19A. This vector contains the SV40 origin of replication, which enables replication of the vector in large numbers within the COS-7 mammalian cell line due to expression of the viral SV40 large T antigen. In addition the cytomegalovirus (CMV) promoter sequence is also located within the vector, which initiates translation along with a FLAG tag coding sequence. This is a hydrophobic sequence located adjacent to the cloned insert so that the LEPR protein product produced is a fusion protein that contains the FLAG tag which enables affinity purification using monoclonal antibodies to the FLAG tag. The expression vector was digested with the restriction enzymes, *BamHI* and *SalI*, which are located within the multiple cloning site of the vector (Figure 3.1.19B) to provide the same sticky ends as the LEPR product so they could be ligated together. The combined size of the vector (4.3 kb) and LEPR insert (2457 bp) would be 6757 bp. Transformants were spread on to LB agar plate with kanamycin to detect those that had incorporated the vector.

The ligation and transformation of the LEPR insert into pCMV-Tag 2B vector was also problematic. Five ligations were set up, three of which were control reactions to determine the effectiveness of the double restriction digest of the vector (control one), to determine whether any uncut vector plasmid remained (control two) and to verify if any of the original plasmid remained (control three). Two experimental reactions were used to determine the best insert to vector ratio for ligation between the LEPR insert and the expression vector. The volume of LEPR extract to be added to the vector for a ratio of 1:1

was 58.12 μg. The concentration of LEPR extracted from the agarose gel after double digest was determined and found to be 4925 μg/ml. As with the ligations for the pGEM-T vector, an additional set of reactions was set up with pCMV-Tag 2B so that they could be electrophoresed to ensure the ligation had been successful before transforming competent *E. coli*. However, this showed that the concentration of the LEPR insert was actually low as the LEPR band wasnot visible and therefore the number of ligations into the pCMV-Tag 2B would be expected to be low (Figure 3.1.20). In addition the bands visible from the two experimental ligations seemed to relate to the pCMV-Tag 2B vector from control ligation one which contained no control DNA. The ligation reactions were still used to transform competent *E.coli* as it was possible the vector, which contains the insert, produced the 3 kb band upon electrophoresis, because the vector and insert ligated together would be expected to be smaller than the expected band of 6.8 kb, as it would be circular and supercoiled therefore it would run further. Despite this, no transformations were detected. The LB agar plates were checked so that the correct volume of Kanamycin was added to produce a final concentration of 50 μg/ml. The volume of LEPR insert ligated into the vector was increased, to increase the concentration further, with a decrease in the volume of distilled water so that the total volume of the ligation remained at 10 μl, however this still failed to produce any colonies. The LEPR was ligated straight into the pCMV-Tag 2B vector, bypassing the sub-cloning into the pGEM-T vector by digesting the LEPR product extracted from the gel with *BamHI* and *SalI* to produce the same sticky ends however this again resulted in no colonies.

A)



B)

pCMV-Tag 2 Multiple Cloning Site Region
(sequence shown 620? 43)

T3 promoter

AA TTA ACC CTC ACT AAA GGG AAC AAA AGC TGG AGC TCC ACC GCG GTG GCG GCC GCC ACC ATG...
KOZAK
M

FLAG tag
D Y K D D D D K
...GAT TAC AAG GAT GAC GAC GAT AAG * GCC CGG GCG GAT CCC CCG GGC TGC AGG AAT TC...
Srf I
BamH I
Pst I
EcoR I

EcoR V
Hind III
Acc I/Sal I
Xho I
Apa I
...G ATA TCA AGC TTA TCG ATA CCG TCG ACC TCG AGG GGG GGC CCG GTA CCT...

T7 promoter
...TAATTAATTAAGGTACCAGGTAAGTGTACCCAATTCGCCCTATAGTGAGTCGTATTA
MULTIPLE STOP CODONS

Figure 3.1.19: Diagrammatic representation of the pCMV-Tag 2B expression vector. Part A) shows the map of the vector with SV40 origin of replication for large copy number in COS-7 mammalian cells, neomycin and kanamycin resistance gene and the multiple cloning site (MCS), adjacent to the FLAG tag. Part B) shows the multiple cloning site region and the location of the two enzymes, *BamHI* and *SalI* highlighted. Assessed from the pCMV-Tag expression vector technical manual (Stratagene).

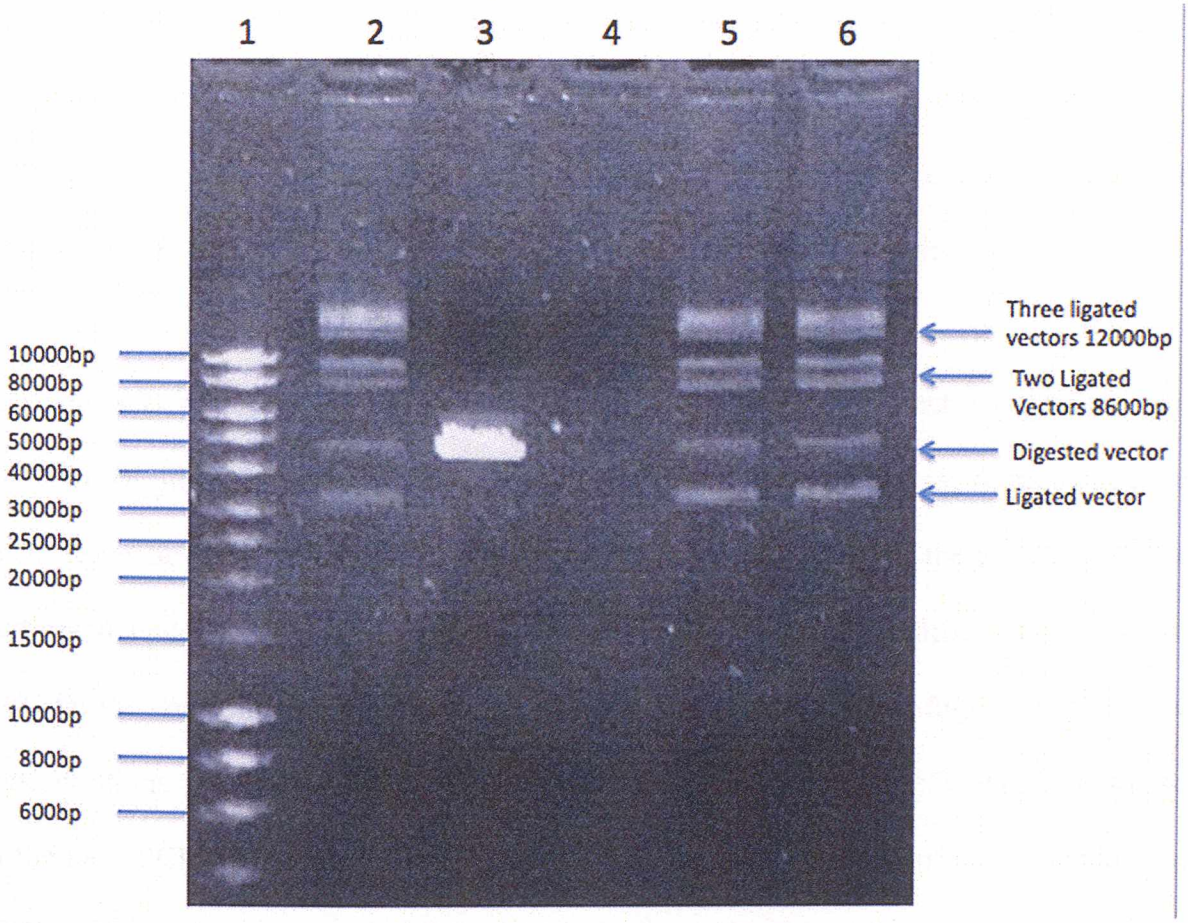


Figure 3.1.20: 1% (w/v) agarose gel electrophoresis of control and experimental ligations between pCMV-Tag 2B expression vector and LEPR insert. Lane 1: DNA ladder (Hyperladder I), Lane 2: control ligation 1 (digested pCMV-Tag 2B vector with ligase); Lane 3: control ligation 2 (digested pCMV-Tag 2B vector); Lane 4: control ligation 3 (LEPR insert); Lane 5: experimental ligation 1; Lane 6: experimental ligation 2. Linear pCMV-Tag 2B expression vector can be seen at 4.3 kb in Lanes 2, 3, 5 and 6.

3.1.5 Nested LEPR PCR

As the ligations with both vectors and subsequent transformations of *E.coli* were unsuccessful, an additional set of primers was designed so that a nested PCR could be performed (Appendix VI). This was to increase the concentration of LEPR product extracted after the second PCR was performed to ensure there was a high concentration of LEPR product to enable successful ligations into pGEM-T vector, in the correct orientation and subsequent transformations into the pCMV-Tag 2B expression vector.

The new sets of primers were designed further out from the original set (LEPRF and LEPRR). During the reverse transcription step the newly designed reverse primer was used to create the new cDNA, using placental RNA (see section 2.4.10) and the nested PCR was performed under standard conditions (see section 2.4.11) but with different conditions, 52°C for the annealing temperature and at 2 mM, 2.5 mM and 3 mM Magnesium Chloride concentrations. The 3 mM reaction was subsequently diluted to 1:20, 1:50 and 1:200 to use in the next PCR step with the original set of LEPR primers at the optimised conditions, 56°C annealing temperature and 2.5 mM Magnesium Chloride. All of these produced the expected band of 2457 bp (Figure 3.1.21). The remaining PCR products, from the two dilutions were electrophoresed using a new agarose gel over seven wells and extracted to be used in the ligation of the pGEM-T vector. Concentration of these seven extracts were determined and found to be between 300 ng/µl and 975 ng/µl. For a 1:3 vector to insert ratio 125 ng of insert was needed, and even though the concentration of the LEPR extract was above this, 3 µl was added to the experimental ligation to increase the number of ligations with the correct orientation of the LEPR insert with the vector. Three separate transformations were performed in which fifty-nine white colonies were produced on LB

agar supplemented with X-Gal, IPTG and Ampicillin, and subsequently grown as overnight cultures. However when these were digested using *BamHI* and *SalI* none of the mini preparations produced the two bands of vector and LEPR insert (Figure 3.1.22). Most of these seemed to be empty vector or contain the LEPR insert in the wrong orientation. The correct orientation was important so that the LEPR insert had the same sticky-ends for the enzymes *BamH*I and *SalI*, as the pCMV-Tag 2B vector so that they could be ligated together. As a result there was no LEPR to be ligated into the pCMV-Tag 2B expression vector. This meant that the subsequent stages in the functional study could not be completed, in particular stage seven which involved sequencing the LEPR insert. This would have identified colonies containing LEPR insert in the correct orientation in pCMV Tag 2B expression vector and also identify any *de novo* polymerase base misincorporation mistakes. In hindsight, sequencing after each cloning stage could have been performed to immediately identify any of these *de novo* mistakes and not wait until stage seven as ligation into both pGEM-T and pCMV Tag 2B vectors was problematic and these mistakes may have been the cause.

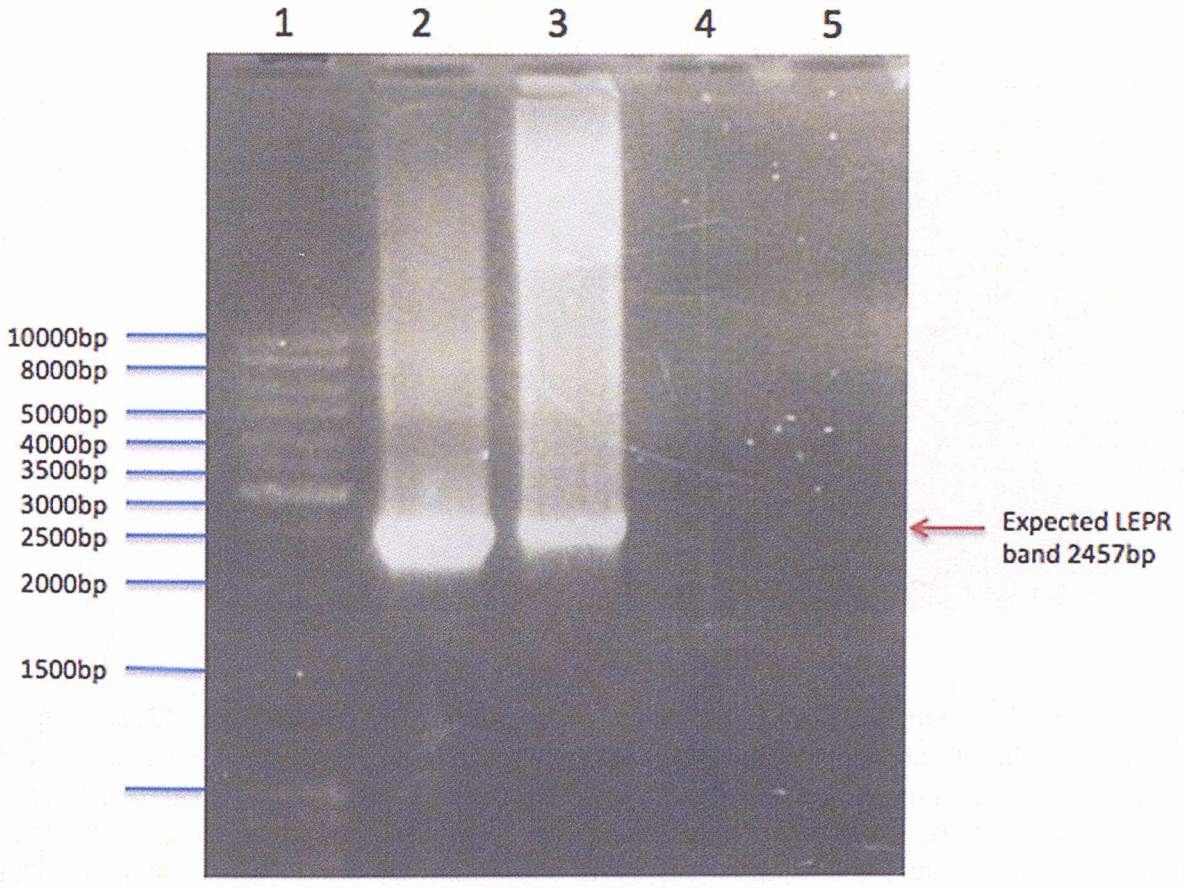


Figure 3.1.21: 1% (w/v) agarose gel electrophoresis of LEPR PCR using the original LEPR primers from different dilutions of the nested RT-PCR. Lane 1: DNA ladder (GeneRuler), Lane 2: a 1:20 dilution of the nested PCR, Lane 3: a 1:50 dilution of the nested PCR, Lane 4: a 1:200 dilution of the nested PCR, Lane 5: water blank. 10 µl of the PCR was loaded into the corresponding lanes.

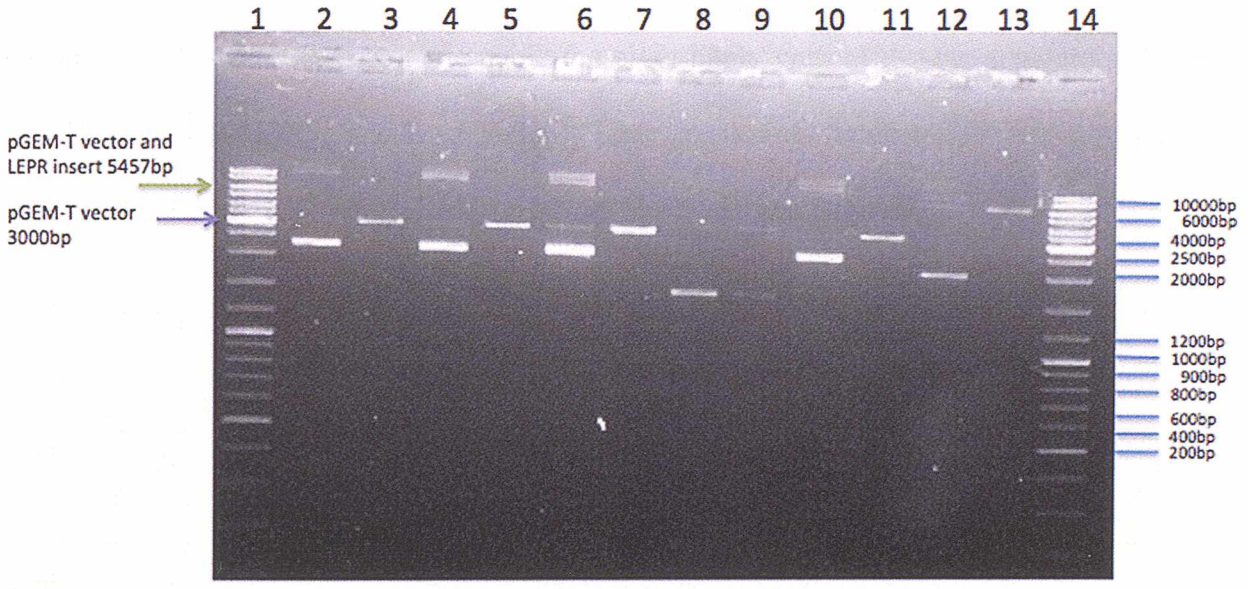


Figure 3.1.22: 1% (w/v) agarose gel electrophoresis of restriction double digests (*BamHI* and *SalI*) of six mini preparations, with undigested mini preparations loaded adjacently. Lanes 1 and 14: DNA ladder (GeneRuler), Lane 2: undigested mini preparation 23, Lane 3: digested mini preparation 23, Lane 4: undigested mini preparation 24, Lane 5: digested mini preparation 24, Lane 6: undigested mini preparation 25, Lane 7: digested mini preparation 25, Lane 8: undigested mini preparation 26, Lane 8: digested mini preparation 26, Lane 9: undigested mini preparation 27, Lane 10: digested mini preparation 27, Lane 11: undigested mini preparation 28, Lane 12: digested mini preparation 28. Lanes 3, 5, 7, 9 and 11 show empty vector at 3 kb when cut with the two enzymes whilst lane 13 shows the LEPR in the wrong orientation within the vector. 5 µl of undigested mini preparation and 10 µl of digested mini preparations were loaded into the corresponding lanes.

In summary, the LEPR extracellular region was successfully PCR-amplified and sub-cloned into the pGEM-T vector however the LEPR insert was not successfully sub-cloned into the pCMV-Tag 2B vector to enable the eight haplotypes of the three polymorphisms (Lys109Arg, Gln223Arg and Lys656Asn). Time constraints prevented further exploration of this line of work and a modelling '*in silico*' approach was adopted.

3.2.0 Computer Modelling of the Leptin Receptor

Leptin is a key hormone involved in energy regulation as well as being apart of the immune, cardiovascular and reproductive systems. However, the precise mechanics driving the interaction between leptin and its receptor, and the complex it forms, remains elusive (Peelman *et al.,* 2006b). Leptin is key in energy regulation (Figure 1.9.1) but proved refractory to cloning (see section 3.1). We decided to use computer modelling methods instead, to see if we could identify whether three SNPs located within the extracellular region of the leptin receptor, Lys109Arg, Gln223Arg and Lys656Asn, widely identified as important in genetic studies (see section 1.9.11) could be shown to alter binding affinity with leptin either isolated or as haplotypes. Some data suggests that this is the case: LEPRb with ^{223}Arg shows lower affinity for Leptin than Gln^{223} (Mechan, Pers Com).

In order to predict whether different leptin receptor haplotypes, defined by the three specific polymorphismsnamed above affects its conformation and thus affects binding of its ligand, leptin, the I-TASSER online server (http://zhanglab.ccmb.med.umich.edu/I-TASSER) was used to predict the 2D and 3D structures including the eight haplotypes (Table 3.2.1; Appendix VII). The server uses a 'sequence to structure to function paradigm' that includes using different servers and additional software was used to identify regions of homology and identity and any differences between the predicted 3D structures of the LEPR that contain the eight haplotypes.

Table 3.2.1: The eight haplotypes within the extracellular region of the *LEPR* defined by the three specific polymorphisms at Lys109Arg, Gln223Arg and Lys656Asn. Haplotype sequence is shown as the amino acid sequence and with respect to the codon sequence at each polymorphism. The SNP is shown in bold.

Haplotype Number	Codon Polymorphisms			Amino Acid Haplotype Sequence
	Lys^{109}	Gln^{223}	Lys^{656}	
Haplotype 1	A**A**G	C**A**G	AA**G**	KQK
Haplotype 2	A**G**G	C**A**G	AA**G**	RQK
Haplotype 3	A**A**G	C**G**G	AA**G**	KRK
Haplotype 4	A**A**G	C**A**G	AG**C**	KQN
Haplotype 5	A**G**G	C**G**G	AA**G**	RRK
Haplotype 6	A**A**G	C**G**G	AG**C**	KRN
Haplotype 7	A**G**G	C**G**G	AG**C**	RRN
Haplotype 8	A**G**G	C**A**G	AG**G**	RQN

3.2.1 I-TASSER Server

The I-TASSER server (Roy *et al.,* 2010) has been a recent development in predicting protein structure. It enables the user to predict protein structure using a 'sequence-to-structure-to-function paradigm'. The prediction of the protein is estimated over four different stages; a) template identification, b) structural assembly, c) model selection and refinement and d) structure-based functional annotation.

3.2.1.1 Template Identification

During template identification, proteins are identified from a PDB structure library, which contains protein structures that have already been solved by a process of threading by a

local meta threading server (LOMETS) (Wu *et al.,* 2007), in which template proteins are identified by using a sequence-structure alignment procedure. It selects proteins that show similarities with the query protein in overall structure or for a particular motif or region. A sequence profile is developed from this as the homologous sequences of the template proteins are aligned together. LOMETS utilises different threading methodologies therefore combines several algorithms to maximise the number of protein templates homologous to the query protein. These include the PSI-BLAST profiles (Altschul *et al.,* 1997) and the Needleman-Wunsch (Needleman and Wunsch, 1970) and Smith-Waterman (Smith and Waterman, 1981) algorithms. As the scoring system and sensitivity of the alignment vary between each algorithm, the quality of the alignments for each program is assessed using a normalised Z-score.

3.2.1.2 Structure Alignment

During structure assembly, continuous fragments from the threading alignment are used from the templates to make a conformational structure of the query protein that have aligned well. The unaligned regions are built by *ab initio* modelling. Two different approaches are applied as a result. Aligned regions are modelled using an off-lattice approach to maximise the accuracy of the secondary structure whereas unaligned regions are modelled on a cubic lattice system. This process is performed by replica-exchange Monte Carlo simulation, which determines the best arrangement of the continuously aligned fragments. These are guided by a knowledge-based force field that involves using different aspects: general knowledge-based statistics terms from PDB (using hydrogen bonding, hydrophobicity, C-alpha/side chain correlations), spatial restraints from threading templates and sequence-based contact predictions from SVMSEQ (a support-vector-machine (SVM) based residue-residue contact predictor that uses sequence information)

(Zhang, 2009). The structure trajectories are then clustered by SPICKER to identify structures that have a low free energy state and by averaging the positions of all the clustered structures the cluster centroids are obtained.

3.2.1.3 Model selection and refinement

During this stage, cluster centroids are used during the fragment assembly simulation to refine the models and remove any steric clashes. The LOMETS treading alignments and the PDB structures that were identified as structurally closest to the cluster centroids were used as constraints within this second round of selection. The structures are also clustered for a second time and those with a low energy state are selected from the different cluster centroids. The cluster centroids are used to make the final structural models by using the REMO (Li and Zhang, 2009) program that builds all atom models by using the C alpha trace and hydrogen bonding.

3.2.1.4 Structure-based functional annotation

The 3D predicted models are used to determine protein function by matching them structurally to proteins of known structure and function, using three methodologies. The topology of the predicted models is used in the alignment of functional templates using TM-align (Zhang and Skolnick, 2005) program. A TM-score is used to rank those depending on the level of structural similarity and given a value of between zero and one. Those closest to one have a high degree of structural similarity. Proteins with similar global structure to the predicted models are selected from the PDB along with those that are similar in structure and sequence surrounding the active/binding site region of the

predicted models. In each of these search results, the matched proteins are measured and then ranked based on the level of structural and sequence similarity.

The accuracy of the predicted protein structure is estimated using the confidence score (C-score). This ranges from -5 to 2, where the higher the score indicates a better quality model. The score is calculated using the significance (Z-score) of the threading alignments in LOMETS and the convergence (cluster density) of the I-TASSER simulations.

3.2.2 Analysis of I-TASSER Results

I-TASSER was used to analyse the wild type LEPR amino acid sequence, with the three mutations individually at Lys109Arg, Gln223Arg and Lys656Arg.The LEPR protein sequence was downloaded from the UNIPROT Knowledgebase, accession number (P48357) termed the native wild type sequence (see section 2.8.1), and then edited and each leptin receptor haplotype sequence uploaded on to the I-TASSER server (see section 2.8.2). Once the structure had been determined the results were displayed using a web browser. They were displayed and analysed with regards to the following criteria: regions of predicted secondary structure, regions of predicted 3D structure, identification of threading templates i.e. finding identity/similarity in fold recognition with the protein of interest compared to that which exists in solved protein structures and identifying structural analogues i.e. structural motifs from known proteins are compared to the protein of interest, with those with a best fit being analogous in structure and thus producing a predicted model.

3.2.2.1 Sequence and predicted secondary structure

Figure 3.2.1 shows the results of submitted amino acid sequence in the FASTA format and the predicted secondary structure shown below with the confidence score for each residue. The secondary structure prediction is given as either H for alpha helix, S for beta pleated sheet or C forrandom coil. The confidence score (C-score) ranges from zero to nine, with nine being the highest and indicating a higher confidence in the prediction and is based on the significance between the threading template alignments and the convergence parameters of the structure assembly simulations (Zhang, 2008).

The four haplotypes sequences, which includes the wild type LEPR and the single polymorphisms at Lys^{109}, Gln^{223} and Lys^{656}, were submitted to the server separately, as more than one protein sequence cannot be submitted at one time. Results of the individual predicted models were viewed but they were also compared when all four predictions of the protein structure were complete. In the first instance, the predicted secondary sequence was viewed. The best model, 'model one', was compared between the wild type LEPR predicted protein and the models with the individual mutations, producing four models in total (Table 3.2.1). These werenamed Model A which was Native LEPR (haplotype sequence 1), Model B which has a single change LEPR Lys^{109} to Arginine (haplotype sequence 2, Model C which has a single change of LEPR Gln^{223} to Arginine (haplotype sequence 3) and Model D which has a single change LEPR Lys^{656} to Asparagine (haplotype sequence 4).

Table 3.2.2: Results from the I-TASSER protein prediction. C-score and TM-scores indicating the best model are given for the 'first model' for the full-length native LEPR protein sequence and for the mutations, ^{109}Arg, ^{223}Arg and ^{656}Asn.

Protein	C-score	TM-score
Native LEPR (Model A)	-1.30	0.55±0.15
LEPR 109Arg (Model B)	-2.79	0.39±0.13
LEPR 223Arg (Model C)	-2.67	0.41±0.14
LEPR 656Asn (Model D)	-2.04	0.47±0.15

Submitted Sequence

```
>protein
MICQKFCVVLLHWEFIYVITAFNLSYPITPWRFKLSCMPPNSTYDYFLLPAGLSKNTSNS
NGHYETAVEPKFNSSGTHFSNLSKTTFHCCFRSEQDRNCSLCADNIEGKTFVSTVNSLVF
QQIDANWNIQCWLKGDLKLFICYVESLFKNLFRNYNYKVHLLYVLPEVLEDSPLVPQKGS
FQMVHCNCSVHECCECLVPVPTAKLNDTLLMCLKITSGGVIFQSPLMSVQPINMVKPDPP
LGLHMEITDDGNLKISWSSPPLVPFPLQYQVKYSENSTTVIREADKIVSATSLLVDSILP
GSSYEVQVRGKRLDGPGIWSDWSTPRVFTTQDVIYFPPKILTSVGSNVSFHCIYKKENKI
VPSKEIVWWMNLAEKIPQSQYDVVSDHVSKVTFFNLNETKPRGKFTYDAVYCCNEHECHH
RYAELYVIDVNINISCETDGYLTKMTCRWSTSTIQSLAESTLQLRYHRSSLYCSDIPSIH
PISEPKDCYLQSDGFYECIFQPIFLLSGYTMWIRINHSLGSLDSPPTCVLPDSVVKPLPP
SSVKAEITINIGLLKISWEKPVFPENNLQFQIRYGLSGKEVQWKMYEVYDAKSKSVSLPV
PDLCAVYAVQVRCKRLDGLGYWSNWSNPAYTVVMDIKVPMRGPEFWRIINGDTMKKEKNV
TLLWKPLMKNDSLCSVQRYVINHHTSCNGTWSEDVGNHTKFTFLWTEQAHTVTVLAINSI
GASVANFNLTFSWPMSKVNIVQSLSAYPLNSSCVIVSWILSPSDYKLMYFIIEWKNLNED
GEIKWLRISSSVKKYYIHDHFIPIEKYQFSLYPIFMEGVGKPKIINSFTQDDIEKHQSDA
GLYVIVPVIISSSILLLGTLLISHQRMKKLFWEDVPNPKNCSWAQGLNFQKPETFEHLFI
KHTASVTCGPLLLEPETISEDISVDTSWKNKDEMMPTTVVSLLSTTDLEKGSVCISDQFN
SVNFSEAEGTEVTYEDESQRQPFVKYATLISNSKPSETGEEQGLINSSVTKCFSSKNSPL
KDSFSNSSWEIEAQAFFILSDQHPNIISPHLTFSEGLDELLKLEGNFPEENNDKKSIYYL
GVTSIKKRESGVLLTDKSRVSCPFPAPCLFTDIRVLQDSCSHFVENNINLGTSSKKTFAS
YMPQFQTCSTQTHKIMENKMCDLTV
```

Predicted Secondary Structure

```
                    20        40        60        80        100       120
                    |         |         |         |         |         |
Sequence    MICQKFCVVLLHWEFIYVITAFNLSYPITPWRFKLSCMPPNSTYDYFLLPAGLSKNTSNSNGHYETAVEPKFNSSGTHFSNLSKTTFHCCFRSEQDRNCSLCADNIEGKTFVSTVNSLVFQ
Prediction  CCSCSHHHHHHHHHHHHHHHHHHCCCCCCCCCCSSSSCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCSSSSSSCCCCCSSSCCCCCSSSSSCCCCCCCCCCCCCCCCCCCCCSSCSSSCCCCCS
Conf.Score  9510276888999999999987604114850225775037634442247466521045787565416777645887346326773789972455445776667985470150217624627
            Secondary structure elements are shown as H for Alpha helix,S for Beta sheet & C for Coil
```

Figure 3.2.1: I-TASSER results showing the submitted sequence in FASTA format and the predicted secondary sequence for the complete native protein sequence of LEPR (accession number P48357).

The predicted secondary structure was viewed and this showed that the majority of the predicted protein consisted of random coil and beta sheets and a small amount of alpha helix, for each of the models A to D (Figure 3.2.2). The confidence score for these secondary structures varied. Some regions had a high confidence score of seven to nine, other regions were low, <5. Beta pleated sheet regions are better predicted than random coil, which would explain the higher confidence score shown for beta pleatedsheets compared to random coil. There also appeared to be slight variations between each of the four models, in particular the length of some of the coil and beta pleated sheet structures vary between them. At position 160, Histidine residue, (Figure 3.2.2) Model C is predicted as having a shorter beta pleated sheet structure compared to the other three predicted models and also at amino acid position 182, Glutamine residue. The confidence scores also vary at these residues between the four models.

The four haplotypes were then viewed with respect to the predicted 3D structure to identify any subtle difference in patterns of helix, beta sheet and coil from the native conformation.

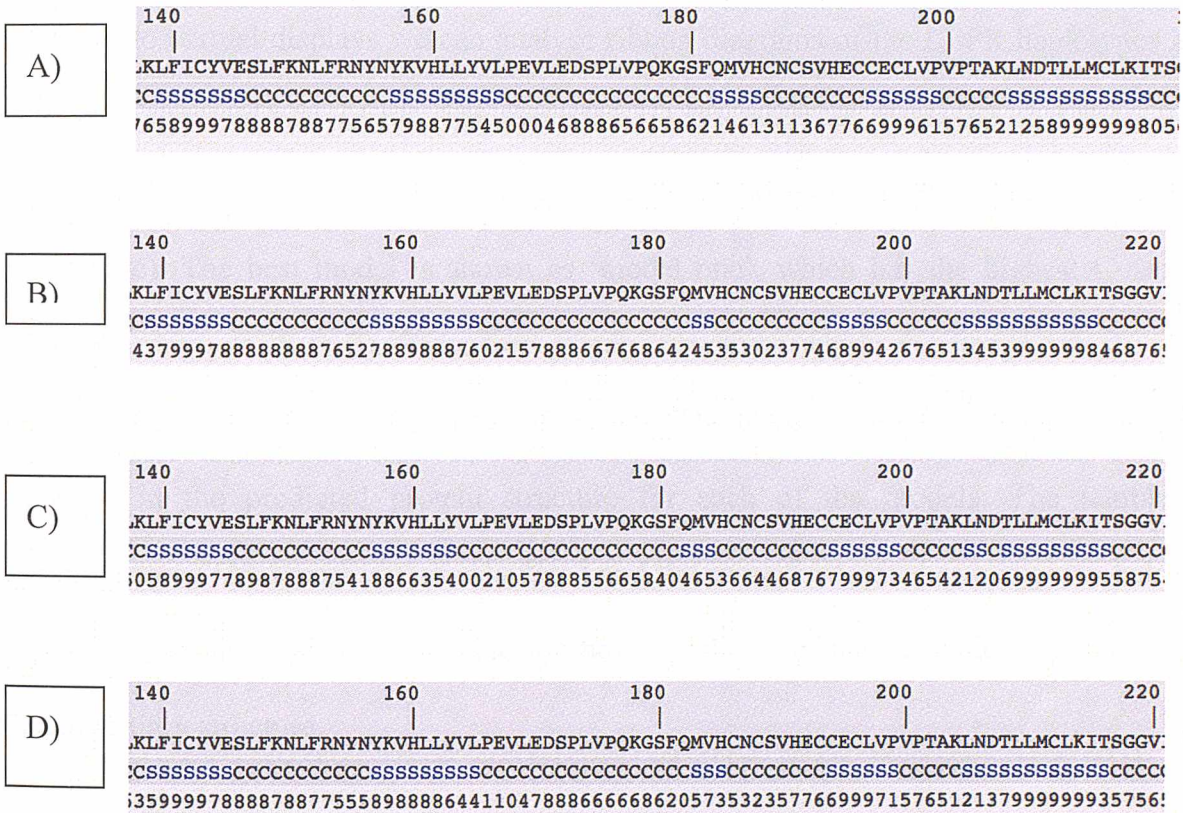


Figure 3.2.2: A part of the predicted secondary structure from amino acid positions 138 to 227 each of the four models. A) Model A; B) Model B; C) Model C and D) Model D. The top row displays the amino acid sequence, middle row indicates where secondary structure is predicted, in this case either a random coil (C) or a beta pleated sheet (S). The last row gives the confidence score. The length of the four predicted beta pleated sheet regions varies between the four models between the amino acids 157 and 216.

3.2.2.2 Predicted 3D Model

In addition to textual displays we can analyse ribbon diagrams of the LEPR haplotypes as compared to the wild type LEPR. The predicted 3D structures are displayed, that show the five best models as determined from the query amino acid sequence (see Figure 3.2.3 for one example). The best model is shown as 'model one', which has the highest C-score (within the range of -5 to 2). Secondary protein structures within the image are shown, which includes alpha helix and beta pleated sheets. The C-score for each image shows the confidence in the predicted protein structure for each of the models. The template modelling score (TM-score) is also calculated and the score for 'model one'is given below the predicted models. This score indicates how similar two protein structures are with different tertiary structure.

The C-scores for the four models were relatively low, -1.30, -2.79, -2.67 and -2.04 for models A, B, C and D respectively. The higher the value, i.e. closer to two, the better the confidence in the predicted structure, that it is the more likely structure of the query protein is to be real. However, when the three mutations were created, separately, the C-score decreased. This suggests that the confidence in all four predicted proteins is relatively low indicating the LEPR protein structure is refractory to this type of computer determination. Interestingly, the image of 'model one' (the best predicted model by I-TASSER) varied between each of the four models A, B, C and D, which showed that that the different alleles did appear to have a structural impact, at least with this prediction package(Figure 3.2.4).

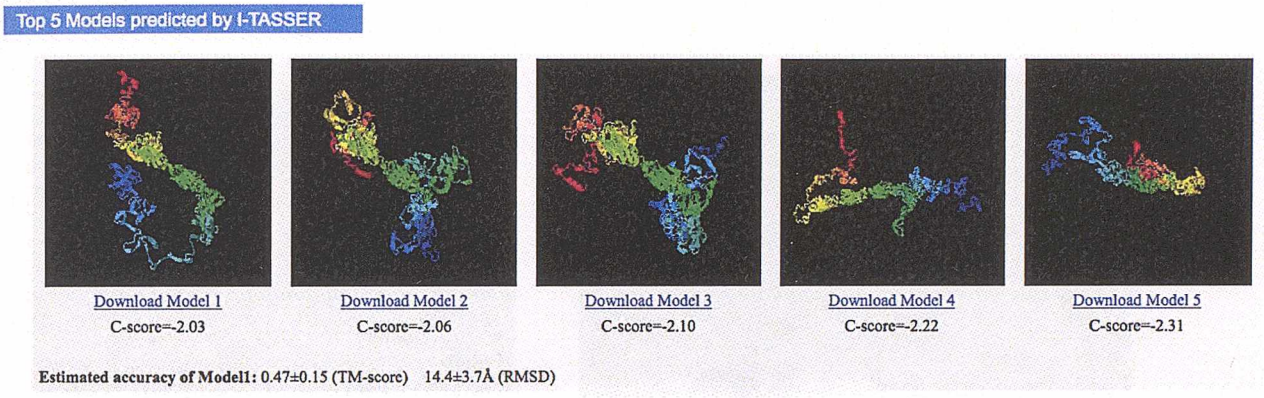


Figure 3.2.3: The I-TASSER predicted five models for the leptin receptor with the Lys656Asn polymorphism. C-score is given at the bottom of each model, which indicates the confidence in the prediction of the protein structure, with the TM-score given for 'model one'.

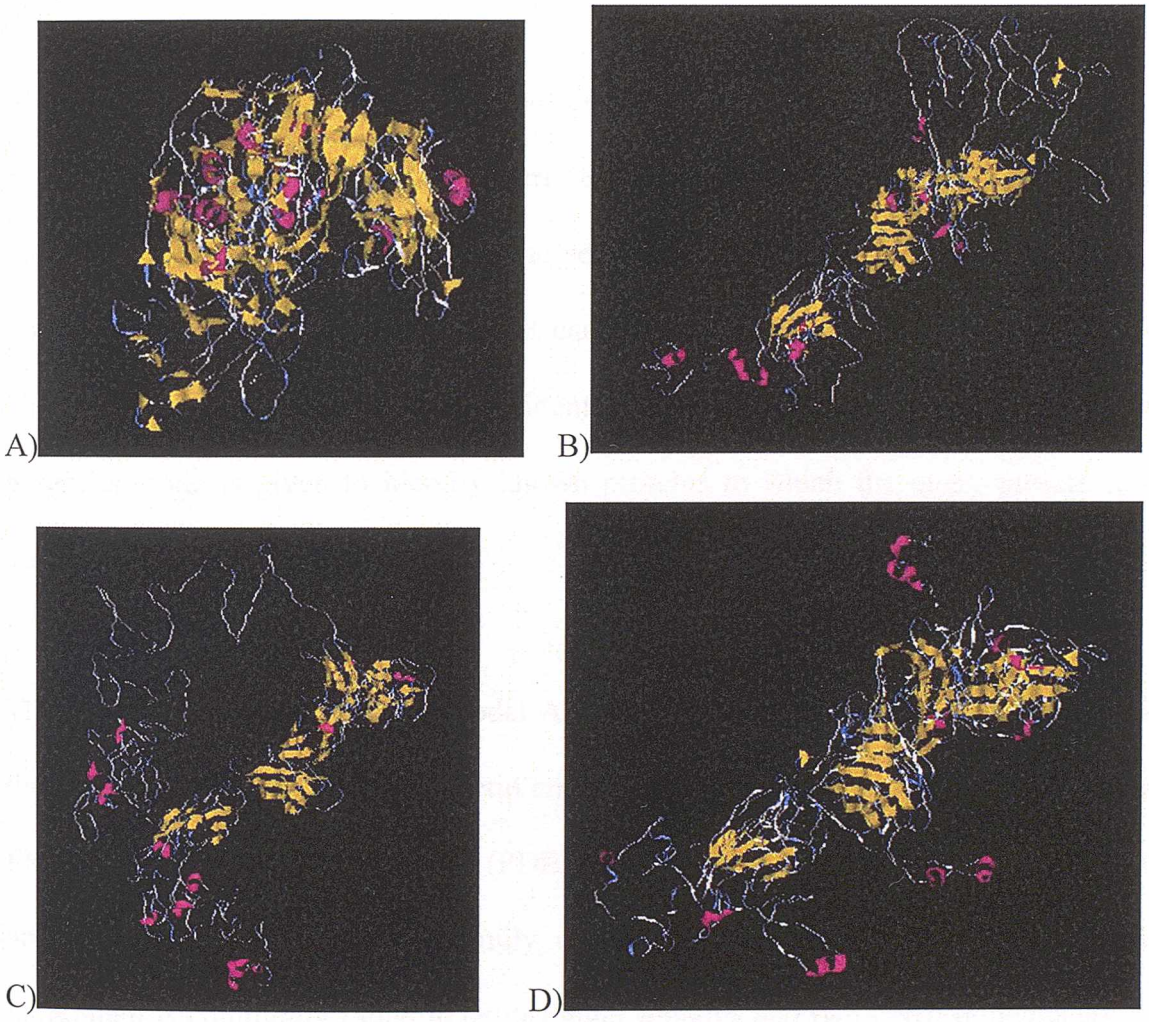


Figure 3.2.4: Predicted protein structures of model 1 from four models a) the complete region of the leptin receptor (LEPR native) b) LEPR complete sequence with ^{109}Arg polymorphism c) LEPR complete sequence with ^{223}Arg polymorphism and d) LEPR complete sequence with ^{656}Asn polymorphism.

3.2.2.3 Threading Alignments

The threading template lists ten templatesthat were used for the query protein, as identified by the LOMETS threading program, part of the I-TASSER suite. The alignments of each of the templates are shown along side the sequence of the query protein (Figure 3.2.5). A normalised Z-score is provided against each template, which indicates the quality of the alignment. A value >1 indicates a confident alignment. A Protein Data Bank (PDB) unique accession code is given to identify known proteins to which the query protein has been aligned, for further information

The ten best templates used for model A included the ectodomain of gp130 (PDB 3I5H), the fragment of the human fibronectin encompassing type III repeats (PDB IFNF) and the FNIII domains of the human gp130 (PDB 3L5I). These are expected as gp130 is a cytokine and belongs to the cytokine superfamily, of which LEPR is a member. The fibronectin type III domain is commonly found in extracellular proteins and two FNIII domains are located within the LEPR (Figure 1.9.3). This was also true for models B, C and D, which would be expected as the query protein sequences are essentially the same but with only one amino acid difference.

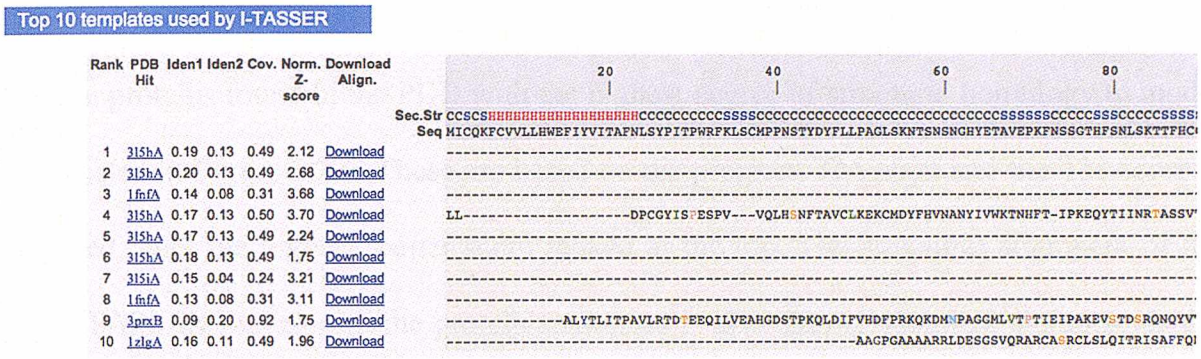


Figure 3.2.5: A exert of the I-TASSER analysis showing ten templates located by the I-TASSER server, from the protein database. These known structures are utilised to predict the protein structure of the query protein. Results shown are for the complete sequence of the native LEPR accession number P48357. The unique PDB reference number represents each template that is used by I-TASSER. Each template sequence is aligned beneath the query protein sequence where regions of homology have been identified.

3.2.2.4 Structural Analogues

The ten proteins found in the PDB with the highest degree of structural homology to model 1 are listed in Figure 3.2.6. These are based on the program TM-align and the TM-score is provided, with the highest, better score ranked at the top. The structural alignment of the best I-TASSER model with the protein analogues is also shown. Identical amino acids are colour matched, with mismatches differentiated using a black typeface.

The ten best proteins used that are structurally closest to model A included the complement component-3 glycoprotein (2B39), human complement component-5 glycoprotein (3CU7), P-glycoprotein (3G6L), Botulinium neurotoxin serotype A (3BTA) and serotype B (1F31), *E. coli* RecA-ssDNA/dsDNA complex (3CMU), endo-alpha-N-acetylgalactosaminidase (2ZXQ), Cobra Venom Factor and human factor B complex (3HS0) and Thioester containing proteins isoform 1 (2PN5). Of these ten proteins only one of these proteins, the P-glycoprotein, is a receptor that can transport a variety of substrates across the intra and extracellular membranes. This indicates that the LEPR is structurally similar to a range of other proteins that do not belong to the cytokine family. However it would be expected that within these results, the LEPR would be structural similar to other type 1 cytokine receptors as they share critical regions of similarity (Iserentant *et al.,* 2005, Stratigopoulos *et al.,* 2009).

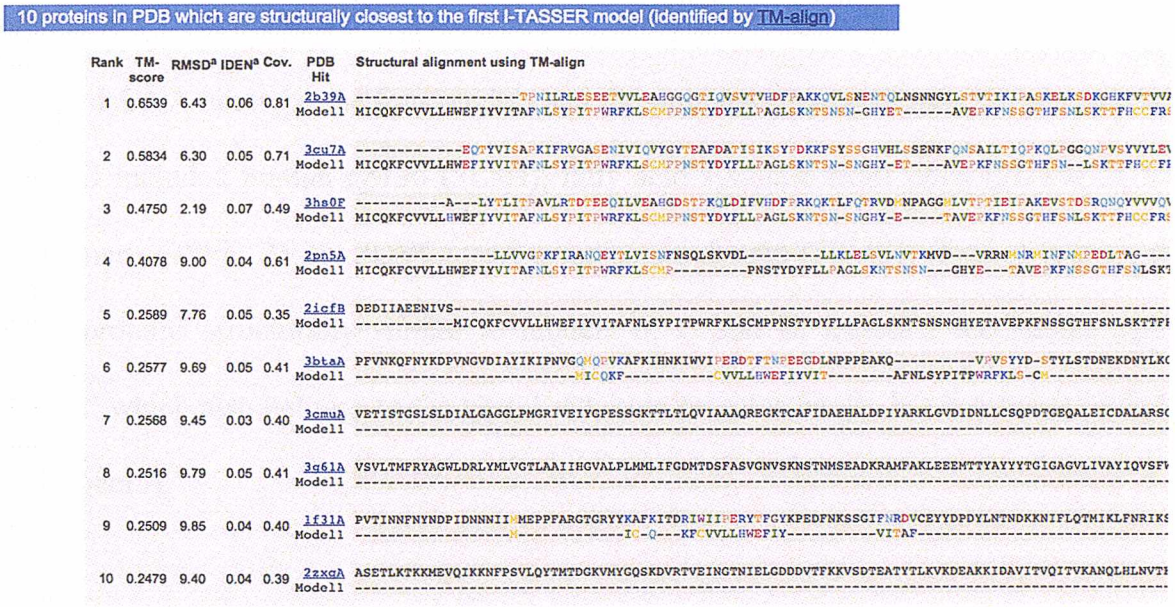


Figure 3.2.6: I-TASSER results showing the ten proteins that are structurally closest to the first model predicted by I-TASSER, using TM-align. Results shown are for the complete protein sequence of the native LEPR (accession number P48357).

When the ten best structural proteins were viewed with models B, C and D there were notable differences with those from model A. Only three proteins were the same, complement component-3 glycoprotein, Botulinum neurotoxin B and RecA-ssDNA/dsDNA amongst the four models and the proteins structurally similar to models B, C and D included human gp130 (3L5H), fatty acid synthase from yeast (2UV8; 2UVA) and humans (2PFF; 2V28; 2UVC), and cytoplasmic dynein (3PMZ). It is of interest that of the proteins structurally similar to models B, C and D, included the human gp130 receptor, which also belongs to the type 1 cytokine receptor family but it was not included with model A.

3.2.3 Extracellular LEPR Protein Structure Prediction

As the Models A to D had a low C-score it was determined that these were not good models at predicting the structure of LEPR. I-TASSER is mainly used to predict proteins that contain one domain and LEPR contains three domains: extracellular, transmembrane and intracellular. Therefore the protein sequence was split so that only the extracellular region was submitted to the server, i.e. amino acid positions 1 to 830. Mutations were also created at positions 109, 223 and 656 as before. These models were designated models E-H respectively. However, again, the four models produced poor C-score values, indicating that the models predicted were not accurate for the LEPR protein (Table 3.2.2).

Table 3.2.3: I-TASSER results for the extracellular region of the leptin receptor. Native structure was predicted along with the polymorphisms at the positions Lys109Arg, Gln223Arg and Lys656Asn.

Protein Name	Model	C-Score	TM-Score
Extracellular LEPR native	E	-2.84	0.39±0.13
Extracellular LEPR 109Arg	F	-2.95	0.38±0.13
Extracellular LEPR223Arg	G	-2.41	0.43±0.14
Extracellular LEPR656Arg	H	-2.30	0.44±0.14

The predicted secondary structure was viewed for each of the four models, E to H. Again the query proteins were predicted to contain beta pleated sheets, alpha helices and random coils and there were slight variations between each of the models. For example at positions 107 and 108 in the amino acid sequence, for Model F, there was indicated alpha helix within the structure followed by beta sheet but this was not predicted in the other three models (Figure 3.2.7). In addition, the length of the beta sheet varied between each of the four models, between three and seven, starting at residue 109 which is lysine to residue 115, valine.

The templates used to predict the protein structure for each of these four models were also viewed, to see if there was any similarity in the length of homology, along with the proteins identified by the programme as having a similar structure. These were both similar to the full-length protein models A to D, gp130 ectodomain and the fragment of human fibronectin encompassing type III repeats. There were also differences in the structural images of the four models E to F, as with the earlier models (Figure 3.2.8).

Figure 3.2.7: Predicted secondary protein structure of the four models: A) = E, B) = F, C) = G and D) = H. Highlighted by the arrows are positions 107 and 108 in which alpha helix is predicted in Model F but is absent in the other three models. Highlighted in green are the predicted beta sheets, which vary in length between the four models.

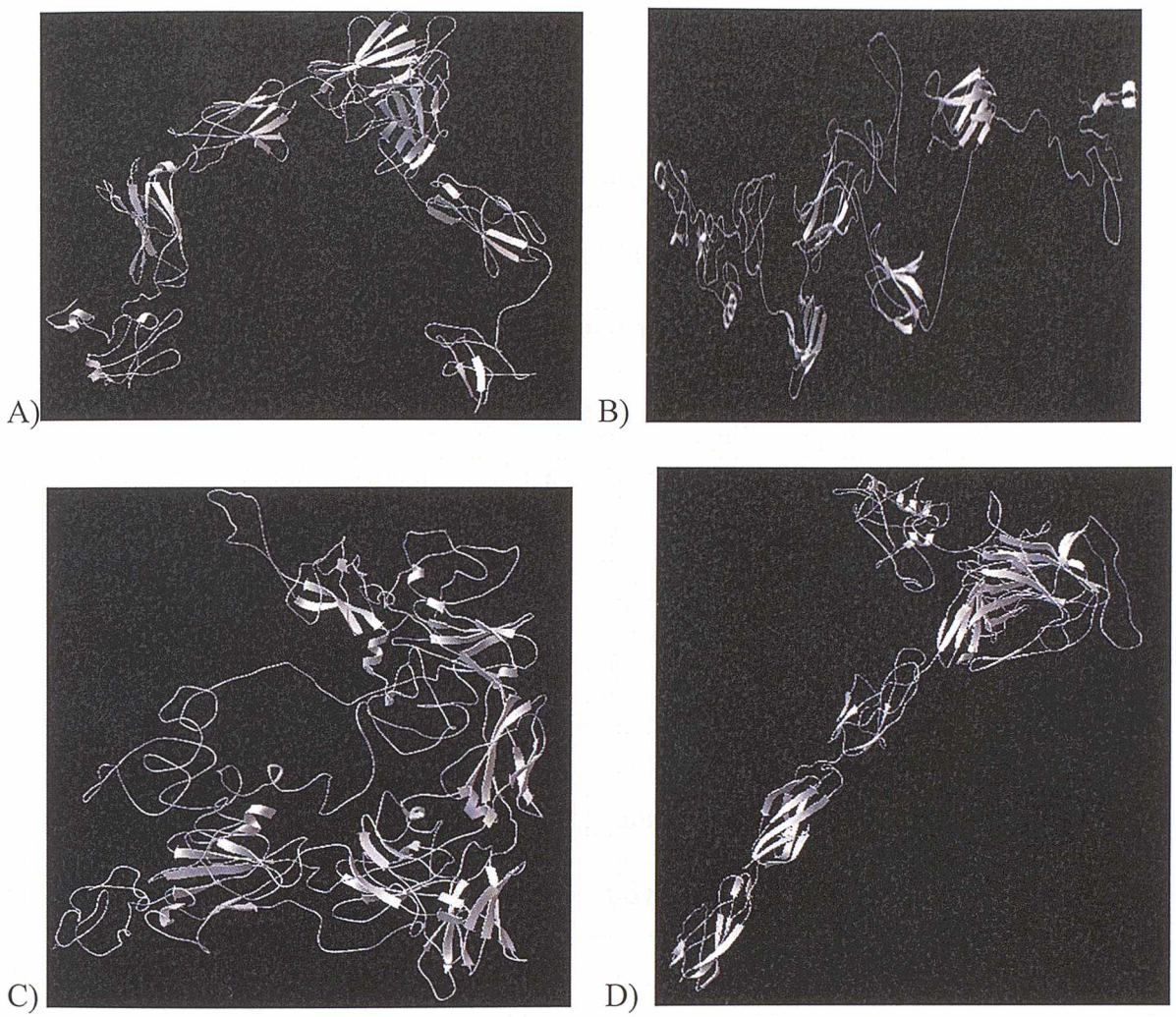


Figure 3.2.8: Predicted protein structures of 'model one' for the four models E, F, G and H. A) the native extracellular region of the leptin receptor, B) the extracellular region with ^{109}Arg polymorphism C) the extracellular region with ^{223}Arg polymorphism and D) the extracellular region with the ^{656}Asn polymorphism.

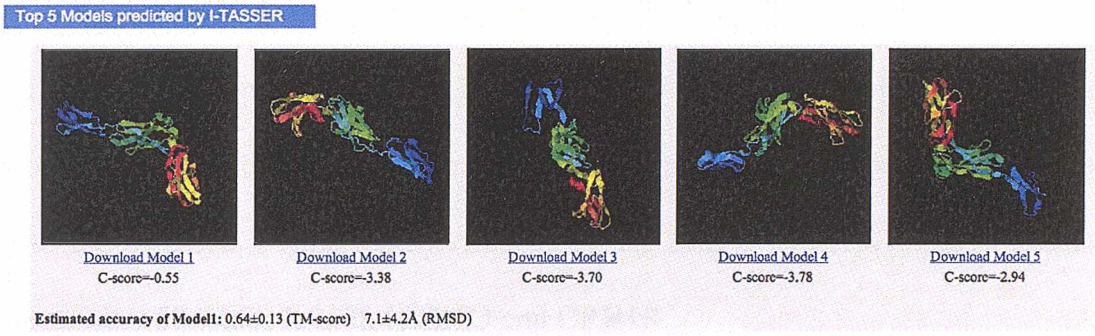
The extracellular region contains four different domains, which might explain why I-TASSER was not able to predict the structure of the protein with strong confidence. The different domains within the extracellular region include CRH1, Ig-like domain, CRH2 and fibronectin type III domain. Out of the three polymorphisms of interest, only the Gln223Arg mutation was located within one of these domains, the CRH1 domain, therefore the amino acid sequence of the CRH1 domain was submitted to the I-TASSER server (266 amino acids in length). The native sequence was initially submitted followed by the version with the amino acid change from Glu^{223} to ^{223}Arg, which corresponded toposition 162 in the CRH1 sequence.

3.2.4 CRH1 I-TASSER Results

The results of modelling the native CRH1 domain were viewed as before and also compared with the version with the ^{223}Arg peptide. The native protein was given the name CRH1Q and ^{223}Arg peptide, CRH1R.

Compared to the previously described, more complex structure predictions, CRH1Q and CRH1R had relatively high 'model one' C-score values, -0.47 and -0.55 and TM-score values 0.65±0.13 and 0.64±0.13 indicating a higher confidence in the predicted structures. The predicted protein structures of both domains also showed the expected high degree of similarity as viewed by the images of 'model 1' for CRH1Q and CRH1R (Figure 3.2.9).

A)



B)

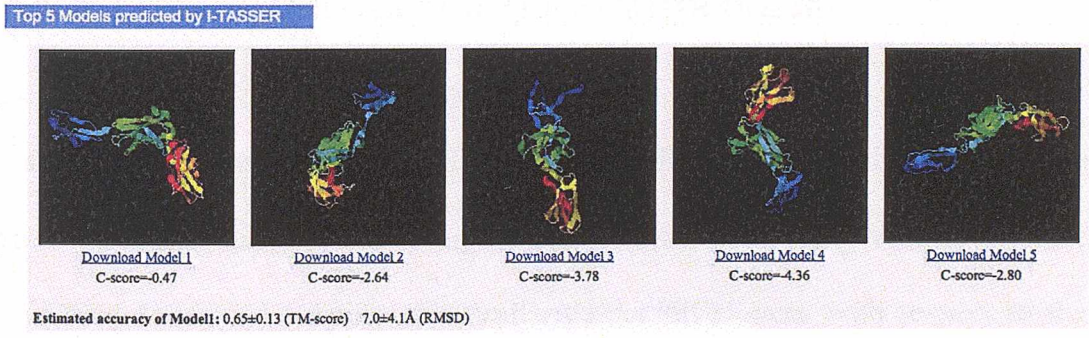


Figure 3.2.9: The five best models predicted by I-TASSER of the CRH1 domain with (A) the glutamine residue at position 223 (CRH1Q) and (B) the arginine residue at the same position (CRH1R). 'Model one' for the two predicted domains are considered the best model of the five predicted.

The threading templates used by LOMETS were viewed for the two models. Proteins of similar structure to the CRH1 domain were used as templates to determining the structure of both the query peptides, CRH1Q and CRH1R. The templates included the extracellular domain of the interleukin 6 receptor alpha chain, the human G-CSF receptor-signalling complex, Leukaemia inhibitory factor (LIF) in complex with LIF receptor domain and gp130 cytokine receptor complex. These are all cytokines with extracellular receptors, belonging to the cytokine 1 superfamily (Zabeau *et al.,* 2003). These peptides were also found to be structurally close to both CRH1Q and CRH1R.

3.2.4.1 Structural Alignment between CRH1Q and CRH1R Domains

The two best models for the peptides CRH1Q and CRH1R from I-TASSER results were downloaded as PDB files then viewed using the DeepView/SWISS-PDB Viewer software (v4.0.4). This was done to determine any similarities or differences with respect to the amino acid change between the two protein structures.

The CRH1Q native protein shows three regions of beta pleated sheet formation of differing sizes (Figure 3.2.10). The smallestis at the start of the domain and contains 56 amino acids, this is connected to the second beta sheet structure, amino acids 61 to 174, by four amino acids, and this is then connected to a third beta sheet structure, amino acids 175 to 266. CRH1R shows the same structure, with three beta sheet structures (Figure 3.2.10).

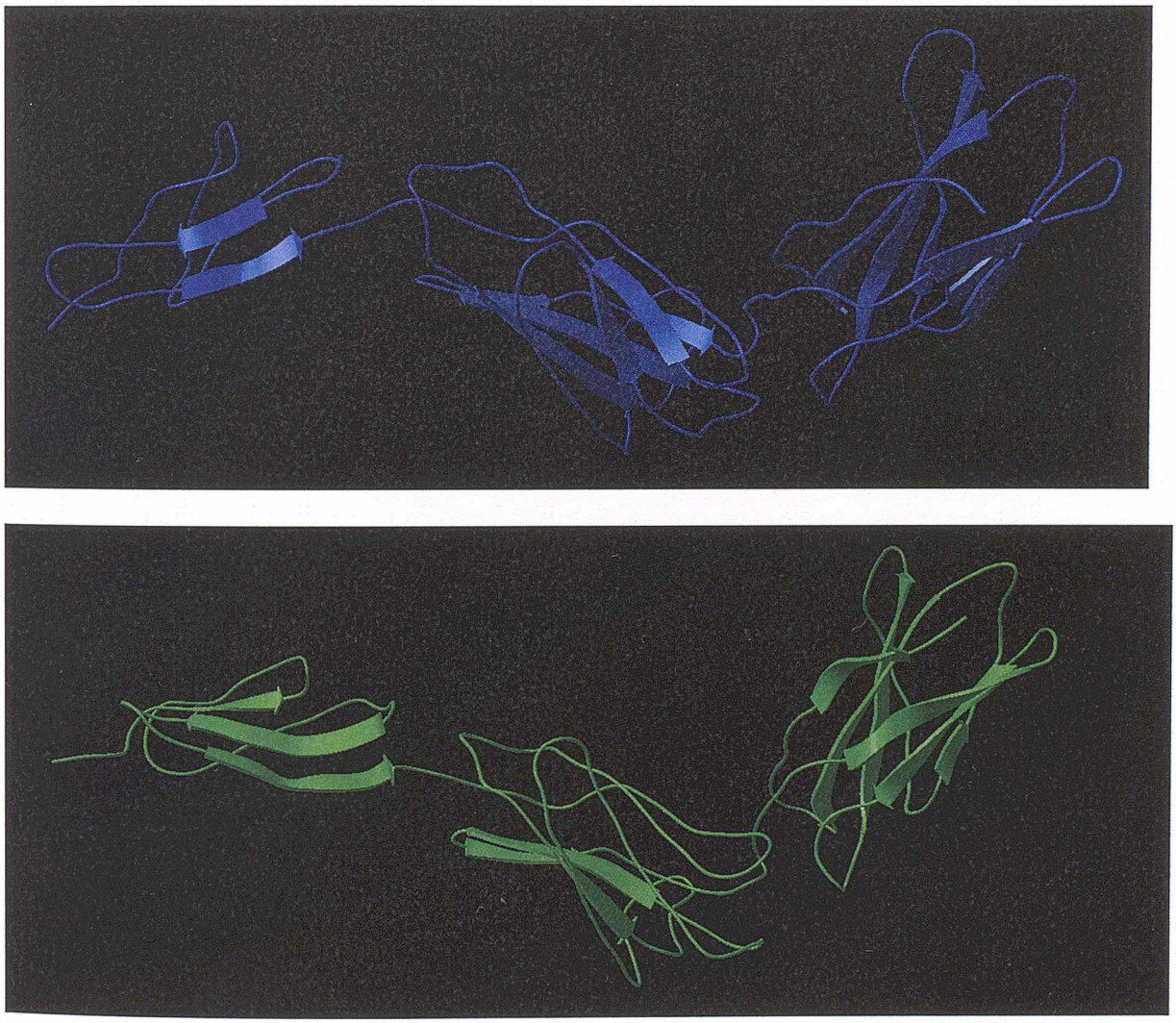


Figure 3.2.10: CHR1Q (Blue) CHR1R (Green) predicted protein structure as a ribbon model showing three beta sheet formations.

The DeepView software was used to align CRH1R to the native structure, CRH1Q. They showed very similar structures however there are some slight differences between them, as the alignment did not show a perfect fit (Figure 3.2.11A). This was noticed when the improved fit option was used, as it aligned the two protein structures together based on the second and third beta sheet formation (Figure 3.2.11B). This showed that these two beta pleated sheet structures were highly similar between the two predicted structures of the CHR1 domain, with the pivot point between the first and second beta sheet being poorly defined.

The smallest of the three beta pleated sheet structures in the models of CRH1Q and CRH1R was also aligned and showed a high degree of similarity (Figure 3.2.12). However, again there were slight unexpected differences although these can probably be attributed to the prediction algorithm. When the two predicted secondary structures were compared between the CRH1Q and CRH1R peptides there were seven differences between them, where either a beta sheet was predicted or a random coil (Figure 3.2.13).

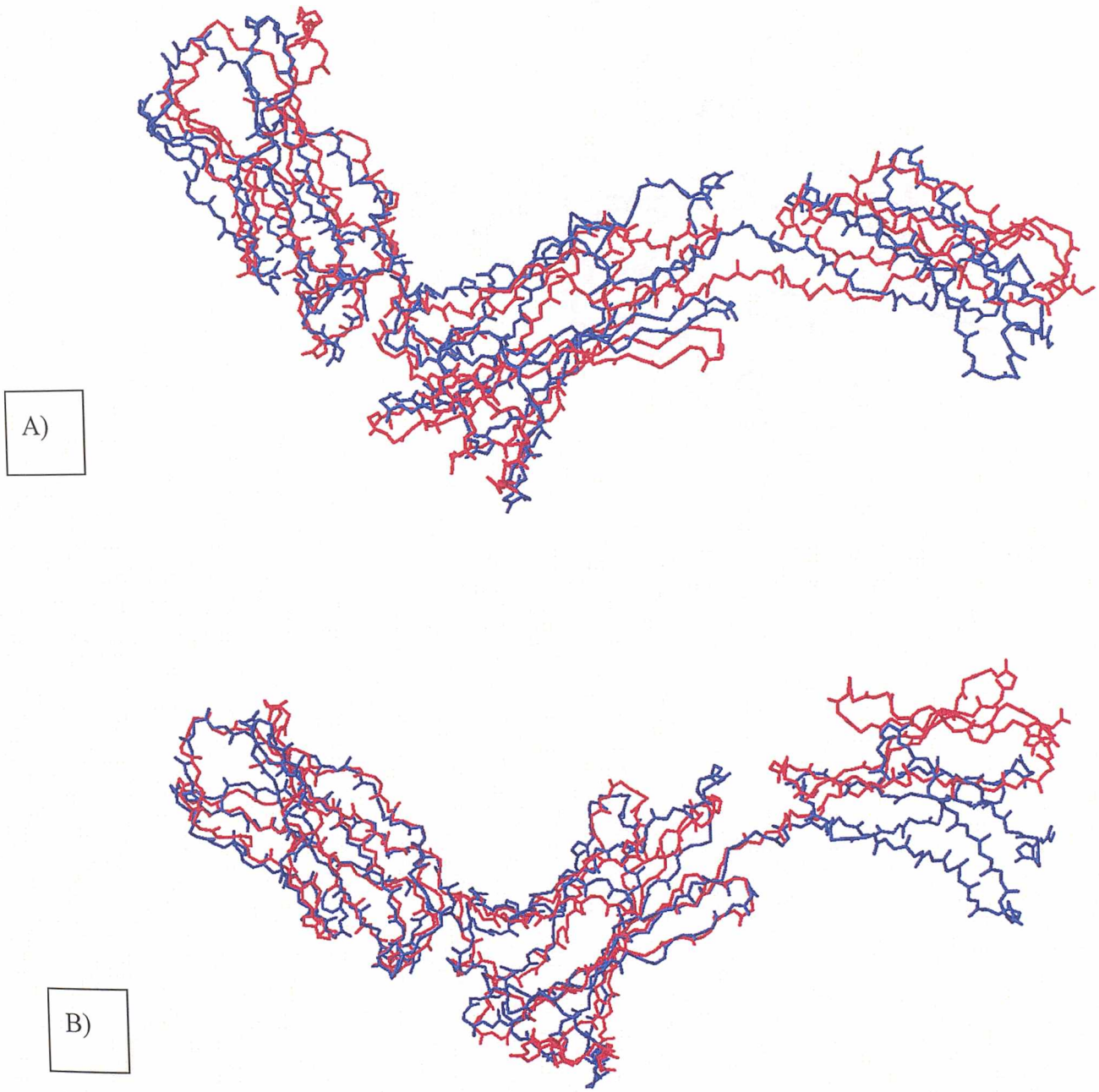


Figure 3.2.11: Alignment of CRH1Q (blue) and CRH1R (red) as a ball and stick model (A), improved alignment of the two larger beta sheet structures within the domain (B). CRH1Q was used as the reference strand.

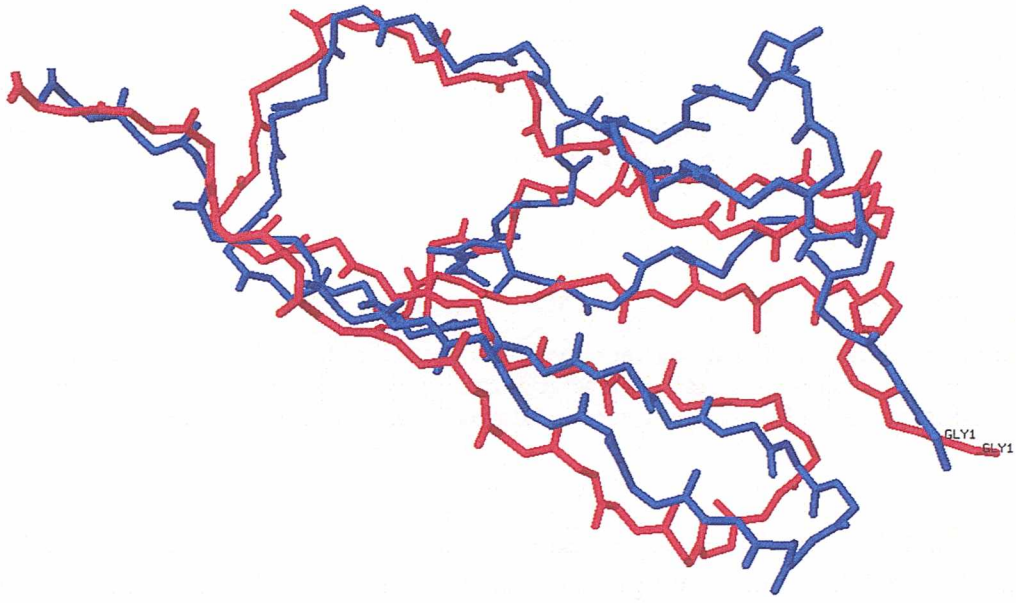


Figure 3.2.12: A close up view of the alignment between the smaller beta sheet structure (amino acids 1-60) in CRH1Q (blue) and CRH1R (red), shown as a ball and stick model. The amino acid Glycine is shown at position one.

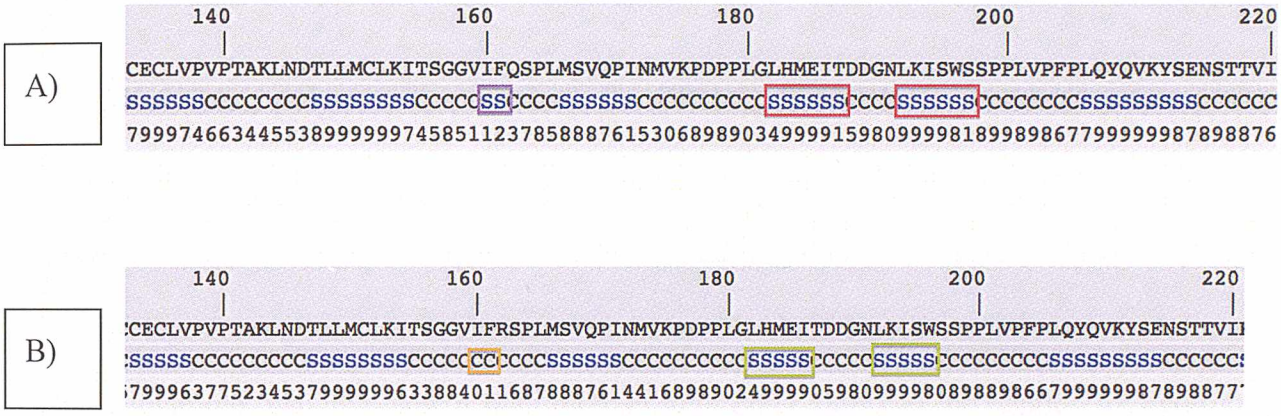


Figure 3.2.13: A section of the predicted secondary structure of the CRH1 domain with the glutamine residue at 223 (A) (CRH1Q) and the arginine residue (B) (CRH1R), from amino acid positions 133 to 220 (corresponding to 112 to 159 in the complete long isoform of the leptin receptor). The top row displays the amino acid sequence, the middle row indicates the secondary structure that has been predicted, either a random coil (C) or a beta-pleated sheet (S). The bottom row indicates the confidence score of the prediction. Differences between the two proteins can be seen at position 160 as two amino acids have been predicted to form a beta pleated sheet in CRH1Q (highlighted in purple) while the same amino acids are predicted to form a random coil in CRH1R (highlighted in orange). The length of the beta-pleated sheet *also* varies between the two predicted proteins between amino acids 180 and 200, with five in CRH1Q peptide (highlighted in red) and six in CRH1R peptide (highlighted in green).

3.2.4.2 Hydrogen Bonding in CRH1Q and CRH1R

The Gln223Arg polymorphism, in both the CRH1Q and CRH1R peptide structures, was investigatedby aligning the two models of the peptides and viewing the amino acid R groups to highlight possible steric differences that might occur between the two peptides. The change from glutamine to arginine results in side chain length. Arginine is much longer than glutamine, which means it may interact with other atoms in the peptide by hydrogen bonding. To determine whether there were alterations in hydrogen bonding patterns associated with the change in amino acid, the hydrogen bonding option was selected. This was carried outindividuallyfor each peptide: CRH1Q and CRH1R.

The results showed that the peptide containing the glutamine residue showed no hydrogen bonding with any other atom within the peptide. However, the peptide with the arginine residue forms a hydrogen bond with the R group of the cysteine residue at position 151, which equates to position 213 in the CRH1 domain within the complete LEPR sequence (Figure 3.2.14). The distance between the side chains of the amino acids at positions 162 and 151 in CRH1Q was measured using the Deepview software and found to be 6.63Å. The same distance was measured in CRH1R and found to be much shorter, 3.11Å, which indicates that hydrogen bonds are more likely between the two amino acid residues in CRH1R than in the native CRH1 domain, as the gap is so much shorter. Hydrogen bonds are typically between 2.6 and 3.5Å.

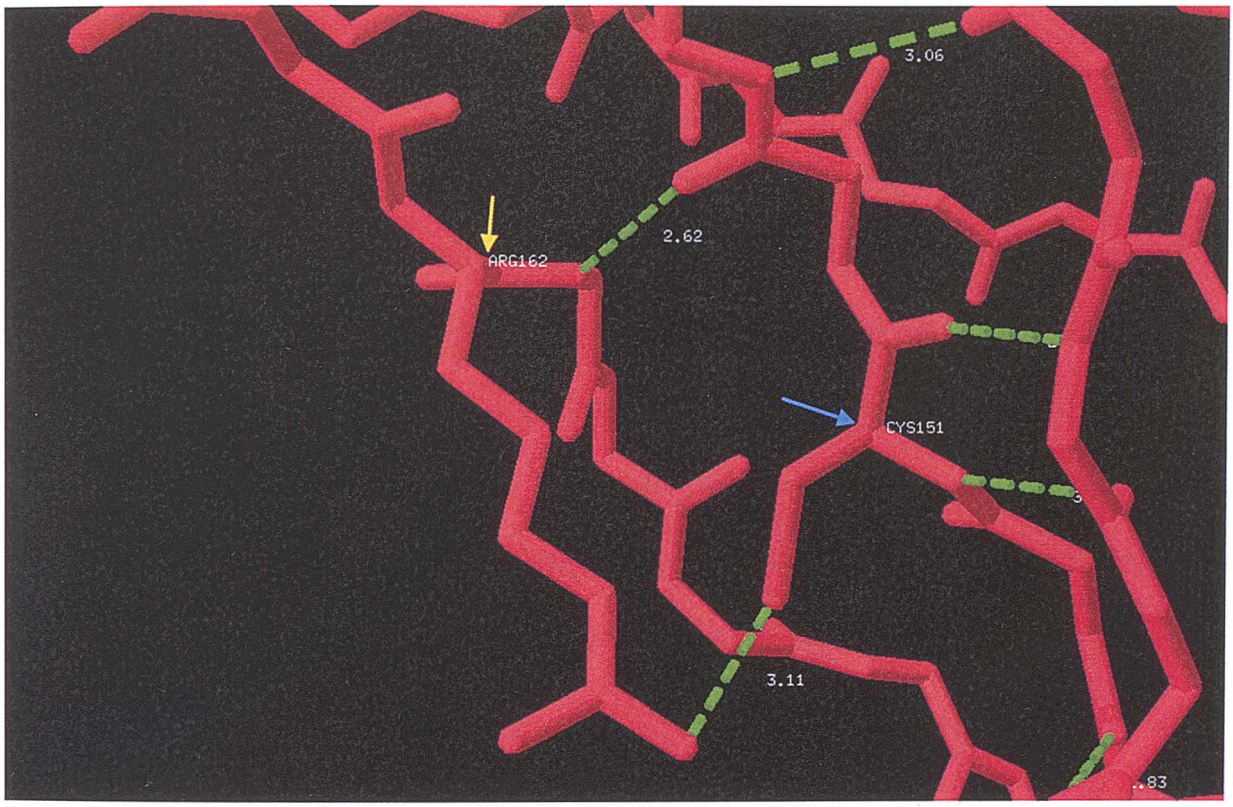


Figure 3.2.14: A close up of the hydrogen bond between the arginine residue at position 162 (highlighted in yellow) and the cysteine residue at position 151 (highlighted in blue), within CHR1R, as a ball and stick model.

3.2.4.3 Electrostatic potential in CRH1Q and CRH1R

Both the CRH1Q and the CRH1R peptide were viewed using the Python Molecular Viewer, which is used to determineelectrostatic potential, which is mapped to the surface of the two domains. Thisis then used to establish any differences or similarities. Electrostatic potential is based on the electron density of particles within a molecule and how these interact withcharged particles. Areas of negative electrostatic potential interact with a positive charge as an attractive force, which increases the negative electrostatic potential whereas positive electrostatic potential interacts as a repulsive force, which increases the negative electrostatic potential.

The electrostatic potential of the query peptides was measured using a macromolecular electrostatistics calculation program called the Adaptive Poisson-Boltzmann Solver (APBS) that the Python Molecular Viewer displays as electrostatic potential molecular surface (Figure 3.2.14). This is colour coded: areas of negative electrostatic potential are shown in red and areas of positive electrostatic potential are shown in blue. The intensity of the colour indicates areas of either strong negative (dark red) or positive (dark blue) electrostatic potential. Peptide regions withdifferences in negative and positive electrostatic potential, i.e. differences in red and blue colouring indicates more polarity if the surface of the protein is mainly white or lighter shades of blue and less polarity if there is morered. The electrostatic potential was determined for both CRH1Q and CRH1R individually.

The CRH1Q peptide was found to have both electronegative and electropositive regions. Some areas have strong electronegative potential, which equates to -11.3kT/e, whereas other areas show strong electropositive potential, with a high potential value of up to 11.3

kT/e (Figure 3.2.15). However, most of the CRH1Q peptide showed only low electronegative or electropositive potential, between 5.63 and -5.63 kT/e, or zero potential (white surface). The white and lighter colour shades between the two potentials indicates that the CRH1Q molecule is mainly non-polar. The CRH1R peptide produced similar results (Figure 3.2.16), mainly areas of low electronegativity, electropositvity orzero potential throughout the molecule. Interestingly, there does appear to be slight differences between the twopeptide allelesas evidenced by the electronegativity mapping indicated by intensity of the blue/red colouring. Figure 3.2.15 (A) shows that the centralregion of the CRH1Q peptide has strong negative electronegative potential whereas Figure 3.2.16 (A) shows the CRH1R protein to have a more positive electronegative potential in the same region. Mirror images in the figure give a better view of the appropriate residues.

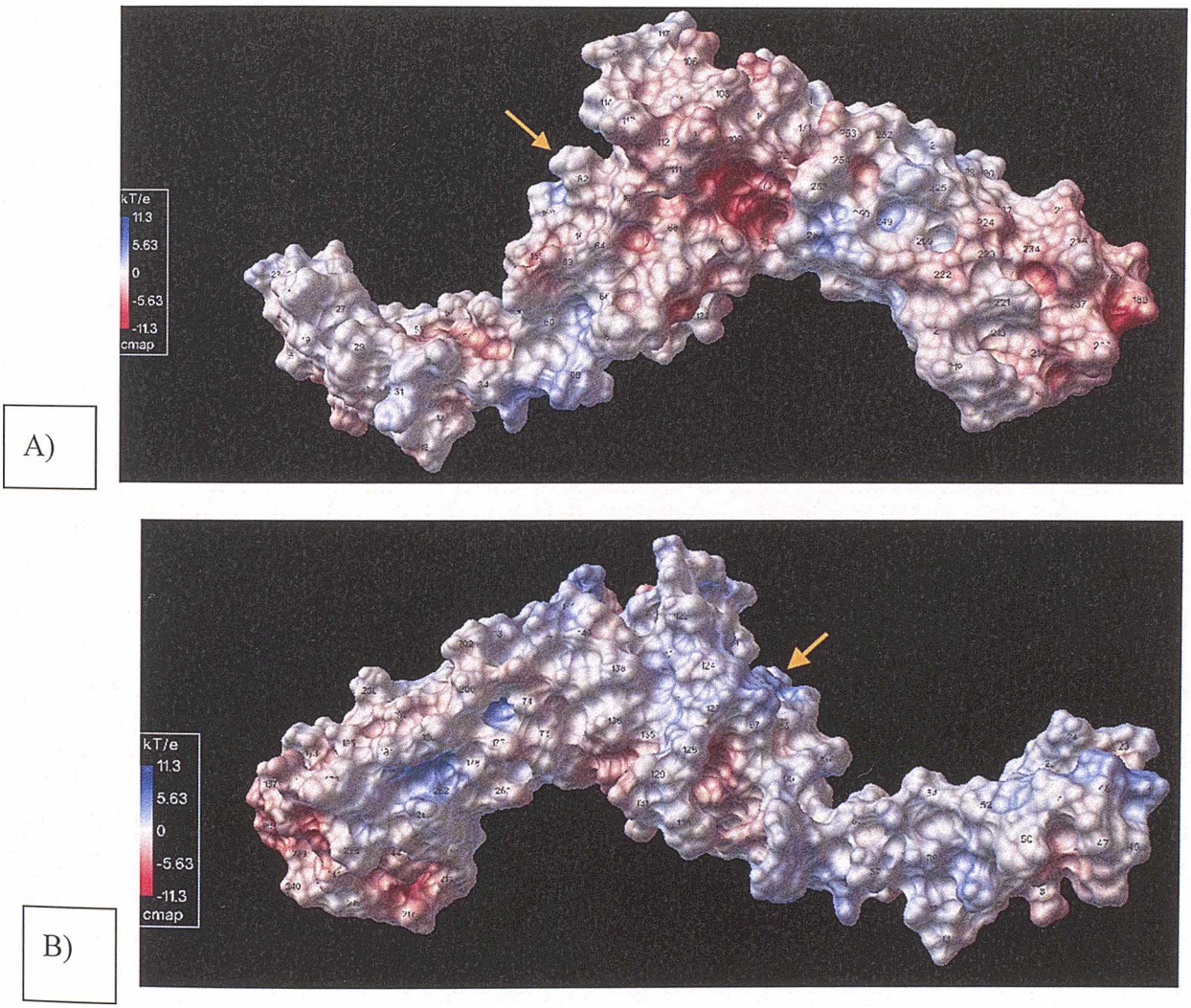


Figure 3.2.15: Electrostatic potential of the CRH1Q predicted domain, mapped to the surface of the protein. A) This view of the protein shows more areas of electronegativity than electropositive whereas B) shows the opposite side to the protein indicating more electropositive potential. The arrows indicate the position of the glutamine amino acid, which is situated at position 162.

With respect to the Gln223Arg amino acid substitution at position 162, there was a noticeable difference in the effect of the residue on the charge on the surface of the protein when the residues were substituted. The glutamine, in the CRH1Q peptide was found to be mostly electropositive (Figure 3.2.17B) with a small region of electronegativity (Figure 3.2.17A). The glutamine residue is in close proximity to amino acids at positions 153 (Lysine) and 151 (Cysteine), which are electropositive and the amino acids at positions 165 (Leucine), 164 (Proline) and 160 (Isoleucine), which confer electronegativity on the surface. Arginine in the sequence CRH1R was shown to be electropositive and Valine at 153, 164 and 159 Serine at 163 165 and Proline at 115 were also found to be confer more electropositivity on the surface rather than electronegative (Figure 3.2.18).

The change in amino acid caused a subtle change in the structure of the protein around the region of interest. The glutamine residue in CRH1Q protrudes from the surface of the protein (Figure 3.2.17A) whereas the surface of the protein is much flatter with the arginine amino acid in CRH1R (Figure 3.2.18 A and B).

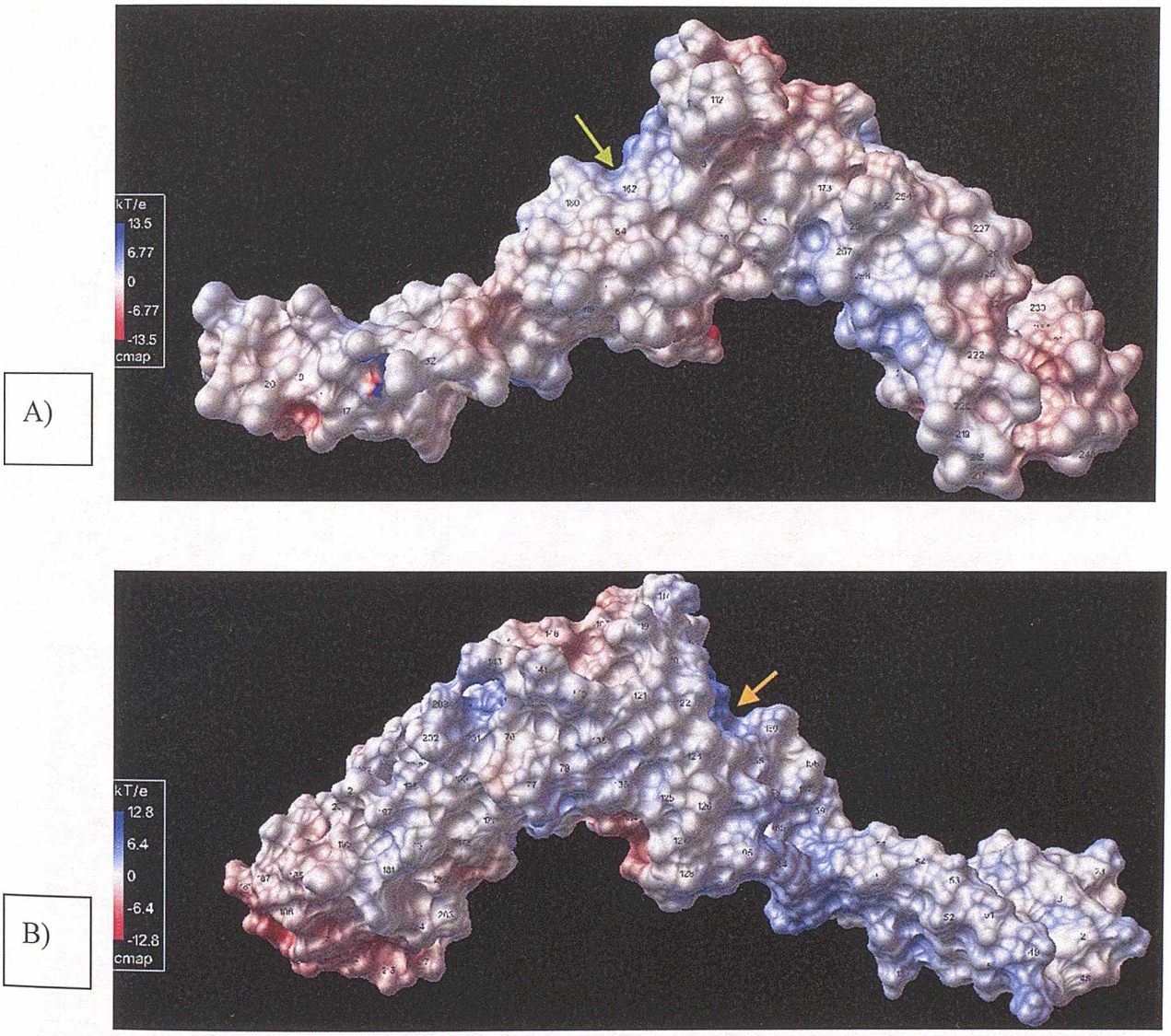
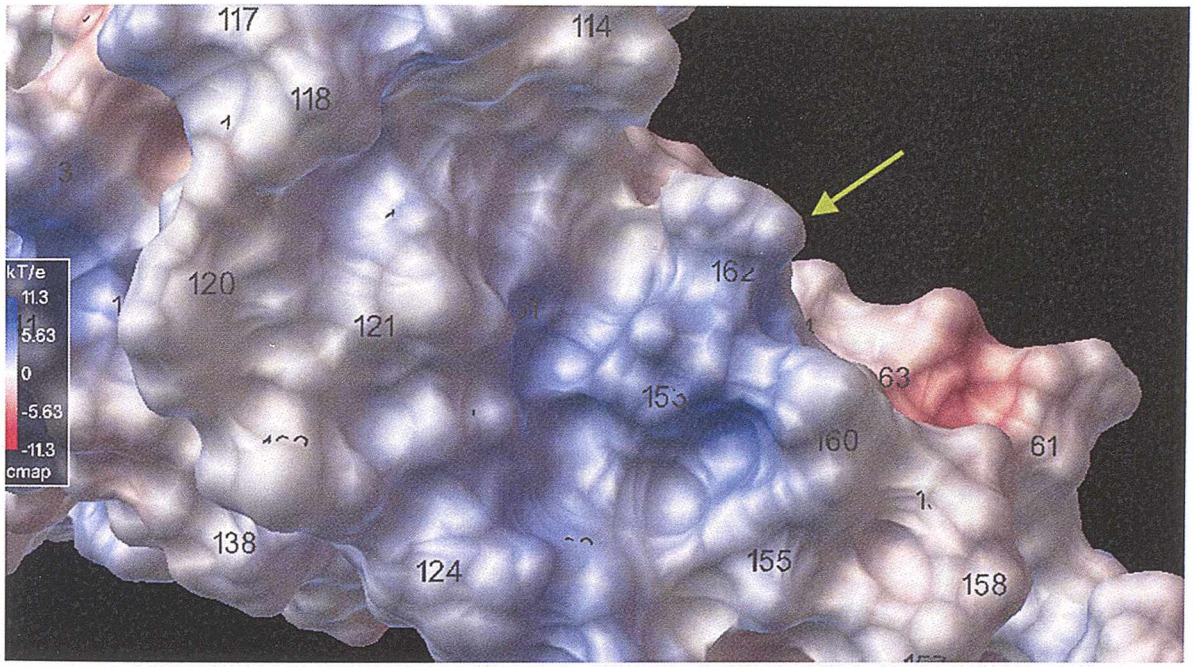


Figure 3.2.16: The overall surface electrostatic potential of the CRH1R peptide. Part A) displays one side of the protein with areas of low electronegative and low electropositive potential and white areas of zero potential. Part B) displays the opposite side of the CRH1R peptide, which shows more areas of low electropositive potential. The arginine residue at position 162 is, indicated by the arrow.

A)



B)

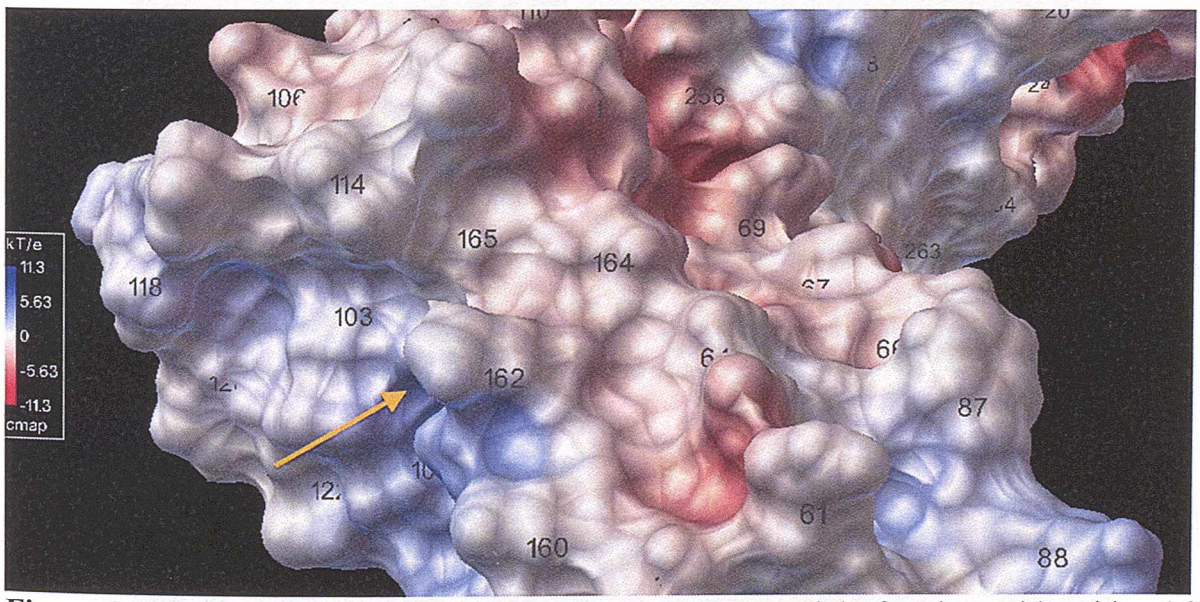


Figure 3.2.17: A close up of the surface electrostatic potential of amino acid residue 162 and surrounding residues in CRH1Q peptide, as indicated by the arrows. A) The amino acids at positions 165, 164 and 64 show a high electropositive potential whereas amino acids at positions 103, 118 and 122 are higher electropositive. B) The amino acids at positions 153, 155, 160, 121, 124 are shown to be electropositive.

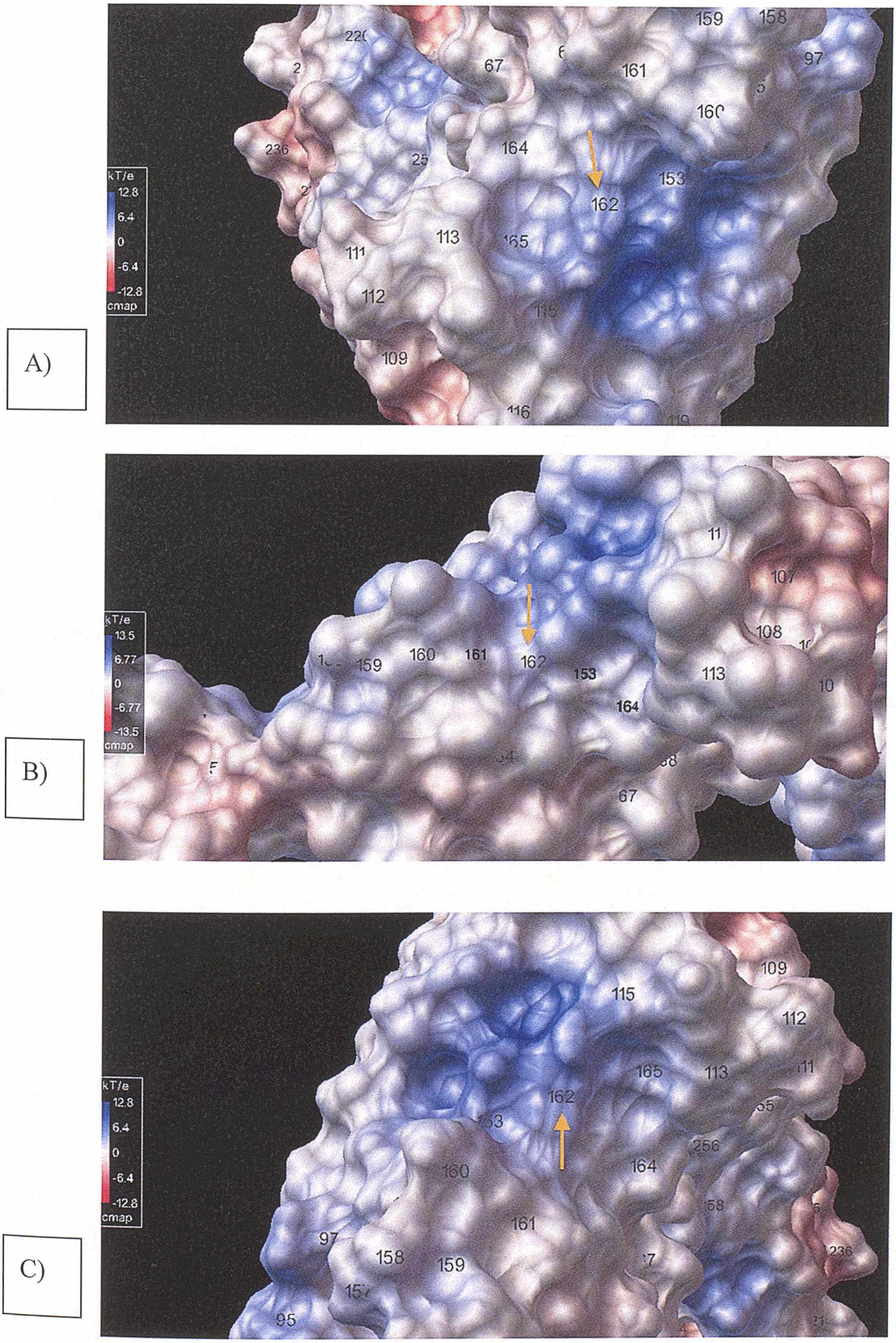


Figure 3.2.18: A close up of the electrostatic potential of the arginine residue at position 162 in the CHR1R peptide as indicated by the arrows. Both views A) and B) show arginine residue to have electropositive potential, along with the residues at positions 153, 161, 164,165 and 115.

3.3.5 CRH2 Domain I-TASSER Results

The second CRH domain (CRH2) peptide sequence was also uploaded on to the I-TASSER server for protein structure prediction. This was carried out to determine whether there were similarities with the CRH1 domain as they both contain two FNIII folds. The C-score for the protein was high, 0.35 (TM score 0.76±0.10), which indicates that there was high confidence in the prediction of its structure.

Using DeepView software the CRH2 domain showed two beta pleated sheet formations (Figure 3.2.19), of similar size. Upon alignment with the CRH1Q peptide sequence, the two larger beta sheet structures of CRH1Q aligned with the CRH2 (Figure 3.2.20) indicating that the two domains containedhomologous structures and amino acid sequences although there were areas in which the two domains have not aligned well.

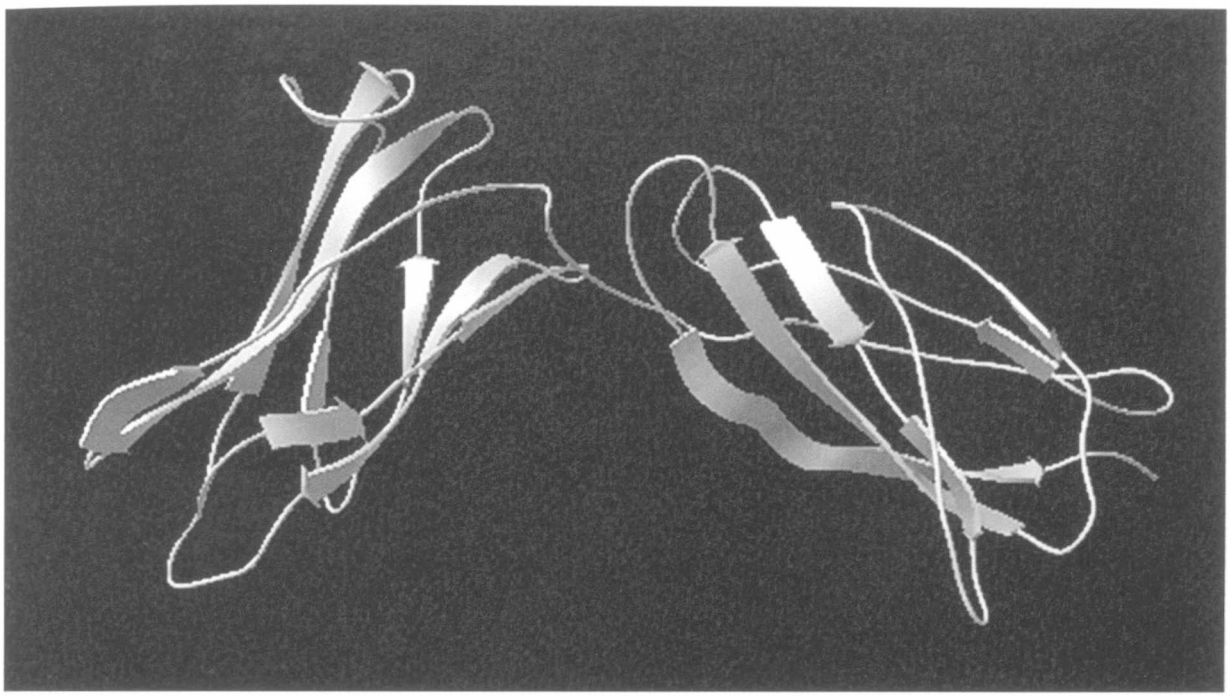


Figure 3.2.19: CRH2 domain predicted protein structure depicted as a ribbon model showing two anti-parallel beta pleated sheet structures.

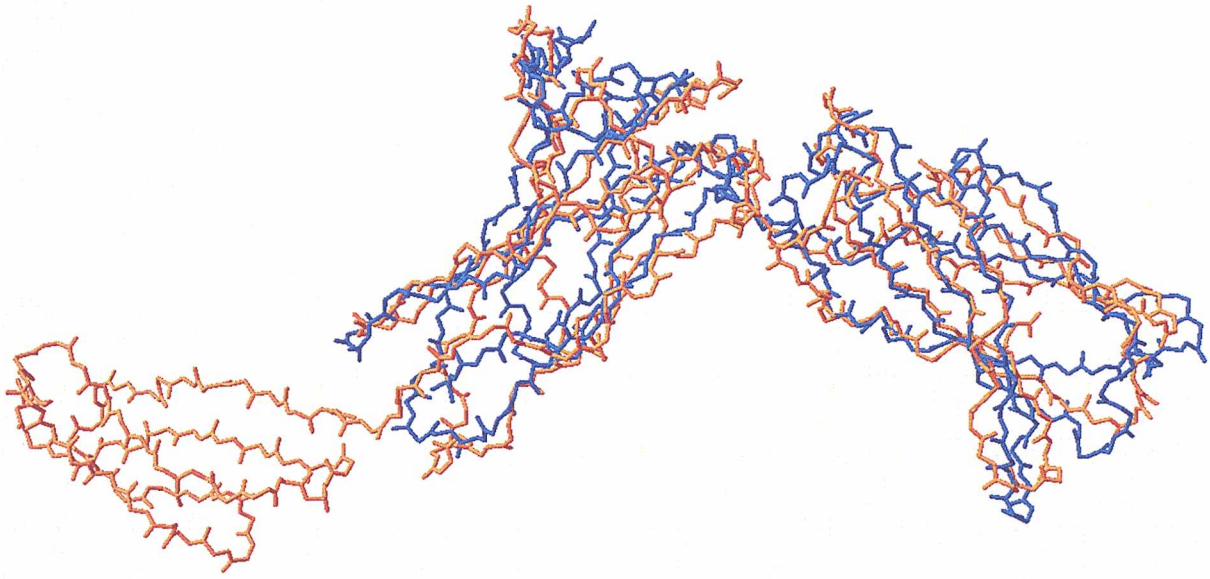


Figure 3.2.20: The alignment of the two CRH peptide domains, CRH1 (Red) and CRH2 (blue), shown as a ball and stick model. The two larger beta pleated sheet structures of CRH1 are similar to the two beta pleated sheet structures of CRH2.

3.3.5.1 Electrostatic Potential in CRH2

As with the CRH1Q and CRH1R peptide sequences, the electrostatic potential was determined in CRH2 using the Python Molecular Viewer software. The majority of the protein displayed no overall positive or negative electrostatic potential, depictedby the white colour on the model, but areas of intense red and blue colour around the central area of the model which indicated strong negative and positive areas of electrostatic potential respectively (Figure 3.2.21). These areas of strong potential were found to be -32.1 kT/e and 32.1 k/Te. These resultsare similar to those found with the CRH1Q and CRH1R peptides (Figure 3.2.14; Figure 3.2.15) as these showed areas of positive and negative electrostatic potential over the whole surface of the domain. However there were differences in the intensity of the positive and negative electrostatic potential, along with the location of these over the surface of the peptides.

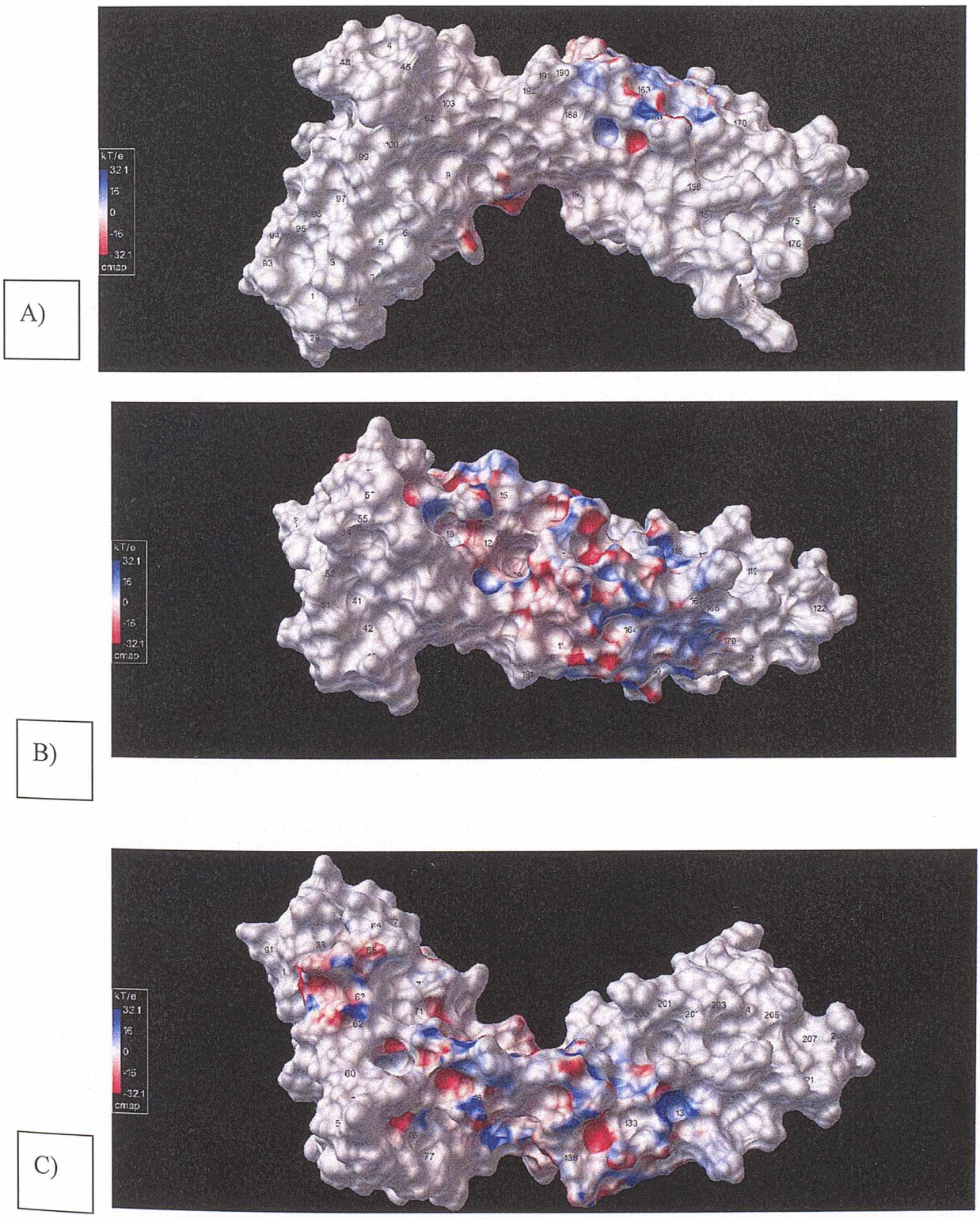


Figure 3.2.21: Electrostatic potential mapped to the surface of the CRH2 domain. Differential areas of negative, positive and neutral electrostatic potential are displayed on the surface throughout the domain structure as can be seen in A), B) and C).

Key residues within the CRH2 domain were viewed; Ile^{501}, Phe^{502}, Leu^{503}, Leu^{504}, Ser^{505} and Asp^{615} which are important for binding of leptin (Iserentant *et al.,* 2005). The amino acid $Phe^{502}$was found to result in electrostatic negative potential on the surface while the other residues conferred a neutral effect. Ile^{501} and Asp^{615} were hidden from the surface of the domain so did not affect the surface (Figure 3.2.22).

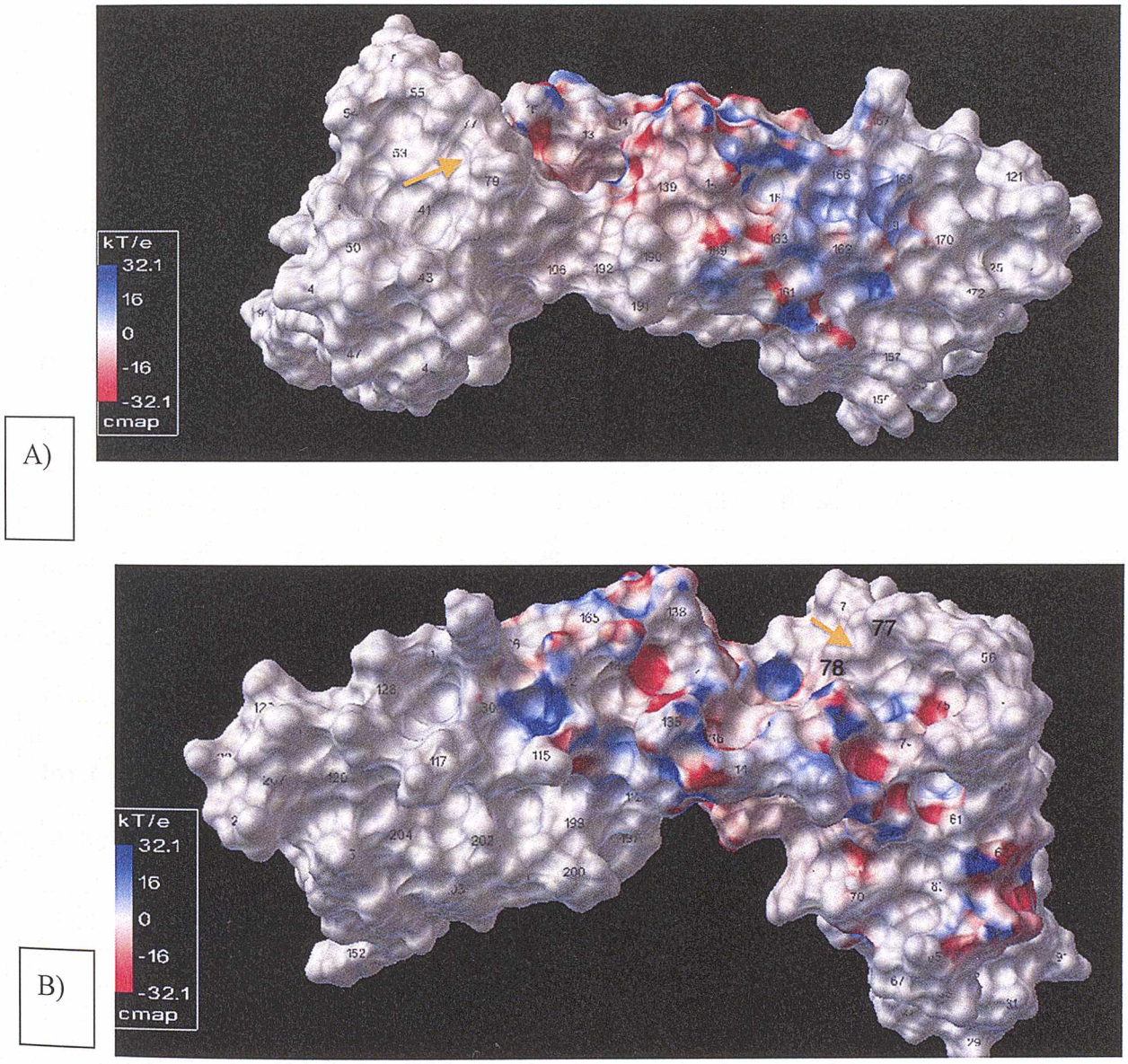


Figure 3.2.22: The overall surface electrostatic potential of CRH2 domain. A) The arrow displays the location of the residues Ile^{501} (77) and Leu^{503} (79), which did not changethe electrostatic potential on the surface. B) The arrow displays the location of the residues Ile^{501} (77) and Leu^{503} (79), where Leu^{503} has slight positive electrostatic potential.

3.3.6 Electrostatic Potential in Leptin

The known leptin structure was downloaded from the PDB and the Python Molecular Viewer software (accession number P41159) and was used to determine its surface electrostatic potential. The overall surface electrostatic potential of the protein was found to be negative and the intensity of the potential varied across the whole of the molecule (Figure 3.2.23). One side of the molecule was found to have the highest intensity (-0.45 kT/e), which seemed to decrease across the molecule to the opposite end (-0.150 kT/e). Key residues on the surface of leptin were viewed more closely as these are important in binding of its receptor. Leu^{13}, Leu^{86}, Leu^{89}, Phe^{92} are important in binding site II, while Phe^{41}, Phe^{43}, Val^{36}, Thr^{37} are important in binding site I. Leu^{13}, Leu^{86}, Leu^{89}, Phe^{92}, Phe^{41} and Phe^{43}conferred negative surface electrostatic potential while Val^{36} and Thr^{37} were hidden from view and so did not alter the surface charge. These residues are located in the region of the wild type molecule, which has the highest intensity of electrostatic potential.

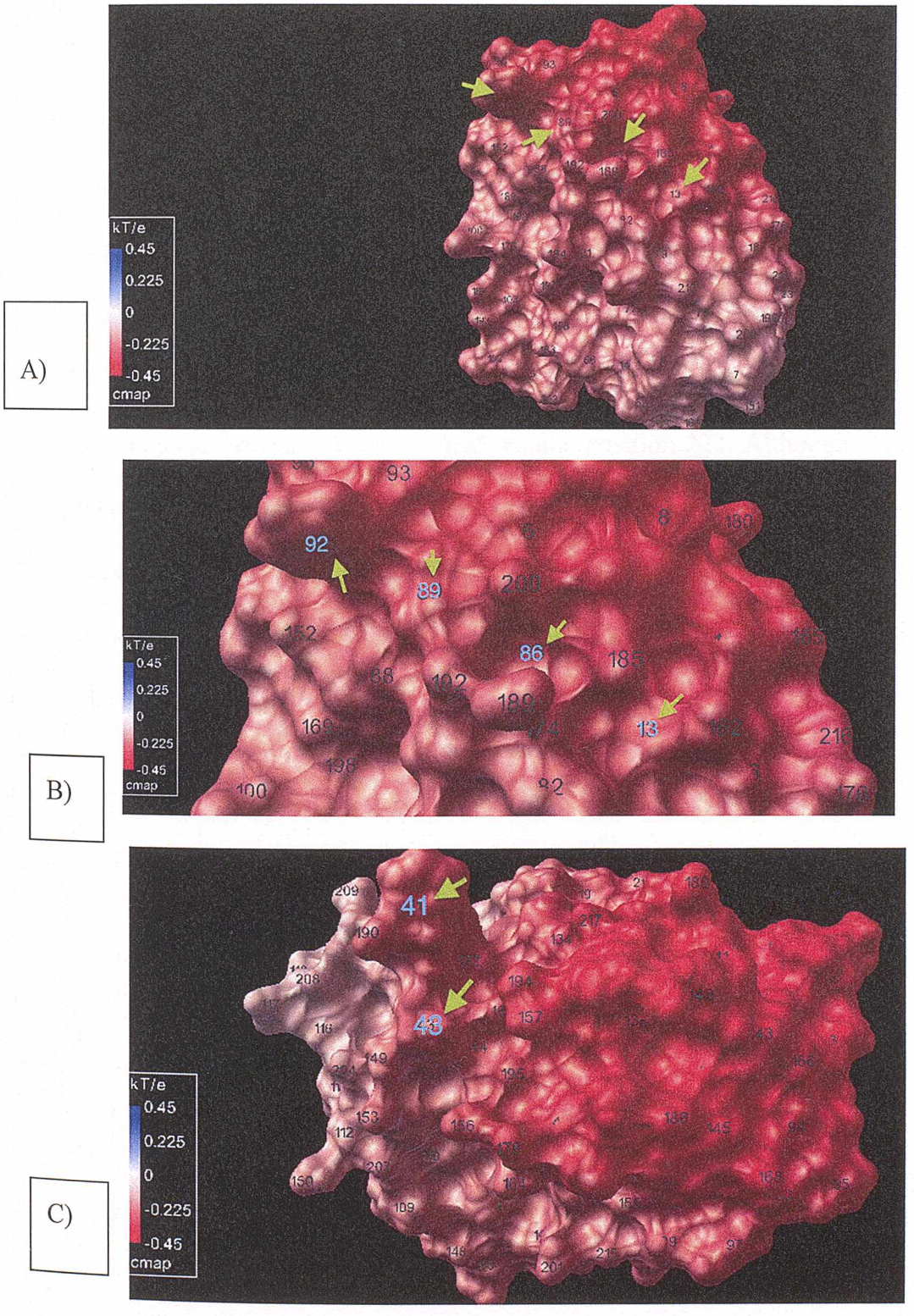


Figure 3.2.23: The positive surface electrostatic potential of leptin protein. A) The overall leptin molecule displaying the position of the four residues Leu^{13}, Leu^{86}, Leu^{89} and Phe^{92} B) Indicated are the residues, Leu^{13}, Leu^{86}, Leu^{89} and Phe^{92}conferring strong positive electrostatic potential. Part C) displays the overall leptin protein indicating the position of the residues Phe^{41} and Phe^{43} showing strong positive electrostatic potential.

In summary, the 2D and 3D protein structure of the native CRH1 peptide sequence and the same domain with the Gln^{223} replaced by Arg^{223} (CRH1R) were successfully predicted using the I-TASSER server software suite. Upon comparison of these two structures, hydrogen bonding was seen in the CRH1R 3D peptide sequence structure between the Arg^{223} residue and Cys^{212} residue, which was not present in the native structure. Differences in the overall surface electrostatic potential were also noted between the two protein structures and at amino acid residue position 223. Although the complete 2D and 3D protein structures of the leptin receptor were not successfully predicted due to their size and limitations of the software the CRH2 domain structure in the extracellular region of the leptin receptor also successfully predicted and was found to be similar in structure to the CRH1 domain.

4.0 Investigating Association of Maternal T2D and Obesity Gene SNPs with Fetal Birth Weight

The aim of this study was to genotype variants within two genes (*FTO* and *ADCY5*) and one gene region (*HHEX-IDE*) in our maternal Gujarati Indian DNA samples, to investigate if they are associated with lower or higher fetal birth weight. *ADCY5* and *HHEX-IDE* have previously been strongly associated with T2D (Grarup *et al.,* 2007; Horikawa *et al.,* 2008; Sladek *et al.,* 2007; Cauchi *et al.,* 2008; Rees *et al.,* 2011; Vasam *et al.,* 2011; Windholz *et al.,* 2011; Dupuis *et al.,* 2010), while *FTO* has previously been associated with obesity (Dina *et al.,* 2007; Frayling *et al.,* 2007; Scuteri *et al.,* 2007; Heid *et al.,* 2010; Hong and Oh, 2011) and more recently fetal genotype of variants within *ADCY5* and *HHEX-IDE* have been associated with lower fetal birth weight (Freathy *et al.,* 2009; Zhao *et al.,* 2009; Andersson *et al.,* 2010; Ryckman *et al.,* 2010), while *FTO* expression in the placenta has been associated with fetal growth (Bassols *et al.,* 2010; Sébert *et al.,* 2010). Predominately these studies have been performed in European Caucasian populations and ethnic minorities in Western countries are understudied therefore these loci were genotyped in our maternal Gujarati Indian cohort, to ascertain whether the same loci were also associated along with fetal genotype from published research.

Two different PCR based methodologies were used to genotype variants in *FTO*, *ADCY5* and *HHEX-IDE* as the number of SNPs genotyped was different for each gene. The Sequenom iPLEX assay enables up to forty SNPs in one assay to be genotyped at one time, therefore this will be used to genotype a selected number of SNPs across the *HHEX-IDE*and *FTO*gene regions, while the TaqMan Real Time genotyping assay was used to easily genotype one SNP in the *ADCY5* gene.

4.1.0 Genotyping of *HHEX-IDE* in Guarajti Indian Women using the Sequenom iPLEX assay

Both the NCBI website and published data was used to identify potential SNPs in the *HHEX-IDE* gene locus. Twenty-seven were found in total, twenty-two of which were identified from NCBI and the remaining five were selected from published data associating these variants with T2D (rs1111875, rs5015480, rs7923837, rs1544210, rs10509646) (Cauchi *et al.,* 2008, Grarup *et al.,* 2007, Horikawa *et al.,* 2008, Horikoshi *et al.,* 2007, Omori *et al.,* 2008, Pascoe *et al.,* 2007, Saxena *et al.,* 2007, Schulze *et al.,* 2007, Sladek *et al.,* 2007, Staiger *et al.,* 2008, van Vilet-Ostaptchouk, 2008, Zeggini *et al.,* 2007). Rs1418387 and rs17851141 are mis-sense mutations that occur in exon one and exon two respectively. As the *HHEX-IDE* gene region is in high LD and covers over 200 kb a selected number of SNPs were chosen to capture all the genetic variation within this region using the Tagger module in Haploview. The Haploview software enables the user to download phased genotype data from the International HapMap project (http://www.hapmap.org) to visualise LD within a particular region of the genome (Barrett *et al.,* 2005). Data was downloaded from HapMap using Haploview, by selecting positions 94427 and 94520 on chromosome 10, where the *HHEX-IDE* gene region is situated. The previously associated SNPs with T2D and birth weight are located downstream of the *HHEX* gene therefore this region was selected to be investigated within our SA cohort. From this further variants were identified and the total number of SNPs was increased to fifty-six. In this study, for QC purposes, the HWE cut-off was set at $p<0.05$, the genotype success rate cut-off was set at 80%, Mendelian errors were left as the default and the minor allele frequency was set to 5%.

Using CEU data for the *HHEX-IDE* region, fifty-six SNPs were found to be in high LD (r^2 = 71-94), with two main haplotype blocks (as shown in Figure 4.1.1). The Tagger option within the Haploview software was used to generate a set of thirty-one tag-SNPs, using a minimum minor allele frequency of 0.001 and an r^2 value of 0.80. The maximum number of SNPs captured by a single tag-SNP (rs12765131) was eight (see Table 4.1.1). However Haploview did not select SNPs within the *HHEX* gene itself so these were added when designing primers using the MassARRAY software. The thirty-seven SNPs were split into two assays, a 20-plex assay and a 17-plex assay.

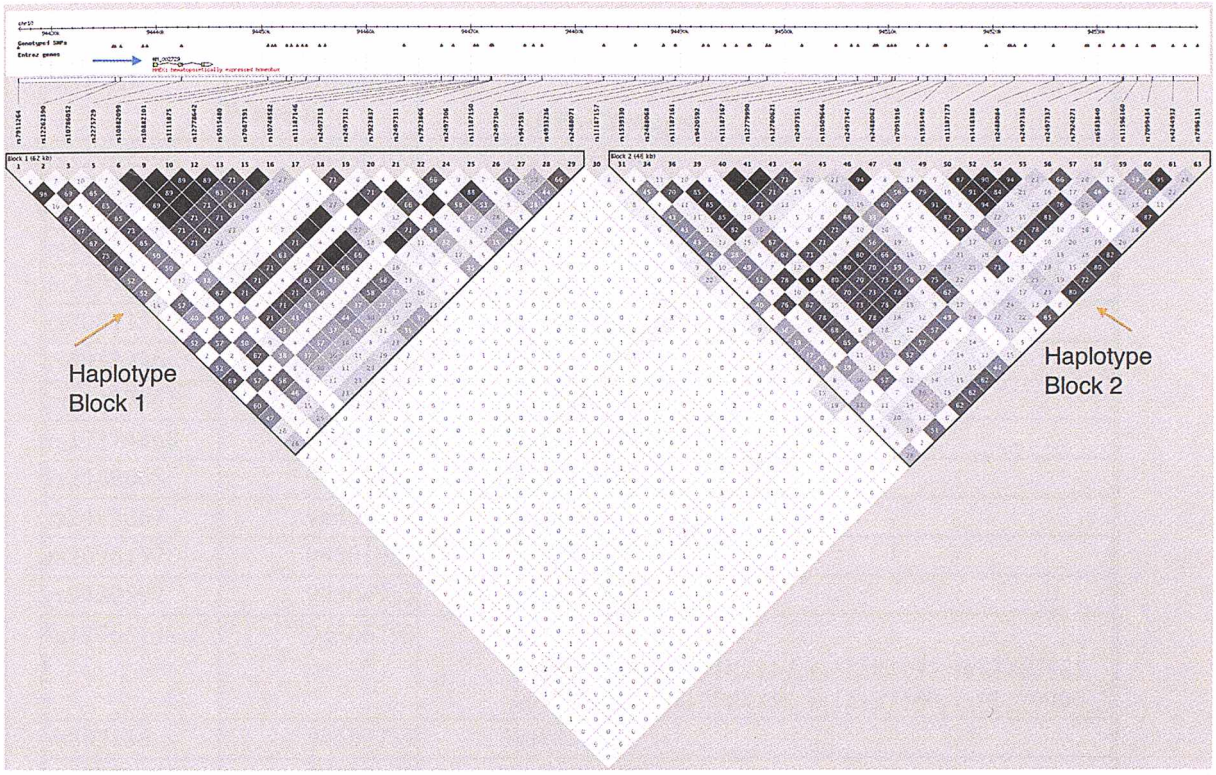


Figure 4.1.1: A map of R^2 across the *HHEX* region of interest using HapMap CEU data. The yellow arrows indicate two Haplotype blocks (high LD as represented by the black colouring), across the *HHEX* loci and the blue arrow indicates the location of the *HHEX* gene.

Table 4.2.1: The thirty-one SNPs within the *HHEX* locus genotyped, using the Sequenom iPLEX assay. The chromosome position of each of these SNPs is shown along with the assay number and the SNPs captured by each of the SNPs.

SNP ID	Chromosomal Position	Assay	SNPs Captured
rs12262390	94436103	1	rs12262390
rs10882102	94456475	1	rs5015480, rs12778642, rs1111875
rs11187146	94468335	1	rs11187146
rs2497311	94471975	1	rs2497313, rs2488079
rs7081351	94474989	1	rs7081035, rs7081294
rs4933236	94488416	1	rs4933236
rs17107841	94488811	1	rs17107841
rs17374868	94491883	1	rs17374868
rs2488068	94496608	1	rs2488068
rs2488067	94496718	1	rs2488067
rs2096177	94501663	1	rs2096177
rs9420592	94504955	1	rs9420592
rs12765131	94506793	1	rs12763786, rs11187167, rs12784232, rs12763565, rs12780621, rs12764758, rs11187169, rs12779990
rs882136	94507739	1	rs882136
rs10509646	94510649	1	rs10509646
rs2497350	94511239	1	rs2497350
rs2497349	94511545	1	rs2497349
rs2488087	94436021	2	rs10882100
rs11187144	94459960	2	rs10748582, rs7087591, rs7923837, rs7923866
rs12780253	94472115	2	rs12780253
rs2497309	94473956	2	rs2497312, rs2497307
rs11597458	94476016	2	rs11597458
rs2497304	94482696	2	rs2488075
rs947591	94485733	2	rs947591
rs2488073	94488955	2	rs2488073
rs11187152	94490091	2	rs11187152
rs11187157	94492224	2	rs11187157
rs11187158	94492409	2	rs11187158
rs1539330	94492718	2	rs1539330
rs11187161	94498323	2	rs2488066, rs2497351
rs10437472	94509969	2	rs10437472

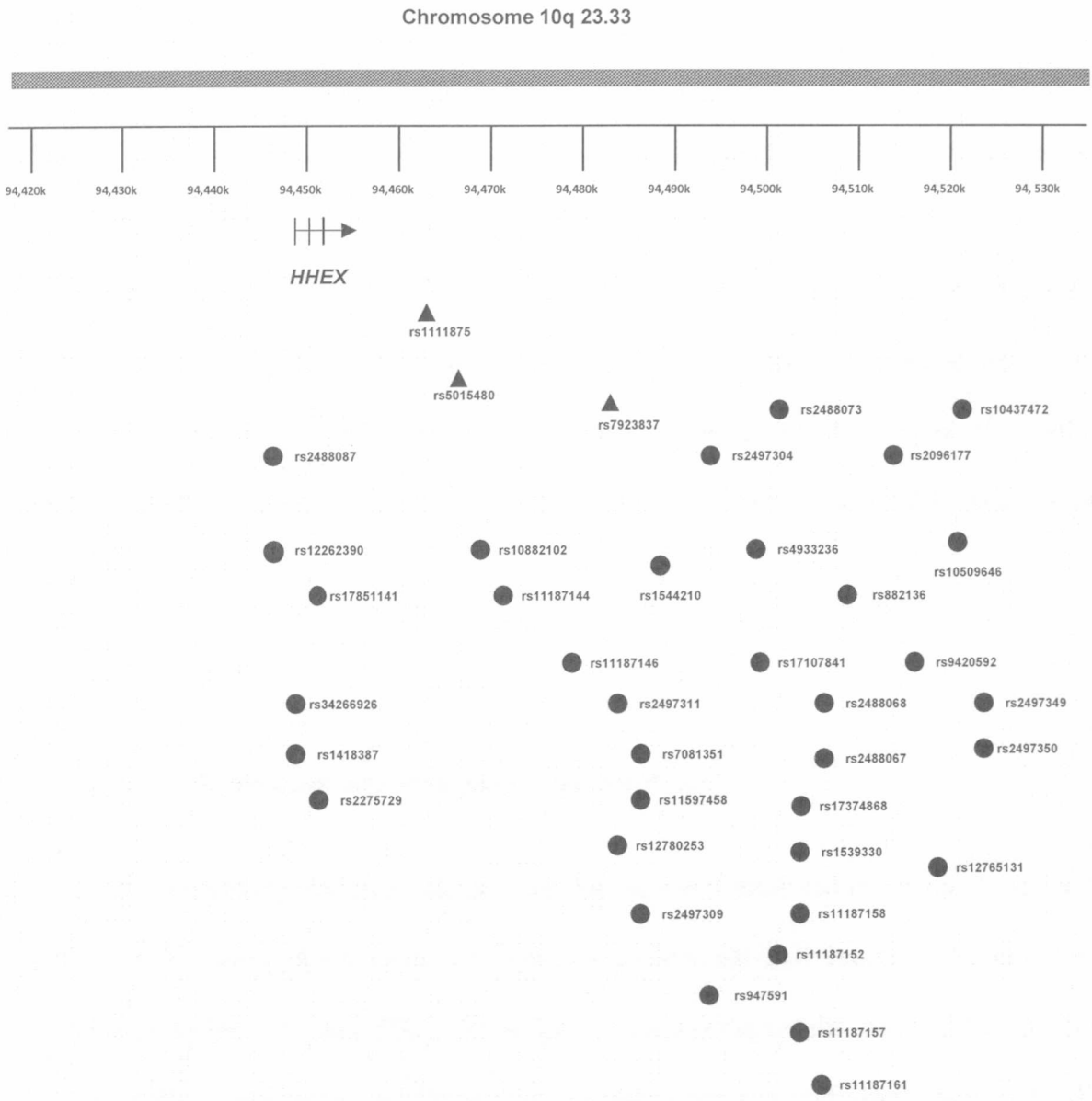


Figure 4.1.2: A graphical representation of the location of the *HHEX* gene and the position of the thirty-seven SNPs genotyped in this thesis indicated by the circle alongside the SNP ID. SNPs represented by a triangle are those that have been significantly associated with fetal phenotype in previous research (Table 1.7.11)

4.1.1 MassARRAY Software

The Sequeom MassARRAY assay design software was used to design a set of three primes for each of the thirty-seven SNPs for use in a multiplex assay (see section 2.5.4). An additional six SNPs were included into the assay, as they were not included when using Haploview; rs34266926, rs2229328, rs1418387, rs17851141, rs2275729 which are located across the *HHEX* gene, and rs1544210 which have been previously associated with T2D (Sladek *et al.,* 2007; Saxena *et al.,* 2007; Scott *et al.,* 2007; Zeggini *et al.,* 2007). The position of the thirty-seven SNPs to be genotyped in relation to the *HHEX* gene is shown in Figure 4.1.2.

4.1.2 Typer Software for Genotyping Data Analysis

Once the genotyping had been carried out, the data was analysed using the Typer software provided with the Sequenom iPLEX Assay. Post MALDI-TOF reaction, the samples were given a genotype for every SNP. The software presents the results in five different views as a: spectrum, a histogram, a cluster plot, overall plate run and single well data (Figure 4.1.4). As well as each genotype, the software indicates whether it is conservative, moderate, low probability, aggressive, bad spectrum, no alleles or user call, to distinguish between those that are good and bad quality genotypes. Those that are of good quality are given conservative and moderate status or those that have been changed by the user, which is then denoted as a user call. However those that are of poor quality are denoted aggressive or low probability depending on quality of the spectrum. If a genotype cannot be established they are described as no alleles. Using the histogram, and the yield versus skew and log height cluster plots the success rates of each SNP and the frequency of each

genotype was checked at a glance. Careful user calling was carried out to ensure that as many of the samples as possible had a successful genotype recorded. With the cluster plots the three genotypes were situated in distinct groups and labelled as blue, yellow or green. Those that had an overall yield below 50% or where the genotype clusters were indistinct were changed to no calls. In addition the cluster plots were checked so that three distinct clusters were shown representing the three genotypes.

For those SNPs where $p<0.05$ the cluster plots were rechecked to determine if there was an error in the calling of the genotypes made by the software, as indicated by more than three distinct genotype clusters or poorly-defined boundary regions between the clusters.

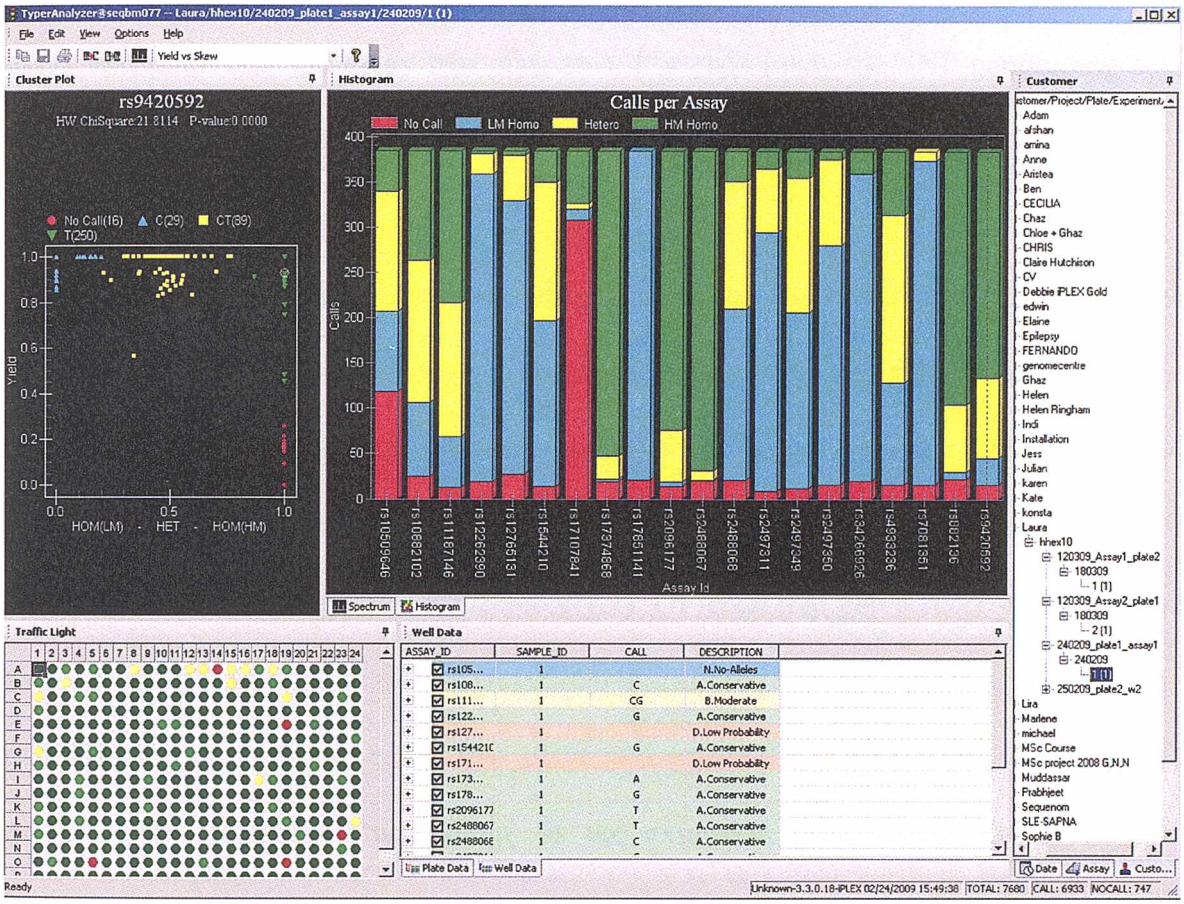


Figure 4.1.4: Screen shot of assay one plate one of the Typer software used for the genotyping analysis. The histogram in the centre depicts all the SNPs within the assay along with the total number of the genotype calls (colour coded). The top left hand corner shows the distribution of genotypes as a cluster plot for a single SNP, the bottom left displays the 384 well plate with the overall success of the individual samples and underneath the histogram is listed all the genotypes and the quality of the call for each SNP for an individual sample that has been highlighted in the plate display. Key: Blue = low mass homozygote, Yellow = heterozygote, Green = high mass homozygote and Red = No calls where there was insufficient spectrum data to call a genotype.

4.1.3 *HHEX* Genotyping Success Rates

The thirty-seven SNPs were split into two assays: a 20-plex assay (assay one) and a 17-plex assay (assay two). The South Asian DNA samples were split over two 384-well plates therefore both assays were performed twice; assay one plate one, assay one plate two, assay two plate one and assay two plate two. For each assay the SNP genotyping success rate was calculated, with expected results of ≥90%.

Of the thirty-seven SNPs genotyped, three had a success rate below 90%, rs10509646 (42.73% and 66.67% from plates one and two, respectively) and rs17107841 (13.95% and 15.89% from plates one and two respectively) from assay one and rs11187152 (88.40% and 42.04% from plates one and two respectively) from assay two. Another five SNPs were between 80-90% (Tables 4.1.2 and 4.1.3). Of those that had a success rate ≤90% five were rerun on assay one plate one (rs10509646, rs17107841, rs11187146, rs12765131 and rs10882102) and four on assay two plate two (rs2488087, rs1539330, rs11187152 and rs11597458). Rs1539330 was rerun, as the extension primer was not added for the Sequenom reaction in the first reaction due to human error, as indicated by the low percentage success rate within Table 4.1.3.

Table 4.1.2: Percentage genotyping success rates for each of the SNPs in Assay one.

SNP ID	ASSAY 1					
	Plate 1			Plate 2		
	N° of Calls	N° of User Calls	Percentage Success	N° of Calls	N° of User Calls	Percentage Success
rs882136	319	6	96.44	348	8	92.71
rs17851141	324	0	96.14	363	0	94.53
rs2497349	317	1	94.36	343	20	89.32
rs12262390	315	0	93.18	242	13	89.06
rs10509646	144	0	42.73	163	93	66.67
rs1544210	306	8	93.47	342	15	92.97
rs34266926	317	0	94.07	363	0	94.53
rs2488068	313	5	94.36	346	9	90.10
rs2497311	312	2	92.58	359	8	95.57
rs17107841	47	0	13.95	61	6	15.89
rs2497350	321	4	95.25	350	12	91.15
rs9420592	311	1	92.28	352	8	93.75
rs2488067	323	0	95.85	355	1	92.45
rs11187146	302	0	89.61	296	9	79.43
rs7081351	321	0	95.25	366	1	95.31
rs4933236	305	12	94.07	340	17	92.97
rs17374868	319	0	94.66	358	6	93.23
rs12765131	312	4	93.77	309	29	80.47
rs10882102	305	5	92.00	335	15	87.20
rs2096177	320	0	94.96	361	7	94.01

After the five SNPs were rerun, the percentage success rates were checked again and there was an improvement in the number of samples genotyped. For assay one, three of the SNPs had ≥90% genotyped, one SNP with 81.43% and one with just 9.43%. For assay two all of the four SNPs were successfully genotyped above 90% (Table 4.1.4).

Table 4.1.3: Percentage success rates for each SNP in assay two.

	ASSAY 2					
	Plate 1			Plate 2		
SNP ID	N° of Calls	N° of User Calls	Percentage Success	N° of Calls	N° of User Calls	Percentage Success
rs2229328	358	0	98.90	304	1	97.13
rs2497304	345	9	97.79	306	5	99.04
rs2488087	335	15	96.69	239	14	80.57
rs2488073	337	8	95.03	288	17	97.13
rs1418387	355	0	98.07	307	0	97.77
rs2497309	348	9	98.62	290	12	96.18
rs11187144	330	15	95.30	249	23	86.62
rs1539330	351	4	98.07	0	0	0
rs2275729	317	2	88.12	301	4	97.13
rs11187161	351	4	96.96	309	2	99.04
rs10437472	356	0	98.34	287	0	91.40
rs11187158	340	1	93.92	275	23	94.90
rs11187152	320	2	88.40	132	0	42.04
rs11597458	351	2	97.51	175	64	76.11
rs947591	315	12	90.33	279	17	94.27
rs12780253	354	1	98.07	284	0	90.45
rs11187157	325	23	96.13	236	28	84.08

The final sets of results indicate that SNPs rs10509646 and rs17107841 failed within the assay design. The genotyping also revealed five SNPs were actually non polymorphic within this population (from assay one rs17851141, rs7081351 and from assay two s10437472, rs2229328, rs1418387). Of the remaining thirty SNPs, twenty-two were found to be in HWE ($p>0.05$) leaving these to be analysed further (Table 4.1.5).

Table 4.1.4: Genotype success rates for the SNPs that were rerun from both assays

	ASSAY 1			ASSAY 2		
	Plate 1			Plate 2		
SNP ID	N° of Calls	N° of User Calls	Percentage Success	N° of Calls	N° of User Calls	Percentage Success
rs10509646	273	12	81.43			
rs17107841	33	0	9.43			
rs11187146	324	7	94.57			
rs12765131	312	3	90.00			
rs10882102	316	4	91.43			
rs2488087				316	10	93.41
rs1539330				329	7	96.28
rs11187152				332	4	96.28
rs11597458				333	1	95.70

Table 4.1.5: Numbers of each genotype for each SNP and the corresponding Hardy-Weinberg values where p>0.05.

	GENOTYPE			
SNP ID	11	12	22	HWE
rs2488087	188	257	111	0.175
rs2275729	422	42	1	0.423
rs10882102	213	287	136	0.299
rs11187144	165	280	134	0.470
rs12780253	612	0	6	0.903
rs1544210	332	261	60	0.170
rs2497304	391	229	40	0.188
rs947591	336	205	51	0.409
rs4933236	202	308	132	0.428
rs2488073	324	251	62	0.116
rs11187152	393	32	3	0.130
rs17374868	607	55	3	0.923
rs11187157	307	206	54	0.372
rs1539330	571	85	6	0.161
rs2488068	322	248	51	0.105
rs2488067	548	21	0	0.175
rs2096177	556	103	4	0.744
rs9420592	444	176	19	0.759
rs12765131	512	100	8	0.100
rs882136	514	118	12	0.234
rs2497350	477	164	7	0.482
rs2497349	345	250	49	0.600

4.1.4. Normalisation of the *HHEX-IDE* phenotype data

Prior to analysing, the genotype data was combined with the phenotype data and the variables of fetal birth weight, maternal booking weight and gestational age were individually checked for normal distribution using SPSS (v16) analysis software. Histograms showed maternal booking weight and gestational age to be positively skewed whilst fetal birth weight was negatively skewed, therefore they needed to be transformed using the logarithm function and the square root function, respectively, whereas fetal birth weight was transformed using the reflect and log function. Subjects that had a gestational age <36 weeks were removed and only singleton births were used in the analysis.

The mean fetal birth weight was 3143.31 g ± 490.17 (SD), median gestation was thirty-nine weeks (38-40), mean number of previous children was 1.79 ± 0.88, median maternal age was thirty-one years (28-33) and the median maternal booking weight was 59.00 kg (53-66).

4.1.5 Analysis of Maternal Genotype and Fetal Birth Weight

To investigate if the maternal genotype, within the *HHEX-IDE* locus has an effect on fetal birth weight, ANCOVA analysis was used with each SNP separately. The genotype data was recoded into the additive model format, which is based on the number of risk alleles; homozygous for the major allele was changed to zero, heterozygous was changed to one and homozygous for the minor allele was changed to two. The covariates were gestational age, parity, maternal booking weight with fetal sex and the SNP as fixed factors. Associations were considered significant for $p<0.05$ and trend associated for $0.05<p<0.1$.

ANCOVA analysis revealed one SNP significantly associated with fetal birth weight rs12765131 (p=0.002) and two SNPs were found to be trend associated rs2275729 (p=0.063) and rs1539330 (p=0.076) (Table 4.1.6). As expected, from previous research, significant associations were also found between fetal birth weight and the covariates of gestational age ($p=8.06x10^{-19}$), parity (p=0.0011), maternal booking weight ($3.18x10^{-12}$) and fixed factor fetal sex (P=0.004).

Table 4.1.6: ANCOVA analysis of the association of fetal birth weight with maternal SNP genotype, with the covariates parity, gestational age, fetal sex and maternal booking weight. P values, where they are associated are significant at $p \leq 0.05$.

SNP ID	P VALUE
rs2488087	0.408
rs2275729	0.063
rs10882102	0.131
rs11187144	0.292
rs12780253	0.761
rs1544210	0.256
rs2497304	0.202
rs947591	0.260
rs4933236	0.304
rs2488073	0.208
rs11187152	0.526
rs17374868	0.849
rs11187157	0.975
rs1539330	0.076
rs2488068	0.440
rs2488067	0.186
rs2696177	0.703
rs9420592	0.890
rs12765131	0.002
rs882136	0.591
rs2497350	0.791
rs2497349	0.845

Rs12765131 was investigated further, due to its high significance. The TT genotype was found to confer a lower birth weight, 3017 g ± 36.13 compared with 3273 g ± 46.58 and 3123 g ± 39.20 for the CT and CC genotypes respectively. The genotype frequencies for this SNP are 84.9% CC, 13.9% CT and 1.2% TT.

4.1.6 Multiple Testing Correction

Correction for multiple testing was performed as separate ANCOVA tests were carried out for each of the twenty-two SNPs. Bonferroni correction is one of the widely used methods, where the result is taken as significant when the *P*-value is $\leq \alpha/n$ (α is the original significance threshold and n is the number of separate tests performed). However as many of these SNPs are in LD, the use of Bonferroni can be overly harsh, therefore results were corrected for multiple testing using the spectral decomposition (SpD) of matrices of pairwise linkage disequilibrium between SNPs (Nyholt, 2004), this gave a threshold of p=0.003 for significance. A web interface was used to calculate the corrected significance value where a pedigree and a map file were uploaded to the online server to perform the spectral decomposition (http://gump.qimr.edu.au/general/daleN/SNPSpD/).

4.1.7 Haploview Analysis to Investigate LD across *HHEX-IDE* region in SA cohort

Using Haploview, linkage disequilibrium across the *HHEX-IDE* region was estimated using R^2 values from the experimental data (Figure 4.1.5). The results of which show two LD blocks, which lie across 75 kb. Ten SNPs are located within Haplotype block one which is 52 kb in length and eight SNPs are located within Haplotype block two, which is 14 kb in length. This correlates with the LD plot generated from HapMap CEU data (Figure 4.1.1). The rs12765131 SNP is in strong LD with six others ($r^2>0.8$), which lie across 5 kb.

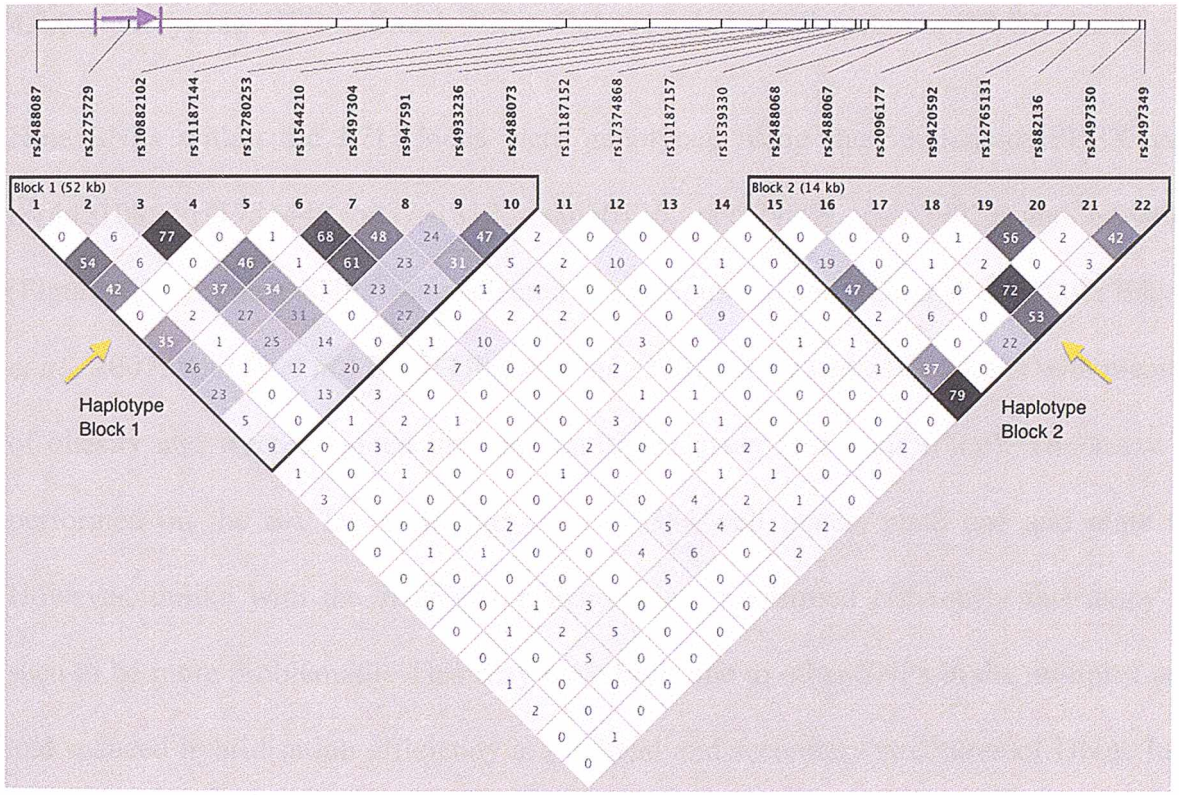


Figure 4.1.5: Map of R^2 using the twenty-two SNPs from this study. Two blocks of high LD can been seen as indicated by the yellow arrows. SNPs numbered one to ten are within Haplotype block one and SNPs numbered fifteen to twenty-two are within Haplotype block two. The location of the *HHEX* gene is indicated by the purple arrow, which the SNP rs2275729 is located within it.

4.2.0 Genotyping *FTO* in South Asian Cohort using the Sequenom iPLEX Assay

Nine SNPs within the *FTO* locus were genotyped using the Sequenom iPLEX assay: rs14421085, rs1781449, rs3751812, rs8050136, rs9939609, rs9939973 and rs9940128 (Figure 4.2.1). These SNPs were selected as previous research (Dina *et al.,* 2007; Frayling *et al.,* 2007; Scuteri *et al.,* 2007; Hinney *et al.,* 2007), indicated associations with measures of obesity and were designed into one multiplex assay (see section 2.6). The assay was performed on the SA DNA samples on two 384-well plates, plate one and plate two. However, unlike with the *HHEX-IDE* genotyping, performed previously this assay was seen to be more problematic. This may have been due to other SNPs in the muliplex assay and reduced hybridisation efficiency at the local and secondary structures of DNA. Initial run (run one), results viewed using the Typer software produced a high number of no calls for five of the nine *FTO* SNPs, either through; a bad spectrum, low probability or poor-quality aggressive genotype calling. As a result, the assay for both plates was run a second time (run two). Despite a second run, the genotyping success rate was still below 80% for most of the SNPs. After four runs, success rates were still found to vary between each of the nine SNPs (Table 4.2.1).

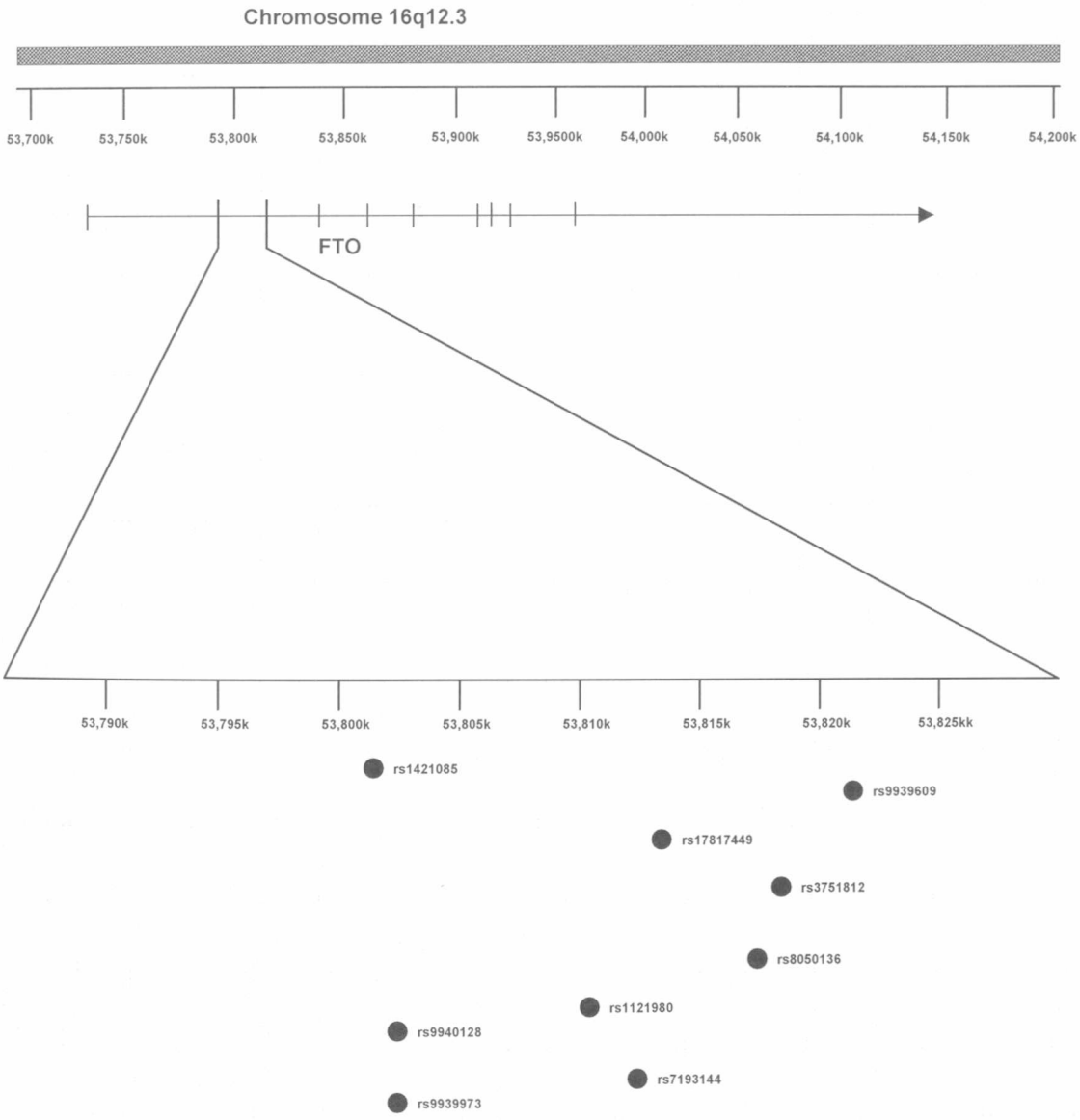


Figure 4.2.1: Graphic representation of the location of the *FTO* gene on chromosome 16q12.3 and the position of the SNPs which were genotyped.

Table 4.2.1: Genotyping combined success rates for both plates for the five runs of the FTO assay.

SNP	Run One	Run Two	Run Three	Run Four	Run Five
Rs1421085	35.94%	77.13%	77.08%	60.11%	94.09%
Rs7817449	63.38%	84.99%	71.63%	78.19%	97.81%
Rs9939609	17.56%	25.71%	4.61%	21.23%	93.42%
Rs3751812	67.26%	92.21%	83.81%	82.07%	
Rs8050136	76.02%	75.68%	99.86%	71.16%	
Rs9939973	73.43%	90.52%	92.69%	81.64%	
Rs9940128	67.62%	91.39%	87.39%	82.35%	

The SNP rs9939609 had success rate continuingly below <30%, the SNP rs1421085 varied between 35 and 78%. One SNP, rs17817449, had a success rate between 63-85% across the four runs of the assay. For one of the assays the SNPs rs3751812, rs8050136 and rs9940128 were >90% but between 67-88% for the other three assays and rs9939973 had a success rate between 73-81% for two assays and >90% for the other two assays. To improve the genotyping success rates for these SNPs, a fifth run of the assay was carried out. SNPs that had previously had a success rate of >90% (rs3751812, rs9939973, rs9940128 and rs8050136) in the previous four runs of the assay were removed. In addition the dilution of the DNA was changed to 1 in 25, from 1 in 100. When the results of both plates were viewed using the Typer software, it was seen that three of the SNPs, rs1421085, rs9939609 and rs7817449, improved, with more of the samples assigned conservative or moderate genotype calls, improving the overall success rate (Table 4.2.1). Combining the genotype data for all the samples from the five runs further improved the success rates for seven out of the nine SNPs. Two SNPs, rs1121980 and rs7193144, failed

across all five runs of the assay, with no genotypes assigned to the former, and a success rate below the 90% cut off value for the latter. The other seven SNPs were found to have an overall success rate >82% and were analysed further.

4.2.1 Combing *FTO* Genotyping Data

Before the data from each run was combined, plates one and two for each SNP, for each of the five runs were rechecked using the Typer Assay software, to ensure that the maximum number of conservative, moderate or user calls possible were achieved and that each SNP was in HWE. Each run of the assay was taken in turn, and data was sorted according to SNP and sample number. Aggressive genotypes were deleted for each sample and those samples that were given a genotype for less than 50% of the SNPs, were also deleted. A note of the sample number that was deleted was recorded. Before the genotype data for each SNP from each run was combined, it was assigned a different colour. The combined genotype data was sorted according to SNP and sample ID. The genotype for each sample was checked from all the five assay runs. Samples were deleted if they had no genotype or had two or more genotypes that were different. If all the genotypes were the same or if there was a majority genotype given the sample was kept. For example if there were three cases of one genotype and one case of another genotype it was kept and if there were two cases of one genotype and one case of another genotype but not all five runs produced a genotype for one sample. Samples that were deleted were recorded. Percentage success rate for each SNP was calculated from the samples that remained after the data was combined from the total number of samples originally genotyped (Table 4.2.2). Four out of the seven SNPs had a success rate above 90% and three SNPs were >80%. HWE was calculated (http://ihg.gsf.de/cgi-bin/hw/hwa1.pl) using Pearson's goodness-of-fit Chi-square and these were all above p=0.05, apart from one (Table 4.2.2). As seven tests were

performed the HWE cut off was changed to p=0.0071 to correct for multiple testing, therefore rs9939973 is within HWE.

Table 4.2.2: Percentage success rates for the seven *FTO* SNPs after combining the data for the five runs.

SNP	Percentage Success	Genotype 11	12	22	HWE
rs1421085	91.94	341	266	66	0.1852
rs17817449	91.94	327	278	69	0.3852
rs3751812	90.27	341	241	49	0.4843
rs8050136	88.97	293	275	53	0.3048
rs9939609	82.67	295	246	64	0.0755
rs9939973	89.88	299	293	105	0.0192
rs9940128	90.78	225	302	102	0.9687

4.2.2 Normalisation of the *FTO* Data

Before any association with birth weight was investigated, the genotype data was added to the phenotype data and the variables fetal birth weight, maternal booking weight and gestational age were checked for normality. Phenotype data was checked for normality, transformed where appropriate and removed as previously described in section 4.1.3.

4.2.3 Maternal Phenotype and Fetal Birth Weight

Analysing the maternal characteristics in more detail, the mean maternal booking BMI was 23.8 kg/m^2 (±3.84 SD), mean maternal age was 30.36 years (±4.53 SD), mean height was 158.8 cm (±10.51 SD) and the mean booking weight was 60.4 kg (±11.08 SD). The mean gestational age was 39.28 weeks (±1.44 SD), mean previous number of pregnancies was 1.78 (±0.88 SD) and the mean fetal birth weight was 3160.46 g (±480.76 SD). Just over half of the offspring were male (52.4%) with 47.6% female offspring. Male offspring had a slightly higher mean birth weight when compared to females, 3160.53 g (±488.65 SD) and

3152.93 g (±470.95 SD) respectively but they were not significantly different ($p>0.05$). This data differs from previous maternal data in the *HHEX-IDE* genotyping study as not all the same DNA samples would produce a genotype for each SNP therefore there are differences in the total number of raw data analysed.

The parity of the women, 295, first pregnancy, 244, second pregnancy, ninety, third pregnancy, twenty-three, their fourth, two, fifth and sixth was three. Even though fewer women were found to be having more than one child, the mean birth weight tended to increase (Table 4.2.3), although there was a decrease in birth weight for a third and fifth pregnancy.

The majority of subjects (423) within this study had a booking BMI within the normal range of 18.00-24.99 kg/m^2. Twenty-three were classified as underweight (<18.00), 160 as overweight (25.00-29.99 kg/m^2), thirty-six as obese class I (30.00-34.99 kg/m^2) and seven being classified as obese class II (35.00-39.99 kg/m^2). It was found that as maternal booking BMI increased from underweight to obese, the mean fetal birth weight also increased (Table 4.2.4). The mean fetal birth weight was highest in women who were classified as obese (Class II).

Table 4.2.3: Mean fetal birth weight classified by pregnancy number.

Parity	Mean Birth Weight (95% CI)
1	3076.19 (3022.71-3129.66)
2	3221.35 (3164.36-3278.34)
3	3182.57 (3067.64-3298.09)
4	3474.35 (3276.84-3671.86)
5	3070.00 (1926.44-4213.56)
6	3540.00 (2513.36-4566.64)

Table 4.2.4: Mean fetal birth weight classified by maternal BMI.

Maternal BMI	Mean Birth Weight (95% CI)
Underweight	2903.70 (2754.77-3052.62)
Normal	3105.87 (3062.23-3149.52)
Overweight	3281.91 (3201.89-3361.93)
Obese Class I	3299.44 (3122.89-3475.99)
Obese Class II	3612.86 (3182.47-4043.24)

4.2.4 Linear regression analysis of maternal genotype at the *FTO* locus with fetal birth weight

Genotype data was recoded into additive model format, which is based on the number of risk alleles; homozygous for the major allele was changed to zero, heterozygous changed to one and homozygous for the minor allele was changed to two.

Data was analysed to find the mean birth weight for each of the seven SNPs and the number of risk alleles (Table 4.2.5) and box plots were created, to view this graphically

(Figures 4.2.1, 4.2.2, 4.2.3 and 4.2.4). Five SNPs showed that as the number of risk alleles increased the mean birth weight also increased; rs17817449, rs8050136, rs9939609, rs1421085 and rs9940128, with the biggest increase occurring from one to two risk alleles. Three of these five SNPs were found to be significantly different using ANOVA analysis: rs17817449 (p=0.025), rs8050136 (p=0.004) and rs9939609 (p=0.002). LSD post-hoc tests revealed that the significant associations were between one and two risk alleles (p=0.011, p=0.002 and p=0.006, respectively) and zero and two risk alleles (p=0.009, p=0.003 and p=0.001). There was no significant different between zero and one risk allele (p=0.928, p=0.219 and p=0.272). The SNPs rs1421085 and rs99401085 were not found to be significantly different (p=0.718 and p=0.636).

Linear regression analysis was performed, using the additive model, to determine whether there were any significant associations between number of risk alleles for each SNP and fetal birth weight, with parity, sex, gestational age, maternal booking BMI as covariates. All the covariates apart from sex showed significant associations with birth weight ($p<0.05$). Three of the seven SNPs were found to be significantly associated with high birth weight: rs17817449 (p=0.0273), rs8050136 (p=0.0231) and rs9939609 (p=0.001) (Table 4.2.5). Each addition of the risk allele for the significant SNPs increased fetal birth weight by 59.31 g (±26.81 g SD), 63.80 g (±28.01 g SD) and 90.16 g (±27.35 g SD) respectively (Table 4.2.6).

Correction for multiple testing was performed as separate linear regression analyses were carried out for each of the seven SNPs. This was performed as with *HHEX-IDE* genotyping, using the spectral decomposition (SpD) of matrices of pairwise linkage disequilibrium between SNPs (Nyholt, 2004). The significance threshold of p=0.01 was

given to keep Type I errors at 5%. Only one SNP survived correction for multiple testing, rs9939609.

Table 4.2.5: Mean birth weights for the seven *FTO* SNPs and the number of risk alleles for each of them and p values are given for those that were found to be significant ($p<0.05$). Mean birth weight is given in grams with 95% CI. The genotype for each of the SNPs is represented by the number of risk alleles of the minor allele.

SNP	Genotype			P value
	1	2	3	
	Mean Birth Weight			
rs17817449	3136.10 (3081.54-3190.66)	3151.93 (3088.40-3215.47)	3326.72 (3200.87-3452.58)	0.025
rs3751812	3151.10 (3105.07-3209.13)	3143.29 (3073.81-3212.78)	3233.72 (3109.06-3358.38)	0.248
rs8050136	3132.99 (3076.90-3189.08)	3171.17 (3110.23-3232.10)	3299.56 (3150.21-3448.90)	0.004
rs9939609	3114.90 (3058.77-3171.15)	3178.13 (3111.68-3244.58)	3372.04 (3247.17-3496.90)	0.002
rs9939973	3147.06 (3084.49-3209.63)	3191.94 (3129.95-3253.94)	3162.55 (3068.53-3256.57)	0.767
rs9940128	3145.45 (3082.46-3208.45)	3182.48 (3120.90-3244.06)	3186.09 (3093.75-3278.42)	0.636
rs1421085	3153.16 (3098.74-3207.58)	3163.34 (3097.64-3229.05)	3202.67 (3094.94-3310.39)	0.718

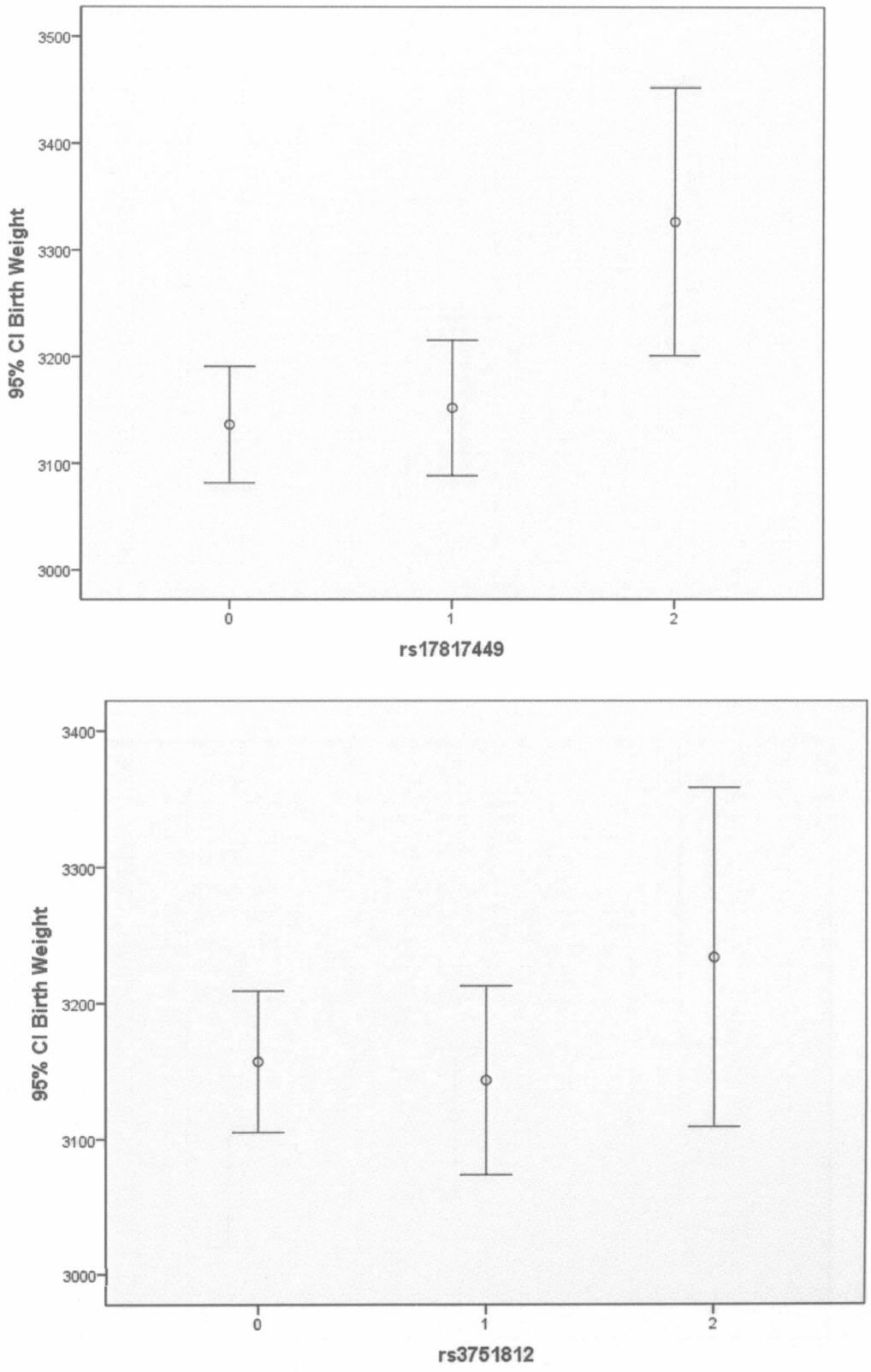
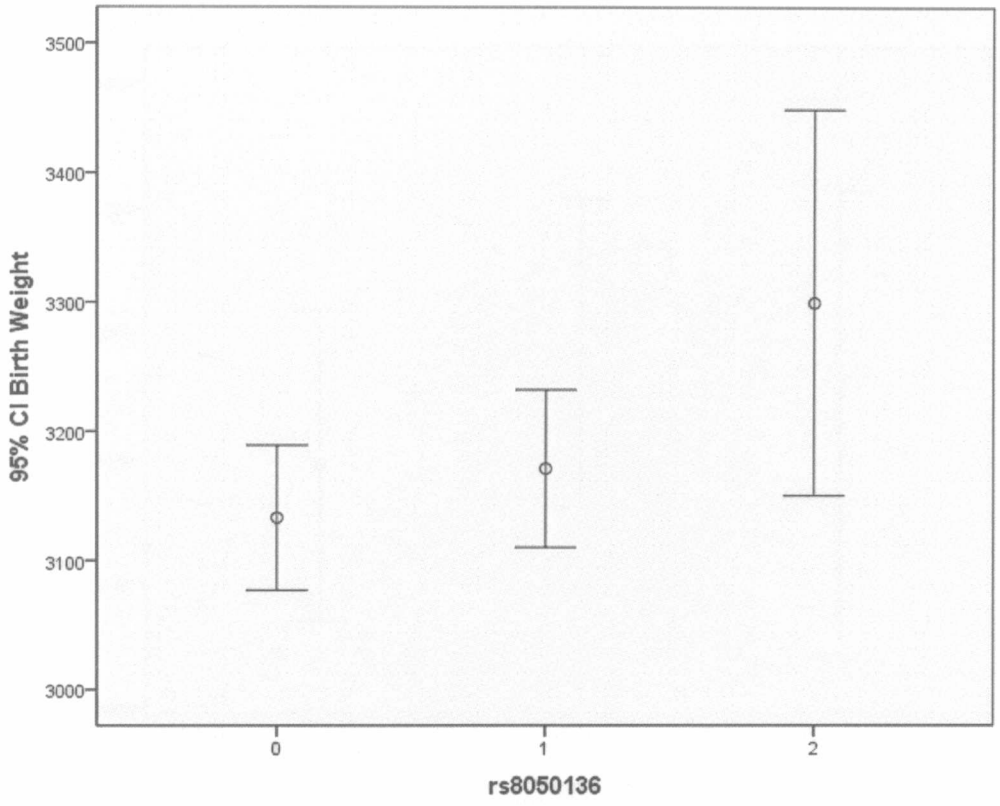


Figure 4.2.2: Graphic representation of the mean birth weights for the SNPs rs17817449 and rs3751812. The zero, one and two denotes the number of risk alleles. Mean birth weight is represented by the circle with the error bars representing 95% confidence interval and was found to be significantly associated with the number of risk alleles for rs17817449 (p=0.025) but not rs3751812 (p=0.248).



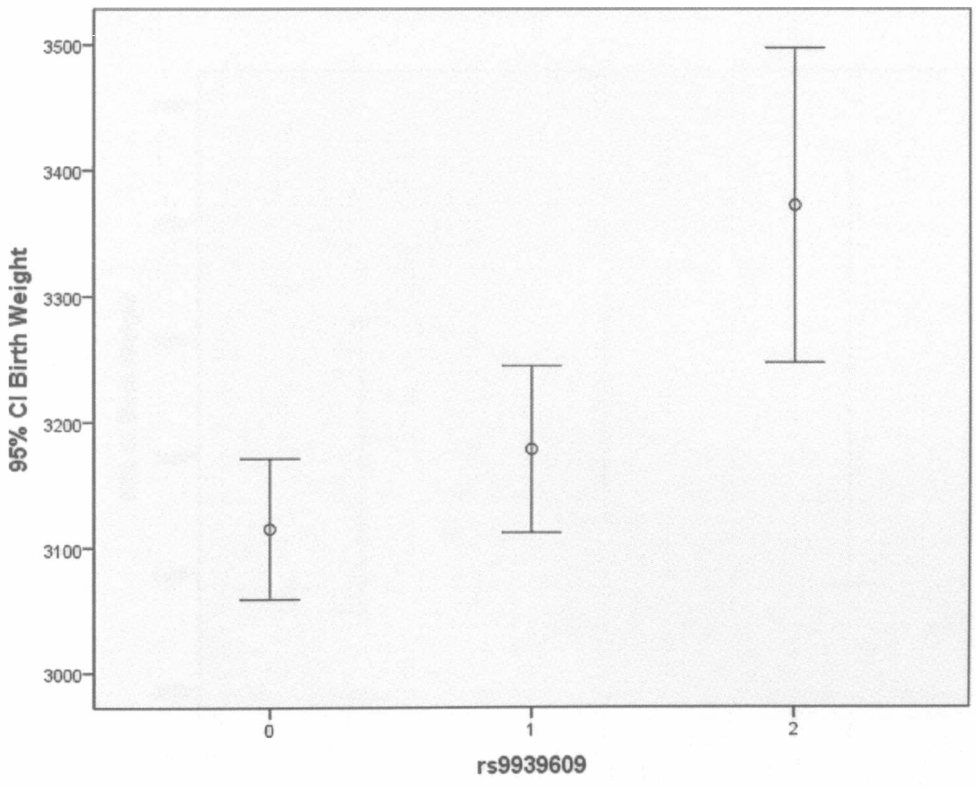
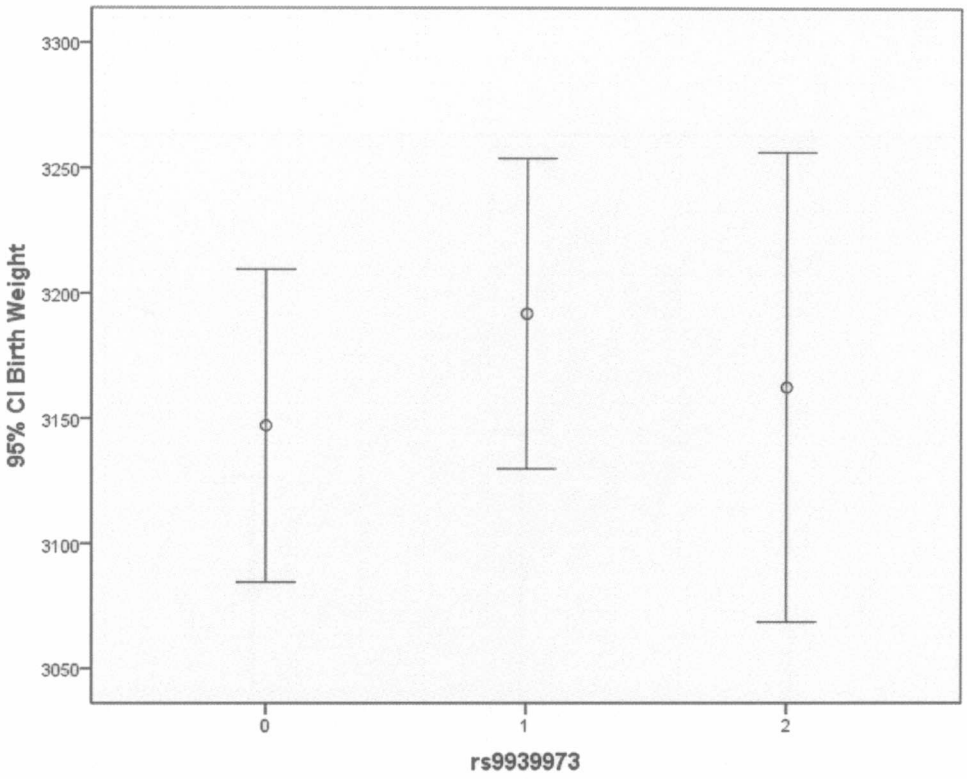


Figure 4.2.3: Graphical representation of the mean birth weights for the SNPs rs8050136 and rs9939609. The zero, one and two denotes the number of risk alleles. Mean birth weight is represented by the circle with the error bars representing 95% confidence interval and was found to be significantly associated with number of risk alleles for both SNPs, p=0.004 and p=0.002 respectively.



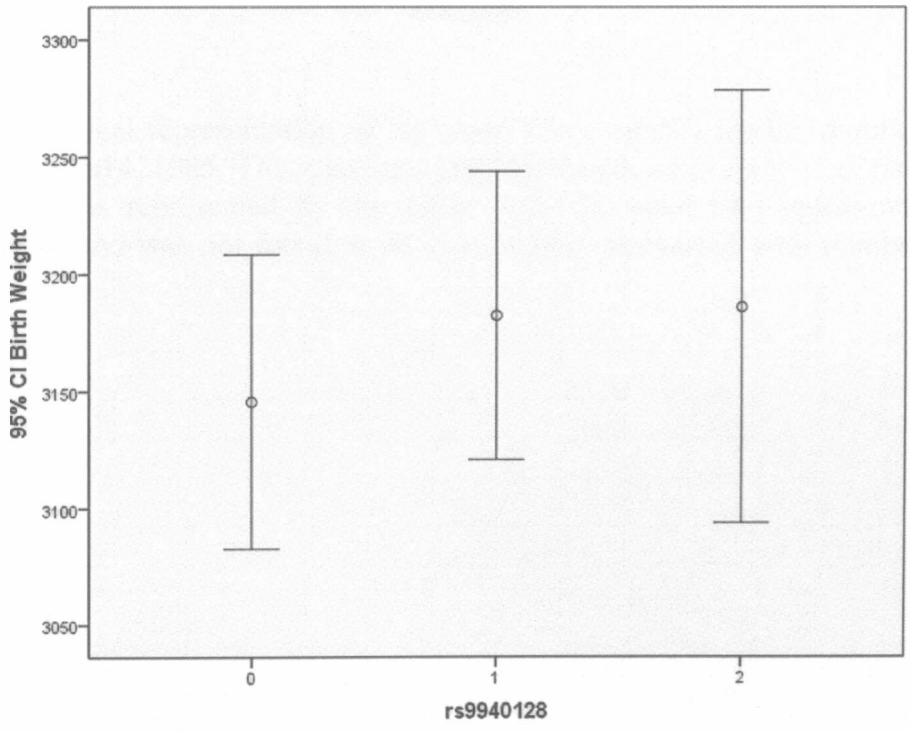


Figure 4.2.4: Graphical representation of the mean birth weights for each of the number of risk alleles for the SNPs rs9939973 and rs9940128. The zero, one and two denotes the number of risk alleles. Mean birth weight is represented by the circle with the error bars representing 95% confidence interval and was not found to be significantly associated with number of risk alleles for either SNP, p=0.767 and p=0.636 respectively.

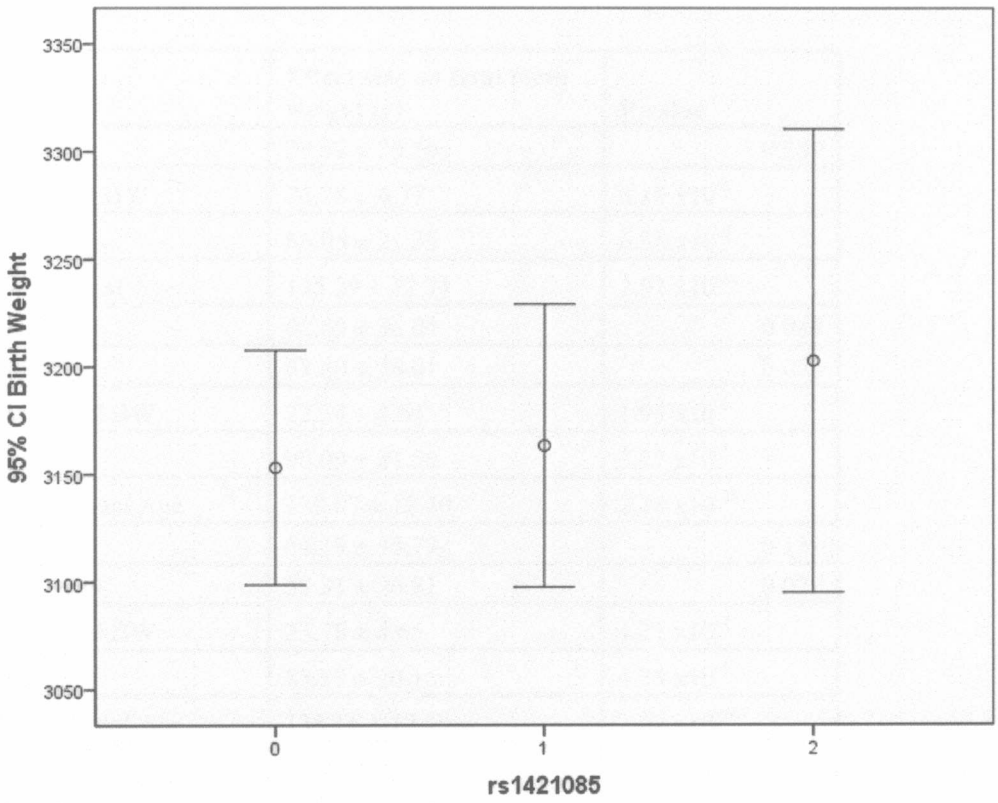


Figure 4.2.5: Graphical representation of the mean birth weights for the number of risk alleles for the SNP rs1421085. The zero, one and two denotes the number of risk alleles. Mean birth weight is represented by the circle with the error bars representing 95% confidence interval and was not found to be significantly associated with number of risk alleles p=0.718.

Table 4.2.6: The effect size of the nine SNPs on fetal birth weight with the covariates parity, maternal weight, gestational age and fetal sex and their p values (p=0.05)

	Effect size on fetal birth weight (g)	P value
rs9940128	29.52 ± 25.56	0.249
Maternal BW	26.78 ± 4.77	3.18×10^{-8}
Parity	86.05 ± 21.35	6.38×10^{-5}
Gestational Age	135.39 ± 12.71	2.91×10^{-21}
Sex	62.69 ± 36.04	0.083
rs8050136	63.80 ± 28.01	0.023
Maternal BW	22.34 ± 4.64	1.93×10^{-6}
Parity	90.09 ± 21.50	3.27×10^{-5}
Gestational Age	139.07 ± 12.40	2.28×10^{-26}
Sex	54.19 ± 35.79	0.131
rs17817449	59.31 ± 26.81	0.027
Maternal BW	23.78 ± 4.65	4.21×10^{-7}
Parity	83.17 ± 20.16	4.24×10^{-5}
Gestational Age	138.16 ± 12.39	2.62×10^{-26}
Sex	58.12 ± 35.54	0.103
rs9939609	90.46 ± 27.35	0.001
Maternal BW	26.39 ± 4.79	5.87×10^{-8}
Parity	81.49 ± 20.99	0.0001
Gestational Age	136.10 ± 12.81	6.01×10^{24}
Sex	83.61 ± 36.76	0.023
rs3751812	19.50 ± 28.20	0.490
Maternal BW	23.86 ± 4.80	8.92×10^{-7}
Parity	91.88 ± 21.19	1.73×10^{-5}
Gestational Age	133.06 ± 12.60	7.44×10^{-24}
Sex	53.73 ± 36.37	0.140
rs1421085	8.19 ± 26.26	0.755
Maternal BW	24.07 ± 4.58	2.07×10^{-7}
Parity	80.1018 ± 20.5790	0.0001
Gestational Age	141.1099 ± 12.1607	4.0282×10^{-28}
Sex	60.6093 ± 35.3963	0.0873
rs9939973	16.0979 ± 25.2736	0.5244
Maternal BW	26.2571 ± 4.7339	4.5651×10^{-8}
Parity	85.0619 ± 21.0880	6.2832×10^{-5}
Gestational Age	138.9484 ± 12.5669	9.3172×10^{-26}
Sex	58.0444 ± 35.8514	0.106

No significant associations with fetal sex and birth weight was observed, but both female and male offspring were found to have on average larger birth weights with the homozygous genotype of the minor alleles, with four out of the seven SNPs, rs17817449, rs3751812, rs8050136 and rs99399609. The increase in the number of risk alleles was only found to be significantly different for the SNPs rs99399609 and rs17817449, with females (p=0.010 and p=0.0310, respectively) but not males (p=0.085 and p=0.419), from zero to two (p=0.004 and p=0.010) and one to two (p=0.012 and p=0.016) but not from zero to one risk allele (p=0.900 and p=0.716). However, only female babies were found to have on average larger birth weights with the homozygous genotype of the minor alleles for the SNPs rs1421085, rs9939973 and rs9940128 not with male babies, who had larger birth weight with the heterozygous genotype. In addition with the latter three SNPs the heterozygous genotype was found to confer a lower birth weight at term with female babies whereas it was found to confer a lower birth weight in male offspring (Table 4.2.7).

Of the significant associations on birth weight with the covariates gestational age was found to have the greatest effect on the increase in birth weight, ranging from 133.06-141.11 g and was also found to be the most significantly associated with birth fetal weight ($p=2.2810^{-21}$-$4.03x10^{-28}$). The second largest effect size on fetal birth weight was parity, for each of the SNPs, ranging from 80.10-91.87 g, ($p=0.0001$-$6.38x10^{-5}$). Maternal BMI was found to have a lesser effect on fetal birth weight, compared with the other covariates, 22.34-26.78 g, the associations were still strongly significant, $p=1.93x10^{-6}$-$4.57x10^{-8}$.

Table 4.2.7: Mean fetal birth weight grouped by genotype for each of the seven *FTO* SNPs and fetal sex

SNP	Genotype		
	11	12	22
	Mean Birth Weight (±SD)		
rs1421085			
Female	3150.62 (472.98)	3123.55 (546.67)	3306.78 (444.35)
Male	3155.36(497.268)	3165.82 (521.69)	3111.56 (375.02)
rs17817449			
Female	3148.85 (456.06)	3107.36 (571.10)	3384.64 (452.63)
Male	3141.44 (489.79)	3159.26 (504.42)	3272.67 (503.28)
rs3751812			
Female	3168.76 (455.07)	3109.35 (557.63)	3284.35 (386.07)
Male	3146.52 (471.79)	3149.91 (542.39)	3175.50 (428.27)
rs8050136			
Female	3138.98 (544.24)	3133.15 (457.59)	3290.00 (545.25)
Male	3174.01 (489.63)	3127.31 (464.23)	3311.50 (557.98)
rs9939609			
Female	3108.50 (545.32)	3116.75 (443.18)	3413.21 (444.41)
Male	3202.64 (516.07)	3113.35 (474.89)	3327.69 (475.85)
rs9939973			
Female	3147.91 (552.70)	3142.50 (447.76)	3199.30 (427.90)
Male	3146.20 (520.23)	3210.93 (454.22)	3131.57 (485.80)
rs9940128			
Female	3227.21 (387.74)	3132.39 (561.92)	3142.47 (453.96)
Male	3150.00 (492.43)	3200.19 (515.04)	3148.32 (447.38)

4.3.0 Genotyping the *ADCY5* SNP, rs11708067, in the Maternal SA Cohort with Respect to Fetal Birth Weight

The Life Technologies TaqMan genotyping assay was used to genotype one SNP within the *ADCY5* locus, rs11708067.

4.3.0 Post-read Step

Once the PCR was conducted on the SA samples, on both 348 well plates, the post-read step was performed to enable genotyping of the samples, which uses the fluorescent signals measured during the RT-PCR for each allele for each sample. After the post-read step was conducted, both plates were initially checked as a quality control step using the SDS software. Results can be viewed either using the allelic discrimination plot, results grid, raw data plot or via a results table, which are all synchronised (Figure 4.3.1). The results grid, which showed the whole plate, was green, which meant that the all of the samples could be read by the machine and were called and given a genotype. The results table showed all the samples being called as either LT-1 (A allele), LT-2 (G allele), both, or undetermined, with additional data on call type (either by the user or automatic), quality value (the estimate of the probability that the assigned genotype for a particular sample is correct, with respect to the other samples in the dataset) and normalised fluorescence values for allele A and allele G, when the dye is cleaved from the probe, which are represented as allele X Rn and allele Y Rn respectively. The raw data plot displays the raw reporter fluorescent signal for the selected samples. Each genotype has a distinct fluorescent signal. The x-axis of the plot represents the spectral bin number where each bin corresponds to a 5 nm window. The allelic discrimination plot graphically shows the reporter dye fluorescence of allele A (x-axis) against allele G (y-axis), which are shown as three clusters representing the three genotypes, AA, AG and GG, (colour coded, see key).

This plot demonstrates graphically see how all the samples had been called. All the samples are expected to fall within one of three genotypes and any of those that do not are selected and viewed using the raw data plot. Samples that fall directly in between two genotypes are changed to undetermined, as they cannot be reliably confirmed as one genotype or the other. Samples that are close to one genotype are viewed using the raw data plot with another confirmed sample to look at the fluorescence of the two, to see if they are similar and decide whether to call the uncertain sample as a particular genotype or undetermined. Wells that contained water and acted as negative controls were inspected to ensure they were described as undetermined. Duplicate samples were also checked to make sure they had been called the same as each other and given the same genotype.

4.3.2 Quality Control of Genotyping Data

Both plates were viewed using the SDS software and showed similar results, i.e. three clusters on the allelic discrimination plot however the cluster for GG had shifted slightly towards the AG cluster than would be expected but the genotyping was still considered successful (Figures 4.3.2 and 4.3.3). The plate grids were both green, which meant all samples were successfully read by the plate reader and given a genotype. Success rates for plates one and two were 92.41% and 99.46% respectively and overall 95.93%, with 708 out of the total 738 samples successfully genotyped. Pipetting error for plate one prevented the success rate to be similar to that of plate two so many samples within that plate failed.

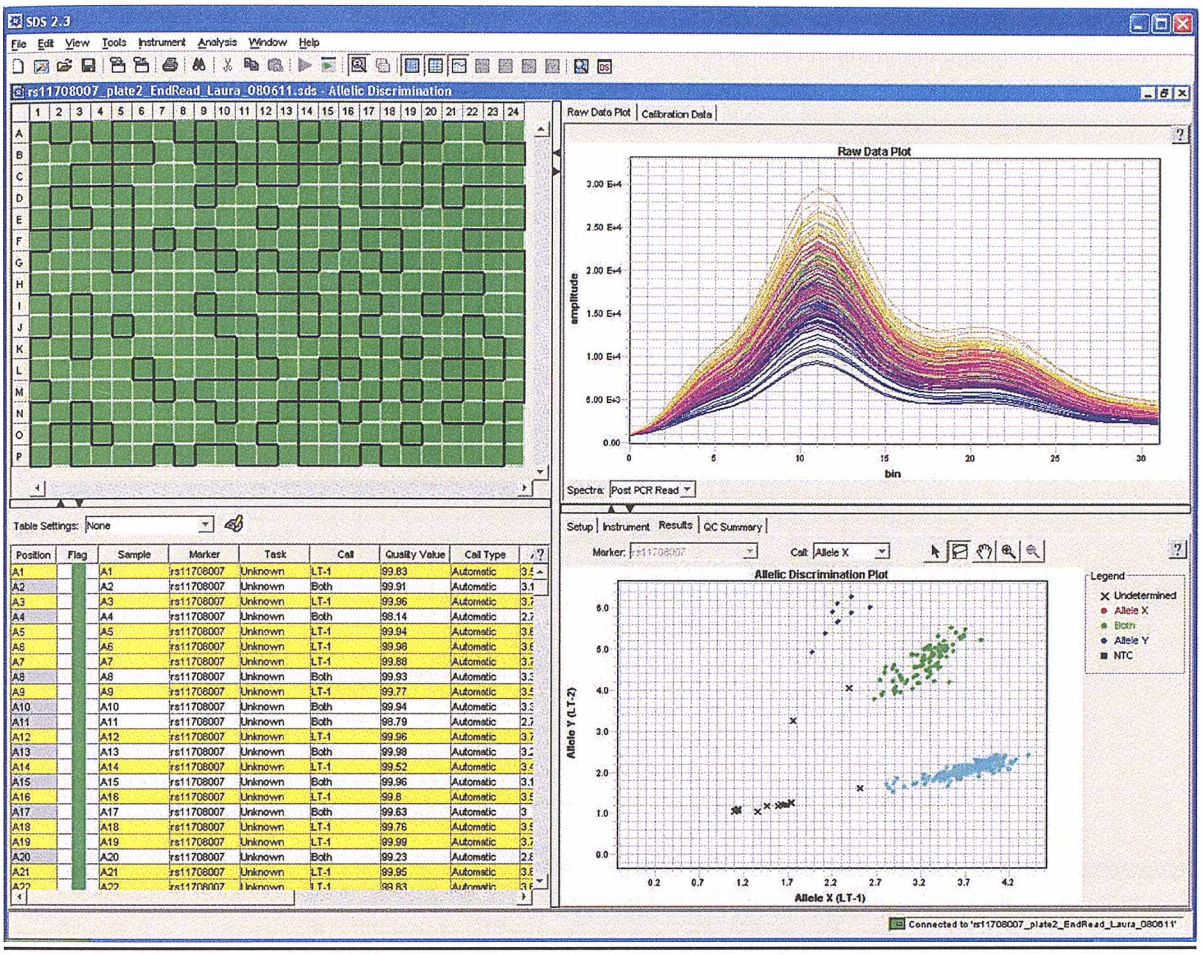


Figure 4.3.1: Results of the TaqMan SNP genotyping assay for plate two, represented in plots, tables or grids using the SDS software.Key: Genotypes AA = light blue, AG = green and GG = dark blue.

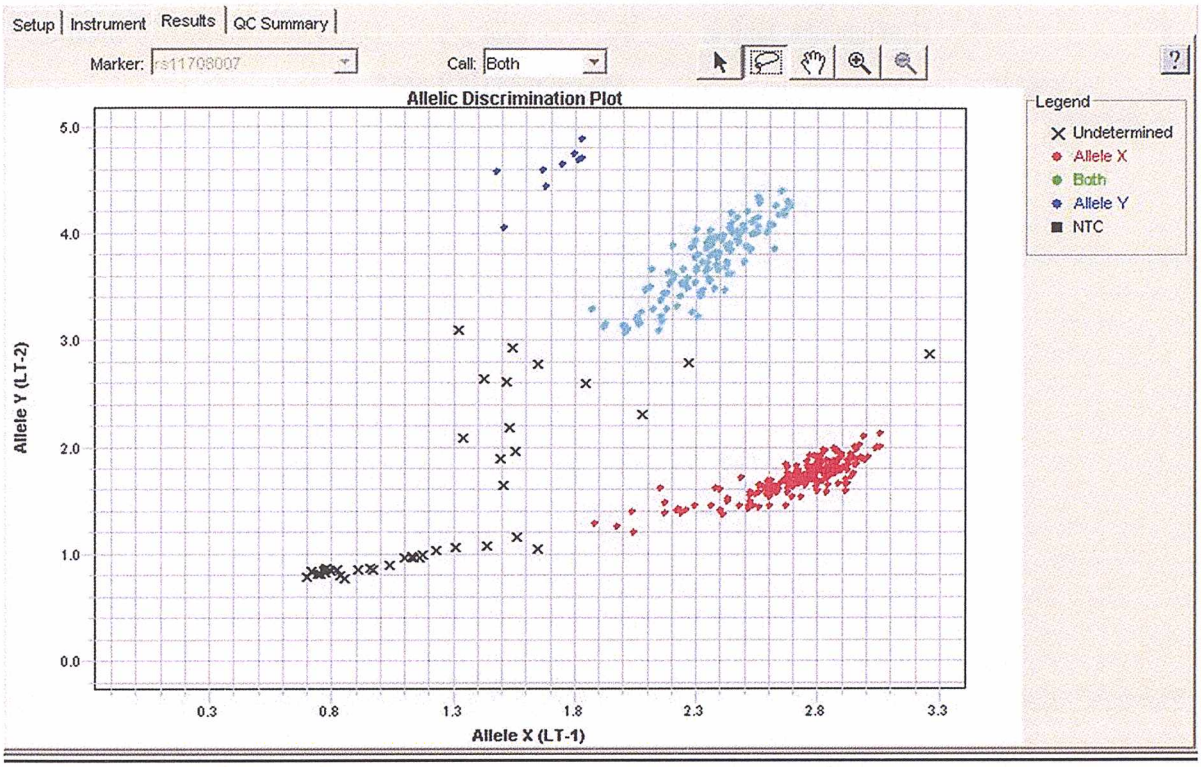


Figure 4.3.2: Allelic discrimination plot of Plate one. Fluorescence of the reporter dye for allele A on the x-axis against the fluorescence of the reporter dye for allele G on the y axis. The three genotypes of the SNP are represented by three different colours, AA equals red, AG equals pale blue/green and GG is given dark blue. Samples in which genotype is undetermined are represented by X.

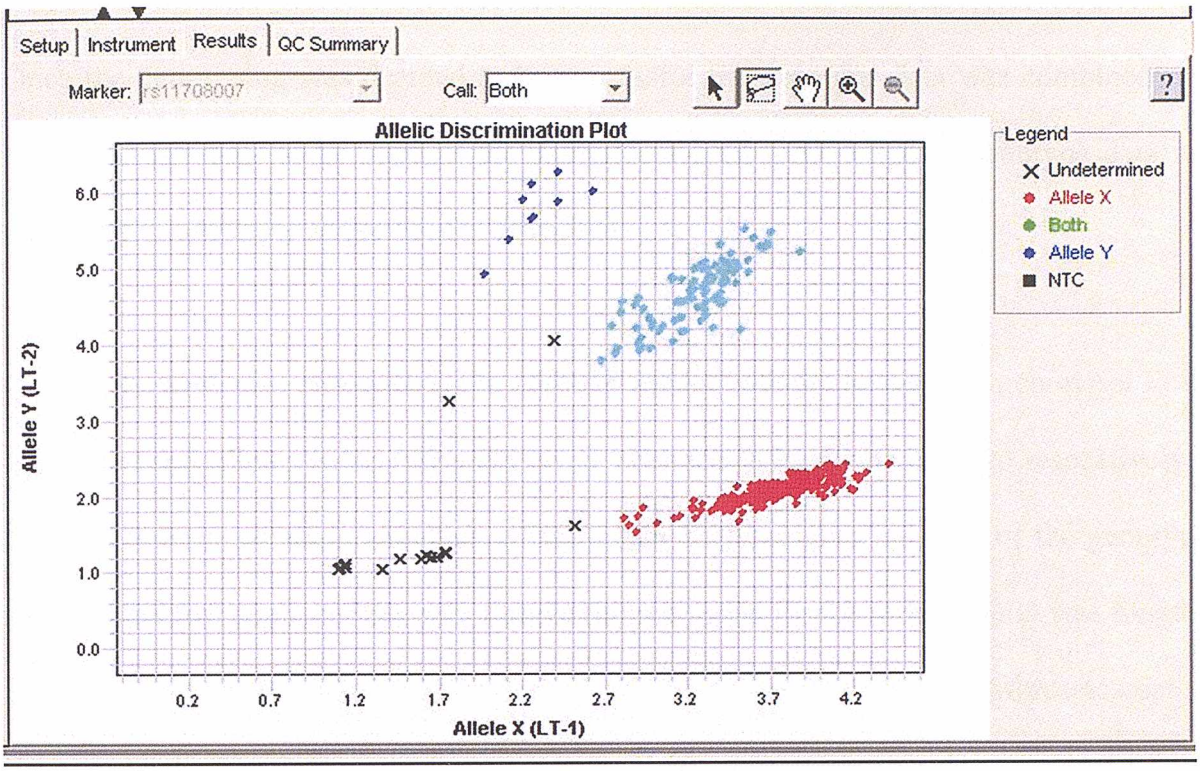


Figure 4.3.3: Allelic discrimination plot of plate two. Fluorescence of the reporter dye for allele A on the x-axis against the fluorescence of the reporter dye for allele G on the y axis. The three genotypes of the SNP are represented by three different colours, AA equals red, AG equals pale blue/green and GG is given dark blue. Samples in which genotype is undetermined are represented by X.

4.3.3 Analysis of Maternal Genotype at rs11708067 with Fetal Birth Weight

Once the genotyping was deemed to be successful the genotype data was combined with the phenotype data to be used in the analysis of fetal birth weight. Phenotype data was checked for normality as previously described.

Genotypes were also changed to the additive model, which changes the genotypes to the number of risk alleles so A, AG and G were changed to zero, one and two respectively as G is the risk/minor allele. Analysis of the mean birth weights for each of the genotypes showed that two copies of the minor allele (homozygous for the minor allele) was found to confer a higher birth weight, although the lower birth weight was found with one copy of the risk allele (the heterozygous genotype) (Table 4.3.1). These were not found to be significantly different from one another (p=0.707).

Analysing fetal sex separately, fetal birth weight was found to be higher with no copies of the minor allele, (homozygous for the major allele, AA) in both male and female offspring (3165.29 g and 3148.53 g respectively) whereas birth weight was lowest with two copies of the minor allele (3137.00 g for both). Descriptive statistics on the other phenotype data are as previously described as rs11708067 was analysed along with the *FTO* SNPs.

Table 4.3.1: Mean birth weight for the different genotypes of SNP rs11708067. Birth weight is given in grams with 95% CI. The genotype, zero, one and two represents the number of risk alleles of the minor allele. Mean birth weight for each genotype was not found to be significantly different (p=0.05).

SNP	Genotype			P Value
	0	1	2	
	Mean Birth Weight			
rs11708067	3164.43 (3117.68-3211.18)	3144.61 (3076.30-3212.92)	3248.24 (2960.74-3535.73)	0.707

Linear regression was performed to investigate any significant association with fetal birth weight with the covariates parity, gestational age, booking weight and sex.

No significant associations were found with the higher birth weight and the SNP (p=0.510), however significant associations were found with the covariates maternal booking BMI ($6.93x10^{-8}$), parity (p=0.0058) and gestational age ($4.31x10^{-28}$). Fetal sex was found to be trend associated with birth weight (p=0.0732).

Gestational age was found to have the highest effect on the increase in birth weight, 138.34 g (± 20.07 SD), and the most strongly significantly associated ($p=4.31x10^{-28}$). Parity and maternal BMI had a lesser effect on birth weight compared with gestational age, but they still increased fetal birth weight by 69.45 g and 24.92 g respectively and are strongly associated (0.00058 and $6.93x10^{-8}$ respectively) (Table 4.3.2). The effect of fetal sex on birth weight was found to be high, 61.90 g, despite being only trend associated with birth weight (p=0.073).

Table 4.3.2: Linear regression results for the *ADCY5* SNP rs11708067 with the covariates, booking BMI, parity, gestational age and sex

	Effect size (g ± SD)	**P value**
rs11708067	21.14 ± 32.04	0.510
Booking BMI	24.92 ± 4.56	6.93×10^{-8}
Parity	69.45 ± 20.07	0.0058
Gestational age	138.34 ± 11.95	4.31×10^{-28}
Sex	61.90 ± 34.49	0.073

In summary, SNPs within three maternal T2D susceptibility and obesity genes were investigated for association with fetal birth weight in a Gujarati Indian population: thirty-seven SNPs in *HHEX-IDE*, nine SNPs in *FTO* and one SNP in *ADCY5*. Significant associations were found with rs12765131 in *HHEX-IDE* with lower birth weight and rs9939609 in *FTO* with higher birth weight but no significant associations with rs11768067 in *ADCY5* and higher birth weight. These suggest that both the maternal *HHEX-IDE* and *FTO* loci are involved in birth weight in Gujarati South Asian population, but have opposing effects, whereas the maternal *ADCY5* is not involved in birth weight within this population.

5.0 Discussion

Birth weight and obesity are complex traits. Their aetiologies are seemingly intertwined, as birth weight is a significant risk factor in the development of adult onset obesity, T2D and the metabolic syndrome. Understanding the role of genetic involvement in both may elucidate their interactions within the current obesogenic environment. Two different methodological approaches were undertaken in order to investigate three variants of the leptin receptor, (located in the extracellular region) using molecular cloning and computer modelling and to investigatespecific candidate genes for fetal birth weight inthe South Asian population.

5.1.0 Functional analysis of the Leptin Receptor

The aim of this analysis was to investigate the binding affinity between leptin and the eight haplotypes at three polymorphic sites Lys109Arg, Gln223Arg and Lys656Asp in the leptin receptor. These were chosen since they were identified in published studies as being associated with obesity and measures of adiposity (Yiannalcouris *et al.,* 2001; Quinton *et al.,* 2001; Mattevi *et al.,* 2002; Fairbrother *et al.,* 2007; Ben Ali *et al.,* 2009). The cloning process required many consecutive stages (Figure 3.1.1) and stage one to four were completed. However, the long-term aim of this research, the analysis of the effect of published mutations on the binding affinity of the receptor for leptin, could not be determined due to the unsuccessful sub-cloning of the LEPR insert into the mammalian expression vector (stage five).

5.1.1 Methodological Improvements

There are a number of common problems that are encountered when sub-cloning these were borne in mind when trying to trouble shoot. The competent *E.coli* cells used, JM109,

(Stage two) were chosen to be recombination deficient (*recA*$^-$), however there are also different commercially available strains of *E.coli* e.g. Mach1, DH5α, and INVαF′ cells, which are competent and *recA*$^-$. These could be used in future experiments to successfully sub-clone the *LEPR* insert into the pGEMT vector.

One of the most important aims of sub-cloning the *LEPR* was to insert it into the vector in the correct orientation. Multiple mini-preparations were performed to detect colonies that contained the insert in the correct orientation. The correct orientation was important so that subsequent sub-cloning into the pCMV2B Tag expression vector could be carried out with the use of two restriction enzymes, *BamHI* and *SalI*. This step was not reached but it could have been replaced since many other vectors and restriction enzymes are also commercially available, e.g. the pCR 2.1 TOPO cloning vector (Invitrogen), the pEF6/V5-His TOPO cloning vector (Invitrogen) instead of the pGEM-T vector. Additionally, the restriction enzymes *HindIII,* and *XhoI* or *ApaI* could be used to cleave the *LEPR* insert from the pCR 2.1 TOPO cloning vector before inserting into the pCMV Tag 2B expression vector, as it contains the same restriction sites (Figure 3.1.20). Alternatively the Membrane-Pro Functional protein expression (FPE) system (Invitrogen) could have been used or a pGEX vector (GE Healthcare Life Sciences), which uses glutathione S-transferase as the fusion protein, as opposed to the pCMV-Tag 2B expression vector. An alternative vector may expedite the future cloning of the *LEPR* insert and enable the completion of the functional analysis of the three *LEPR* variants of interest. These might be more compatibile with the large size of the LEPR fragment of interest. In addition, it should be noted that approximately 5% of sequences cannot be cloned into bacteria, which could have been an issue with the LEPR insert (Godiska *et al.,* 2010).

5.2.0 Computer Modelling of the Leptin Receptor

As the functional study was unsuccessful, computer modelling methods were used to investigate the three polymorphisms in the LEPR and if they alter the binding affinity with leptin either in isolation or as haplotypes. The 3D structure of the native CRH1 domain located in the extracellular region of the leptin receptor was successfully predicted using the I-TASSER server suite of computer programmes, with a high confidence score (of -0.63). In addition, the 2D and 3D protein structure of the CRH1 domain containing the amino acid change from glutamine to arginine at position 223 was also successfully predicted, (C-score -0.91). Both of these 3D structures were predicted to have three beta pleated sheet structures and aligned well together indicating regions where the structures were likely to homologous. Further analysis of these two 3D protein structures showed that as a result of the change in the amino acid, hydrogen bonding occurred between the arginine residue at position 162 and the cysteine residue at 151. This does not occur in the native 3D structure with the glutamine residue, which suggests this may have structural implications. Indeed, the majority of the associations between LEPR and obesity are with the Gln223Arg mutation (Chagnon *et al.,* 2000; Mattevi *et al.,* 2002; Fairbrother *et al.,* 2007; Ben Ali *et al.,* 2009). Arginine residues are known to interact with negatively charged amino acids to produce hydrogen bonds, which are important to the stability of the overall secondary and tertiary structure of the protein and this may have altered the shape of the binding domain so that leptin may bind with altered affinity.

However, the function of the CRH1 domain remains elusive: it is less well conserved than the other extracellular domains, Ig-like and FNIII, while the Gln223Arg polymorphism lies in a region that is unique to the CRH1 domain. Upon deletion in mouse models shows no obvious detrimental effect (Peelman *et al.,* 2006), thus it is currently only possible to

speculate on the role of this domain. Evidence suggests that it could enhance leptin signalling levels (Zabeau *et al.*, 2004) and could play a role in the formation of the leptin/leptin receptor complex (Carpenter *et al.*, 2012). Our novel data suggest the change to arginine could subtlety alter the folding of the CRH1 domain by introducing additional H bonding between amino acid, R groups not found in the native conformation. This could impact on the overall 3D protein structure in a subtle way, suggesting a small effect of this polymorphism.and may at least in part explain the alteration in binding avidity in the ^{125}I study (Mechan and Blakemore, *Pers Com*), and the contradictory evidence for an association with obesity (Gotoda *et al.,* 1997; Silver *et al.,* 1997; Pyrzak *et al.,* 2009; Dias *et al.,* 2012; Kumsu-Ornek *et al.,* 2012) and why large GWAS (Dina *et al.,* 2007; Frayling *et al.,* 2007; Scuteri *et al.,* 2007; Loos *et al.,* 2008; Chambers *et al.,* 2008; Lindgren *et al.,* 2009; Thorleifsson *et al.,* 2009; Hong and Oh, 2011; Zhao *et al.,* 2011; Heid *et al.,* 2012; Fox *et al.,* 2012; Wang *et al.,* 2012) have been unable to corroborate the findings demonstrated from functional (Chua *et al.*, 1996, Kimber *et al.*, 2008, Peelman *et al.*, 2004, White *et al.*, 1997, Zabeau *et al.*, 2004) and candidate association studies (Ben Ali *et al.,* 2009; Chagnon *et al.,* 2000; Mattevi *et al.,* 2002; Quinton *et al.,* 2001; Wauters *et al.,* 2001b; Yiannakouris *et al.,* 2001).

Furthermore, it is possible that the Gln223Arg polymorphism could be in LD with a causative SNP as the *LEPR* gene region is in high LD (Figure 5.1.1) (Wauters *et al.,* 2001; Fairbrother *et al.,* 2007; Stratigopoulos *et al.,* 2009). Rodent models of obesity have identified a different mutation in the CRH1 domain, Glu269Pro that causes the obese phenotype in the Zucker (fatty) rat (Iida *et al.*, 1996). Functional investigation of this polymorphism demonstrated that there was reduced binding of leptin at the cell surface (Chua *et al.,* 1996) and a reduced cell surface expression of the receptor (White *et al.,*

1997). The reduced cell surface expression was correlated with a reduced level of signal transducing activity of STAT1 and STAT3 transcription factors involved in the JAK/STAT signalling pathway (Figure 1.9.10), by 2-3 fold compared to the wild type receptor (White *et al.,* 1997). Furthermore, the change in amino acid constitutively activated STAT1 and STAT3 but could still be activated by the binding of leptin, yet interestingly this was not the case with STAT5. What is more, as there are distinct domain elements that are involved in the activation of STAT1, STAT3 and STAT5 it has also been suggested that the conformational change in the protein results in the phosphorylation of the tyrosine residue 1138 that recruits STAT3 and STAT1 but not STAT5. The phosphorylation of the tyrosine residue may prevent recruitment and activation of STAT5. This altered signalling could lead to dysregulation of weight homeostasis and leptin resistance leading to the obese phenotype (Chua *et al.,* 1996; White *et al.,* 1997). The alteration in conformational structure was proposed to affect the Trp-Ser-Xaa-Trp-Ser (WSXWS) consensus motif located within the CRH1 domain that could change ligand binding or protein/protein interactions of the receptor (White *et al.,* 1997). Molecular modelling of the mouse CRH1 domain confirmed that steric clashes occurred in the Trp-Ser-Xaa-Trp-Ser motif, when the proline residue was introduced, affecting the stability and correct folding of the domain (Peelman *et al.,* 2006), therefore it could play a role in the formation of the extracellular region of the receptor (Carpenter *et al.,* 2012). Upon binding of leptin, the leptin receptor undergoes dimerisation leading to activation of the JAK/STAT pathway (see section 1.9.10.1). The structure of the extracellular region, notably the sub-domains located within this region are important for leptin binding thus differences in these domains, could affect the dimerisation of the leptin receptor. Even subtle changes to the conformational structure of the domains could impact on dimerisation. As Gln223Arg is located within the CRH1 domain it could also affect the Trp-Ser-Xaa-Trp-Ser motif in the same way as the

Gln269Pro polymorphism. However, the Gln^{223} is not in close proximity to the motif, unlike Gln^{269} (Figure 5.2.1), therefore it is unlikely that the change in amino acid at Gln^{223} affects the motif in the same way. Therefore it is possible the Gln223Arg polymorphism is in LD with Gln269Pro so an indirect association with obesity has been found with Gln223Arg and the Gln269Pro polymorphism could be the causative SNP. As there is a lack of information regarding the Gln269Pro polymorphism in humans, possibly due to low MAF, further investigation is needed to determine if this is the case. If Gln223Arg is not in LD with Gln269Pro then the close proximity between them could indicate that Gln223Arg may cause subtle changes to the CRH1 that is different from that of the Gln269Pro polymorphism, affecting leptin signalling leading to the leptin resistant state.

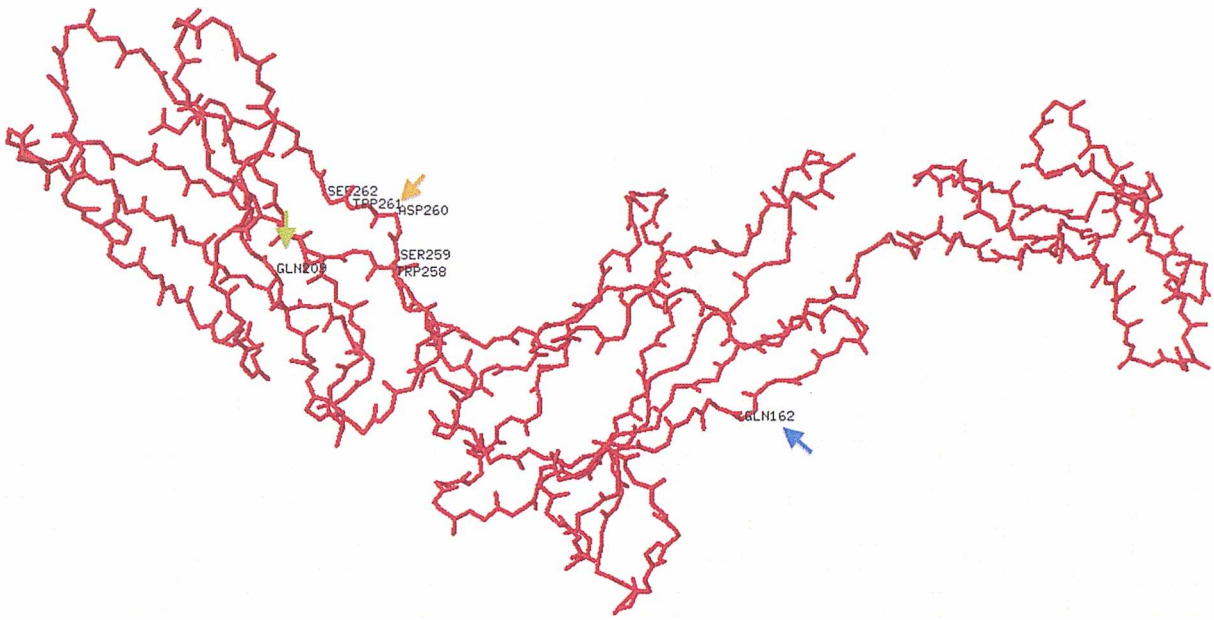


Figure 5.2.1: A ball and stick model of the CRH1 domain with the Gln^{223} amino acid shown at position 162 (Blue arrow). The close proximity between the Gln^{269} amino acid (position 201) and the Trp-Ser-Xaa-Trp-Ser motif (positions 258 to 262) is also shown (green and orange arrows respectively).

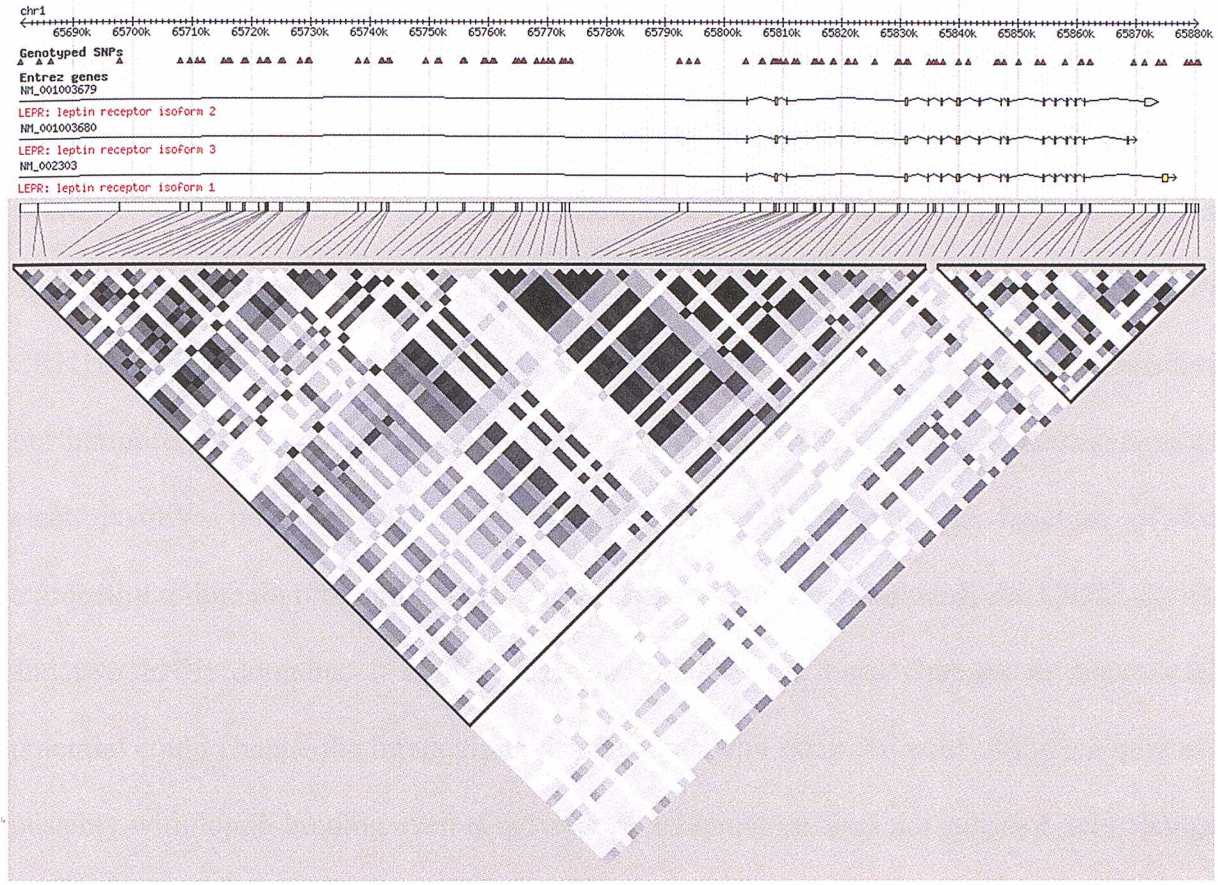


Figure 5.2.2: Map of R^2 across the *LEPR* gene (covering 210 kb) using HapMap CEU data. The location of the gene is shown at the top of the map. The dark areas of grey and black indicate high regions of LD, while the lighter shades of grey indicate low regions of LD. There are two regions of high LD covering 156 kb and 38 kb.

Further modelling analysis of the CRH1 domain showed that the change in amino acid residue at position 223, from glutamine to arginine, caused a slight change in the overall electrostatic surface potential of the domain and at the position of the residue change. The overall electrostatic surface potential of the native CRH1 domain (CRH1Q) showed that one side was electronegative and the opposite side was electropositive/neutral, whereas the CRH1R domain showed that both sides of the domain were neutral to electropositive. Furthermore, the change in amino acid residue at position 162 showed an increase/stronger electropositive potential, within the CRH1R domain. Despite the function of this extracellular domain remaining elusive it lies beside four other domains; the CRH2, Ig-like and two FNIII domains. It is therefore possible that the local changes in electrostatic potential could change the interaction of these areas with other domains of the receptor and possibly with leptin binding even if the overall 3D structure does not substantively change, as leptin is proposed to form a hexameric complex with its receptor (Zabeau *et al.,* 2004; Peelman *et al.,* 2006). Plus, even though this residue appears at the surface of the domain, where it can interact with other amino acid side chains and atoms, the arginine amino acid residue is polar and hydrophobic and therefore is more likely to be buried inside the CRH1 domain itself. Therefore, the change to arginine amino acid could change the interactivity of this region of the domain. There is a strong binding affinity (K_i 200 pM) between leptin and its receptor, using mouse and human ^{125}I-leptin (Lui *et al.,* 1997), with the wild type sequence for each of the three polymorphisms, (A**A**G for Lys^{109}, C**G**G for Gln^{223} and AA**G** for Lys^{656}, with the position of the SNP in bold) (Fong *et al.*, 1998, Liu *et al.*, 1997). Key residues in the other domains in the extracellular region; Ile^{501}, Phe^{502}, Leu^{503}, Leu^{504}, Ser^{505}, Asp^{615} in CRH2 (Figure 1.9.2) are located at the centre of binding site II while Leu^{370}, Ala^{407}, Tyr^{409}, His^{417} and His^{418} form a cluster and are important in binding site III within the Ig-like domain. Mutations within these are known to reduce leptin binding and

cellular signalling (Iserentant *et al.*, 2005, Peelman *et al.*, 2006b). Thus, the conformational structure of the membrane-bound receptor is important and subtle differences in any of the domains of the extracellular region could impact on the overall structure of the receptor (Quinton *et al.*, 2001).

One of the main objectives of this study was to create a computer-based structural prediction of the complete native leptin receptor and then extend this model to include the individual changes in amino acids due to the three polymorphisms Lys109Arg, Glu223Arg and Lys656Asp, creating different haplotypes and determining if there were any structural differences between them. However, the entire leptin receptor protein structure could not be successfully predicted with good confidence and reliability, despite the I-TASSER programme suite being ranked one of the best computer-modelling program by recent critical assessment of protein structure prediction experiments (Kryshtafovych *et al.*, 2011). This is perhaps why the structure has not yet been published. It is likely from the investigations carried out in this study that this program is not good at predicting proteins that are membrane bound or contain more than one domain (Zhang, 2009); the leptin receptor is a large protein and has three domains: intracellular, transmembrane, and extracellular domains. The results reflected this in the low confidence scores (-1.30) and the unexpected differences in structure prediction when the polymorphisms were separately introduced. However, the structural prediction of just the extracellular domain of the leptin receptor also produced the same low-confidence score results, (which is where the three polymorphisms, Gln223Arg, Lys109Arg and Lys656Asp, are located), compared with the native structure and when the individual polymorphisms were introduced. Therefore, the CRH1 domain alone was used as a basis for structural predictions, as the Gln223Arg polymorphism is located in this domain whereas the other two polymorphisms

are not located in any of the structural domains in the extracellular region of the leptin receptor. This resulted in the successful prediction of the CRH1 domain, in both allelic forms. The C-score was not as high as 2 (a high-confidence protein structure prediction) but after the first beta sheet structure there was a group of four amino acids with low-confidence scores for structure prediction and this indicated that this region of the structure was difficult to predict and may have strongly affected the overall confidence score. Furthermore, research suggests that the CRH1 domain actually contains two beta sheet structures, which represent the two FNIII domains (residues 62-178 and 235-328), rather than three and that the Gln223Arg polymorphism occurs in the loop between them (Stratigopoulos *et al.,* 2008). However, the loop region, where Gln223Arg is located, within the CRH1 is unique to this domain (Stratigopoulos *et al.,* 2008). Therefore, the native CRH1 domain (CRH1Q) and the CRH1R domain predicted by the I-TASSER suite of programmes have evident errors, so it is unlikely to be the actual 3D structure of the domain. Despite this, the change in amino acid at Gln223Arg could still be important in the overall 3D structure of the protein since it occurs in the loop region of the domain and because arginine has a longer side chain than glutamine it can form additional hydrogen bonds consequently affecting the structure of this important domain.

The second CRH domain (CRH2) that occurs in the extracellular region of the leptin receptor was also successfully predicted using the I-TASSER suite of programmes and was shown to have two anti-parallel beta pleated sheet structures, creating a V type structure (Figures 3.3.16 and 3.3.18). This corroborates previous research, which also predicted this structure in humans and mice (Iserentant *et al.,* 2005; Peelman *et al.,* 2006), Our work improves their work as receptors which contain the CRH2 domain did not have to be preselected and aligned with the human and mouse LEPR CHR2 domain to create the

homology models, using a different program, t-coffee. I-TASSER does this as part of the protein prediction. This domain is required for high affinity binding of leptin (Fong *et al.*, 1998, Sandowski *et al.*, 2002, Zabeau *et al.*, 2004) and is involved with the binding sites I and II, in the hexameric receptor/ligand complex. Much functional research has been carried out to determine the residues of both leptin and the leptin receptor that are involved with these binding sites including using leptin from chickens and the bovine leptin receptor (Fong *et al.,* 1998, Niv-Spector *et al.,* 2005). The residues, Val^{36}, Thr^{37}, Phe^{41} and Phe^{43} within leptin and Ile^{501}, Phe^{502}, Leu^{503}, Leu^{504}, Ser^{505} and Asp^{615} in CRH2 are thought to be important in binding site I while Leu^{13}, Leu^{86}, Leu^{89} and Phe^{92} residues in leptin and Leu^{469}, Leu^{528}, Leu^{617} in CRH2 are thought to be involved with binding site II (Iserentant *et al.,* 2005; Niv-Spector *et al.,* 2005; Peelman *et al.,* 2006). Those residues in CRH2 thought to be important in binding of leptin were viewed, in our model of CRH2, to determine whether they demonstrated the presence of any strong negative or positive areas of surface electrostatic potential, as a possible mechanism to affect the functional interactions with leptin.

Ile^{501}, Phe^{502}, Leu^{503}, Leu^{504}, Ser^{505} and $Asp^{615\ lie}$ within the V shape region of the domain, where leptin binds but have different effects. Results showed Phe^{502} to confer a negative surface charge (Figure 3.3.2), with the majority of the residues conferring a neutral state on the surface while Ile^{501} and Asp^{615} were buried beneath the surface of the domain. The four relevant residues in leptin, Leu^{13}, Leu^{86}, Leu^{89}, Phe^{92}, were viewed using the known leptin structure, (downloaded from the PDB, accession number P41157). The overall surface electrostatic potential of leptin was negative and the intensity varied on the surface in different regions of the molecule (Figure 3.3.2). This supports findings by Carpenter *et al.,* (2012) which found a mouse anti-leptin monoclonal antibody, which acts as an antagonist

is also negatively charged. Residues Leu^{86} and Leu^{89} showed the strongest negative electrostatic potential while Leu^{13} and Phe^{92} were less so (Figure 3.3.2). In our CRH2 model, there was a neutral electrostatic potential on the surface at this residue while Phe^{502} conferred a positive electrostatic potential. This suggests that the electrostatic interactions are unlikely to occur between these residues, which is actually what is expected.

The four residues in leptin, Val^{36}, Thr^{37}, Phe^{41} and Phe^{43}, important in binding site I of the CRH2 domain, were also investigated and compared to the potential with the CRH2 domain with which they interact. Phe^{41} and Phe^{43} showed negative electrostatic potential while Val^{36} and Thr^{37} were hidden below the interacting surface (Figure 3.3.2). Leu^{469}, Leu^{528}, Leu^{617} residues in CRH2 had neutral electrostatic potential, which suggests there is no electrostatic interaction between these residues to aid in binding.

The minimal electrostatic potential in key residues indicates that there may be other interactions occurring between CRH2 and leptin. Investigation by Iserentant *et al.,* (2005) proposed that hydrophobic bonding occurs between the six residues (Ile^{501}, Phe^{502}, Leu^{503}, Leu^{504}, Ser^{505}, Asp^{615}) in CRH2 and four residues, Leu^{13}, Leu^{86}, Leu^{89}, Phe^{92}, within helix C of leptin forming a hydrophobic cleft, after mutating all of them to alanine, separately, which reduced the binding affinity of leptin and leptin signalling. Alanine is also a non-polar amino acid, which does not change the charge for the first four residues in CRH2 but does change the charge for the latter two. This suggests that the side chains of these six amino acids of CRH2 are also important in the binding between CRH2 and leptin. After mutating the six residues in CRH2 to hydrophilic serine residues individually, only Ile^{501} and Leu^{504} showed a more dramatic change in signalling capacity compared to the other four residues, further supporting the hydrophobic interaction as important, suggesting that

these two residues are more central to binding than the other four residues. It has been proposed that the Leu^{504} residue is key in binding site II and forms the centre of this binding site with Leu^{13} in leptin (Peelman *et al.,* 2006) while Iserentant *et al.,* (2005) found Leu^{86} in leptin to also be important in binding site II. With respect to binding site II hydrophobic interactions are also thought to occur between the key residues in leptin and the CRH2 domain (Peelman *et al.,* 2004; Peelman *et al.,* 2006), which supports our findings that residues Val^{36} and Thr^{37} wereneutralandhydrophobic. Furthermore, the positive and negative electrostatic potentials at other residues surrounding the key residues in the CRH1 and CRH2 domains could indicate that additional interactions are involved not only in the binding of leptin, but also in the overall hexameric complex. The stociochiometry of binding is that for two leptin molecules there are four leptin receptors, (two as dimers), and each leptin molecule interacts with three leptin receptor molecules therefore additional residues may be important for stability of the overall structure through electrostatic interactions. Research by Carpenter *et al.,* (2012) involving the CRH2 domain found that different residues in CRH2 interacted with the mouse monoclonal antibody, Arg^{465}, His^{467}, Arg^{468}, Ser^{469}, Ser^{470} and Glu^{484}, supporting the view that these additional residues may be involved in leptin binding. These were viewed in our CRH2 molecule but were found to have no positive or negative electrostatic potential, which could indicate a different, hydrophobic, interaction of these residues.

The lack of strong reliability in the prediction of the structures of the leptin receptor and the separate domains from I-TASSER is supported by the lack of available structural data for the leptin receptor and the structure of the exact binding complex with leptin remains undetermined (Carpenter *et al.,* 2012). The X-ray crystallography and nuclear magnetic resonance spectroscopy methods in determining protein structure have failed so far because

of the large size of the leptin receptor and the different domains that it contains (Carpenter *et al.,* 2012, Mancour *et al.,* 2012). I-TASSER uses a combination of different techniques to determine the structure of the query protein including *ab initio* modelling threading and sequence homology comparison. The leptin receptor shows low sequence homology to other cytokines, of 25-40% (Siddiqui *et al.,* 2010) but critical areas are conserved between them (Stratigopoulos *et al.,* 2009; Iserentant *et al.,* 2005), and structurally leptin receptor is similar to G-CSF, IL-6 and gp130 (Zabeau *et al.,* 2003). However, the leptin receptor only has two FNIII domains compared to G-CSF and gp130, which have three (Zabeau *et al.,* 2003). Therefore, the majority of research into the structure of the leptin receptor has focused on the CRH2 and Ig-like domains as a single protein structure to determine the interaction with leptin (Peelman *et al.,* 2004; Iserentant *et al.,* 2005; Peelman *et al.,* 2006), This is pertinent since the CRH2 domain, is key for the high affinity binding of leptin (Fong *et al.,* 1998; Sandowski *et al.,* 2002; Zabeau *et al.,* 2004) and the Ig-like domain along with FNIII domain are important for LEPR activation (Fong *et al.,* 1998; Zabeau *et al.,* 2004).

To determine the leptin and leptin receptor complex structure, the IL6/IL6α-receptor/gp130 complex has been used, with modifications, to take into account the slight differences between the LEPR and leptin structures (Peelman *et al.,* 2006; Carpenter *et al.,* 2012). CRH1 domain is not thought to contribute to these functions (Fong *et al.,* 1998) but one residue within the CRH1 has been shown to be important, as a mutation in this locus affects protein stability (Peelman *et al.,* 2006). These domains are found in the extracellular region of IL-6, gp130 and G-CSF and have three binding sites in their extracellular region, (I, II and III) which form a complex with their respective ligands. This has also caused problems when identifying the structure of the receptor and its complex

with leptin, as several regions from both proteins are required. Binding site II occurs between the CRH2 domain of the receptor and the A and C helices of leptin whereas binding site III occurs between the Ig-like domain of the receptor and N-terminus of helix D of leptin. The structural complex with leptin has been inferred from IL6 and G-CSF receptor complexes and is thought to form a hexameric complex like IL-6 (Peelman *et al.,* 2006). Leptin receptor is usually a monomer, which forms a dimer upon binding to leptin (Couturier and Jockers, 2003, Devos *et al.,* 1997, Nakashima *et al.,* 1997), while two leptin molecules are required to form the complex with four molecules of the leptin receptor. However, determination of the crystal structure of leptin and in complex with its receptor is important to clarify this formation of the complex (Carpenter *et al.,* 2012).

For our study it was considered the stronger approach to model shorter specific regions of interest with greatest accuracy using the best that was available, I-TASSER. A limitation of I-TASSER server, suite of programmes was that it was not able to predict the complete LEPR protein structure. Although it may have been possible to have used the individual extracellular domains and predicted them separately finally combining to form an overall 3D structure of the receptor. This would not have been desirable since there are regions within the LEPR structure that are not within a domain, but linking regions which would have been missed; the predicted structure of these are also important in the overall 3D structure of the protein and if these are not determined correctly it could adversely affect the 3D structure prediction. Two of the polymorphisms of interest within the LEPR (Lys109Arg and Ly656Asp) are contained within these regions. Militating against this, the IL6/IL6α-receptor/gp130 complex could have been used as a template and the nine haplotypes concerning the three polymorphisms created separately. However, since there are significant differences between the IL6/IL6α-receptor/gp130 complex and the leptin

and LEPR complex these would need to be taken into consideration (Iserentant *et al.,* 2005; Peelman *et al.,* 2006). The four helices in leptin are much shorter than those within IL6, therefore, the Ig-like domain has to be rotated so that residues that interact in binding site II are closer to each other (Peelman *et al.,* 2005).

As there is a lack of research into the electrostatic potential of this domain, and the effect of the change in residue, our results are novel and interesting but further analysis is needed. Furthermore, research is needed to fully understand the role of the CRH1 domain in the structure of the receptor, its involvement in the binding of leptin and how the Gln223Arg polymorphism affects these two functions. It has been suggested that altered binding of leptin with leptin receptor could result from this polymorphism, reducing leptin signalling which is thought to lead to the loss of function of leptin and leptin resistance leading to the obese phenotype (Anubhuti and Arora, 2008; Myers *et al.,* 2012). It is important to note, however that significant associations within the *LEPR* locus have also been found with two other SNPs, Lys109Arg and Lys656Asn. Therefore even though they are not located within one of the domains in the extracellular region it would still important to determine the effect of these SNPs on the structure of the receptor as it is thought to form a hexameric complex with leptin, and the interactive effect of these three polymorphisms i.e. the eight different possible haplotypes. Previous published research investigating different LEPR haplotypes have found significant associations with obesity (Liu *et al.,* 2004; Qu *et al.,* 2008) but have not included Gln223Arg even though the majority of the associations with obesity have been found with this SNP. Therefore this aspect would be of interest to complete using computer-modeling techniques to identify possible differences in the overall structure of the LEPR and in the hexameric complex with leptin. Furthermore in addition functional studies would also be as important to either confirm or dispute the

findings. The methodology of using a mouse model to 'knock in' human LEPR exon 4 as conducted by Stratigopoulos *et al.,* (2009) could be useful to generate the eight haplotypes of the *LEPR* to investigate their effect on weight gain, lean mass and fat mass in mice and the expression level, despite no associations being found with the Gln223Arg polymorphism. Their finding does support the view that Gln223Arg polymorphism does produce a small effect and additional SNPs may be involved. Therefore, elucidation of the functional effect of this amino acid change would be important to treat or ameliorate leptin resistance in individuals with obesity, as leptin resistance is seen as a key contributing factors in the development of the obese state. A vast number of the population are either overweight or obese, which is expected to rise over the coming years, thus identifying the mechanism involved in leptin resistance would lead to possible drug treatments, targeting those at increased risk.

Prediction of 2D and 3D structures of proteins is important in elucidation of functionality and possible interactions with other proteins and molecules. Advances have been made in computer modelling over the last twenty years and the programs designed have increased reliability in their 3D structures preditions. However, further development is needed to enable accurate determination of the structure of membrane bound proteins and those that contain more than one domain such as the leptin receptor. Further investigation of this receptor will help in the understanding of the complex it forms with its ligand leptin, and in particular the role CRH1 domain has within this, which is the site of the Gln223Arg polymorphism a locus, which has been repeatedly associated with obesity and measures of adiposity. This could ultimately contribute to determining its role in the development of leptin resistance, along with additional polymorphisms in leptin and the leptin receptor, which are known to be important in the development of obesity and sequelae. Our study

has found that one polymorphism is predicted to be involved in additional hydrogen bonding, which may have functional implications in leptin binding and leptin resistance. The combination of both computer modelling and functional studies of the leptin receptor would be expected to be synergistic in the elucidation of the structure-function relationships of the leptin receptor.

5.3.0 Investigation into T2D Risk Alleles with Respect to Fetal Birth Weight

Three maternal T2D susceptibility loci were investigated for an association with fetal birth weight and one SNP in *HHEX-IDE* was significantly associated with lower birth weight, one SNP in *FTO* was associated with higher birth weight while there was no association with the SNP genotyped in *ADCY5*. These results are relevant as it suggests that maternal loci previously associated with T2D and obesity are additionally involved with fetal growth, in South Asians, which further supports the link between size at birth and future risk of developing obesity and T2D in later life. However, it is clear that further investigation is needed to fully understand the role of these SNPs and genes in fetal growth as well as T2D and obesity. With the development of next generation sequencing (NGS), (Davey *et al.*, 2011), it can be used for target region sequencing to study genes of interest rather than the whole genome, to reduce the cost further and reduce the amount of data processing needed (Liu *et al.*, 2012). This would allow the simultaneous investigation of all the genomic regions, where positive associations have been reported, many of which are in high LD. This would be very likely to lead to the identification of a number of new SNPs that could then be investigated further as possible candidate accounting for a functional connection between the maternal genes investigated and fetal birth weight.

5.3.1 Candidate Genes: *HHEX-IDE*, *ADCY5* and *FTO*

5.3.1.1 *HHEX-IDE* Locus

In this study thirty-seven SNPs within the maternal *HHEX-IDE* gene region were investigated and rs12765131 was significantly associated with lower fetal birth weight ($p=0.002$), in 663 South Asian subjects. This SNP was investigated further and it was found that subjects, who were homozygous for the rare allele, were at greater risk of a lower fetal birth weight but still within the normal range. This suggests that there is a recessive effect, with a proposal that the effect is involved in energy regulation due to

previous associations of this locus with T2D. *HHEX-IDE* covers a genomic region of 200kb, which is in high LD. As a result, when the Sequenom iPLEX assay was designed, rs12765131 was a suitable single tag SNP, capturing eight others ($r^2 > 0.7$), which lie across a region of 5 kb. Therefore, this SNP implicates a 5 kb region of the maternal *HHEX-IDE* gene locus, in lowering fetal birth weight. Further investigation is needed, as rs12765131 may not be the causative SNP. The association found could be indirect due to the high LD within this region and the causative SNP/haplotype lies somewhere within this candidate region. Therefore in depth sequencing of the *HHEX-IDE* region in SA subjects would be of use to identify additional polymorphisms that might be the causative SNP.

Copy number polymorphisms overlap *HHEX* and have been found in Asian populations, a loss around the 5′ end of the gene (Conrad *et al.,* 2010) and a gain across a larger region, which covers the whole of the gene (Park *et al.,* 2010). Indian Asian populations were not included in these studies, therefore further analysis of this CNV in that cohort is needed to confirm or refute a contribution to both birth weight and T2D. Researchers have investigated copy number variation in relation to the risk of developing T2D but *HHEX* has not been among those significantly associated (Bae *et al.,* 2011b).

Our association with fetal birth weight is specific to the South Asian population and differs from previous studies in European Caucasians, which found associations with the SNP rs1111875 (Freathy *et al.,* 2009, Andersson *et al.,* 2010) (see section 1.10.1). In our study this SNP was not genotyped directly but it was in LD with rs10882102, which was not found to be significantly associated with birth weight (p=0.131). In addition, in a study by Zhao *et al.,* (2009) two SNPs, rs1111875 and rs7923837, were genotyped in a European American cohort, where no significant association was found with fetal birth weight

(p=0.8315 and p=0.778 respectively). Therefore specific polymorphisms within this locus could influence the regulation of fetal birth weight differently in different populations, which could explain the different SNPs that have been found to be significant. Interestingly, the *HHEX-IDE* locus has not previouslybeen investigated with respect to birth weight in the SA population therefore replication is needed to confirm our results.

The studies performed in European Caucasians analysed one (Freathy *et al.,* 2009; Andersson *et al.,* 2010) or two (Zhao *et al.,* 2009) SNPs within the *HHEX-IDE* region,notably those that had previously been associated with T2D: rs1111875 and rs7923837 (Grarup *et al.,* 2007; Saxena *et al.,* 2007; Sladek *et al.,* 2007; Scott *et al.,* 2007; Zeggini *et al.,* 2007). In contrast, in our study twenty-two SNPs were analysed across 75kb of the *HHEX-IDE* region. The two SNPs, rs1111875 and rs7923837, cover a region of 22 kb therefore compared to the published literature our study investigated more variants across alarger region of *HHEX-IDE* with respect to fetal birth weight. No previous research has been carried out analysing this locus in the SA population therefore our novel study cannot be corroborated by published findings in other cohorts (Freathy *et al.,* 2009; Andersson *et al.,* 2010). However, a limitation to our study was that only twenty-two of the thirty-seven SNPs genotyped were analysed for an association with birth weight. Two SNPs failed and five were uninformative in our population and the typing of the other eight SNPs was not in HWE i.e. data was not correctly called. The copy number variant overlapping the *HHEX* gene may have affected the miscalling of samples. One study found a gain in CNV number in Korean, Japanese and Chinese populations across the whole of the *HHEX* gene region (Park *et al.,* 2010).

Research has shown that the *HHEX-IDE* locus is associated with fetal birth weight but only one study has investigated the maternal genotype specifically, as our study did (Freathy *et al.,* 2009). Interestingly, the Freathy study found no significant association between rs1111875 and fetal birth weight (p=0.5), even when fetal genotype was adjusted for. Similarly, we found no significant association with fetal birth weight with one SNP, rs10882102: which is in LD with the SNP they tested rs1111875. In contrast, we did identify an effect of the maternal genotype on fetal birth weight, with a different SNP (rs12765131), which was significantly associated (p=0.002). We may then be able to narrow the region down if this data was repeated in both SA and European Caucasian cohorts.

The mechanism by which sequence variants within the *HHEX-IDE* gene region might be directly responsible for lower birth weight is yet to be determined. One possibility is through influencing insulin secretion, since insulin is an important intrauterine growth factor. Genetic studies have shown that this locus has been previously associated with T2D in many different populations; European (Sladek *et al.,* 2007; Scott *et al.,* 2007; Saxena *et al.,* 2007; Zeggini *et al.,* 2007; Chauchi *et al.,* 2008; Ruchat *et al.,* 2009; Dupuis *et al.,* 2010), Japanese (Omori *et al.,* 2008; Furukawa *et al.,* 2008), Chinese (Wu *et al.,* 2008; Zhou *et al.,* 2010), South Asian (Chauhan *et al.,* 2008; Ng *et al.,* 2008; Rees *et al.,* 2011) and Korean (Lee *et al.,* 2008). Functional studies looking at *HHEX-IDE* variants have previously been associated with decreased beta cell function (Grarup *et al.,* 2007; Pascoe *et al.,* 2007). Mouse studies have shown that *Hhex* is expressed in the embryo during development of the pancreas and strikingly mice carrying a *Hhex* null mutation lack the ventral pancreas completely (Bort *et al.,* 2004). Therefore variants in *HHEX-IDE* may

influence fetal insulin secretion *in utero* resulting in a decrease in secretion, leading to lower birth weight, as proposed by the 'fetal insulin hypothesis' (see section 1.7.12).

Analysing further data, *HHEX* is in high LD with the genes *IDE* and *KIF11* and a small part of this LD region (namely the *HHEX* gene itself and the SNPs that have been associated with T2D and birth weight) overlaps with a genomic regulatory block that covers 390 kb (Ragvin *et al.,* 2010). These genomic regulatory blocks are chromosomal domains that control expression of developmental regulator genes (termed target genes) and can include unrelated genes (termed bystander genes) that are located on the same chromosome. This is recognised by shared conserved regions in these two types of genes and expression is maintained by an enhancer, which is located in the introns of the bystander gene that regulates the target gene. SNPs either intronic in the bystander gene or intergenic can affect the target gene. *HHEX* is in conserved synteny with its neighbouring gene *EXOC6,* (the genes *IDE* and *KIF11* are outside this conserved synteny block), it suggests the *cis*-regulatory elements regulate *HHEX* i.e. is the target gene, rather than *IDE* and *KIF11.* The SNPs associated with T2D are located within intergenic DNA, between *HHEX* and *EXOC6.* Since *HHEX* is expressed in the pancreas in early development, it is possible that the variants within the LD block can affect regulation of *HHEX* (Ragvin *et al.,* 2009). Further analysis is needed to determine if the SNPs significantly associated with T2D do actually affect regulation.

5.3.1.2 *FTO* Locus

In this study nine SNPs within the maternal *FTO* locus (Figure 5.3.1) were genotyped and three were found to be significantly associated with higher fetal birth weight, with one (rs9939609) surviving correction for multiple testing (p=0.001). Each copy of the A risk

allele increased birth weight by 90.46 g. This indicates that the maternal *FTO* locus has a role in fetal birth weight, which is a novel finding in the SA population. One similar study agrees with our findings and has found a significant association between the fetal genotype of rs1421085 and higher Z-score of BMI (p=0.02) and ponderal index (p=0.005) in a European Caucasian population (Cauchi *et al.,* 2009). The rs1421085 SNP was genotyped in our study but it was not significantly associated with higher birth weight (p=0.75). However, other investigations of offspring rs9939609 genotype in relation to fetal birth weight, birth length and ponderal index did not replicate our findings in European (Elks *et al.,* 2012, Elks *et al.,* 2010, Frayling *et al.,* 2007b, Hakanen *et al.,* 2009, Jess *et al.,* 2008, Kilpelainen *et al.,* 2011b, Labayen *et al.,* 2012, Lawlor *et al.,* 2011, Lopez-Bermejo *et al.,* 2008, Andersson *et al.,* 2010b), or in American (Choh *et al.,* 2011), Chinese (Xi *et al.,* 2010), African (Hennig *et al.,* 2009), and American Indian (Seal *et al.,* 2011) populations. Additional SNPs within the *FTO* locus were similarly found not to be significantly associated with birth weight rs6499640, rs8050136, rs11642776, rs8047473 and rs1558902 (Figure 4.2.1 and Figure 4.2.2) in American and Chinese populations (p>0.05) (Hong *et al.,* 2012, Mei *et al.,* 2012, Mei *et al.,* 2010).

Of these additional SNPs, rs8050136 was genotyped in our study and initially found to be associated with higher birth weight but this result did not survive correction for multiple testing. The lack of replication of the significant association with SNPs in the *FTO* locus and birth weight from our study and published work could be due to the maternal locus being investigated in the first instance as compared to the fetal locus in published research. However, the study by Cauchi *et al.,* (2009) in which an association with rs1421085 at the fetal *FTO* locus was analysed, with respect to fetal BMI suggests that different variants at both the fetal and maternal *FTO* locus maybe involved in influencing birth weight. The

size of the cohort is important to ensure a high statistical power to detect an effect. There were differences in cohort size between these studies, ranging below 1000 (Seal *et al.,* 2011; Lopez-Bermejo *et al.,* 2008; Hakanen *et al.* 2009; Mei *et al.,* 2010; Mei *et al.,* 2012; Choh *et al.,* 2011), between 1,000 and 10,000 (Labayen *et al.,* 2012; Elks *et al.,* 2010; Elks *et al* 2012; Andersson *et al.,* 2010; Hennig *et al.,* 2009; Hong *et al.,* 2012) and more than 10,000 (Kilpelainen *et al.,* 2011; Frayling *et al.,* 2007; Lawlor *et al.,* 2011) which meant that it is possible the effect was not detectable in studies with a smaller cohort size. Ethnicity is additionally important indicating that our results and those by Cauchi *et al.,* (2009) could bepopulation specific with different SNPs within the *FTO* region having a role in fetal birth weight in different populations but further investigation of this locus is needed to determine their effect.

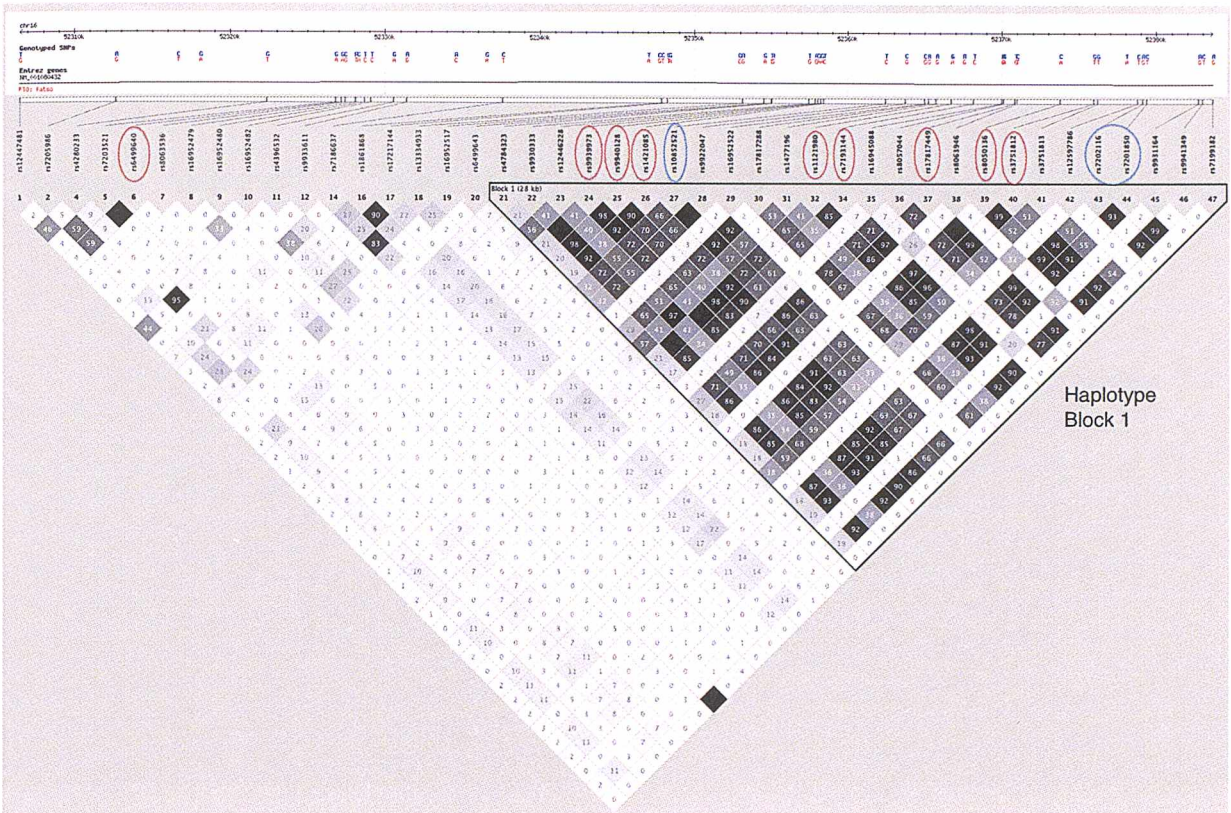


Figure 5.3.1: A map of R^2 across part of the *FTO* region, 52300 kb-52385 kb using HapMap CEU data, which shows one haplotype block. The SNPs highlighted in red are those that have been found to be associated with birth weight in our study and published research; rs6499640 (position 6), rs9939973 (position 24), rs9940128 (investigated position 25), rs1421085 (position 26), rs1121980 (position 32), rs7193144 (position 34), rs17817449 (position 37), rs8050136 (position 39) and rs3751812 (position 40). The SNPs highlighted in blue show the positions of the SNPs rs1558902 and rs99399609, with the former close to the SNP rs10852521 and the latter close to the SNPs rs7202116 and rs7201850.

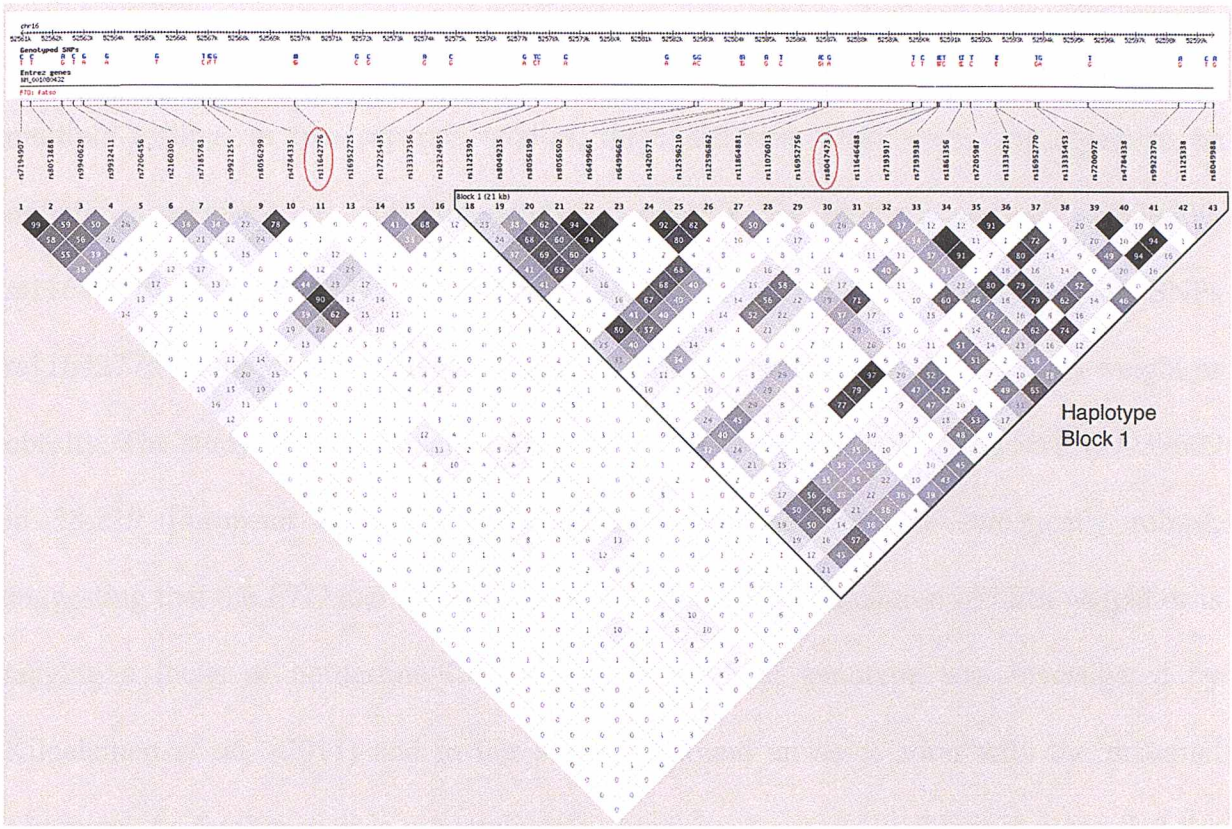


Figure 5.3.2: A map of R^2 across the *FTO* region, 52560kb-52600 kb using HapMap CEU data, which shows one haplotype block. The SNPs highlighted in red are rs11642776 (position 11) and rs8047473 (position 30), which have been studied by Mei *et al.,* (2010) for an association with fetal birth weight.

In our study, an additional SNP rs17817449 was also initially found to be significantly associated with fetal birth weight, however this result did not survive correction for multiple testing. In two separate studies, three different SNPs were also found to be initially associated but did not survive correction for multiple testing, rs9939609, rs11642776 and rs8047473 (Kilpelainen *et al.,* 2011; Mei *et al.,* 2010). The SNPs rs11642776 and rs8047473 have not been previously associated with fetal birth weight or obesity. The study by Kilpelainen *et al.,* (2011) was large, genotyping the SNP rs9939609 in 28,219 European Caucasian subjects, therefore the result is statistically robust, suggesting that the *FTO* loci either has a small effect in the regulation of birth weight or is tagging a locus of borderline LD. However, the fetal genotype was investigated by Kilpelainen *et al.,* (2011) and in our study we found an association with the maternal genotype. As fetal growth is predominately determined maternally it might mean that the fetal *FTO* genotype has little influence in the regulation of fetal growth (Lawlor *et al.,* 2011). Two studies have genotyped both the maternal and offspring rs9939609 locus, but no significant associations were found with either (Lawlor *et al.,* 2011; Seal *et al.,* 2011).

The studies conducted by Lawlor *et al.,*(2011) and Seal *et al.,* (2011) used two different ethnic populations, European Caucasian and American Indian respectively and the population studied in this instance was Gujarati SA. Therefore the difference in the results indicates that the effect of rs9939609 on fetal growth and birth weight could be population specific. The different roles of the *FTO* locus in fetal growth between different populations could be important in the risk of developing obesity, T2D and the metabolic syndrome in later life as ethnic minorities are at increased risk of developing these disease in an obesogenic environment (Jenum *et al.,* 2012, Kurian and Cardarelli, 2007). Fetal growth and birth weight have been found to be an important risk factor in the development of these

complex diseases therefore the identification of different genetic loci influencing birth weight between different ethnic groups could be important vary. The contribution of *FTO* variants could be more influential in the SA population than in European populations, thus explaining the lack of association found.

Within the design of the assay nine SNPs previously associated with BMI were chosen to ascertain whether they were significantly associated with fetal birth weight. These SNPs are within intron one of the *FTO* gene in which there is high LD. Therefore, as this gene covers such a large region further investigation is needed to determine if rs9939609, significantly associated with higher birth weight is the common causal variant that directly has a role in birth weight or if the causal variant is in LD with this SNP. The study by Mei *et al.,* (2010), found an association with two SNPs (rs11642776 and rs8047473) and birth weight and although they did not survive correction for multiple testing, it does suggest that additional variants may also be involved in fetal birth weight. These two SNPs were among thirty designed as tag SNPs across the *FTO* region and are located towards the 3' end of the gene. However, these thirty tag SNPs only capture 57.6% of the common genetic variation with a mean maximum r^2 value of 0.61, which is quite low. Tag SNPs are useful to reduce the number of SNPs to be genotyped particularly if the region is large, however 100% coverage of the region of interest is desirable to ensure that the causal variant is in LD with at least one of the SNPs that is genotyped.

Despite the lack of robust association between variants and fetal birth weight, mRNA expression analysis has indicated that *FTO* has a role in fetal growth and birth weight. *FTO* is expressed in placental tissue, which is related to fetal weight, length and placental weight (Bassols *et al.,* 2010). The placenta expresses high levels of *FTO* and expression

levels are significantly associated with birth weight (p=0.005) and birth length (p=0.0001). This suggests that *FTO* induces fetal growth. However, the expression levels of *FTO* were also related to the number of previous pregnancies; in women with no previous children smaller placentas expressed higher levels of *FTO* and also higher levels were found in larger placentas in women who have previously had children. In addition, the mRNA expression was highest in women who have previously had more than one child ($p<0.05$). As offspring tend to be smaller in women who have not previously had children while offspring are much larger in women that have, it is suggested that *FTO* influences fetal growth in both an indirect and direct mechanism; indirectly through an increase in placental growth and directly through an independent regulation of placental size, possibly through regulating portioning of placenta and fetal growth (Bassols *et al.,* 2009). Therefore the effects of *FTO* could be dependent on the tissue in which *FTO* is expressed (Mei *et al.,* 2010) however it could not be determined whether maternal or fetal expression of FTO contributed to fetal growth (Bassols *et al.,* 2009). This should be explored further, as research into this area is limited. The function of *FTO* is yet to be elucidated and this needs to be investigated further; possible roles include regulation of gene expression (Wu *et al.,* 2010), DNA repair and nucleic acid modification (Gerken *et al.,* 2007; Sanchez-Pulido and Andrade-Navarro, 2007), which maybe important in the regulation of birth weight. The SNPs within intron 1 of *FTO* are within a genomic regulatory block that regulates the gene *IRX3,* which is a transcriptional regulator and expressed during early development in the kidney, hypothalamus, notochord and forebrain (Houweling *et al.,* 2001). Therefore, risk alleles in intron 1 may affect regulation of this gene and affect early development (Ragvin *et al.,* 2009).

It has been suggested that as *FTO* has been associated with BMI and weight and is expressed in adipocytes (Frayling *et al.,* 2007) it could also be influential in gestational weight gain (Lawlor *et al.,* 2011). Both high and low levels of weight gain during pregnancy can influence the growth of the fetus; too little reduces fetal growth and too much increases fetal growth. A part of the gain in weight is adipose tissue, occurring both maternally (between 2 and 6 kg) (Nelson *et al.,* 2010) and within the fetus (Rasmussen *et al.,* 2009). Gain in maternal adipose tissue occurs at the start of the pregnancy until the middle of the second trimester (Nelson *et al.,* 2010), in which the fetus starts to accumulate adipose tissue, around 2.5 g to 4 g per day (Rasmussen *et al.,* 2009). Therefore the expression levels and/or variants within *FTO* could play a role in increasing weight gain, through increase in maternal and/or fetal adipose tissue in pregnancy however when the *FTO* SNP rs9939609, along with three others in *MC4R* (17782313), *TMEM18* (rs6548238) and *GNPDA2* (rs10938397), was investigated there was an inverse association between the combined maternal risk allele score of the four loci with gestational weight gain (Lawlor *et al.,* 2011). The increase in the risk allele score (between zero and eight) was associated with a decrease in weight gain within the first eighteen weeks of pregnancy ($p=0.002$). Offspring risk allele score was not significantly associated with gestational weight gain. Maternal weight gain is highly correlated to maternal gain in adipose tissue (Pearson's correlation = 0.8) therefore this result was surprising and suggests that these loci previously associated with BMI are not involved with gestational weight gain. Gestational weight gain is a complex phenotype, thus environmental factors and different genetic pathways may have more of an influence, than adiposity-related genetic variants (Lawlor *et al.,* 2011). However, many loci have been associated with BMI and weight through GWA studies therefore it is possible that many loci are involved with gestational weight gain and these need to be investigated. Two T2D susceptibility loci have been associated with

greater gestational weight gain (Stuebe *et al.,* 2010) indicating these loci also need to be considered for a role in gaining adipose tissue during pregnancy which affects fetal growth.

5.3.1.3 *ADCY5*

The SNP rs1708067 within the maternal *ADCY5* locus was genotyped within the SA cohort butthe minor allele was not found to be significantly associated with either higher or lower fetal birth weight. However, analysis showed that women with two copies of the G risk allele were found to confer a higher fetal birth weight and those with one copy were found to have lower birth weight. This result differs from previous research, as having the minor allele of this SNP was found to significantly lower fetal birth weight (p=0.004) in European Caucasian populations (Andersson *et al.,* 2010). Interestingly, previous research performed in Filipino and African-descended Moroccan and Turkish populations (Freathy *et al.,* 2010) and SA population (Vasan *et al.,* 2011), which found no association of this allele with fetal birth weight. However, in these two studies the rs1108067 SNP was not genotyped directly, the SNP rs9883204 was used as a proxy due to the high LD between them (r^2=0.75). This could explain the different result but the rs9883204 was found to be the most significantly associated with lower fetal birth weight $p=7x10^{-15}$ in European population (Freathy *et al.,* 2010). Although failure to detect association in *ADCY5* in non-European populations for the SNP mitigates against this locusbeing involved with birth weight in these populations; in our study homozygotes carrying the G allele did seem to show an effect. With respect to the SA population environmental factors may be more influential in regulation of birth weight or more likely alternative genetic loci are involved (Vasan *et al.,* 2011). This could explain the different phenotypes observed between SA and European babies; SA babies are much lighter and smaller and have a higher amount of adipose tissue than lean tissue. However these two SNPs are in a region of high LD,

therefore the true causative SNP still needs further investigation not only in European populations but also other ethnic populations to determine the role of *ADCY5* in fetal growth.

A limitation to the study by Vasan *et al.,* (2011) was that only birth weight was used as a reflection of fetal growth, but the study by Freathy *et al.,* (2010) measured birth weight as well as birth length, birth head circumference and ponderal index to investigate any association with rs9883204, which were also significantly associated, $p=4x10^{-5}$, p=0.030 and p=0.006 respectively. Different measurements of size at birth can reflect fetal growth therefore investigation of these with respect to genetic variants is needed as additional research indicates *ADCY5* may have a role in fetal growth itself (Mook-Kanamori *et al.,* 2011). The SNP rs9883204 has been significantly associated with the fetal growth characteristics abdominal circumference, femur length, and estimated fetal weight, with an estimated combined effect of -16.9 g ($p=4.2x10^{-5}$) for each risk allele, in the 3rd trimester of pregnancy. During this period of time the fetus grows in size and gains weight therefore the *ADCY5* could reduce fetal growth but in an asymmetric manner as the SNP was not associated with head circumference (Mook-Kanamori *et al.,* 2011). Further investigation of the mechanism involved is needed as well as replication in additional studies and ethnic populations.

Surprisingly there is a lack of published research investigating the maternal *ADCY5* locus with regards to fetal birth weight therefore our finding in the SA population is novel in this respect. One study by Freathy *et al.,* (2010) who investigated the fetal ADCY5 locus and found a significant association with birth weight ($5x10^{-15}$), additionally adjusted for the maternal genotype, along with fetal sex, gestational age, where data was available

(n=7,910). Significant associations were still found ($p=5x10^{-6}$), indicating that there is a direct effect between fetal genotype and fetal growth. Despite our results of the maternal genotypenot being significant it suggests that there is a recessive effect of the minor allele to increase fetal birth weight. Therefore there is an opposite effect between fetal and maternal genotype where a regulatory mechanism is involved in controlling fetal growth within normal maternal constraints.

Nonetheless the exact mechanism remains unknown. It has been suggested, that along with variants in the *HHEX-IDE* locus, that variants in *ADCY5* affect insulin secretion *in utero* since *ADCY5* is expressed in pancreatic islets and beta cells (Dupuis *et al.,* 2010; Leech *et al.,* 1999). It has been associated with increased fasting glucose level and increased risk of T2D in adults and children (Rees *et al.,* 2011; Vasan *et al.,* 2011; Windholz *et al.,* 2011; Barker *et al.,* 2011; Dupuis *et al.,* 2010). The fetal genotype could decrease insulin secretion since fetal insulin is a key fetal growth factor and a reduced level could reduce fetal growth, as indicated by lower birth weight (Freathy *et al.,* 2010). This supports the 'fetal insulin hypothesis' (Hattersley and Tooke, 1999). Alternatively additional research suggests that it could be involved with regulation of placental glucose transporter expression (Ogura *et al.,* 2000), the structure and permeability of the materno-fetal placental barrier (Leach, 2002) and the uptake of vitamin B_2 by the placenta (D'Souza *et al.,* 2006). Certainly this seems likely as *ADCY5* encodes adenylate cyclase 5, that catalyses the generation of cAMP, from ATP (Hanoune *et al.,* 1997). cAMP is involved in signal transduction as it activates protein kinase A which regulates ion channels and can affect gene expression by binding to DNA (de Rooij *et al.,* 1998) (see section 1.12.3).

5.3.2 Limitations of Candidate Gene Analysis

Overall, our study is limited in the respect that fetal DNA was not available for each investigation of the genetic loci, *HHEX-IDE*, *FTO* and *ADCY5*, and so the direct effect of fetal genotype on fetal birth weight could not be determined. Previous investigations of these loci on birth weight have used fetal genotype and found significant associations with lower fetal birth weight in European Caucasians (Freathy *et al.,* 2009; Zhao *et al.,* 2009; Andersson *et al.,* 2010; Freathy *et al.,* 2010). Thus investigation of these loci in the SA population would be important to determine if the fetal genotypes are also influential in birth weight as well as the maternal. Identifying the influence of fetal genotype is important as previous research in two additional T2D susceptibility loci have shown differing results. Fetal genotype in *GCK* was not significantly associated (Weedon *et al.,* 2006) while fetal genotype in *TCF7L2* was associated with increased fetal birth weight (Freathy *et al.,* 2007). Furthermore our phenotype data of the offspring lacks additional birth characteristics that reflect fetal growth: birth length, head circumference and ponderal index. Birth length and ponderal index have previously been associated with variants in *ADCY5* and *CDKAL1* (Freathy *et al.,* 2010; Andersson *et al.,*2010) but not *HHEX-IDE* (Andersson *et al.,* 2010) or *FTO* (Labayen *et al.,* 2012; Choh *et al.,* 2011).

Adjusting for covariates is also important when analysing birth weight as environmental and maternal factors can increase or decrease birth weight. Babies born <36 weeks of age were removed from the analysis, while those >37 weeks of age were included as covariates along with parity and fetal sex. In addition mothers with gestational diabetes were not included as this can increase fetal birth weight. However, smoking is another environmental factor that can decrease birth weight but this information was not available

to us. Including this in future studies would be useful to control for this factor so the effect of genetic loci on birth weight can be determined more rigorously.

A limitation of the *HHEX-IDE* genotyping assay was that the thirty-seven tag SNPs/SNPs in the assay covered half of the *HHEX-IDE* locus, therefore 100 kb upstream of the *HHEX* gene including SNPs in the adjacent genes, namely Kinesin family member 11 (*KIF11*) and insulin degrading enzyme (*IDE*) were not covered in this study (Figure 1.10.1; Figure 3.2.1). There was an error in the design of the iPLEX assay as at the time of assay design the HapMap data indicated that *KIF11* and *IDE* were downstream of *HHEX* and our original design would have captured the genetic information in the whole region. IDE is involved in the degradation of insulin (Affholter *et al.,* 1990, Seta and Roth, 1997) and variants in *IDE* have previously been associated with T2D and impaired glucose and insulin homeostasis (Karamohamed *et al.,*2003; Gu *et al.,*2004; Kwak *et al.,*2008) therefore additional variants within this region may also be implicated in birth weight. Alternatively an unidentified causal variant could lie within this region but upstream of *HHEX*. However, published associations with SNPs rs1111875, rs7923837, rs1544210, rs10509646 and rs5015480 with birth weight and T2D (Garaup *et al.,*2007; Saxena *et al.,*2007; Zeggini *et al.,*2007; Cauchi *et al.,*2008; van Vilet-Ostaptchouk *et al.,*2008; Horikoshi *et al.,*2007; Horikawa *et al.,*2008; Freathy *et al.,*2009; Zhao *et al.,*2009; Andersson *et al.,*2010), are all located downstream of the *HHEX* gene (Table 3.2.1; Figure 3.2.3), a region that this project did cover. Future work should explore the region upstream of *HHEX*, which covers *KIF11* and *IDE*, to ascertain whether this region is associated with fetal birth weight and/or T2D.

5.3.3 Environmental and Genetic Factors in Birth Weight Variation and Risk of Obesity, T2D and the Metabolic Disease in Adulthood

Our data demonstrates the complex interaction between birth weight and obesity and T2D, as variants in susceptibility loci for both complex diseases were significantly associated with normal birth weight variation. However, the risk of obesity, T2D and the metabolic syndrome develops over the course of an individual's lifetime. Research has implicated over 100 genes with the obese phenotype (see section 1.5.4) and investigations into adult onset and childhood obesity have found notable differences (see section 1.5.4.1). Certain candidate loci seem to have a more pronounced effect on weight gain during childhood compared to adulthood (Zhao *et al.,* 2009; Elks *et al.,* 2010; den Hoed *et al.,* 2010; Hardy *et al.,* 2010). The combined effects of eleven variants, previously associated with adult BMI and some of which lacked association with fetal birth weight, showed strong significant associations with weight gain in childhood (den Hoed *et al.,* 2010; Elks *et al.,* 2012). This suggests that weight gain in early childhood is influenced more strongly by the effects of certain genes (Beardsall *et al.,* 2009). Weight gain during early postnatal life and during early infancy, namely catch-up growth, has a detrimental effect, increasing the risk of obesity and its sequelae in later life (see section 1.7.8 and 1.7.9). Therefore a longitudinal study design would be very useful to assess weight gain during the early postnatal period, through childhood and, ideally, on into adulthood, to investigate whether risk alleles increase risks further in the SA population specifically.

In addition, within a longitudinal study design, it would be very useful to conduct an assessment of infant feeding regimes as research has shown that infants bottle-fed with formula have a higher BMI (Bergmann *et al.,* 2003, Li *et al.,* 2012, Ong *et al.,* 2002) and a

higher fat free mass at three months, eight months and twelve months (Gale *et al.,* 2012) and are at increased risk of being overweight and obese in childhood (Gillman *et al.,* 2001, Gooze *et al.,* 2011). Furthermore, infants who are given breast milk directly, as compared to breast milk given via a bottle, and in a baby led manner, (where they regulate their own feeding), have a better satiety responsiveness from six months of age (Brown and Lee, 2012) and in early childhood, at three and six years of age (Disantis *et al.,* 2011). Current advice from midwives and antenatal clinics is that breast milk is given to the infant rather than formula based products, due to it being nutritionally better and in terms of increasing immunity. Differences in protein and fat content have been noted, where breast milk has a lower protein content and infant formula has a higher content. This is thought to be responsible for the increased growth that occurs in formula fed infants. A protective effect of breast milk has also been observed in relation to obesity and type 2 diabetes, namely that it induces a lower plasma insulin level leading to decreased fat storage and preventing early adipocyte development (Oddy, 2012). However, catch up growth has long-term consequences on the adult health of the offspring, therefore a change in the content of infant formula, to one with a lower protein and fat content and in the current advice of midwives and antenatal clinics to encourage rapid weight gain in newborns up to normal weight needs to occur to reduce this effect.

It is becoming increasingly evident that epigenetics may be just as important as genetics and the environment in the risk of obesity, T2D and the metabolic syndrome. Interestingly, the fetal epigenome is established during critical periods of development (see section 1.7.11.1). The adaptations the fetus undergoes, in terms of changes in metabolism and organ structure, as a result of a poor nutritional environment have been attributed to modifications in the epigenome (Herrera *et al.,* 2011) and this evidence has formed the

basis of the 'fetal origins of adult disease' hypothesis and PAR theory. These changes could permanently change the epigenome and alter fetal growth in the first instance to aid in survival in the postnatal environment, but then be detrimental by increasing the risk of obesity, T2D and the metabolic syndrome in later life (Waterland and Jirtle, 2004; Sullivan and Grove, 2010). As the current environment is obesogenic, maternal overnutrition has additionally been found to be important in these critical periods of fetal growth (Dyer and Rosenfeld, 2011, Wu *et al.,* 2004, Sullivan and Grove, 2010). A high fat diet can change the expression of leptin (Milagro *et al.,* 2009), *POMC* (Plagemann *et al.,* 2009) and *MC4R* (Widiker *et al.,* 2010). These are important in energy regulation and involved in the leptin-melanocortin pathway (see section 1.9.6). Therefore, as expression of these genes can change they may have a role in establishment of the fetal epigenome, which can have long-term consequences.

It has been suggested that hypothalamic imprinting is important in the later development of obesity where changes in the hormone levels including leptin and insulin are likely to have a role (Bouret, 2009). Mouse and rat models have supported this theory, where an increase in postnatal leptin levels affects the development of the hypothalamic neuronal connections (Djiane and Attig, 2008). In humans, there is likely to be a different mechanism involved, as circulating leptin levels do not increase until the third trimester when hypothalamic development occurs (Davidowa and Plagemann, 2000). This signifies that the nutritional environment as a whole interacts with the epigenome of the fetus and has long-term effects, which is of interest considering that many imprinted genes are important in fetal growth (*IGF2, IGF2R, PHLDA2* and *PLAGL1* (McMinn *et al.,* 2006; Adkins *et al.,* 2009; Frost and Moore *et al.,* 2010)), fetoplacental development (*ZNF331* and *SLC22A18* (Lambertini *et al.,* 2012)) and metabolic function or regulation of these factors (Herrera *et*

al., 2011). Changes in expression of these genes have been associated with birth weight (Johnston *et al.,* 2003; Salas *et al.,* 2004; Katu *et al.,* 2007; Adkins *et al.,* 2009; Koutsaki *et al.,* 2011; Petry *et al.,* 2011) and loss of imprinting of some genes and from some parental origins results in the obese phenotype, as found in PWS (see section 1.5.3). Imprinted genes are expressed in fetal tissues and placenta, which is vital to provide the fetus with nutrients and are essential for normal placental development. Interestingly, some loci, *IGF2*, *LEP*, *GNASAS*, *ABCA1,* can change their epigenetic status in response to environmental cues to adapt the placental phenotype, indicating they have a major role in fetal programming and its consequences of which (Fowden *et al.,* 2011, Tobi *et al.,* 2009). As a result it is of great interest to investigate the changes in the human fetal epigenome that may result directly from environmental cues such as maternal overnutrition and undernutrition and to then assess the resultant risk of developing obesity and its sequelae. However, establishment of this type of study is difficult for many reasons, not least due to ethical considerations, therefore animal models might need to be used instead.

Genetic factors influencing the phenotype of birth weight were investigated, specifically T2D sequence variants to determine any association with size at birth, as it is a commonly used measurement to reflect fetal growth. However, it would be of interest to determine whether additional loci, previously associated with birth weight, are also associated with fetal growth characteristics throughout gestation, as with *ADCY5* and *CCNL/LERK1* (Mook-Kanamori *et al.,* 2011) (see section 1.7.11). Previous associations have been found between the *ADCY5, HHEX-IDE and CDKAL1* loci and birth length, birth head circumference and ponderal index (Freathy *et al.,* 2010; Andersson *et al.,* 2010), suggesting these loci may have a role in fetal growth. It would further our understanding of these loci in birth weight if the study by Mook-Kanamori *et al.,* (2011) was replicated in

additional European Caucasians cohorts as well as ethnic minorities including the SA population as they are understudied with respect to obesity and T2D but they have a higher propensity to develop these diseases in the obesogenic environment. SAs are known to have a different body composition, compared to European Caucasians, termed 'thin-fat' phenotype (see section 1.2.8). This phenotype is thought to develop *in utero* as SA babies are 0.8 kg lighter at birth compared to European Caucasians (Yajnik *et al* 2002; Modi *et al.,* 2009) and they also have a significantly higher visceral and subcutaneous adipose tissue level but relatively less muscle and viscera (Modi *et al.,* 2009). Therefore identifying genetic loci that are involved in birth weight would further our understanding of their interactions with the environment, in all ethnicities.

5.4 Conclusion

Sequence variations at specific candidate genes were investigated with respect to the related phenotypes: birth weight, obesity and T2D using genetic and functional approaches. Our results demonstrate an underlying genetic connection between birth weight and obesity and its sequelae. Molecular cloning and computer modelling of three variants within the leptin receptor revealed possible structural and functional effects. Further work is needed to complete the long-term aims of this project: to fully determine the effect of the three variants on the binding avidity between leptin and the LEPR both using molecular modelling and cloning and expression of the protein variants. Multiple obesity and T2D susceptibility loci have been identified but we are only just beginning to unravel their involvement in energy regulation and metabolism. Our data, in addition to previous research, shows that they can exert their effect during different and multiple periods of an individual's lifetime, during fetal life, childhood or adulthood. In addition, even subtle structural changes may produce a functional effect and influence the obese

phenotype during these points in time. Therefore, functional, genetic and expression approaches are of interest to examine their effects to ultimately more fully understand the genetic aetiology of birth weight and obesity and its complex interactions with the current obesogenic environment. Recent advances in sequencing the human genome, through NGS, is an important aspect in this regard, not least as it has the potential to be developed for personalised genomics, where risk factors can be identified at an individual level leading to tailored treatments and/or preventative measures.

References

ABDULLA, J., KOBER, L., ABILDSTROM, S. Z., CHRISTENSEN, E., JAMES, W. P. & TORP-PEDERSEN, C. 2008. Impact of obesity as a mortality predictor in high-risk patients with myocardial infarction or chronic heart failure: a pooled analysis of five registries. *Eur Heart J,* 29, 594-601.

ADAIR, L. S. & COLE, T. J. 2003. Rapid child growth raises blood pressure in adolescent boys who were thin at birth. *Hypertension,* 41, 451-6.

ADKINS, R. M., SOMES, G., MORRISON, J. C., HILL, J. B., WATSON, E. M., MAGANN, E. F. & KRUSHKAL, J. 2010. Association of birth weight with polymorphisms in the IGF2, H19, and IGF2R genes. *Pediatric Research,* 68, 429-34.

AFFHOLTER, J. A., CASCIERI, M. A., BAYNE, M. L., BRANGE, J., CASARETTO, M. & ROTH, R. A. 1990. Identification of residues in the insulin molecule important for binding to insulin-degrading enzyme. *Biochemistry,* 29, 7727-33.

AL-ATTAR 2008. Association between the FTO rs9939609 polymorphism and the metabolic syndrome in a non-Caucasian multi-ethnic sample. *Cardiovascular Diabetology,* 13.

ALBERTI, K. G., ECKEL, R. H., GRUNDY, S. M., ZIMMET, P. Z., CLEEMAN, J. I., DONATO, K. A., FRUCHART, J. C., JAMES, W. P., LORIA, C. M. & SMITH, S. C., JR. 2009. Harmonizing the metabolic syndrome: a joint interim statement of the International Diabetes Federation Task Force on Epidemiology and Prevention; National Heart, Lung, and Blood Institute; American Heart Association; World Heart Federation; International Atherosclerosis Society; and International Association for the Study of Obesity. *Circulation,* 120, 1640-5.

ALBERTI, K. G., ZIMMET, P. & SHAW, J. 2006. Metabolic syndrome--a new world-wide definition. A Consensus Statement from the International Diabetes Federation. *Diabet Med,* 23, 469-80.

ALKAN, C., COE, B. P. & EICHLER, E. E. 2011. Genome structural variation discovery and genotyping. *Nature Reviews Genetics,* 12, 363-76.

ALLISON, D. B., HESHKA, S., NEALE, M. C., LYKKEN, D. T. & HEYMSFIELD, S. B. 1994. A genetic analysis of relative weight among 4,020 twin pairs, with an emphasis on sex effects. *Health Psychol,* 13, 362-5.

ALMIND, K., MANIERI, M., SIVITZ, W. I., CINTI, S. & KAHN, C. R. 2007. Ectopic brown adipose tissue in muscle provides a mechanism for differences in risk of metabolic syndrome in mice. *Proc Natl Acad Sci U S A,* 104, 2366-71.

ALTSCHUL, S. F., MADDEN, T. L., SCHAFFER, A. A., ZHANG, J., ZHANG, Z., MILLER, W. & LIPMAN, D. J. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res,* 25, 3389-402.

ANAZAWA, S., ATSUMI, Y. & MATSUOKA, K. 2003. Low birth weight and development of type 2 diabetes in a Japanese population. *Diabetes Care,* 26, 2210-1.

ANDERSEN, L. G., ANGQUIST, L., ERIKSSON, J. G., FORSEN, T., GAMBORG, M., OSMOND, C., BAKER, J. L. & SORENSEN, T. I. 2010. Birth weight, childhood body mass index and risk of coronary heart disease in adults: combined historical cohort studies. *PLoS One,* 5, e14126.

ANDERSSON, E. A. 2010. Type 2 diabetes risk alleles near ADCY5, CDKAL1 and HHEX-IDE are associated with reduced birth weight. *Diabetologia,* 53, 1908-1916.

ANDERSSON, E. A., PILGAARD, K., PISINGER, C., HARDER, M. N., GRARUP, N., FAERCH, K., POULSEN, P., WITTE, D. R., JORGENSEN, T., VAAG, A., HANSEN, T. & PEDERSEN, O. 2010a. Type 2 diabetes risk alleles near ADCY5, CDKAL1 and HHEX-IDE are associated with reduced birthweight. *Diabetologia,* 53, 1908-16.

ANDERSSON, E. A., PILGAARD, K., PISINGER, C., HARDER, M. N., GRARUP, N., FAERCH, K., SANDHOLT, C., POULSEN, P., WITTE, D. R., JORGENSEN, T., VAAG, A., PEDERSEN, O. & HANSEN, T. 2010b. Do gene variants influencing adult adiposity affect birth weight? A population-based study of 24 loci in 4,744 Danish individuals. *PLoS One,* 5, e14190.

ANGEL-CHAVEZ, L. I., TENE-PEREZ, C. E. & CASTRO, E. 2012. Leptin receptor gene K656N polymorphism is associated with low body fat levels and elevated high-density cholesterol levels in Mexican children and adolescents. *Endocr Res,* 37, 124-34.

ANUBHUTI & ARORA, S. 2008. Leptin and its metabolic interactions: an update. *Diabetes Obes Metab,* 10, 973-93.

APOSTOLIDOU, S., ABU-AMERO, S., O'DONOGHUE, K., FROST, J., OLAFSDOTTIR, O., CHAVELE, K. M., WHITTAKER, J. C., LOUGHNA, P., STANIER, P. & MOORE, G. E. 2007. Elevated placental expression of the imprinted PHLDA2 gene is associated with low birth weight. *J Mol Med (Berl),* 85, 379-87.

ARAVIND, L. & KOONIN, E. V. 2001. The DNA-repair protein AlkB, EGL-9, and leprecan define new families of 2-oxoglutarate- and iron-dependent dioxygenases. *Genome Biol,* 2, RESEARCH0007.

ARCH, J. R., STOCK, M. J. & TRAYHURN, P. 1998. Leptin resistance in obese humans: does it exist and what does it mean? *Int J Obes Relat Metab Disord,* 22, 1159-63.

ASHWELL, M., GUNN, P. & GIBSON, S. 2012. Waist-to-height ratio is a better screening tool than waist circumference and BMI for adult cardiometabolic risk factors: systematic review and meta-analysis. *Obesity Reviews,* 13, 275-86.

ASHWELL, M. A. 1998. The Ashwell Shape Chart - a public health approach to the metabolic risks of obesity. *International Journal of Obesity,* 22 (Supplement 3), S213.

AUGUSTYNIAK, R. A., SINGH, K., ZELDES, D., SINGH, M. & ROSSI, N. F. 2010. Maternal protein restriction leads to hyperresponsiveness to stress and salt-sensitive hypertension in male offspring. *Am J Physiol Regul Integr Comp Physiol,* 298, R1375-82.

BAE, J., PARK, J. H., PARK, Y. K., KIM, J. Y., LEE, S. W. & PARK, S. W. 2011a. Changes in the distribution of maternal age and parity and increasing trends in the low birth weight rate in Korea between 1995 and 2005. *J Prev Med Public Health,* 44, 111-7.

BAE, J. S., CHEONG, H. S., KIM, J. H., PARK, B. L., KIM, J. H., PARK, T. J., KIM, J. Y., PASAJE, C. F., LEE, J. S., PARK, Y. J., PARK, M., PARK, C., KOH, I., CHUNG, Y. J., LEE, J. Y. & SHIN, H. D. 2011b. The genetic effect of copy number variations on the risk of type 2 diabetes in a Korean population. *PLoS One,* 6, e19091.

BAIRD, J., OSMOND, C., MACGREGOR, A., SNIEDER, H., HALES, C. N. & PHILLIPS, D. I. 2001. Testing the fetal origins hypothesis in twins: the Birmingham twin study. *Diabetologia,* 44, 33-9.

BALABAN, G. & SILVA, G. A. 2004. [Protective effect of breastfeeding against childhood obesity]. *J Pediatr (Rio J),* 80, 7-16.

BANCI, M., SACCUCCI, P., DOFCACI, A., SANSONI, I., MAGRINI, A., BOTTINI, E. & GLORIA-BOTTINI, F. 2009. Birth weight and coronary artery disease. The effect of gender and diabetes. *Int J Biol Sci,* 5, 244-8.

BANERJI, M. A., FARIDI, N., ATLURI, R., CHAIKEN, R. L. & LEBOVITZ, H. E. 1999. Body composition, visceral fat, leptin, and insulin resistance in Asian Indian men. *J Clin Endocrinol Metab,* 84, 137-44.

BANIK, N. D., KRISHNA, R., MANE, S. I. & RAJ, L. 1967. A study of birth weight of Indian infants and its relationship to sex, period of gestation, maternal age, parity and socio-economic classes. *Indian J Med Res,* 55, 1378-86.

BANKS, A. S., DAVIS, S. M., BATES, S. H. & MYERS, M. G., JR. 2000. Activation of downstream signals by the long form of the leptin receptor. *Journal of Biological Chemistry,* 275, 14563-72.

BARKER, A., SHARP, S. J., TIMPSON, N. J., BOUATIA-NAJI, N., WARRINGTON, N. M., KANONI, S., BEILIN, L. J., BRAGE, S., DELOUKAS, P., EVANS, D. M., GRONTVED, A., HASSANALI, N., LAWLOR, D. A., LECOEUR, C., LOOS, R. J., LYE, S. J., MCCARTHY, M. I., MORI, T. A., NDIAYE, N. C., NEWNHAM, J. P., NTALLA, I., PENNELL, C. E., ST POURCAIN, B., PROKOPENKO, I., RING, S. M., SATTAR, N., VISVIKIS-SIEST, S., DEDOUSSIS, G. V., PALMER, L. J., FROGUEL, P., SMITH, G. D., EKELUND, U., WAREHAM, N. J. & LANGENBERG, C. 2011. Association of genetic Loci with glucose levels in childhood and adolescence: a meta-analysis of over 6,000 children. *Diabetes,* 60, 1805-12.

BARKER, D. J., BULL, A. R., OSMOND, C. & SIMMONDS, S. J. 1990. Fetal and placental size and risk of hypertension in adult life. *BMJ,* 301, 259-62.

BARKER, D. J., GLUCKMAN, P. D., GODFREY, K. M., HARDING, J. E., OWENS, J. A. & ROBINSON, J. S. 1993. Fetal nutrition and cardiovascular disease in adult life. *Lancet,* 341, 938-41.

BARKER, D. J., OSMOND, C., KAJANTIE, E. & ERIKSSON, J. G. 2009. Growth and chronic disease: findings in the Helsinki Birth Cohort. *Ann Hum Biol,* 36, 445-58.

BARKER, M., ROBINSON, S., OSMOND, C. & BARKER, D. J. 1997. Birth weight and body fat distribution in adolescent girls. *Archives of Disease in Childhood,* 77, 381-3.

BARNETT, A. H., DIXON, A. N., BELLARY, S., HANIF, M. W., O'HARE J, P., RAYMOND, N. T. & KUMAR, S. 2006. Type 2 diabetes and cardiovascular risk in the UK south Asian community. *Diabetologia,* 49, 2234-46.

BARREIRA, T. V., STAIANO, A. E., HARRINGTON, D. M., HEYMSFIELD, S. B., SMITH, S. R., BOUCHARD, C. & KATZMARZYK, P. T. 2012. Anthropometric correlates of total body fat, abdominal adiposity, and cardiovascular disease risk factors in a biracial sample of men and women. *Mayo Clin Proc,* 87, 452-60.

BASSOLS, J., PRATS-PUIG, A., VAZQUEZ-RUIZ, M., GARCIA-GONZALEZ, M. M., MARTINEZ-PASCUAL, M., AVELLI, P., MARTINEZ-MARTINEZ, R., FABREGA, R., COLOMER-VIROSTA, C., SORIANO-RODRIGUEZ, P., DIAZ, M., DE ZEGHER, F.,

IBANEZ, L. & LOPEZ-BERMEJO, A. 2010. Placental FTO expression relates to fetal growth. *Int J Obes (Lond),* 34, 1365-70.

BASTARD, J. P., MAACHI, M., LAGATHU, C., KIM, M. J., CARON, M., VIDAL, H., CAPEAU, J. & FEVE, B. 2006. Recent advances in the relationship between obesity, inflammation, and insulin resistance. *European Cytokine Network,* 17, 4-12.

BAUER, F., ELBERS, C. C., ADAN, R. A., LOOS, R. J., ONLAND-MORET, N. C., GROBBEE, D. E., VAN VLIET-OSTAPTCHOUK, J. V., WIJMENGA, C. & VAN DER SCHOUW, Y. T. 2009. Obesity genes identified in genome-wide association studies are associated with adiposity measures and potentially with nutrient-specific food preference. *American Journal of Clinical Nutrition,* 90, 951-9.

BAVDEKAR, A., YAJNIK, C. S., FALL, C. H., BAPAT, S., PANDIT, A. N., DESHPANDE, V., BHAVE, S., KELLINGRAY, S. D. & JOGLEKAR, C. 1999. Insulin resistance syndrome in 8-year-old Indian children: small at birth, big at 8 years, or both? *Diabetes,* 48, 2422-9.

BAYOL, S. A., SIMBI, B. H., FOWKES, R. C. & STICKLAND, N. C. 2010. A maternal "junk food" diet in pregnancy and lactation promotes nonalcoholic Fatty liver disease in rat offspring. *Endocrinology,* 151, 1451-61.

BEARDSALL, K., ONG, K. K., MURPHY, N., AHMED, M. L., ZHAO, J. H., PEETERS, M. W. & DUNGER, D. B. 2009. Heritability of childhood weight gain from birth and risk markers for adult metabolic disease in prepubertal twins. *J Clin Endocrinol Metab,* 94, 3708-13.

BECKERS, S., ZEGERS, D., VAN GAAL, L. F. & VAN HUL, W. 2011. Replication of the SH2B1 rs7498665 association with obesity in a Belgian study population. *Obes Facts,* 4, 473-7.

BELFIORE, F., IANNELLO, S. & RABUAZZO, A. M. 1979. Insulin resistance in obesity: a critical analysis at enzyme level. A review. *Int J Obes,* 3, 301-23.

BELKACEMI, L., NELSON, D. M., DESAI, M. & ROSS, M. G. 2010. Maternal undernutrition influences placental-fetal development. *Biol Reprod,* 83, 325-31.

BELL, C. G., MEYRE, D., SAMSON, C., BOYLE, C., LECOEUR, C., TAUBER, M., JOURET, B., JAQUET, D., LEVY-MARCHAL, C., CHARLES, M. A., WEILL, J., GIBSON, F., MEIN, C. A., FROGUEL, P. & WALLEY, A. J. 2005a. Association of melanin-concentrating hormone receptor 1 5' polymorphism with early-onset extreme obesity. *Diabetes,* 54, 3049-55.

BELL, C. G., WALLEY, A. J. & FROGUEL, P. 2005b. The genetics of human obesity. *Nature Reviews Genetics,* 6, 221-34.

BELTOWSKI, J., G, W. J., GORNY, D. & MARCINIAK, A. 2002. Human leptin administered intraperitoneally stimulates natriuresis and decreases renal medullary Na+, K+-ATPase activity in the rat -- impaired effect in dietary-induced obesity. *Med Sci Monit,* 8, BR221-9.

BEN ALI, S., KALLEL, A., SEDIRI, Y., FTOUHI, B., FEKI, M., SLIMENE, H., JEMAA, R. & KAABACHI, N. 2009. LEPR p.Q223R Polymorphism influences plasma leptin levels and body mass index in Tunisian obese patients. *Arch Med Res,* 40, 186-90.

BENNETT, R. G., DUCKWORTH, W. C. & HAMEL, F. G. 2000. Degradation of amylin by insulin-degrading enzyme. *Journal of Biological Chemistry,* 275, 36621-5.

BERG, A. H. & SCHERER, P. E. 2005. Adipose tissue, inflammation, and cardiovascular disease. *Circ Res,* 96, 939-49.

BERGMANN, K. E., BERGMANN, R. L., VON KRIES, R., BOHM, O., RICHTER, R., DUDENHAUSEN, J. W. & WAHN, U. 2003. Early determinants of childhood overweight and adiposity in a birth cohort study: role of breast-feeding. *Int J Obes Relat Metab Disord,* 27, 162-72.

BERRAHO, M., EL ACHHAB, Y., BENSLIMANE, A., EL RHAZI, K., CHIKRI, M. & NEJJARI, C. 2012. Hypertension and type 2 diabetes: a cross-sectional study in Morocco (EPIDIAM Study). *Pan Afr Med J,* 11, 52.

BIRD, A. 2007. Perceptions of epigenetics. *Nature,* 447, 396-8.

BJORGE, T., ENGELAND, A., TVERDAL, A. & SMITH, G. D. 2008. Body mass index in adolescence in relation to cause-specific mortality: a follow-up of 230,000 Norwegian adolescents. *Am J Epidemiol,* 168, 30-7.

BLAKEMORE, A. I. & FROGUEL, P. 2008. Is obesity our genetic legacy? *J Clin Endocrinol Metab,* 93, S51-6.

BLAKEMORE, A. I. & FROGUEL, P. 2010. Investigation of Mendelian forms of obesity holds out the prospect of personalized medicine. *Ann N Y Acad Sci,* 1214, 180-9.

BLANGY, A., LANE, H. A., D'HERIN, P., HARPER, M., KRESS, M. & NIGG, E. A. 1995. Phosphorylation by p34cdc2 regulates spindle association of human Eg5, a kinesin-related motor essential for bipolar spindle formation in vivo. *Cell,* 83, 1159-69.

BLICKSTEIN, I. 2003. Motherhood at or beyond the edge of reproductive age. *Int J Fertil Womens Med,* 48, 17-24.

BLOOMFIELD, F. H., OLIVER, M. H., GIANNOULIAS, C. D., GLUCKMAN, P. D., HARDING, J. E. & CHALLIS, J. R. 2003. Brief undernutrition in late-gestation sheep programs the hypothalamic-pituitary-adrenal axis in adult offspring. *Endocrinology,* 144, 2933-40.

BOBBIONI-HARSCH, E., PATAKY, Z., MAKOUNDOU, V., KOZAKOVA, M., DEKKER, J. & GOLAY, A. 2009. Fat distribution influences the cardio-metabolic profile in a clinically healthy European population. *Eur J Clin Invest,* 39, 1055-64.

BODARY, P. F., WESTRICK, R. J., WICKENHEISER, K. J., SHEN, Y. & EITZMAN, D. T. 2002. Effect of leptin on arterial thrombosis following vascular injury in mice. *JAMA,* 287, 1706-9.

BODURTHA, J. N., MOSTELLER, M., HEWITT, J. K., NANCE, W. E., EAVES, L. J., MOSKOWITZ, W. B., KATZ, S. & SCHIEKEN, R. M. 1990. Genetic analysis of anthropometric measures in 11-year-old twins: the Medical College of Virginia Twin Study. *Pediatric Research,* 28, 1-4.

BOESGAARD, T. W., GRARUP, N., JORGENSEN, T., BORCH-JOHNSEN, K., HANSEN, T. & PEDERSEN, O. 2010. Variants at DGKB/TMEM195, ADRA2A, GLIS3 and C2CD4B loci are associated with reduced glucose-stimulated beta cell function in middle-aged Danish people. *Diabetologia,* 53, 1647-55.

BOGUE, C. W., GANEA, G. R., STURM, E., IANUCCI, R. & JACOBS, H. C. 2000. Hex expression suggests a role in the development and function of organs derived from foregut endoderm. *Dev Dyn,* 219, 84-9.

BORT, R., MARTINEZ-BARBERA, J. P., BEDDINGTON, R. S. & ZARET, K. S. 2004. Hex homeobox gene-dependent tissue positioning is required for organogenesis of the ventral pancreas. *Development,* 131, 797-806.

BOUBRED, F., DANIEL, L., BUFFAT, C., FEUERSTEIN, J. M., TSIMARATOS, M., OLIVER, C., DIGNAT-GEORGE, F., LELIEVRE-PEGORIER, M. & SIMEONI, U. 2009. Early postnatal overfeeding induces early chronic renal dysfunction in adult male rats. *Am J Physiol Renal Physiol,* 297, F943-51.

BOUCHARD, C. 2008. Gene-environment interactions in the etiology of obesity: defining the fundamentals. *Obesity (Silver Spring),* 16 Suppl 3, S5-S10.

BOURET, S. G. 2009. Early life origins of obesity: role of hypothalamic programming. *J Pediatr Gastroenterol Nutr,* 48 Suppl 1, S31-8.

BRAMBILLA, P., CRINO, A., BEDOGNI, G., BOSIO, L., CAPPA, M., CORRIAS, A., DELVECCHIO, M., DI CANDIA, S., GARGANTINI, L., GRECHI, E., IUGHETTI, L., MUSSA, A., RAGUSA, L., SACCO, M., SALVATONI, A., CHIUMELLO, G. & GRUGNI, G. 2011. Metabolic syndrome in children with Prader-Willi syndrome: the effect of obesity. *Nutr Metab Cardiovasc Dis,* 21, 269-76.

BRANUM, A. M. & SCHOENDORF, K. C. 2002. Changing patterns of low birthweight and preterm birth in the United States, 1981-98. *Paediatr Perinat Epidemiol,* 16, 8-15.

BROWN, A. & LEE, M. 2012. Breastfeeding during the first year promotes satiety responsiveness in children aged 18-24 months. *Pediatr Obes,* 7, 382-90.

BROWNING, L. M., MUGRIDGE, O., DIXON, A. K., AITKEN, S. W., PRENTICE, A. M. & JEBB, S. A. 2011. Measuring abdominal adipose tissue: comparison of simpler methods with MRI. *Obes Facts,* 4, 9-15.

BUCHBINDER, S., BARTSCH, U., MULLER, M., ZORN, M., NAWROTH, P. P. & SCHILLING, T. 2011. A novel missense mutation T101N in the melanocortin-4 receptor gene associated with obesity. *Genet Mol Res,* 10, 1042-9.

BUCK, J., SINCLAIR, M. L., SCHAPAL, L., CANN, M. J. & LEVIN, L. R. 1999. Cytosolic adenylyl cyclase defines a unique signaling molecule in mammals. *Proc Natl Acad Sci U S A,* 96, 79-84.

BULIK, C. M. & ALLISON, D. B. 2001. The genetic epidemiology of thinness. *Obesity Reviews,* 2, 107-15.

BURKE, J. P., FORSGREN, J., PALUMBO, P. J., BAILEY, K. R., DESAI, J., DEVLIN, H. & LEIBSON, C. L. 2004. Association of birth weight and type 2 diabetes in Rochester, Minnesota. *Diabetes Care,* 27, 2512-3.

BYBERG, L., MCKEIGUE, P. M., ZETHELIUS, B. & LITHELL, H. O. 2000. Birth weight and the insulin resistance syndrome: association of low birth weight with truncal obesity and raised plasminogen activator inhibitor-1 but not with abdominal obesity or plasma lipid disturbances. *Diabetologia,* 43, 54-60.

CALKINS, K. & DEVASKAR, S. U. 2011. Fetal origins of adult disease. *Curr Probl Pediatr Adolesc Health Care,* 41, 158-76.

CALLE, E. E., RODRIGUEZ, C., WALKER-THURMOND, K. & THUN, M. J. 2003. Overweight, obesity, and mortality from cancer in a prospectively studied cohort of U.S. adults. *N Engl J Med,* 348, 1625-38.

CALLE, E. E., THUN, M. J., PETRELLI, J. M., RODRIGUEZ, C. & HEATH, C. W., JR. 1999. Body-mass index and mortality in a prospective cohort of U.S. adults. *N Engl J Med,* 341, 1097-105.

CARDON, L. R., CARMELLI, D., FABSITZ, R. R. & REED, T. 1994. Genetic and environmental correlations between obesity and body fat distribution in adult male twins. *Hum Biol,* 66, 465-79.

CAREY, D. G., NGUYEN, T. V., CAMPBELL, L. V., CHISHOLM, D. J. & KELLY, P. 1996. Genetic influences on central abdominal fat: a twin study. *Int J Obes Relat Metab Disord,* 20, 722-6.

CARLSSON, S., PERSSON, P. G., ALVARSSON, M., EFENDIC, S., NORMAN, A., SVANSTROM, L., OSTENSON, C. G. & GRILL, V. 1999. Low birth weight, family history of diabetes, and glucose intolerance in Swedish middle-aged men. *Diabetes Care,* 22, 1043-7.

CARNELL, S., HAWORTH, C. M., PLOMIN, R. & WARDLE, J. 2008. Genetic influence on appetite in children. *Int J Obes (Lond),* 32, 1468-73.

CARNELL, S. & WARDLE, J. 2008. Appetitive traits and child obesity: measurement, origins and implications for intervention. *Proc Nutr Soc,* 67, 343-55.

CARPENTER, B., HEMSWORTH, G. R., WU, Z., MAAMRA, M., STRASBURGER, C. J., ROSS, R. J. & ARTYMIUK, P. J. 2012. Structure of the human obesity receptor leptin-binding domain reveals the mechanism of leptin antagonism by a monoclonal antibody. *Structure,* 20, 487-97.

CARROLL, J. F., CHIAPA, A. L., RODRIQUEZ, M., PHELPS, D. R., CARDARELLI, K. M., VISHWANATHA, J. K., BAE, S. & CARDARELLI, R. 2008. Visceral fat, waist circumference, and BMI: impact of race/ethnicity. *Obesity (Silver Spring),* 16, 600-7.

CATALANO, P. M., PRESLEY, L., MINIUM, J. & HAUGUEL-DE MOUZON, S. 2009. Fetuses of obese mothers develop insulin resistance in utero. *Diabetes Care,* 32, 1076-80.

CAUCHI, S., PROENCA, C., CHOQUET, H., GAGET, S., DE GRAEVE, F., MARRE, M., BALKAU, B., TICHET, J., MEYRE, D., VAXILLAIRE, M. & FROGUEL, P. 2008. Analysis of novel risk loci for type 2 diabetes in a general French population: the D.E.S.I.R. study. *J Mol Med (Berl),* 86, 341-8.

CAUCHI, S., STUTZMANN, F., CAVALCANTI-PROENCA, C., DURAND, E., POUTA, A., HARTIKAINEN, A. L., MARRE, M., VOL, S., TAMMELIN, T., LAITINEN, J., GONZALEZ-IZQUIERDO, A., BLAKEMORE, A. I., ELLIOTT, P., MEYRE, D., BALKAU, B., JARVELIN, M. R. & FROGUEL, P. 2009. Combined effects of MC4R and FTO common genetic variants on obesity in European general populations. *J Mol Med,* 87, 537-46.

CEDDIA, R. B. 2005. Direct metabolic regulation in skeletal muscle and fat tissue by leptin: implications for glucose and fatty acids homeostasis. *Int J Obes (Lond),* 29, 1175-83.

CHAGNON, Y. C., WILMORE, J. H., BORECKI, I. B., GAGNON, J., PERUSSE, L., CHAGNON, M., COLLIER, G. R., LEON, A. S., SKINNER, J. S., RAO, D. C. & BOUCHARD, C. 2000. Associations between the leptin receptor gene and adiposity in middle-aged Caucasian males from the HERITAGE family study. *J Clin Endocrinol Metab,* 85, 29-34.

CHALDAKOV, G. N., FIORE, M., STANKULOV, I. S., HRISTOVA, M., ANTONELLI, A., MANNI, L., GHENEV, P. I., ANGELUCCI, F. & ALOE, L. 2001. NGF, BDNF, leptin, and mast cells in human coronary atherosclerosis and metabolic syndrome. *Arch Physiol Biochem,* 109, 357-60.

CHAMBERS, J. C., EDA, S., BASSETT, P., KARIM, Y., THOMPSON, S. G., GALLIMORE, J. R., PEPYS, M. B. & KOONER, J. S. 2001. C-reactive protein, insulin resistance, central obesity, and coronary heart disease risk in Indian Asians from the United Kingdom compared with European whites. *Circulation,* 104, 145-50.

CHAMBERS, J. C., ELLIOTT, P., ZABANEH, D., ZHANG, W., LI, Y., FROGUEL, P., BALDING, D., SCOTT, J. & KOONER, J. S. 2008. Common genetic variation near MC4R is associated with waist circumference and insulin resistance. *Nature Genetics,* 40, 716-8.

CHAMBERS, J. C., OBEID, O. A., REFSUM, H., UELAND, P., HACKETT, D., HOOPER, J., TURNER, R. M., THOMPSON, S. G. & KOONER, J. S. 2000. Plasma homocysteine concentrations and risk of coronary heart disease in UK Indian Asian and European men. *Lancet,* 355, 523-7.

CHAMBERS, J. C., ZHANG, W., ZABANEH, D., SEHMI, J., JAIN, P., MCCARTHY, M. I., FROGUEL, P., RUOKONEN, A., BALDING, D., JARVELIN, M. R., SCOTT, J., ELLIOTT, P. & KOONER, J. S. 2009. Common genetic variation near melatonin receptor MTNR1B contributes to raised plasma glucose and increased risk of type 2 diabetes among Indian Asians and European Caucasians. *Diabetes,* 58, 2703-8.

CHANDALIA, M., LIN, P., SEENIVASAN, T., LIVINGSTON, E. H., SNELL, P. G., GRUNDY, S. M. & ABATE, N. 2007. Insulin resistance and body fat distribution in South Asian men compared to Caucasian men. *PLoS One,* 2, e812.

CHAUHAN, G., SPURGEON, C. J., TABASSUM, R., BHASKAR, S., KULKARNI, S. R., MAHAJAN, A., CHAVALI, S., KUMAR, M. V., PRAKASH, S., DWIVEDI, O. P., GHOSH, S., YAJNIK, C. S., TANDON, N., BHARADWAJ, D. & CHANDAK, G. R. 2010. Impact of common variants of PPARG, KCNJ11, TCF7L2, SLC30A8, HHEX, CDKN2A, IGF2BP2, and CDKAL1 on the risk of type 2 diabetes in 5,164 Indians. *Diabetes,* 59, 2068-74.

CHEHAB, F. F., LIM, M. E. & LU, R. 1996. Correction of the sterility defect in homozygous obese female mice by treatment with the human recombinant leptin. *Nature Genetics,* 12, 318-20.

CHEN, A., KLEBANOFF, M. A. & BASSO, O. 2009a. Pre-pregnancy body mass index change between pregnancies and preterm birth in the following pregnancy. *Paediatr Perinat Epidemiol,* 23, 207-15.

CHEN, J. H., MARTIN-GRONERT, M. S., TARRY-ADKINS, J. & OZANNE, S. E. 2009b. Maternal protein restriction affects postnatal growth and the expression of key proteins involved in lifespan regulation in mice. *PLoS One,* 4, e4950.

CHEUNG, C. Y., TSO, A. W., CHEUNG, B. M., XU, A., ONG, K. L., FONG, C. H., WAT, N. M., JANUS, E. D., SHAM, P. C. & LAM, K. S. 2010. Obesity susceptibility genetic variants identified from recent genome-wide association studies: implications in a chinese population. *J Clin Endocrinol Metab,* 95, 1395-403.

CHEUNG, O. 2007. The impact of fat distribution on the severity of non-alcoholic fatty liver disease and metabolic syndrome. *Hepatology,* 46, 1091 - 1100.

CHIU, K. C., CHU, A., CHUANG, L. M. & SAAD, M. F. 2004. Association of leptin receptor polymorphism with insulin resistance. *Eur J Endocrinol,* 150, 725-9.

CHOH, A. C., CURRAN, J. E., ODEGAARD, A. O., NAHHAS, R. W., CZERWINSKI, S. A., BLANGERO, J., TOWNE, B. & DEMERATH, E. W. 2011. Differences in the heritability of growth and growth velocity during infancy and associations with FTO variants. *Obesity (Silver Spring),* 19, 1847-54.

CHRISTOFORIDIS, A., DIMITRIADOU, M., PAPADOPOLOU, E., STILPNOPOULOU, D., KATZOS, G. & ATHANASSIOU-METAXA, M. 2011. Defining overweight and obesity among Greek children living in Thessaloniki: International versus local reference standards. *Hippokratia,* 15, 141-6.

CHUA, S. C., JR., WHITE, D. W., WU-PENG, X. S., LIU, S. M., OKADA, N., KERSHAW, E. E., CHUNG, W. K., POWER-KEHOE, L., CHUA, M., TARTAGLIA, L. A. & LEIBEL, R. L. 1996. Phenotype of fatty due to Gln269Pro mutation in the leptin receptor (Lepr). *Diabetes,* 45, 1141-3.

CHURCH, C., LEE, S., BAGG, E. A., MCTAGGART, J. S., DEACON, R., GERKEN, T., LEE, A., MOIR, L., MECINOVIC, J., QUWAILID, M. M., SCHOFIELD, C. J., ASHCROFT, F. M. & COX, R. D. 2009. A mouse model for the metabolic effects of the human fat mass and obesity associated FTO gene. *PLoS Genet,* 5, e1000599.

CINTI, S. 2012. The adipose organ at a glance. *Dis Model Mech,* 5, 588-94.

CLARK, P. M., ATTON, C., LAW, C. M., SHIELL, A., GODFREY, K. & BARKER, D. J. 1998. Weight gain in pregnancy, triceps skinfold thickness, and blood pressure in offspring. *Obstet Gynecol,* 91, 103-7.

CLAUSSON, B., LICHTENSTEIN, P. & CNATTINGIUS, S. 2000. Genetic influence on birthweight and gestational length determined by studies in offspring of twins. *BJOG,* 107, 375-81.

CLEMENT, K., VAISSE, C., LAHLOU, N., CABROL, S., PELLOUX, V., CASSUTO, D., GOURMELEN, M., DINA, C., CHAMBAZ, J., LACORTE, J. M., BASDEVANT, A., BOUGNERES, P., LEBOUC, Y., FROGUEL, P. & GUY-GRAND, B. 1998. A mutation in the human leptin receptor gene causes obesity and pituitary dysfunction. *Nature,* 392, 398-401.

COLE, T. J., BELLIZZI, M. C., FLEGAL, K. M. & DIETZ, W. H. 2000. Establishing a standard definition for child overweight and obesity worldwide: international survey. *BMJ,* 320, 1240-3.

COLLINS, F. S. 1992. Positional cloning: let's not call it reverse anymore. *Nature Genetics,* 1, 3-6.

CONRAD, D. F., PINTO, D., REDON, R., FEUK, L., GOKCUMEN, O., ZHANG, Y., AERTS, J., ANDREWS, T. D., BARNES, C., CAMPBELL, P., FITZGERALD, T., HU, M., IHM, C. H., KRISTIANSSON, K., MACARTHUR, D. G., MACDONALD, J. R., ONYIAH, I., PANG, A. W., ROBSON, S., STIRRUPS, K., VALSESIA, A., WALTER, K., WEI, J., TYLER-SMITH, C., CARTER, N. P., LEE, C., SCHERER, S. W. & HURLES, M. E. 2010. Origins and functional impact of copy number variation in the human genome. *Nature,* 464, 704-12.

CONSIDINE, R. V., SINHA, M. K., HEIMAN, M. L., KRIAUCIUNAS, A., STEPHENS, T. W., NYCE, M. R., OHANNESIAN, J. P., MARCO, C. C., MCKEE, L. J., BAUER, T.

L. & *ET AL.* 1996. Serum immunoreactive-leptin concentrations in normal-weight and obese humans. *N Engl J Med,* 334, 292-5.

COOKE, J. P. & OKA, R. K. 2002. Does leptin cause vascular disease? *Circulation,* 106, 1904-5.

CORVALAN, C., GREGORY, C. O., RAMIREZ-ZEA, M., MARTORELL, R. & STEIN, A. D. 2007. Size at birth, infant, early and later childhood growth and adult body composition: a prospective study in a stunted population. *Int J Epidemiol,* 36, 550-7.

COTE, M., MAURIEGE, P., BERGERON, J., ALMERAS, N., TREMBLAY, A., LEMIEUX, I. & DESPRES, J. P. 2005. Adiponectinemia in visceral obesity: impact on glucose tolerance and plasma lipoprotein and lipid levels in men. *J Clin Endocrinol Metab,* 90, 1434-9.

COTSAPAS, C., PROKUNINA-OLSSON, L., WELCH, C., SAXENA, R., WEAVER, C., USHER, N., GUIDUCCI, C., BONAKDAR, S., TURNER, N., LACROIX, B. & HALL, J. L. 2010. Expression analysis of loci associated with type 2 diabetes in human tissues. *Diabetologia,* 53, 2334-9.

COUPE, B., DELAMAIRE, E., HOEBLER, C., GRIT, I., EVEN, P., FROMENTIN, G., DARMAUN, D. & PARNET, P. 2011. Hypothalamus integrity and appetite regulation in low birth weight rats reared artificially on a high-protein milk formula. *J Nutr Biochem,* 22, 956-63.

COUTURIER, C. & JOCKERS, R. 2003. Activation of the leptin receptor by a ligand-induced conformational change of constitutive receptor dimers. *Journal of Biological Chemistry,* 278, 26604-11.

COUTURIER, C., SARKIS, C., SERON, K., BELOUZARD, S., CHEN, P., LENAIN, A., CORSET, L., DAM, J., VAUTHIER, V., DUBART, A., MALLET, J., FROGUEL, P., ROUILLE, Y. & JOCKERS, R. 2007. Silencing of OB-RGRP in mouse hypothalamic arcuate nucleus increases leptin receptor signaling and prevents diet-induced obesity. *Proc Natl Acad Sci U S A,* 104, 19476-81.

CRIPPS, R. L. & OZANNE, S. E. 2006. The relationship between poor growth rate and increased risk of Type 2 diabetes, insulin resistance and obesity. *Expert Rev Pharmacoecon Outcomes Res,* 6, 79-86.

CROSSROW, N. 2004. Race/ethnic issues in obesity and obesity-related comorbidities. *Journal of Clinical Endocrinology & Metabolism,* 89, 2590 - 2594.

CROWTHER, N. J., CAMERON, N., TRUSLER, J. & GRAY, I. P. 1998. Association between poor glucose tolerance and rapid post natal weight gain in seven-year-old children. *Diabetologia,* 41, 1163-7.

CURHAN, G. C., CHERTOW, G. M., WILLETT, W. C., SPIEGELMAN, D., COLDITZ, G. A., MANSON, J. E., SPEIZER, F. E. & STAMPFER, M. J. 1996. Birth weight and adult hypertension and obesity in women. *Circulation,* 94, 1310-5.

D'SOUZA, V. M., FORAKER, A. B., FREE, R. B., RAY, A., SHAPIRO, P. S. & SWAAN, P. W. 2006. cAMP-Coupled riboflavin trafficking in placental trophoblasts: a dynamic and ordered process. *Biochemistry,* 45, 6095-104.

DABELEA, D., PETTITT, D. J., HANSON, R. L., IMPERATORE, G., BENNETT, P. H. & KNOWLER, W. C. 1999. Birth weight, type 2 diabetes, and insulin resistance in Pima Indian children and young adults. *Diabetes Care,* 22, 944-50.

DALTON, M., CAMERON, A. J., ZIMMET, P. Z., SHAW, J. E., JOLLEY, D., DUNSTAN, D. W. & WELBORN, T. A. 2003. Waist circumference, waist-hip ratio and body mass index and their correlation with cardiovascular disease risk factors in Australian adults. *J Intern Med,* 254, 555-63.

DAVEY, J. W., HOHENLOHE, P. A., ETTER, P. D., BOONE, J. Q., CATCHEN, J. M. & BLAXTER, M. L. 2011. Genome-wide genetic marker discovery and genotyping using next-generation sequencing. *Nature Reviews Genetics,* 12, 499-510.

DAVIDOWA, H. & PLAGEMANN, A. 2000. Different responses of ventromedial hypothalamic neurons to leptin in normal and early postnatally overfed rats. *Neurosci Lett,* 293, 21-4.

DAY, F. R. & LOOS, R. J. 2011. Developments in obesity genetics in the era of genome-wide association studies. *J Nutrigenet Nutrigenomics,* 4, 222-38.

DE LUIS, D. A., ALLER, R. & IZAOLA, O. 2006. Resting energy expenditure and insulin resitance in obese patients, differences in women and men. *Eur Rev Med Pharmacol Sci,* 10, 285-9.

DE LUIS, D. A., GONZALEZ SAGRADO, M., ALLER, R., IZAOLA, O. & CONDE, R. 2008. Influence of Lys656Asn polymorphism of the leptin receptor gene on insulin resistance in nondiabetic obese patients. *J Diabetes Complications,* 22, 199-204.

DE ROOIJ, J., ZWARTKRUIS, F. J., VERHEIJEN, M. H., COOL, R. H., NIJMAN, S. M., WITTINGHOFER, A. & BOS, J. L. 1998. Epac is a Rap1 guanine-nucleotide-exchange factor directly activated by cyclic AMP. *Nature,* 396, 474-7.

DE ROOIJ, S. R., PAINTER, R. C., PHILLIPS, D. I., OSMOND, C., MICHELS, R. P., GODSLAND, I. F., BOSSUYT, P. M., BLEKER, O. P. & ROSEBOOM, T. J. 2006a. Impaired insulin secretion after prenatal exposure to the Dutch famine. *Diabetes Care,* 29, 1897-901.

DE ROOIJ, S. R., PAINTER, R. C., ROSEBOOM, T. J., PHILLIPS, D. I., OSMOND, C., BARKER, D. J., TANCK, M. W., MICHELS, R. P., BOSSUYT, P. M. & BLEKER, O. P. 2006b. Glucose tolerance at age 58 and the decline of glucose tolerance in comparison with age 50 in people prenatally exposed to the Dutch famine. *Diabetologia,* 49, 637-43.

DEFER, N., BEST-BELPOMME, M. & HANOUNE, J. 2000. Tissue specificity and physiological relevance of various isoforms of adenylyl cyclase. *Am J Physiol Renal Physiol,* 279, F400-16.

DEFRIES, J. C. & FULKER, D. W. 1988. Multiple regression analysis of twin data: etiology of deviant scores versus individual differences. *Acta Genet Med Gemellol (Roma),* 37, 205-16.

DEN HOED, M., EKELUND, U., BRAGE, S., GRONTVED, A., ZHAO, J. H., SHARP, S. J., ONG, K. K., WAREHAM, N. J. & LOOS, R. J. 2010. Genetic susceptibility to obesity and related traits in childhood and adolescence: influence of loci identified by genome-wide association studies. *Diabetes,* 59, 2980-8.

DENNISON, B. A., EDMUNDS, L. S., STRATTON, H. H. & PRUZEK, R. M. 2006. Rapid infant weight gain predicts childhood overweight. *Obesity (Silver Spring),* 14, 491-9.

DERIAZ, O., TREMBLAY, A. & BOUCHARD, C. 1993. Non linear weight gain with long term overfeeding in man. *Obes Res,* 1, 179-85.

DESPRES, J. P. & LEMIEUX, I. 2006. Abdominal obesity and metabolic syndrome. *Nature,* 444, 881-7.

DEURENBERG, P., YAP, M. & VAN STAVEREN, W. A. 1998. Body mass index and percent body fat: a meta analysis among different ethnic groups. *Int J Obes Relat Metab Disord,* 22, 1164-71.

DEVOS, R., GUISEZ, Y., VAN DER HEYDEN, J., WHITE, D. W., KALAI, M., FOUNTOULAKIS, M. & PLAETINCK, G. 1997. Ligand-independent dimerization of the extracellular domain of the leptin receptor and determination of the stoichiometry of leptin binding. *Journal of Biological Chemistry,* 272, 18304-10.

DIAS, N. F., FERNANDES, A. E., MELO, M. E., REINHARDT, H. L., CERCATO, C., VILLARES, S. M., HALPERN, A. & MANCINI, M. C. 2012. Lack of mutations in the leptin receptor gene in severely obese children. *Arq Bras Endocrinol Metabol,* 56, 178-83.

DIETZ, W. H. 1998. Childhood weight affects adult morbidity and mortality. *Journal of Nutrition,* 128, 411S-414S.

DINA, C., MEYRE, D., GALLINA, S., DURAND, E., KORNER, A., JACOBSON, P., CARLSSON, L. M., KIESS, W., VATIN, V., LECOEUR, C., DELPLANQUE, J., VAILLANT, E., PATTOU, F., RUIZ, J., WEILL, J., LEVY-MARCHAL, C., HORBER, F., POTOCZNA, N., HERCBERG, S., LE STUNFF, C., BOUGNERES, P., KOVACS, P., MARRE, M., BALKAU, B., CAUCHI, S., CHEVRE, J. C. & FROGUEL, P. 2007. Variation in FTO contributes to childhood obesity and severe adult obesity. *Nature Genetics,* 39, 724-6.

DIPLAS, A. I., LAMBERTINI, L., LEE, M. J., SPERLING, R., LEE, Y. L., WETMUR, J. & CHEN, J. 2009. Differential expression of imprinted genes in normal and IUGR human placentas. *Epigenetics,* 4, 235-40.

DISANTIS, K. I., COLLINS, B. N., FISHER, J. O. & DAVEY, A. 2011. Do infants fed directly from the breast have improved appetite regulation and slower growth during early childhood compared with infants fed from a bottle? *Int J Behav Nutr Phys Act,* 8, 89.

DJIANE, J. & ATTIG, L. 2008. Role of leptin during perinatal metabolic programming and obesity. *J Physiol Pharmacol,* 59 Suppl 1, 55-63.

DUBERN, B. & CLEMENT, K. 2012. Leptin and leptin receptor-related monogenic obesity. *Biochimie,* 94, 2111-5.

DUCKWORTH, W. C., HAMEL, F. G., BENNETT, R., RYAN, M. P. & ROTH, R. A. 1990. Human red blood cell insulin-degrading enzyme and rat skeletal muscle insulin protease share antigenic sites and generate identical products from insulin. *Journal of Biological Chemistry,* 265, 2984-7.

DULLOO, A. G., JACQUET, J., SEYDOUX, J. & MONTANI, J. P. 2006. The thrifty 'catch-up fat' phenotype: its impact on insulin sensitivity during growth trajectories to obesity and metabolic syndrome. *Int J Obes (Lond),* 30 Suppl 4, S23-35.

DUNGER, D. B., ONG, K. K., HUXTABLE, S. J., SHERRIFF, A., WOODS, K. A., AHMED, M. L., GOLDING, J., PEMBREY, M. E., RING, S., BENNETT, S. T. & TODD, J. A. 1998. Association of the INS VNTR with size at birth. ALSPAC Study Team. Avon Longitudinal Study of Pregnancy and Childhood. *Nature Genetics,* 19, 98-100.

DUPUIS, J., LANGENBERG, C., PROKOPENKO, I., SAXENA, R., SORANZO, N., JACKSON, A. U., WHEELER, E., GLAZER, N. L., BOUATIA-NAJI, N., GLOYN, A. L., LINDGREN, C. M., MAGI, R., MORRIS, A. P., RANDALL, J., JOHNSON, T., ELLIOTT, P., RYBIN, D., THORLEIFSSON, G., STEINTHORSDOTTIR, V., HENNEMAN, P., GRALLERT, H., DEHGHAN, A., HOTTENGA, J. J., FRANKLIN, C. S., NAVARRO, P., SONG, K., GOEL, A., PERRY, J. R., EGAN, J. M., LAJUNEN, T., GRARUP, N., SPARSO, T., DONEY, A., VOIGHT, B. F., STRINGHAM, H. M., LI, M., KANONI, S., SHRADER, P., CAVALCANTI-PROENCA, C., KUMARI, M., QI, L., TIMPSON, N. J., GIEGER, C., ZABENA, C., ROCHELEAU, G., INGELSSON, E., AN, P., O'CONNELL, J., LUAN, J., ELLIOTT, A., MCCARROLL, S. A., PAYNE, F., ROCCASECCA, R. M., PATTOU, F., SETHUPATHY, P., ARDLIE, K., ARIYUREK, Y., BALKAU, B., BARTER, P., BEILBY, J. P., BEN-SHLOMO, Y., BENEDIKTSSON, R., BENNETT, A. J., BERGMANN, S., BOCHUD, M., BOERWINKLE, E., BONNEFOND, A., BONNYCASTLE, L. L., BORCH-JOHNSEN, K., BOTTCHER, Y., BRUNNER, E., BUMPSTEAD, S. J., CHARPENTIER, G., CHEN, Y. D., CHINES, P., CLARKE, R., COIN, L. J., COOPER, M. N., CORNELIS, M., CRAWFORD, G., CRISPONI, L., DAY, I. N., DE GEUS, E. J., DELPLANQUE, J., DINA, C., ERDOS, M. R., FEDSON, A. C., FISCHER-ROSINSKY, A., FOROUHI, N. G., FOX, C. S., FRANTS, R., FRANZOSI, M. G., GALAN, P., GOODARZI, M. O., GRAESSLER, J., GROVES, C. J., GRUNDY, S., GWILLIAM, R., GYLLENSTEN, U., HADJADJ, S., *et al.* 2010. New genetic loci implicated in fasting glucose homeostasis and their impact on type 2 diabetes risk. *Nature Genetics,* 42, 105-16.

DYER, J. S. & ROSENFELD, C. R. 2011. Metabolic imprinting by prenatal, perinatal, and postnatal overnutrition: a review. *Semin Reprod Med,* 29, 266-76.

EBRAHIM, S., KINRA, S., BOWEN, L., ANDERSEN, E., BEN-SHLOMO, Y., LYNGDOH, T., RAMAKRISHNAN, L., AHUJA, R. C., JOSHI, P., DAS, S. M., MOHAN, M., DAVEY SMITH, G., PRABHAKARAN, D. & REDDY, K. S. 2010. The effect of rural-to-urban migration on obesity and diabetes in India: a cross-sectional study. *PLoS Med,* 7, e1000268.

ECHWALD, S. M., SORENSEN, T. D., SORENSEN, T. I., TYBJAERG-HANSEN, A., ANDERSEN, T., CHUNG, W. K., LEIBEL, R. L. & PEDERSEN, O. 1997. Amino acid variants in the human leptin receptor: lack of association to juvenile onset obesity. *Biochem Biophys Res Commun,* 233, 248-52.

EDWARDS, L. J. & MCMILLEN, I. C. 2002a. Impact of maternal undernutrition during the periconceptional period, fetal number, and fetal sex on the development of the hypothalamo-pituitary adrenal axis in sheep during late gestation. *Biol Reprod,* 66, 1562-9.

EDWARDS, L. J. & MCMILLEN, I. C. 2002b. Periconceptional nutrition programs development of the cardiovascular system in the fetal sheep. *Am J Physiol Regul Integr Comp Physiol,* 283, R669-79.

EINFELD, S. L., KAVANAGH, S. J., SMITH, A., EVANS, E. J., TONGE, B. J. & TAFFE, J. 2006. Mortality in Prader-Willi syndrome. *Am J Ment Retard,* 111, 193-8.

ELIJAH, I. E., BRANSKI, L. K., FINNERTY, C. C. & HERNDON, D. N. 2011. The GH/IGF-1 system in critical illness. *Best Pract Res Clin Endocrinol Metab,* 25, 759-67.

ELKS, C. E., LOOS, R. J., HARDY, R., WILLS, A. K., WONG, A., WAREHAM, N. J., KUH, D. & ONG, K. K. 2012. Adult obesity susceptibility variants are associated with greater childhood weight gain and a faster tempo of growth: the 1946 British Birth Cohort Study. *American Journal of Clinical Nutrition,* 95, 1150-6.

ELKS, C. E., LOOS, R. J., SHARP, S. J., LANGENBERG, C., RING, S. M., TIMPSON, N. J., NESS, A. R., DAVEY SMITH, G., DUNGER, D. B., WAREHAM, N. J. & ONG, K. K. 2010. Genetic markers of adult obesity risk are associated with greater early infancy weight gain and growth. *PLoS Med,* 7, e1000284.

ELLARD, S. & COLCLOUGH, K. 2006. Mutations in the genes encoding the transcription factors hepatocyte nuclear factor 1 alpha (HNF1A) and 4 alpha (HNF4A) in maturity-onset diabetes of the young. *Hum Mutat,* 27, 854-69.

ENDO, K., YANAGI, H., HIRANO, C., HAMAGUCHI, H., TSUCHIYA, S. & TOMURA, S. 2000. Association of Trp64Arg polymorphism of the beta3-adrenergic receptor gene and no association of Gln223Arg polymorphism of the leptin receptor gene in Japanese schoolchildren with obesity. *Int J Obes Relat Metab Disord,* 24, 443-9.

ENRIORI, P. J., EVANS, A. E., SINNAYAH, P., JOBST, E. E., TONELLI-LEMOS, L., BILLES, S. K., GLAVAS, M. M., GRAYSON, B. E., PERELLO, M., NILLNI, E. A., GROVE, K. L. & COWLEY, M. A. 2007. Diet-induced obesity causes severe but reversible leptin resistance in arcuate melanocortin neurons. *Cell Metab,* 5, 181-94.

ERIKSSON, J., FORSEN, T., TUOMILEHTO, J., OSMOND, C. & BARKER, D. 2000. Fetal and childhood growth and hypertension in adult life. *Hypertension,* 36, 790-4.

ERIKSSON, J., FORSEN, T., TUOMILEHTO, J., OSMOND, C. & BARKER, D. 2001a. Size at birth, childhood growth and obesity in adult life. *Int J Obes Relat Metab Disord,* 25, 735-40.

ERIKSSON, J. G., FORSEN, T., TUOMILEHTO, J., OSMOND, C. & BARKER, D. J. 2001b. Early growth and coronary heart disease in later life: longitudinal study. *BMJ,* 322, 949-53.

ERIKSSON, J. G., FORSEN, T., TUOMILEHTO, J., OSMOND, C. & BARKER, D. J. 2003. Early adiposity rebound in childhood and risk of Type 2 diabetes in adult life. *Diabetologia,* 46, 190-4.

ERIKSSON, J. G., FORSEN, T., TUOMILEHTO, J., WINTER, P. D., OSMOND, C. & BARKER, D. J. 1999. Catch-up growth in childhood and death from coronary heart disease: longitudinal study. *BMJ,* 318, 427-31.

ESPOSITO, K., NICOLETTI, G. & GIUGLIANO, D. 2002. Obesity, cytokines and endothelial dysfunction: a link for the raised cardiovascular risk associated with visceral obesity. *J Endocrinol Invest,* 25, 646-9.

EZZAHIR, N., ALBERTI, C., DEGHMOUN, S., ZACCARIA, I., CZERNICHOW, P., LEVY-MARCHAL, C. & JAQUET, D. 2005. Time course of catch-up in adiposity influences adult anthropometry in individuals who were born small for gestational age. *Pediatric Research,* 58, 243-7.

FAGOT-CAMPAGNA, A. 2000. Emergence of type 2 diabetes mellitus in children: epidemiological evidence. *J Pediatr Endocrinol Metab,* 13 Suppl 6, 1395-402.

FAIRBROTHER, U. L., TANKO, L. B., WALLEY, A. J., CHRISTIANSEN, C., FROGUEL, P. & BLAKEMORE, A. I. F. 2007. Leptin receptor genotype at Gln223Arg is associated with body composition, BMD, and vertebral fracture in postmenopausal Danish women. *Journal of Bone and Mineral Research,* 22, 544-550.

FAKHRAI-RAD, H., NIKOSHKOV, A., KAMEL, A., FERNSTROM, M., ZIERATH, J. R., NORGREN, S., LUTHMAN, H. & GALLI, J. 2000. Insulin-degrading enzyme identified as a candidate diabetes susceptibility gene in GK rats. *Hum Mol Genet,* 9, 2149-58.

FAROOQI, I. S., JEBB, S. A., LANGMACK, G., LAWRENCE, E., CHEETHAM, C. H., PRENTICE, A. M., HUGHES, I. A., MCCAMISH, M. A. & O'RAHILLY, S. 1999. Effects of recombinant leptin therapy in a child with congenital leptin deficiency. *N Engl J Med,* 341, 879-84.

FAROOQI, I. S., KEOGH, J. M., YEO, G. S., LANK, E. J., CHEETHAM, T. & O'RAHILLY, S. 2003. Clinical spectrum of obesity and mutations in the melanocortin 4 receptor gene. *N Engl J Med,* 348, 1085-95.

FARRIS, W., MANSOURIAN, S., CHANG, Y., LINDSLEY, L., ECKMAN, E. A., FROSCH, M. P., ECKMAN, C. B., TANZI, R. E., SELKOE, D. J. & GUENETTE, S. 2003. Insulin-degrading enzyme regulates the levels of insulin, amyloid beta-protein, and the beta-amyloid precursor protein intracellular domain in vivo. *Proc Natl Acad Sci U S A,* 100, 4162-7.

FAWCETT, J., PERMANA, P. A., LEVY, J. L. & DUCKWORTH, W. C. 2007. Regulation of protein degradation by insulin-degrading enzyme: analysis by small interfering RNA-mediated gene silencing. *Arch Biochem Biophys,* 468, 128-33.

FESTA, A., D'AGOSTINO, R., JR., WILLIAMS, K., KARTER, A. J., MAYER-DAVIS, E. J., TRACY, R. P. & HAFFNER, S. M. 2001. The relation of body fat mass and distribution to markers of chronic inflammation. *Int J Obes Relat Metab Disord,* 25, 1407-15.

FINKELSTEIN, E. A., KHAVJOU, O. A., THOMPSON, H., TROGDON, J. G., PAN, L., SHERRY, B. & DIETZ, W. 2012. Obesity and severe obesity forecasts through 2030. *Am J Prev Med,* 42, 563-70.

FINKEN, M. J., KEIJZER-VEEN, M. G., DEKKER, F. W., FROLICH, M., HILLE, E. T., ROMIJN, J. A. & WIT, J. M. 2006. Preterm birth and later insulin resistance: effects of birth weight and postnatal growth in a population based longitudinal study from birth into adult life. *Diabetologia,* 49, 478-85.

FISCHER, J., EMMERLING, C. & RUTHER, U. 2008. On the history of Fto. *Obes Facts,* 1, 43-4.

FISCHER, J., KOCH, L., EMMERLING, C., VIERKOTTEN, J., PETERS, T., BRUNING, J. C. & RUTHER, U. 2009. Inactivation of the Fto gene protects from obesity. *Nature,* 458, 894-8.

FLEGAL, K. M., CARROLL, M. D., OGDEN, C. L. & JOHNSON, C. L. 2002. Prevalence and trends in obesity among US adults, 1999-2000. *JAMA,* 288, 1723-7.

FLODMARK, C. E., SVEGER, T. & NILSSON-EHLE, P. 1994. Waist measurement correlates to a potentially atherogenic lipoprotein profile in obese 12-14-year-old children. *Acta Paediatr,* 83, 941-5.

FONG, T. M., HUANG, R. R., TOTA, M. R., MAO, C., SMITH, T., VARNERIN, J., KARPITSKIY, V. V., KRAUSE, J. E. & VAN DER PLOEG, L. H. 1998. Localization of leptin binding domain in the leptin receptor. *Mol Pharmacol,* 53, 234-40.

FORRESTER, T. E., WILKS, R. J., BENNETT, F. I., SIMEON, D., OSMOND, C., ALLEN, M., CHUNG, A. P. & SCOTT, P. 1996. Fetal growth and cardiovascular risk factors in Jamaican schoolchildren. *BMJ,* 312, 156-60.

FORSEN, T., ERIKSSON, J., TUOMILEHTO, J., REUNANEN, A., OSMOND, C. & BARKER, D. 2000. The fetal and childhood growth of persons who develop type 2 diabetes. *Annals of Internal Medicine,* 133, 176-182.

FORSEN, T., ERIKSSON, J. G., TUOMILEHTO, J., OSMOND, C. & BARKER, D. J. 1999. Growth in utero and during childhood among women who develop coronary heart disease: longitudinal study. *BMJ,* 319, 1403-7.

FORSEN, T., ERIKSSON, J. G., TUOMILEHTO, J., TERAMO, K., OSMOND, C. & BARKER, D. J. 1997. Mother's weight in pregnancy and coronary heart disease in a cohort of Finnish men: follow up study. *BMJ,* 315, 837-40.

FOWDEN, A. L., COAN, P. M., ANGIOLINI, E., BURTON, G. J. & CONSTANCIA, M. 2011. Imprinted genes and the epigenetic regulation of placental phenotype. *Prog Biophys Mol Biol,* 106, 281-8.

FOWDEN, A. L., FORHEAD, A. J., COAN, P. M. & BURTON, G. J. 2008. The placenta and intrauterine programming. *J Neuroendocrinol,* 20, 439-50.

FOX, C. S., LIU, Y., WHITE, C. C., FEITOSA, M., SMITH, A. V., HEARD-COSTA, N., LOHMAN, K., JOHNSON, A. D., FOSTER, M. C., GREENAWALT, D. M., GRIFFIN, P., DING, J., NEWMAN, A. B., TYLAVSKY, F., MILJKOVIC, I., KRITCHEVSKY, S. B., LAUNER, L., GARCIA, M., EIRIKSDOTTIR, G., CARR, J. J., GUDNASON, V., HARRIS, T. B., CUPPLES, L. A. & BORECKI, I. B. 2012. Genome-wide association for abdominal subcutaneous and visceral adipose reveals a novel locus for visceral fat in women. *PLoS Genet,* 8, e1002695.

FRANK, D., MENDELSOHN, C. L., CICCONE, E., SVENSSON, K., OHLSSON, R. & TYCKO, B. 1999. A novel pleckstrin homology-related gene family defined by Ipl/Tssc3, TDAG51, and Tih1: tissue-specific expression, chromosomal location, and parental imprinting. *Mamm Genome,* 10, 1150-9.

FRAYLING, T. M., TIMPSON, N. J., WEEDON, M. N., ZEGGINI, E., FREATHY, R. M., LINDGREN, C. M., PERRY, J. R., ELLIOTT, K. S., LANGO, H., RAYNER, N. W., SHIELDS, B., HARRIES, L. W., BARRETT, J. C., ELLARD, S., GROVES, C. J., KNIGHT, B., PATCH, A. M., NESS, A. R., EBRAHIM, S., LAWLOR, D. A., RING, S. M., BEN-SHLOMO, Y., JARVELIN, M. R., SOVIO, U., BENNETT, A. J., MELZER, D., FERRUCCI, L., LOOS, R. J., BARROSO, I., WAREHAM, N. J., KARPE, F., OWEN, K. R., CARDON, L. R., WALKER, M., HITMAN, G. A., PALMER, C. N., DONEY, A. S., MORRIS, A. D., SMITH, G. D., HATTERSLEY, A. T. & MCCARTHY, M. I. 2007a. A common variant in the FTO gene is associated with body mass index and predisposes to childhood and adult obesity. *Science,* 316, 889-94.

FRAYLING, T. M., TIMPSON, N. J., WEEDON, M. N., ZEGGINI, E., FREATHY, R. M., LINDGREN, C. M., PERRY, J. R. B., ELLIOTT, K. S., LANGO, H., RAYNER, N. W., SHIELDS, B., HARRIES, L. W., BARRETT, J. C., ELLARD, S., GROVES, C. J., KNIGHT, B., PATCH, A. M., NESS, A. R., EBRAHIM, S., LAWLOR, D. A., RING, S. M., BEN-SHLOMO, Y., JARVELIN, M. R., SOVIO, U., BENNETT, A. J., MELZER, D., FERRUCCI, L., LOOS, R. J. F., BARROSO, I., WAREHAM, N. J., KARPE, F., OWEN,

K. R., CARDON, L. R., WALKER, M., HITMAN, G. A., PALMER, C. N. A., DONEY, A. S. F., MORRIS, A. D., SMITH, G. D., HATTERSLEY, A. T., MCCARTHY, M. I. & CONTROL, W. T. C. 2007b. A common variant in the FTO gene is associated with body mass index and predisposes to childhood and adult obesity. *Science,* 316**,** 889-894.

FREATHY, R. M., BENNETT, A. J., RING, S. M., SHIELDS, B., GROVES, C. J., TIMPSON, N. J., WEEDON, M. N., ZEGGINI, E., LINDGREN, C. M., LANGO, H., PERRY, J. R., POUTA, A., RUOKONEN, A., HYPPONEN, E., POWER, C., ELLIOTT, P., STRACHAN, D. P., JARVELIN, M. R., SMITH, G. D., MCCARTHY, M. I., FRAYLING, T. M. & HATTERSLEY, A. T. 2009. Type 2 diabetes risk alleles are associated with reduced size at birth. *Diabetes,* 58**,** 1428-33.

FREATHY, R. M., MOOK-KANAMORI, D. O., SOVIO, U., PROKOPENKO, I., TIMPSON, N. J., BERRY, D. J., WARRINGTON, N. M., WIDEN, E., HOTTENGA, J. J., KAAKINEN, M., LANGE, L. A., BRADFIELD, J. P., KERKHOF, M., MARSH, J. A., MAGI, R., CHEN, C. M., LYON, H. N., KIRIN, M., ADAIR, L. S., AULCHENKO, Y. S., BENNETT, A. J., BORJA, J. B., BOUATIA-NAJI, N., CHAROEN, P., COIN, L. J., COUSMINER, D. L., DE GEUS, E. J., DELOUKAS, P., ELLIOTT, P., EVANS, D. M., FROGUEL, P., GLASER, B., GROVES, C. J., HARTIKAINEN, A. L., HASSANALI, N., HIRSCHHORN, J. N., HOFMAN, A., HOLLY, J. M., HYPPONEN, E., KANONI, S., KNIGHT, B. A., LAITINEN, J., LINDGREN, C. M., MCARDLE, W. L., O'REILLY, P. F., PENNELL, C. E., POSTMA, D. S., POUTA, A., RAMASAMY, A., RAYNER, N. W., RING, S. M., RIVADENEIRA, F., SHIELDS, B. M., STRACHAN, D. P., SURAKKA, I., TAANILA, A., TIESLER, C., UITTERLINDEN, A. G., VAN DUIJN, C. M., WIJGA, A. H., WILLEMSEN, G., ZHANG, H., ZHAO, J., WILSON, J. F., STEEGERS, E. A., HATTERSLEY, A. T., ERIKSSON, J. G., PELTONEN, L., MOHLKE, K. L., GRANT, S. F., HAKONARSON, H., KOPPELMAN, G. H., DEDOUSSIS, G. V., HEINRICH, J., GILLMAN, M. W., PALMER, L. J., FRAYLING, T. M., BOOMSMA, D. I., DAVEY SMITH, G., POWER, C., JADDOE, V. W., JARVELIN, M. R. & MCCARTHY, M. I. 2010. Variants in ADCY5 and near CCNL1 are associated with fetal growth and birth weight. *Nature Genetics,* 42**,** 430-5.

FREEDLAND, E. S. 2004. Role of a critical visceral adipose tissue threshold (CVATT) in metabolic syndrome: implications for controlling dietary carbohydrates: a review. *Nutr Metab (Lond),* 1**,** 12.

FREEDMAN, D. S., DIETZ, W. H., SRINIVASAN, S. R. & BERENSON, G. S. 1999. The relation of overweight to cardiovascular risk factors among children and adolescents: the Bogalusa Heart Study. *Pediatrics,* 103**,** 1175-82.

FREEDMAN, D. S., KHAN, L. K., DIETZ, W. H., SRINIVASAN, S. R. & BERENSON, G. S. 2001. Relationship of childhood obesity to coronary heart disease risk factors in adulthood: the Bogalusa Heart Study. *Pediatrics,* 108**,** 712-8.

FRIEDLANDER, Y., SISCOVICK, D. S., ARBOGAST, P., PSATY, B. M., WEINMANN, S., LEMAITRE, R. N., RAGHUNATHAN, T. E. & COBB, L. A. 2002.

Sudden death and myocardial infarction in first degree relatives as predictors of primary cardiac arrest. *Atherosclerosis,* 162, 211-6.

FROGUEL, P. & VELHO, G. 1993. Non-sense mutation of glucokinase gene. *Lancet,* 341, 385.

FROST, J. M. & MOORE, G. E. 2010. The importance of imprinting in the human placenta. *PLoS Genet,* 6, e1001015.

FRUHBECK, G. 2006. Intracellular signalling pathways activated by leptin. *Biochemical Journal,* 393, 7-20.

FURUKAWA, Y., SHIMADA, T., FURUTA, H., MATSUNO, S., KUSUYAMA, A., DOI, A., NISHI, M., SASAKI, H., SANKE, T. & NANJO, K. 2008. Polymorphisms in the IDE-KIF11-HHEX gene locus are reproducibly associated with type 2 diabetes in a Japanese population. *Journal of Clinical Endocrinology & Metabolism,* 93, 310-314.

FURUSAWA, T., NAKA, I., YAMAUCHI, T., NATSUHARA, K., KIMURA, R., NAKAZAWA, M., ISHIDA, T., INAOKA, T., MATSUMURA, Y., ATAKA, Y., NISHIDA, N., TSUCHIYA, N., OHTSUKA, R. & OHASHI, J. 2010. The Q223R polymorphism in LEPR is associated with obesity in Pacific Islanders. *Hum Genet,* 127, 287-94.

GALE, C., LOGAN, K. M., SANTHAKUMARAN, S., PARKINSON, J. R., HYDE, M. J. & MODI, N. 2012. Effect of breastfeeding compared with formula feeding on infant body composition: a systematic review and meta-analysis. *American Journal of Clinical Nutrition,* 95, 656-69.

GARDNER, D. S., PEARCE, S., DANDREA, J., WALKER, R., RAMSAY, M. M., STEPHENSON, T. & SYMONDS, M. E. 2004. Peri-implantation undernutrition programs blunted angiotensin II evoked baroreflex responses in young adult sheep. *Hypertension,* 43, 1290-6.

GARDNER, D. S. & TAI, E. S. 2012. Clinical features and treatment of maturity onset diabetes of the young (MODY). *Diabetes Metab Syndr Obes,* 5, 101-8.

GAROFALO, C., SISCI, D. & SURMACZ, E. 2004. Leptin interferes with the effects of the antiestrogen ICI 182,780 in MCF-7 breast cancer cells. *Clinical Cancer Research,* 10, 6466-75.

GAROFANO, A., CZERNICHOW, P. & BREANT, B. 1997. In utero undernutrition impairs rat beta-cell development. *Diabetologia,* 40, 1231-4.

GASTALDELLI, A., KOZAKOVA, M., HOJLUND, K., FLYVBJERG, A., FAVUZZI, A., MITRAKOU, A. & BALKAU, B. 2009. Fatty liver is associated with insulin resistance, risk of coronary heart disease, and early atherosclerosis in a large European population. *Hepatology,* 49, 1537-44.

GERKEN, T., GIRARD, C. A., TUNG, Y. C., WEBBY, C. J., SAUDEK, V., HEWITSON, K. S., YEO, G. S., MCDONOUGH, M. A., CUNLIFFE, S., MCNEILL, L. A., GALVANOVSKIS, J., RORSMAN, P., ROBINS, P., PRIEUR, X., COLL, A. P., MA, M., JOVANOVIC, Z., FAROOQI, I. S., SEDGWICK, B., BARROSO, I., LINDAHL, T., PONTING, C. P., ASHCROFT, F. M., O'RAHILLY, S. & SCHOFIELD, C. J. 2007. The obesity-associated FTO gene encodes a 2-oxoglutarate-dependent nucleic acid demethylase. *Science,* 318, 1469-72.

GIELEN, M., LINDSEY, P. J., DEROM, C., SMEETS, H. J., SOUREN, N. Y., PAULUSSEN, A. D., DEROM, R. & NIJHUIS, J. G. 2008. Modeling genetic and environmental factors to increase heritability and ease the identification of candidate genes for birth weight: a twin study. *Behavior Genetics,* 38, 44-54.

GILLMAN, M. W., RIFAS-SHIMAN, S. L., CAMARGO, C. A., JR., BERKEY, C. S., FRAZIER, A. L., ROCKETT, H. R., FIELD, A. E. & COLDITZ, G. A. 2001. Risk of overweight among adolescents who were breastfed as infants. *JAMA,* 285, 2461-7.

GLUCKMAN, P. D., HANSON, M. A., MORTON, S. M. & PINAL, C. S. 2005. Life-long echoes--a critical analysis of the developmental origins of adult disease model. *Biol Neonate,* 87, 127-39.

GODFREY, K. M. 2002. The role of the placenta in fetal programming-a review. *Placenta,* 23 Suppl A, S20-7.

GODFREY, K. M. & BARKER, D. J. 2000. Fetal nutrition and adult disease. *American Journal of Clinical Nutrition,* 71, 1344S-52S.

GODFREY, K. M., BARKER, D. J., ROBINSON, S. & OSMOND, C. 1997. Maternal birthweight and diet in pregnancy in relation to the infant's thinness at birth. *Br J Obstet Gynaecol,* 104, 663-7.

GODFREY, K. M., FORRESTER, T., BARKER, D. J., JACKSON, A. A., LANDMAN, J. P., HALL, J. S., COX, V. & OSMOND, C. 1994. Maternal nutritional status in pregnancy and blood pressure in childhood. *Br J Obstet Gynaecol,* 101, 398-403.

GODISKA, R., MEAD, D., DHODDA, V., WU, C., HOCHSTEIN, R., KARSI, A., USDIN, K., ENTEZAM, A. & RAVIN, N. 2010. Linear plasmid vector for cloning of repetitive or unstable sequences in Escherichia coli. *Nucleic Acids Res,* 38, e88.

GOEL, M. S., MCCARTHY, E. P., PHILLIPS, R. S. & WEE, C. C. 2004. Obesity among US immigrant subgroups by duration of residence. *JAMA,* 292, 2860-7.

GONZALEZ, D. A., NAZMI, A. & VICTORA, C. G. 2010. Growth from birth to adulthood and abdominal obesity in a Brazilian birth cohort. *Int J Obes (Lond),* 34, 195-202.

GOODPASTER, B. H., KRISHNASWAMI, S., HARRIS, T. B., KATSIARAS, A., KRITCHEVSKY, S. B., SIMONSICK, E. M., NEVITT, M., HOLVOET, P. & NEWMAN, A. B. 2005. Obesity, regional body fat distribution, and the metabolic syndrome in older men and women. *Archives of Internal Medicine,* 165, 777-83.

GOODPASTER, B. H., KRISHNASWAMI, S., RESNICK, H., KELLEY, D. E., HAGGERTY, C., HARRIS, T. B., SCHWARTZ, A. V., KRITCHEVSKY, S. & NEWMAN, A. B. 2003. Association between regional adipose tissue distribution and both type 2 diabetes and impaired glucose tolerance in elderly men and women. *Diabetes Care,* 26, 372-379.

GOOZE, R. A., ANDERSON, S. E. & WHITAKER, R. C. 2011. Prolonged bottle use and obesity at 5.5 years of age in US children. *J Pediatr,* 159, 431-6.

GOTODA, T., MANNING, B. S., GOLDSTONE, A. P., IMRIE, H., EVANS, A. L., STROSBERG, A. D., MCKEIGUE, P. M., SCOTT, J. & AITMAN, T. J. 1997. Leptin receptor gene variation and obesity: lack of association in a white British male population. *Hum Mol Genet,* 6, 869-76.

GOYAL, R., GALFFY, A., FIELD, S. A., GHEORGHE, C. P., MITTAL, A. & LONGO, L. D. 2009. Maternal protein deprivation: changes in systemic renin-angiotensin system of the mouse fetus. *Reprod Sci,* 16, 894-904.

GRARUP, N., ROSE, C. S., ANDERSSON, E. A., ANDERSEN, G., NIELSEN, A. L., ALBRECHTSEN, A., CLAUSEN, J. O., RASMUSSEN, S. S., JORGENSEN, T., SANDBAEK, A., LAURITZEN, T., SCHMITZ, O., HANSEN, T. & PEDERSEN, O. 2007. Studies of association of variants near the HHEX, CDKN2A/B, and IGF2BP2 genes with type 2 diabetes and impaired insulin release in 10,705 Danish subjects: validation and extension of genome-wide association studies. *Diabetes,* 56, 3105-11.

GREENBERG, A. S. & OBIN, M. S. 2006. Obesity and the role of adipose tissue in inflammation and metabolism. *American Journal of Clinical Nutrition,* 83, 461S-465S.

GREGOR, M. F. & HOTAMISLIGIL, G. S. 2011. Inflammatory mechanisms in obesity. *Annu Rev Immunol,* 29, 415-45.

GROVES, C. J., WILTSHIRE, S., SMEDLEY, D., OWEN, K. R., FRAYLING, T. M., WALKER, M., HITMAN, G. A., LEVY, J. C., O'RAHILLY, S., MENZEL, S., HATTERSLEY, A. T. & MCCARTHY, M. I. 2003. Association and haplotype analysis of the insulin-degrading enzyme (IDE) gene, a strong positional and biological candidate for type 2 diabetes susceptibility. *Diabetes,* 52, 1300-5.

GRUNDY, S. M. 2004. Obesity, metabolic syndrome, and cardiovascular disease. *Journal of Clinical Endocrinology & Metabolism,* 89, 2595-2600.

GRUNDY, S. M. 2006. Metabolic syndrome: connecting and reconciling cardiovascular and diabetes worlds. *J Am Coll Cardiol,* 47, 1093-100.

GRUNNET, L. G., NILSSON, E., LING, C., HANSEN, T., PEDERSEN, O., GROOP, L., VAAG, A. & POULSEN, P. 2009. Regulation and function of FTO mRNA expression in human skeletal muscle and subcutaneous adipose tissue. *Diabetes,* 58, 2402-8.

GRUPPUSO, P. A., BOYLAN, J. M., ANAND, P. & BIENIEKI, T. C. 2005. Effects of maternal starvation on hepatocyte proliferation in the late gestation fetal rat. *Pediatric Research,* 57, 185-91.

GU, H. F., EFENDIC, S., NORDMAN, S., OSTENSON, C. G., BRISMAR, K., BROOKES, A. J. & PRINCE, J. A. 2004. Quantitative trait loci near the insulin-degrading enzyme (IDE) gene contribute to variation in plasma insulin levels. *Diabetes,* 53, 2137-42.

GUERRA, C., KOZA, R. A., YAMASHITA, H., WALSH, K. & KOZAK, L. P. 1998. Emergence of brown adipocytes in white fat in mice is under genetic control. Effects on body weight and adiposity. *J Clin Invest,* 102, 412-20.

HAFFNER, S. M., GINGERICH, R. L., MIETTINEN, H. & STERN, M. P. 1996. Leptin concentrations in relation to overall adiposity and regional body fat distribution in Mexican Americans. *Int J Obes Relat Metab Disord,* 20, 904-8.

HAJER, G. R., VAN HAEFTEN, T. W. & VISSEREN, F. L. 2008. Adipose tissue dysfunction in obesity, diabetes, and vascular diseases. *Eur Heart J,* 29, 2959-71.

HAKANEN, M., RAITAKARI, O. T., LEHTIMAKI, T., PELTONEN, N., PAHKALA, K., SILLANMAKI, L., LAGSTROM, H., VIIKARI, J., SIMELL, O. & RONNEMAA, T. 2009. FTO genotype is associated with body mass index after the age of seven years but not with energy intake or leisure-time physical activity. *J Clin Endocrinol Metab,* 94, 1281-7.

HALAAS, J. L., BOOZER, C., BLAIR-WEST, J., FIDAHUSEIN, N., DENTON, D. A. & FRIEDMAN, J. M. 1997. Physiological response to long-term peripheral and central leptin infusion in lean and obese mice. *Proc Natl Acad Sci U S A,* 94, 8878-83.

HALES, C. N. & BARKER, D. J. 1992. Type 2 (non-insulin-dependent) diabetes mellitus: the thrifty phenotype hypothesis. *Diabetologia,* 35, 595-601.

HALES, C. N., BARKER, D. J., CLARK, P. M., COX, L. J., FALL, C., OSMOND, C. & WINTER, P. D. 1991. Fetal and infant growth and impaired glucose tolerance at age 64. *BMJ,* 303, 1019-22.

HAMEL, F. G., BENNETT, R. G., UPWARD, J. L. & DUCKWORTH, W. C. 2001. Insulin inhibits peroxisomal fatty acid oxidation in isolated rat hepatocytes. *Endocrinology,* 142, 2702-6.

HAN, J. C., LAWLOR, D. A. & KIMM, S. Y. 2010. Childhood obesity. *Lancet,* 375, 1737-48.

HAN, Z., MULLA, S., BEYENE, J., LIAO, G. & MCDONALD, S. D. 2011. Maternal underweight and the risk of preterm birth and low birth weight: a systematic review and meta-analyses. *Int J Epidemiol,* 40, 65-101.

HANIU, M., ARAKAWA, T., BURES, E. J., YOUNG, Y., HUI, J. O., ROHDE, M. F., WELCHER, A. A. & HORAN, T. 1998. Human leptin receptor. Determination of disulfide structure and N-glycosylation sites of the extracellular domain. *Journal of Biological Chemistry,* 273, 28691-9.

HANOUNE, J., POUILLE, Y., TZAVARA, E., SHEN, T., LIPSKAYA, L., MIYAMOTO, N., SUZUKI, Y. & DEFER, N. 1997. Adenylyl cyclases: structure, regulation and function in an enzyme superfamily. *Mol Cell Endocrinol,* 128, 179-94.

HARADA, S., SMITH, R. M., SMITH, J. A. & JARETT, L. 1993. Inhibition of insulin-degrading enzyme increases translocation of insulin to the nucleus in H35 rat hepatoma cells: evidence of a cytosolic pathway. *Endocrinology,* 132, 2293-8.

HARDIE, L., TRAYHURN, P., ABRAMOVICH, D. & FOWLER, P. 1997. Circulating leptin in women: a longitudinal study in the menstrual cycle and during pregnancy. *Clin Endocrinol (Oxf),* 47, 101-6.

HARDING, S., ROSATO, M. G. & CRUICKSHANK, J. K. 2004. Lack of change in birthweights of infants by generational status among Indian, Pakistani, Bangladeshi, Black Caribbean, and Black African mothers in a British cohort study. *Int J Epidemiol,* 33, 1279-85.

HATTERSLEY, A. T., BEARDS, F., BALLANTYNE, E., APPLETON, M., HARVEY, R. & ELLARD, S. 1998. Mutations in the glucokinase gene of the fetus result in reduced birth weight. *Nature Genetics,* 19, 268-70.

HATTERSLEY, A. T. & TOOKE, J. E. 1999. The fetal insulin hypothesis: an alternative explanation of the association of low birthweight with diabetes and vascular disease. *Lancet,* 353, 1789-92.

HAWORTH, C. M., CARNELL, S., MEABURN, E. L., DAVIS, O. S., PLOMIN, R. & WARDLE, J. 2008a. Increasing heritability of BMI and stronger associations with the FTO gene over childhood. *Obesity (Silver Spring),* 16, 2663-8.

HAWORTH, C. M., PLOMIN, R., CARNELL, S. & WARDLE, J. 2008b. Childhood obesity: genetic and environmental overlap with normal-range BMI. *Obesity (Silver Spring),* 16, 1585-90.

HEARD-COSTA, N. L., ZILLIKENS, M. C., MONDA, K. L., JOHANSSON, A., HARRIS, T. B., FU, M., HARITUNIANS, T., FEITOSA, M. F., ASPELUND, T., EIRIKSDOTTIR, G., GARCIA, M., LAUNER, L. J., SMITH, A. V., MITCHELL, B. D., MCARDLE, P. F., SHULDINER, A. R., BIELINSKI, S. J., BOERWINKLE, E., BRANCATI, F., DEMERATH, E. W., PANKOW, J. S., ARNOLD, A. M., CHEN, Y. D., GLAZER, N. L., MCKNIGHT, B., PSATY, B. M., ROTTER, J. I., AMIN, N., CAMPBELL, H., GYLLENSTEN, U., PATTARO, C., PRAMSTALLER, P. P., RUDAN, I., STRUCHALIN, M., VITART, V., GAO, X., KRAJA, A., PROVINCE, M. A., ZHANG, Q., ATWOOD, L. D., DUPUIS, J., HIRSCHHORN, J. N., JAQUISH, C. E., O'DONNELL, C. J., VASAN, R. S., WHITE, C. C., AULCHENKO, Y. S., ESTRADA, K., HOFMAN, A., RIVADENEIRA, F., UITTERLINDEN, A. G., WITTEMAN, J. C., OOSTRA, B. A., KAPLAN, R. C., GUDNASON, V., O'CONNELL, J. R., BORECKI, I. B., VAN DUIJN, C. M., CUPPLES, L. A., FOX, C. S. & NORTH, K. E. 2009. NRXN3 is a novel locus for waist circumference: a genome-wide association study from the CHARGE Consortium. *PLoS Genet,* 5, e1000539.

HEID, I. M., JACKSON, A. U., RANDALL, J. C., WINKLER, T. W., QI, L., STEINTHORSDOTTIR, V., THORLEIFSSON, G., ZILLIKENS, M. C., SPELIOTES, E. K., MAGI, R., WORKALEMAHU, T., WHITE, C. C., BOUATIA-NAJI, N., HARRIS, T. B., BERNDT, S. I., INGELSSON, E., WILLER, C. J., WEEDON, M. N., LUAN, J., VEDANTAM, S., ESKO, T., KILPELAINEN, T. O., KUTALIK, Z., LI, S., MONDA, K. L., DIXON, A. L., HOLMES, C. C., KAPLAN, L. M., LIANG, L., MIN, J. L., MOFFATT, M. F., MOLONY, C., NICHOLSON, G., SCHADT, E. E., ZONDERVAN, K. T., FEITOSA, M. F., FERREIRA, T., ALLEN, H. L., WEYANT, R. J., WHEELER, E., WOOD, A. R., ESTRADA, K., GODDARD, M. E., LETTRE, G., MANGINO, M., NYHOLT, D. R., PURCELL, S., SMITH, A. V., VISSCHER, P. M., YANG, J., MCCARROLL, S. A., NEMESH, J., VOIGHT, B. F., ABSHER, D., AMIN, N., ASPELUND, T., COIN, L., GLAZER, N. L., HAYWARD, C., HEARD-COSTA, N. L., HOTTENGA, J. J., JOHANSSON, A., JOHNSON, T., KAAKINEN, M., KAPUR, K., KETKAR, S., KNOWLES, J. W., KRAFT, P., KRAJA, A. T., LAMINA, C.,

LEITZMANN, M. F., MCKNIGHT, B., MORRIS, A. P., ONG, K. K., PERRY, J. R., PETERS, M. J., POLASEK, O., PROKOPENKO, I., RAYNER, N. W., RIPATTI, S., RIVADENEIRA, F., ROBERTSON, N. R., SANNA, S., SOVIO, U., SURAKKA, I., TEUMER, A., VAN WINGERDEN, S., VITART, V., ZHAO, J. H., CAVALCANTI-PROENCA, C., CHINES, P. S., FISHER, E., KULZER, J. R., LECOEUR, C., NARISU, N., SANDHOLT, C., SCOTT, L. J., SILANDER, K., STARK, K., TAMMESOO, M. L., *et al.* 2010. Meta-analysis identifies 13 new loci associated with waist-hip ratio and reveals sexual dimorphism in the genetic basis of fat distribution. *Nature Genetics,* 42, 949-60.

HENNIG, B. J., FULFORD, A. J., SIRUGO, G., RAYCO-SOLON, P., HATTERSLEY, A. T., FRAYLING, T. M. & PRENTICE, A. M. 2009. FTO gene variation and measures of body mass in an African population. *BMC Med Genet,* 10, 21.

HERBERT, A., GERRY, N. P., MCQUEEN, M. B., HEID, I. M., PFEUFER, A., ILLIG, T., WICHMANN, H. E., MEITINGER, T., HUNTER, D., HU, F. B., COLDITZ, G., HINNEY, A., HEBEBRAND, J., KOBERWITZ, K., ZHU, X., COOPER, R., ARDLIE, K., LYON, H., HIRSCHHORN, J. N., LAIRD, N. M., LENBURG, M. E., LANGE, C. & CHRISTMAN, M. F. 2006. A common genetic variant is associated with adult and childhood obesity. *Science,* 312, 279-83.

HERMANUSSEN, M., STAUB, K., ASSMANN, C. & VAN BUUREN, S. 2012. Dilemmas in choosing and using growth charts. *Pediatr Endocrinol Rev,* 9, 650-6.

HERRERA, B. M., KEILDSON, S. & LINDGREN, C. M. 2011. Genetics and epigenetics of obesity. *Maturitas,* 69, 41-9.

HERSOUG, L. 2007. The link between the epidemics of obesity and allergic diseases: does obesity induce decreased immune tolerance. *Allergy,* 62, 1205 - 1213.

HEUDE, B., DUBOIS, S., CHARLES, M. A., DEWEIRDER, M., DINA, C., BORYS, J. M., DUCIMETIERE, P. & FROGUEL, P. 2004. VNTR polymorphism of the insulin gene and childhood overweight in a general population. *Obes Res,* 12, 499-504.

HEUDE, B., PETRY, C. J., PEMBREY, M., DUNGER, D. B. & ONG, K. K. 2006. The insulin gene variable number of tandem repeat: associations and interactions with childhood body fat mass and insulin secretion in normal children. *J Clin Endocrinol Metab,* 91, 2770-5.

HEUDE, B., THIEBAUGEORGES, O., GOUA, V., FORHAN, A., KAMINSKI, M., FOLIGUET, B., SCHWEITZER, M., MAGNIN, G. & CHARLES, M. A. 2012. Pre-pregnancy body mass index and weight gain during pregnancy: relations with gestational diabetes and hypertension, and birth outcomes. *Matern Child Health J,* 16, 355-63.

HIGGINS, M. & MC AULIFFE, F. 2010. A review of maternal and fetal growth factors in diabetic pregnancy. *Curr Diabetes Rev,* 6, 116-25.

HINDMARSH, P. C., GEARY, M. P., RODECK, C. H., KINGDOM, J. C. & COLE, T. J. 2002. Intrauterine growth and its relationship to size and shape at birth. *Pediatric Research,* 52, 263-8.

HINNEY, A., NGUYEN, T. T., SCHERAG, A., FRIEDEL, S., BRONNER, G., MULLER, T. D., GRALLERT, H., ILLIG, T., WICHMANN, H. E., RIEF, W., SCHAFER, H. & HEBEBRAND, J. 2007. Genome wide association (GWA) study for early onset extreme obesity supports the role of fat mass and obesity associated gene (FTO) variants. *PLoS One,* 2, e1361.

HINNEY, A., VOGEL, C. I. & HEBEBRAND, J. 2010. From monogenic to polygenic obesity: recent advances. *Eur Child Adolesc Psychiatry,* 19, 297-310.

HIROKAWA, N. 1998. Kinesin and dynein superfamily proteins and the mechanism of organelle transport. *Science,* 279, 519-26.

HONG, J., SHI, J., QI, L., CUI, B., GU, W., ZHANG, Y., LI, L., XU, M., WANG, L., ZHAI, Y., MIAO, L., WANG, R., BI, Y., WANG, W. & NING, G. 2012. Genetic susceptibility, birth weight and obesity risk in young Chinese. *Int J Obes (Lond).*

HONG, K. W. & OH, B. 2011. Recapitulation of genome-wide association studies on body mass index in the Korean population. *Int J Obes (Lond).*

HORIKAWA, Y., MIYAKE, K., YASUDA, K., ENYA, M., HIROTA, Y., YAMAGATA, K., HINOKIO, Y., OKA, Y., IWASAKI, N., IWAMOTO, Y., YAMADA, Y., SEINO, Y., MAEGAWA, H., KASHIWAGI, A., YAMAMOTO, K., TOKUNAGA, K., TAKEDA, J. & KASUGA, M. 2008. Replication of genome-wide association studies of type 2 diabetes susceptibility in Japan. *J Clin Endocrinol Metab,* 93, 3136-41.

HORIKOSHI, M., HARA, K., ITO, C., SHOJIMA, N., NAGAI, R., UEKI, K., FROGUEL, P. & KADOWAKI, T. 2007. Variations in the HHEX gene are associated with increased risk of type 2 diabetes in the Japanese population. *Diabetologia,* 50, 2461-6.

HOTTA, K., NAKATA, Y., MATSUO, T., KAMOHARA, S., KOTANI, K., KOMATSU, R., ITOH, N., MINEO, I., WADA, J., MASUZAKI, H., YONEDA, M., NAKAJIMA, A., MIYAZAKI, S., TOKUNAGA, K., KAWAMOTO, M., FUNAHASHI, T., HAMAGUCHI, K., YAMADA, K., HANAFUSA, T., OIKAWA, S., YOSHIMATSU, H., NAKAO, K., SAKATA, T., MATSUZAWA, Y., TANAKA, K., KAMATANI, N. & NAKAMURA, Y. 2008. Variations in the FTO gene are associated with severe obesity in the Japanese. *Journal of Human Genetics,* 53, 546-553.

HOUWELING, A. C., DILDROP, R., PETERS, T., MUMMENHOFF, J., MOORMAN, A. F., RUTHER, U. & CHRISTOFFELS, V. M. 2001. Gene and cluster-specific expression of the Iroquois family members during mouse development. *Mech Dev,* 107, 169-74.

HU, C., ZHANG, R., WANG, C., WANG, J., MA, X., HOU, X., LU, J., YU, W., JIANG, F., BAO, Y., XIANG, K. & JIA, W. 2010. Variants from GIPR, TCF7L2, DGKB, MADD, CRY2, GLIS3, PROX1, SLC30A8 and IGF1 are associated with glucose metabolism in the Chinese. *PLoS One,* 5, e15542.

HU, F. B., WILLETT, W. C., LI, T., STAMPFER, M. J., COLDITZ, G. A. & MANSON, J. E. 2004a. Adiposity as compared with physical activity in predicting mortality among women. *N Engl J Med,* 351, 2694-703.

HU, G., BARENGO, N. C., TUOMILEHTO, J., LAKKA, T. A., NISSINEN, A. & JOUSILAHTI, P. 2004b. Relationship of physical activity and body mass index to the risk of hypertension: a prospective study in Finland. *Hypertension,* 43, 25-30.

HULT, M., TORNHAMMAR, P., UEDA, P., CHIMA, C., BONAMY, A. K., OZUMBA, B. & NORMAN, M. 2010. Hypertension, diabetes and overweight: looming legacies of the Biafran famine. *PLoS One,* 5, e13582.

HUNG, C. N., POON, W. T., LEE, C. Y., LAW, C. Y. & CHAN, A. Y. 2012. A case of early-onset obesity, hypocortisolism, and skin pigmentation problem due to a novel homozygous mutation in the proopiomelanocortin (POMC) gene in an Indian boy. *J Pediatr Endocrinol Metab,* 25, 175-9.

HUR, Y. M., LUCIANO, M., MARTIN, N. G., BOOMSMA, D. I., IACONO, W. G., MCGUE, M., SHIN, J. S., JUN, J. K., OOKI, S., VAN BEIJSTERVELDT, C. E. & HAN, J. Y. 2005. A comparison of twin birthweight data from Australia, the Netherlands, the United States, Japan, and South Korea: are genetic and environmental variations in birthweight similar in Caucasians and East Asians? *Twin Res Hum Genet,* 8, 638-48.

HUSSLEIN, H., MOSWITZER, B., LEIPOLD, H., MOERTL, M. & WORDA, C. 2012. Low placental weight and risk for fetal distress at birth. *J Perinat Med,* 0, 1-3.

HUTCHEON, J. A., MCNAMARA, H., PLATT, R. W., BENJAMIN, A. & KRAMER, M. S. 2012. Placental weight for gestational age and adverse perinatal outcomes. *Obstet Gynecol,* 119, 1251-8.

HYPPONEN, E., POWER, C. & SMITH, G. D. 2003. Prenatal growth, BMI, and risk of type 2 diabetes by early midlife. *Diabetes Care,* 26, 2512-7.

IBANEZ, L., SUAREZ, L., LOPEZ-BERMEJO, A., DIAZ, M., VALLS, C. & DE ZEGHER, F. 2008. Early development of visceral fat excess after spontaneous catch-up growth in children with low birth weight. *J Clin Endocrinol Metab,* 93, 925-8.

IIDA, M., MURAKAMI, T., ISHIDA, K., MIZUNO, A., KUWAJIMA, M. & SHIMA, K. 1996. Substitution at codon 269 (glutamine --> proline) of the leptin receptor (OB-R) cDNA is the only mutation found in the Zucker fatty (fa/fa) rat. *Biochem Biophys Res Commun,* 224, 597-604.

ILLINGWORTH, R. S. 1950. Birth weight and subsequent weight. *Br Med J,* 1, 96.

ISERENTANT, H., PEELMAN, F., DEFEAU, D., VANDEKERCKHOVE, J., ZABEAU, L. & TAVERNIER, J. 2005. Mapping of the interface between leptin and the leptin receptor CRH2 domain. *J Cell Sci,* 118, 2519-27.

JACOBSSON, J. A., KLOVINS, J., KAPA, I., DANIELSSON, P., SVENSSON, V., RIDDERSTRALE, M., GYLLENSTEN, U., MARCUS, C., FREDRIKSSON, R. & SCHIOTH, H. B. 2008. Novel genetic variant in FTO influences insulin levels and insulin resistance in severely obese children and adolescents. *Int J Obes (Lond),* 32, 1730-5.

JAMES, W. P. 2008. The fundamental drivers of the obesity epidemic. *Obesity Reviews,* 9 Suppl 1, 6-13.

JANSSON, T., CETIN, I., POWELL, T. L., DESOYE, G., RADAELLI, T., ERICSSON, A. & SIBLEY, C. P. 2006. Placental transport and metabolism in fetal overgrowth -- a workshop report. *Placenta,* 27 Suppl A, S109-13.

JAYARAMAN, P. S., FRAMPTON, J. & GOODWIN, G. 2000. The homeodomain protein PRH influences the differentiation of haematopoietic cells. *Leuk Res,* 24, 1023-31.

JENNINGS, B. J., OZANNE, S. E., DORLING, M. W. & HALES, C. N. 1999. Early growth determines longevity in male rats and may be related to telomere shortening in the kidney. *FEBS Lett,* 448, 4-8.

JENUM, A. K., DIEP, L. M., HOLMBOE-OTTESEN, G., HOLME, I. M., KUMAR, B. N. & BIRKELAND, K. I. 2012. Diabetes susceptibility in ethnic minority groups from Turkey, Vietnam, Sri Lanka and Pakistan compared with Norwegians - the association with adiposity is strongest for ethnic minority women. *BMC Public Health,* 12, 150.

JEON, J. P., SHIM, S. M., NAM, H. Y., RYU, G. M., HONG, E. J., KIM, H. L. & HAN, B. G. 2010. Copy number variation at leptin receptor gene locus associated with metabolic traits and the risk of type 2 diabetes mellitus. *BMC Genomics,* 11, 426.

JESS, T., ZIMMERMANN, E., KRING, S. I., BERENTZEN, T., HOLST, C., TOUBRO, S., ASTRUP, A., HANSEN, T., PEDERSEN, O. & SORENSEN, T. I. 2008. Impact on weight dynamics and general growth of the common FTO rs9939609: a longitudinal Danish cohort study. *Int J Obes (Lond),* 32, 1388-94.

JIMENEZ-CHILLARON, J. C. & PATTI, M. E. 2007. To catch up or not to catch up: is this the question? Lessons from animal models. *Curr Opin Endocrinol Diabetes Obes,* 14, 23-9.

JOHN, R. & HEMBERGER, M. 2012. A placenta for life. *Reprod Biomed Online,* 25, 5-11.

JOHN, R. M. & SURANI, M. A. 2000. Genomic imprinting, mammalian evolution, and the mystery of egg-laying mammals. *Cell,* 101, 585-8.

JOHNSTON, L. B., DAHLGREN, J., LEGER, J., GELANDER, L., SAVAGE, M. O., CZERNICHOW, P., WIKLAND, K. A. & CLARK, A. J. 2003. Association between insulin-like growth factor I (IGF-I) polymorphisms, circulating IGF-I, and pre- and postnatal growth in two European small for gestational age populations. *J Clin Endocrinol Metab,* 88, 4805-10.

JOWETT, J. B. M., CURRAN, J. E., JOHNSON, M. P., CARLESS, M. A., GORING, H. H. H., DYER, T. D., COLE, S. A., COMUZZIE, A. G., MACCLUER, J. W., MOSES, E. K. & BLANGERO, J. 2010. Genetic Variation at the FTO Locus Influences RBL2 Gene Expression. *Diabetes,* 59, 726-732.

KAPITEIN, L. C., PETERMAN, E. J., KWOK, B. H., KIM, J. H., KAPOOR, T. M. & SCHMIDT, C. F. 2005. The bipolar mitotic kinesin Eg5 moves on both microtubules that it crosslinks. *Nature,* 435, 114-8.

KARAMOHAMED, S., DEMISSIE, S., VOLCJAK, J., LIU, C., HEARD-COSTA, N., LIU, J., SHOEMAKER, C. M., PANHUYSEN, C. I., MEIGS, J. B., WILSON, P., ATWOOD, L. D., CUPPLES, L. A. & HERBERT, A. 2003. Polymorphisms in the insulin-degrading enzyme gene are associated with type 2 diabetes in men from the NHLBI Framingham Heart Study. *Diabetes,* 52, 1562-7.

KARAOLIS-DANCKERT, N., BUYKEN, A. E., BOLZENIUS, K., PERIM DE FARIA, C., LENTZE, M. J. & KROKE, A. 2006. Rapid growth among term children whose birth weight was appropriate for gestational age has a longer lasting effect on body fat percentage than on body mass index. *American Journal of Clinical Nutrition,* 84, 1449-55.

KARN, M. N. & PENROSE, L. S. 1952. Birth weight and gestation time in relation to maternal age, parity and infant survival. *Ann Eugen,* 16, 147-64.

KATZMARZYK, P. T., BRAY, G. A., GREENWAY, F. L., JOHNSON, W. D., NEWTON, R. L., JR., RAVUSSIN, E., RYAN, D. H. & BOUCHARD, C. 2011. Ethnic-specific BMI and waist circumference thresholds. *Obesity (Silver Spring),* 19**,** 1272-8.

KELLY, T., YANG, W., CHEN, C. S., REYNOLDS, K. & HE, J. 2008. Global burden of obesity in 2005 and projections to 2030. *Int J Obes (Lond),* 32**,** 1431-7.

KENG, V. W., FUJIMORI, K. E., MYINT, Z., TAMAMAKI, N., NOJYO, Y. & NOGUCHI, T. 1998. Expression of Hex mRNA in early murine postimplantation embryo development. *FEBS Lett,* 426**,** 183-6.

KENG, V. W., YAGI, H., IKAWA, M., NAGANO, T., MYINT, Z., YAMADA, K., TANAKA, T., SATO, A., MURAMATSU, I., OKABE, M., SATO, M. & NOGUCHI, T. 2000. Homeobox gene Hex is essential for onset of mouse embryonic liver development and differentiation of the monocyte lineage. *Biochem Biophys Res Commun,* 276**,** 1155-61.

KENSARA, O. A., WOOTTON, S. A., PHILLIPS, D. I., PATEL, M., JACKSON, A. A. & ELIA, M. 2005. Fetal programming of body composition: relation between birth weight and body composition measured with dual-energy X-ray absorptiometry and anthropometric methods in older Englishmen. *American Journal of Clinical Nutrition,* 82**,** 980-7.

KESKI-RAHKONEN, A., BULIK, C. M., PIETILAINEN, K. H., ROSE, R. J., KAPRIO, J. & RISSANEN, A. 2007. Eating styles, overweight and obesity in young adult twins. *Eur J Clin Nutr,* 61**,** 822-9.

KESKITALO, K., TUORILA, H., SPECTOR, T. D., CHERKAS, L. F., KNAAPILA, A., KAPRIO, J., SILVENTOINEN, K. & PEROLA, M. 2008. The Three-Factor Eating Questionnaire, body mass index, and responses to sweet and salty fatty foods: a twin study of genetic and environmental associations. *American Journal of Clinical Nutrition,* 88**,** 263-71.

KHADILKAR, V. V., KHADILKAR, A. V., COLE, T. J., CHIPLONKAR, S. A. & PANDIT, D. 2011. Overweight and obesity prevalence and body mass index trends in Indian children. *Int J Pediatr Obes,* 6**,** e216-24.

KIDD, J. M., COOPER, G. M., DONAHUE, W. F., HAYDEN, H. S., SAMPAS, N., GRAVES, T., HANSEN, N., TEAGUE, B., ALKAN, C., ANTONACCI, F., HAUGEN, E., ZERR, T., YAMADA, N. A., TSANG, P., NEWMAN, T. L., TUZUN, E., CHENG, Z., EBLING, H. M., TUSNEEM, N., DAVID, R., GILLETT, W., PHELPS, K. A., WEAVER, M., SARANGA, D., BRAND, A., TAO, W., GUSTAFSON, E., MCKERNAN, K., CHEN, L., MALIG, M., SMITH, J. D., KORN, J. M., MCCARROLL, S. A., ALTSHULER, D. A., PEIFFER, D. A., DORSCHNER, M., STAMATOYANNOPOULOS, J., SCHWARTZ, D., NICKERSON, D. A., MULLIKIN, J. C., WILSON, R. K., BRUHN, L., OLSON, M.

V., KAUL, R., SMITH, D. R. & EICHLER, E. E. 2008. Mapping and sequencing of structural variation from eight human genomes. *Nature,* 453, 56-64.

KILPELAINEN, T. O., DEN HOED, M., ONG, K. K., GRONTVED, A., BRAGE, S., JAMESON, K., COOPER, C., KHAW, K. T., EKELUND, U., WAREHAM, N. J. & LOOS, R. J. 2011a. Obesity-susceptibility loci have a limited influence on birth weight: a meta-analysis of up to 28,219 individuals. *American Journal of Clinical Nutrition,* 93, 851-60.

KILPELAINEN, T. O., QI, L., BRAGE, S., SHARP, S. J., SONESTEDT, E., DEMERATH, E., AHMAD, T., MORA, S., KAAKINEN, M., SANDHOLT, C. H., HOLZAPFEL, C., AUTENRIETH, C. S., HYPPONEN, E., CAUCHI, S., HE, M., KUTALIK, Z., KUMARI, M., STANCAKOVA, A., MEIDTNER, K., BALKAU, B., TAN, J. T., MANGINO, M., TIMPSON, N. J., SONG, Y., ZILLIKENS, M. C., JABLONSKI, K. A., GARCIA, M. E., JOHANSSON, S., BRAGG-GRESHAM, J. L., WU, Y., VAN VLIET-OSTAPTCHOUK, J. V., ONLAND-MORET, N. C., ZIMMERMANN, E., RIVERA, N. V., TANAKA, T., STRINGHAM, H. M., SILBERNAGEL, G., KANONI, S., FEITOSA, M. F., SNITKER, S., RUIZ, J. R., METTER, J., LARRAD, M. T., ATALAY, M., HAKANEN, M., AMIN, N., CAVALCANTI-PROENCA, C., GRONTVED, A., HALLMANS, G., JANSSON, J. O., KUUSISTO, J., KAHONEN, M., LUTSEY, P. L., NOLAN, J. J., PALLA, L., PEDERSEN, O., PERUSSE, L., RENSTROM, F., SCOTT, R. A., SHUNGIN, D., SOVIO, U., TAMMELIN, T. H., RONNEMAA, T., LAKKA, T. A., UUSITUPA, M., RIOS, M. S., FERRUCCI, L., BOUCHARD, C., MEIRHAEGHE, A., FU, M., WALKER, M., BORECKI, I. B., DEDOUSSIS, G. V., FRITSCHE, A., OHLSSON, C., BOEHNKE, M., BANDINELLI, S., VAN DUIJN, C. M., EBRAHIM, S., LAWLOR, D. A., GUDNASON, V., HARRIS, T. B., SORENSEN, T. I., MOHLKE, K. L., HOFMAN, A., UITTERLINDEN, A. G., TUOMILEHTO, J., LEHTIMAKI, T., RAITAKARI, O., ISOMAA, B., NJOLSTAD, P. R., FLOREZ, J. C., LIU, S., NESS, A., SPECTOR, T. D., TAI, E. S., FROGUEL, P., BOEING, H., LAAKSO, M., MARMOT, M., *et al.* 2011b. Physical activity attenuates the influence of FTO variants on obesity risk: a meta-analysis of 218,166 adults and 19,268 children. *PLoS Med,* 8, e1001116.

KIM, A. J. & ENDOW, S. A. 2000. A kinesin family tree. *J Cell Sci,* 113 Pt 21, 3681-2.

KIM, C. S., PARK, J. S., PARK, J., NAM, J. S., KANG, E. S., AHN, C. W., CHA, B. S., LIM, S. K., KIM, K. R., LEE, H. C., HUH, K. B. & KIM, D. J. 2006. The relation between birth weight and insulin resistance in Korean adolescents. *Yonsei Med J,* 47, 85-92.

KIMBER, W., PEELMAN, F., PRIEUR, X., WANGENSTEEN, T., O'RAHILLY, S., TAVERNIER, J. & FAROOQI, I. S. 2008. Functional characterization of naturally occurring pathogenic mutations in the human leptin receptor. *Endocrinology,* 149, 6043-52.

KLOEK, C., HAQ, A. K., DUNN, S. L., LAVERY, H. J., BANKS, A. S. & MYERS, M. G., JR. 2002. Regulation of Jak kinases by intracellular leptin receptor sequences. *Journal of Biological Chemistry,* 277, 41547-55.

KLOTING, N., SCHLEINITZ, D., RUSCHKE, K., BERNDT, J., FASSHAUER, M., TONJES, A., SCHON, M. R., KOVACS, P., STUMVOLL, M. & BLUHER, M. 2008. Inverse relationship between obesity and FTO gene expression in visceral adipose tissue in humans. *Diabetologia,* 51, 641-647.

KO, G. T. & CHAN, J. C. 2008. Burden of obesity--lessons learnt from Hong Kong Chinese. *Obesity Reviews,* 9 Suppl 1, 35-40.

KOBAYASHI, H., OGAWA, Y., SHINTANI, M., EBIHARA, K., SHIMODAHIRA, M., IWAKURA, T., HINO, M., ISHIHARA, T., IKEKUBO, K., KURAHACHI, H. & NAKAO, K. 2002. A Novel homozygous missense mutation of melanocortin-4 receptor (MC4R) in a Japanese woman with severe obesity. *Diabetes,* 51, 243-6.

KOMSU-ORNEK, Z., DEMIREL, F., DURSUN, A., ERMIS, B., PISKIN, E. & BIDECI, A. 2012. Leptin receptor gene Gln223Arg polymorphism is not associated with obesity and metabolic syndrome in Turkish children. *Turk J Pediatr,* 54, 20-4.

KONSTANTINIDES, S., SCHAFER, K., KOSCHNICK, S. & LOSKUTOFF, D. J. 2001. Leptin-dependent platelet aggregation and arterial thrombosis suggests a mechanism for atherothrombotic disease in obesity. *J Clin Invest,* 108, 1533-40.

KONTOGIANNI, M. D., PANAGIOTAKOS, D. B. & SKOPOULI, F. N. 2005. Does body mass index reflect adequately the body fat content in perimenopausal women? *Maturitas,* 51, 307-13.

KOPELMAN, P. & FORMIGUERA, X. 1999. Obesity and type 2 diabetes. *Int J Obes Relat Metab Disord,* 23 Suppl 7, S1.

KORBEL, J. O., URBAN, A. E., GRUBERT, F., DU, J., ROYCE, T. E., STARR, P., ZHONG, G., EMANUEL, B. S., WEISSMAN, S. M., SNYDER, M. & GERSTEIN, M. B. 2007. Systematic prediction and validation of breakpoints associated with copy-number variants in the human genome. *Proc Natl Acad Sci U S A,* 104, 10110-5.

KORNER, A., KIESS, W., STUMVOLL, M. & KOVACS, P. 2008. Polygenic contribution to obesity: genome-wide strategies reveal new targets. *Front Horm Res,* 36, 12-36.

KOUPIL, I., SHESTOV, D. B., SPAREN, P., PLAVINSKAJA, S., PARFENOVA, N. & VAGERO, D. 2007. Blood pressure, hypertension and mortality from circulatory disease in men and women who survived the siege of Leningrad. *Eur J Epidemiol,* 22, 223-34.

KOUTSAKI, M., SIFAKIS, S., ZARAVINOS, A., KOUTROULAKIS, D., KOUKOURA, O. & SPANDIDOS, D. A. 2011. Decreased placental expression of hPGH, IGF-I and IGFBP-1 in pregnancies complicated by fetal growth restriction. *Growth Horm IGF Res,* 21, 31-6.

KRAMER, M. S., OLIVIER, M., MCLEAN, F. H., DOUGHERTY, G. E., WILLIS, D. M. & USHER, R. H. 1990. Determinants of fetal growth and body proportionality. *Pediatrics,* 86, 18-26.

KRISHNAVENI, G. V. 2005. Truncal adiposity is present at birth and in early childhood in South Asian children. *Indian Pediatrics,* 42, 527-538.

KRUPINSKI, J., COUSSEN, F., BAKALYAR, H. A., TANG, W. J., FEINSTEIN, P. G., ORTH, K., SLAUGHTER, C., REED, R. R. & GILMAN, A. G. 1989. Adenylyl cyclase amino acid sequence: possible channel- or transporter-like structure. *Science,* 244, 1558-64.

KRYSHTAFOVYCH, A., FIDELIS, K. & TRAMONTANO, A. 2011. Evaluation of model quality predictions in CASP9. *Proteins,* 79 Suppl 10, 91-106.

KUCZMARSKI, R. J., OGDEN, C. L., GUO, S. S., GRUMMER-STRAWN, L. M., FLEGAL, K. M., MEI, Z., WEI, R., CURTIN, L. R., ROCHE, A. F. & JOHNSON, C. L. 2002. 2000 CDC Growth Charts for the United States: methods and development. *Vital Health Stat 11,* 1-190.

KUH, D., HARDY, R., CHATURVEDI, N. & WADSWORTH, M. E. 2002. Birth weight, childhood growth and abdominal obesity in adult life. *Int J Obes Relat Metab Disord,* 26, 40-7.

KUO, W. L., MONTAG, A. G. & ROSNER, M. R. 1993. Insulin-degrading enzyme is differentially expressed and developmentally regulated in various rat tissues. *Endocrinology,* 132, 604-11.

KURIAN, A. K. & CARDARELLI, K. M. 2007. Racial and ethnic differences in cardiovascular disease risk factors: a systematic review. *Ethn Dis,* 17, 143-52.

KUROCHKIN, I. V. 1998. Amyloidogenic determinant as a substrate recognition motif of insulin-degrading enzyme. *FEBS Lett,* 427, 153-6.

KWAK, S. H., CHO, Y. M., MOON, M. K., KIM, J. H., PARK, B. L., CHEONG, H. S., SHIN, H. D., JANG, H. C., KIM, S. Y., LEE, H. K. & PARK, K. S. 2008. Association of polymorphisms in the insulin-degrading enzyme gene with type 2 diabetes in the Korean population. *Diabetes Res Clin Pract,* 79, 284-90.

KWONG, W. Y., WILD, A. E., ROBERTS, P., WILLIS, A. C. & FLEMING, T. P. 2000. Maternal undernutrition during the preimplantation period of rat development causes blastocyst abnormalities and programming of postnatal hypertension. *Development,* 127, 4195-202.

LABAYEN, I., RUIZ, J. R., ORTEGA, F. B., GOTTRAND, F., HUYBRECHTS, I., DALLONGEVILLE, J., WIDHALM, K., FERRARI, M., BUYKEN, A., KERSTING, M., MOSCHONIS, G., TURCK, D., GOMEZ, S., SJOSTROM, M., MEIRHAEGHE, A. & MORENO, L. A. 2012. Body size at birth modifies the effect of fat mass and obesity associated (FTO) rs9939609 polymorphism on adiposity in adolescents: the Healthy Lifestyle in Europe by Nutrition in Adolescence (HELENA) study. *Br J Nutr,* 107, 1498-504.

LACOURSIERE, D. Y., BLOEBAUM, L., DUNCAN, J. D. & VARNER, M. W. 2005. Population-based trends and correlates of maternal overweight and obesity, Utah 1991-2001. *Am J Obstet Gynecol,* 192, 832-9.

LAITINEN, J., JAASKELAINEN, A., HARTIKAINEN, A. L., SOVIO, U., VAARASMAKI, M., POUTA, A., KAAKINEN, M. & JARVELIN, M. R. 2012. Maternal weight gain during the first half of pregnancy and offspring obesity at 16 years: a prospective cohort study. *BJOG,* 119, 716-23.

LAKKA, T. A., RANKINEN, T., WEISNAGEL, S. J., CHAGNON, Y. C., LAKKA, H. M., UKKOLA, O., BOULE, N., RICE, T., LEON, A. S., SKINNER, J. S., WILMORE, J. H., RAO, D. C., BERGMAN, R. & BOUCHARD, C. 2004. Leptin and leptin receptor gene polymorphisms and changes in glucose homeostasis in response to regular exercise in nondiabetic individuals: the HERITAGE family study. *Diabetes,* 53, 1603-8.

LAMBERTINI, L., MARSIT, C. J., SHARMA, P., MACCANI, M., MA, Y., HU, J. & CHEN, J. 2012. Imprinted gene expression in fetal growth and development. *Placenta,* 33, 480-6.

LANDMAN, J. & CRUICKSHANK, J. K. 2001. A review of ethnicity, health and nutrition-related diseases in relation to migration in the United Kingdom. *Public Health Nutr,* 4, 647-57.

LANGENDONK, J. G., KOK, P., FROLICH, M., PIJL, H. & MEINDERS, A. E. 2006. Decrease in visceral fat following diet-induced weight loss in upper body compared to lower body obese premenopausal women. *Eur J Intern Med,* 17, 465-9.

LANGLEY-EVANS, S. C. & SCULLEY, D. V. 2006. The association between birthweight and longevity in the rat is complex and modulated by maternal protein intake during fetal life. *FEBS Lett,* 580, 4150-3.

LAW, C. M., SHIELL, A. W., NEWSOME, C. A., SYDDALL, H. E., SHINEBOURNE, E. A., FAYERS, P. M., MARTYN, C. N. & DE SWIET, M. 2002. Fetal, infant, and childhood growth and adult blood pressure: a longitudinal study from birth to 22 years of age. *Circulation,* 105, 1088-92.

LAWLOR, D. A., DAVEY SMITH, G. & EBRAHIM, S. 2004. Birth weight is inversely associated with coronary heart disease in post-menopausal women: findings from the British women's heart and health study. *J Epidemiol Community Health,* 58, 120-5.

LAWLOR, D. A., FRASER, A., MACDONALD-WALLIS, C., NELSON, S. M., PALMER, T. M., DAVEY SMITH, G. & TILLING, K. 2011. Maternal and offspring adiposity-related genetic variants and gestational weight gain. *American Journal of Clinical Nutrition,* 94, 149-55.

LAWLOR, D. A. & LEON, D. A. 2005. Association of body mass index and obesity measured in early childhood with risk of coronary heart disease and stroke in middle age: findings from the aberdeen children of the 1950s prospective cohort study. *Circulation,* 111, 1891-6.

LAYTON, J. E., HALL, N. E., CONNELL, F., VENHORST, J. & TREUTLEIN, H. R. 2001. Identification of ligand-binding site III on the immunoglobulin-like domain of the granulocyte colony-stimulating factor receptor. *Journal of Biological Chemistry,* 276, 36779-87.

LE GUELLEC, R., PARIS, J., COUTURIER, A., ROGHI, C. & PHILIPPE, M. 1991. Cloning by differential screening of a Xenopus cDNA that encodes a kinesin-related protein. *Mol Cell Biol,* 11, 3395-8.

LEACH, L. 2002. The phenotype of the human materno-fetal endothelial barrier: molecular occupancy of paracellular junctions dictate permeability and angiogenic plasticity. *J Anat,* 200, 599-606.

LEAR, S. A., HUMPHRIES, K. H., KOHLI, S. & BIRMINGHAM, C. L. 2007. The use of BMI and waist circumference as surrogates of body fat differs by ethnicity. *Obesity (Silver Spring),* 15, 2817-24.

LEE, Y. H., KANG, E. S., KIM, S. H., HAN, S. J., KIM, C. H., KIM, H. J., AHN, C. W., CHA, B. S., NAM, M., NAM, C. M. & LEE, H. C. 2008. Association between polymorphisms in SLC30A8, HHEX, CDKN2A/B, IGF2BP2, FTO, WFS1, CDKAL1,

KCNQ1 and type 2 diabetes in the Korean population. *Journal of Human Genetics,* 53, 991-8.

LEECH, C. A., CASTONGUAY, M. A. & HABENER, J. F. 1999. Expression of adenylyl cyclase subtypes in pancreatic beta-cells. *Biochemical and Biophysical Research Communications,* 254, 703-706.

LEISSRING, M. A., FARRIS, W., WU, X., CHRISTODOULOU, D. C., HAIGIS, M. C., GUARENTE, L. & SELKOE, D. J. 2004. Alternative translation initiation generates a novel isoform of insulin-degrading enzyme targeted to mitochondria. *Biochemical Journal,* 383, 439-46.

LEON, D. A., KOUPILOVA, I., LITHELL, H. O., BERGLUND, L., MOHSEN, R., VAGERO, D., LITHELL, U. B. & MCKEIGUE, P. M. 1996. Failure to realise growth potential in utero and adult obesity in relation to blood pressure in 50 year old Swedish men. *BMJ,* 312, 401-6.

LEON, D. A., LITHELL, H. O., VAGERO, D., KOUPILOVA, I., MOHSEN, R., BERGLUND, L., LITHELL, U. B. & MCKEIGUE, P. M. 1998. Reduced fetal growth rate and increased risk of death from ischaemic heart disease: cohort study of 15 000 Swedish men and women born 1915-29. *BMJ,* 317, 241-5.

LEONG, N. M., MIGNONE, L. I., NEWCOMB, P. A., TITUS-ERNSTOFF, L., BARON, J. A., TRENTHAM-DIETZ, A., STAMPFER, M. J., WILLETT, W. C. & EGAN, K. M. 2003. Early life risk factors in cancer: the relation of birth weight to adult obesity. *International Journal of Cancer,* 103, 789-91.

LEVINE, J. A. 2002. Non-exercise activity thermogenesis (NEAT). *Best Pract Res Clin Endocrinol Metab,* 16, 679-702.

LEVINE, J. A. 2007. Nonexercise activity thermogenesis--liberating the life-force. *J Intern Med,* 262, 273-87.

LEVINE, J. A. & KOTZ, C. M. 2005. NEAT--non-exercise activity thermogenesis--egocentric & geocentric environmental factors vs. biological regulation. *Acta Physiol Scand,* 184, 309-18.

LEVINE, J. A., VANDER WEG, M. W., HILL, J. O. & KLESGES, R. C. 2006. Non-exercise activity thermogenesis: the crouching tiger hidden dragon of societal weight gain. *Arterioscler Thromb Vasc Biol,* 26, 729-36.

LEVITT, N. S., LAMBERT, E. V., WOODS, D., HALES, C. N., ANDREW, R. & SECKL, J. R. 2000. Impaired glucose tolerance and elevated blood pressure in low birth

weight, nonobese, young south african adults: early programming of cortisol axis. *J Clin Endocrinol Metab,* 85, 4611-8.

LEWIS, R. M., CLEAL, J. K. & HANSON, M. A. 2012. Review: Placenta, evolution and lifelong health. *Placenta,* 33 Suppl, S28-32.

LEYVA, F., GODSLAND, I. F., GHATEI, M., PROUDLER, A. J., ALDIS, S., WALTON, C., BLOOM, S. & STEVENSON, J. C. 1998. Hyperleptinemia as a component of a metabolic syndrome of cardiovascular risk. *Arterioscler Thromb Vasc Biol,* 18, 928-33.

LI, R., MAGADIA, J., FEIN, S. B. & GRUMMER-STRAWN, L. M. 2012. Risk of bottle-feeding for rapid weight gain during the first year of life. *Arch Pediatr Adolesc Med,* 166, 431-6.

LI, Y. & ZHANG, Y. 2009. REMO: A new protocol to refine full atomic protein models from C-alpha traces by optimizing hydrogen-bonding networks. *Proteins,* 76, 665-76.

LINDGREN, C. M. 2009. Genome-wide association scan meta-analysis identifies three loci influencing adiposity and fat distribution. *PLoS Genet,* 5.

LINDGREN, C. M., HEID, I. M., RANDALL, J. C., LAMINA, C., STEINTHORSDOTTIR, V., QI, L., SPELIOTES, E. K., THORLEIFSSON, G., WILLER, C. J., HERRERA, B. M., JACKSON, A. U., LIM, N., SCHEET, P., SORANZO, N., AMIN, N., AULCHENKO, Y. S., CHAMBERS, J. C., DRONG, A., LUAN, J., LYON, H. N., RIVADENEIRA, F., SANNA, S., TIMPSON, N. J., ZILLIKENS, M. C., ZHAO, J. H., ALMGREN, P., BANDINELLI, S., BENNETT, A. J., BERGMAN, R. N., BONNYCASTLE, L. L., BUMPSTEAD, S. J., CHANOCK, S. J., CHERKAS, L., CHINES, P., COIN, L., COOPER, C., CRAWFORD, G., DOERING, A., DOMINICZAK, A., DONEY, A. S., EBRAHIM, S., ELLIOTT, P., ERDOS, M. R., ESTRADA, K., FERRUCCI, L., FISCHER, G., FOROUHI, N. G., GIEGER, C., GRALLERT, H., GROVES, C. J., GRUNDY, S., GUIDUCCI, C., HADLEY, D., HAMSTEN, A., HAVULINNA, A. S., HOFMAN, A., HOLLE, R., HOLLOWAY, J. W., ILLIG, T., ISOMAA, B., JACOBS, L. C., JAMESON, K., JOUSILAHTI, P., KARPE, F., KUUSISTO, J., LAITINEN, J., LATHROP, G. M., LAWLOR, D. A., MANGINO, M., MCARDLE, W. L., MEITINGER, T., MORKEN, M. A., MORRIS, A. P., MUNROE, P., NARISU, N., NORDSTROM, A., NORDSTROM, P., OOSTRA, B. A., PALMER, C. N., PAYNE, F., PEDEN, J. F., PROKOPENKO, I., RENSTROM, F., RUOKONEN, A., SALOMAA, V., SANDHU, M. S., SCOTT, L. J., SCUTERI, A., SILANDER, K., SONG, K., YUAN, X., STRINGHAM, H. M., SWIFT, A. J., TUOMI, T., UDA, M., VOLLENWEIDER, P., WAEBER, G., WALLACE, C., WALTERS, G. B., WEEDON, M. N., *et al.* 2009. Genome-wide association scan meta-analysis identifies three Loci influencing adiposity and fat distribution. *PLoS Genet,* 5, e1000508.

LITHELL, H. O., MCKEIGUE, P. M., BERGLUND, L., MOHSEN, R., LITHELL, U. B. & LEON, D. A. 1996. Relation of size at birth to non-insulin dependent diabetes and insulin concentrations in men aged 50-60 years. *BMJ,* 312, 406-10.

LIU, C., LIU, X. J., BARRY, G., LING, N., MAKI, R. A. & DE SOUZA, E. B. 1997. Expression and characterization of a putative high affinity human soluble leptin receptor. *Endocrinology,* 138, 3548-54.

LIU, L., LI, Y., LI, S., HU, N., HE, Y., PONG, R., LIN, D., LU, L. & LAW, M. 2012. Comparison of next-generation sequencing systems. *J Biomed Biotechnol,* 2012, 251364.

LIU, Y. J., LIU, X. G., WANG, L., DINA, C., YAN, H., LIU, J. F., LEVY, S., PAPASIAN, C. J., DREES, B. M., HAMILTON, J. J., MEYRE, D., DELPLANQUE, J., PEI, Y. F., ZHANG, L., RECKER, R. R., FROGUEL, P. & DENG, H. W. 2008. Genome-wide association scans identified CTNNBL1 as a novel gene for obesity. *Hum Mol Genet,* 17, 1803-13.

LIU, Y. J., ROCHA-SANCHEZ, S. M., LIU, P. Y., LONG, J. R., LU, Y., ELZE, L., RECKER, R. R. & DENG, H. W. 2004. Tests of linkage and/or association of the LEPR gene polymorphisms with obesity phenotypes in Caucasian nuclear families. *Physiol Genomics,* 17, 101-6.

LLEWELLYN, C. H., VAN JAARSVELD, C. H., BONIFACE, D., CARNELL, S. & WARDLE, J. 2008. Eating rate is a heritable phenotype related to weight in children. *American Journal of Clinical Nutrition,* 88, 1560-6.

LLOYD, L. J., FOSTER, T., RHODES, P., RHIND, S. M. & GARDNER, D. S. 2012. Protein-energy malnutrition during early gestation in sheep blunts fetal renal vascular and nephron development and compromises adult renal function. *J Physiol,* 590, 377-93.

LONG, N. M., TOUSLEY, C. B., UNDERWOOD, K. R., PAISLEY, S. I., MEANS, W. J., HESS, B. W., DU, M. & FORD, S. P. 2012. Effects of early- to mid-gestational undernutrition with or without protein supplementation on offspring growth, carcass characteristics, and adipocyte size in beef cattle. *Journal of Animal Science,* 90, 197-206.

LOOS, R. J., BEUNEN, G., FAGARD, R., DEROM, C. & VLIETINCK, R. 2001. Birth weight and body composition in young adult men--a prospective twin study. *Int J Obes Relat Metab Disord,* 25, 1537-45.

LOOS, R. J., BEUNEN, G., FAGARD, R., DEROM, C. & VLIETINCK, R. 2002. Birth weight and body composition in young women: a prospective twin study. *American Journal of Clinical Nutrition,* 75, 676-82.

LOOS, R. J., LINDGREN, C. M., LI, S., WHEELER, E., ZHAO, J. H., PROKOPENKO, I., INOUYE, M., FREATHY, R. M., ATTWOOD, A. P., BECKMANN, J. S., BERNDT, S. I., JACOBS, K. B., CHANOCK, S. J., HAYES, R. B., BERGMANN, S., BENNETT, A. J., BINGHAM, S. A., BOCHUD, M., BROWN, M., CAUCHI, S., CONNELL, J. M., COOPER, C., SMITH, G. D., DAY, I., DINA, C., DE, S., DERMITZAKIS, E. T., DONEY, A. S., ELLIOTT, K. S., ELLIOTT, P., EVANS, D. M., SADAF FAROOQI, I., FROGUEL, P., GHORI, J., GROVES, C. J., GWILLIAM, R., HADLEY, D., HALL, A. S., HATTERSLEY, A. T., HEBEBRAND, J., HEID, I. M., LAMINA, C., GIEGER, C., ILLIG, T., MEITINGER, T., WICHMANN, H. E., HERRERA, B., HINNEY, A., HUNT, S. E., JARVELIN, M. R., JOHNSON, T., JOLLEY, J. D., KARPE, F., KENIRY, A., KHAW, K. T., LUBEN, R. N., MANGINO, M., MARCHINI, J., MCARDLE, W. L., MCGINNIS, R., MEYRE, D., MUNROE, P. B., MORRIS, A. D., NESS, A. R., NEVILLE, M. J., NICA, A. C., ONG, K. K., O'RAHILLY, S., OWEN, K. R., PALMER, C. N., PAPADAKIS, K., POTTER, S., POUTA, A., QI, L., RANDALL, J. C., RAYNER, N. W., RING, S. M., SANDHU, M. S., SCHERAG, A., SIMS, M. A., SONG, K., SORANZO, N., SPELIOTES, E. K., SYDDALL, H. E., TEICHMANN, S. A., TIMPSON, N. J., TOBIAS, J. H., UDA, M., VOGEL, C. I., WALLACE, C., WATERWORTH, D. M., WEEDON, M. N., WILLER, C. J., WRAIGHT, YUAN, X., ZEGGINI, E., HIRSCHHORN, J. N., STRACHAN, D. P., OUWEHAND, W. H., CAULFIELD, M. J., *et al*. 2008. Common variants near MC4R are associated with fat mass, weight and risk of obesity. *Nature Genetics,* 40**,** 768-75.

LOPEZ-BERMEJO, A., PETRY, C. J., DIAZ, M., SEBASTIANI, G., DE ZEGHER, F., DUNGER, D. B. & IBANEZ, L. 2008. The association between the FTO gene and fat mass in humans develops by the postnatal age of two weeks. *J Clin Endocrinol Metab,* 93**,** 1501-5.

LOUIS, G. W. & MYERS, M. G. 2007. The role of leptin in the regulation of neuroendocrine function and CNS development. *Reviews in Endocrine & Metabolic Disorders,* 8**,** 85-94.

LUBRANO-BERTHELIER, C., CAVAZOS, M., DUBERN, B., SHAPIRO, A., STUNFF, C. L., ZHANG, S., PICART, F., GOVAERTS, C., FROGUEL, P., BOUGNERES, P., CLEMENT, K. & VAISSE, C. 2003. Molecular genetics of human obesity-associated MC4R mutations. *Ann N Y Acad Sci,* 994**,** 49-57.

LUDWIG, M. G. & SEUWEN, K. 2002. Characterization of the human adenylyl cyclase gene family: cDNA, gene structure, and tissue distribution of the nine isoforms. *J Recept Signal Transduct Res,* 22**,** 79-110.

MA, L., TATARANNI, P. A., BOGARDUS, C. & BAIER, L. J. 2004. Melanocortin 4 receptor gene variation is associated with severe obesity in Pima Indians. *Diabetes,* 53**,** 2696-9.

MACHIN, G. A. 1996. Placental vascular anatomy and twin transfusion syndrome. *Am J Obstet Gynecol,* 174**,** 799-800.

MACLAUGHLIN, S. M., WALKER, S. K., ROBERTS, C. T., KLEEMANN, D. O. & MCMILLEN, I. C. 2005. Periconceptional nutrition and the relationship between maternal body weight changes in the periconceptional period and feto-placental growth in the sheep. *J Physiol,* 565, 111-24.

MAES, H. H., NEALE, M. C. & EAVES, L. J. 1997. Genetic and environmental factors in relative body weight and human adiposity. *Behavior Genetics,* 27, 325-51.

MAGNUS, P., BERG, K. & BJERKEDAL, T. 1985. The association of parity and birth weight: testing the sensitization hypothesis. *Early Human Development,* 12, 49-54.

MALIS, C., RASMUSSEN, E. L., POULSEN, P., PETERSEN, I., CHRISTENSEN, K., BECK-NIELSEN, H., ASTRUP, A. & VAAG, A. A. 2005. Total and regional fat distribution is strongly influenced by genetic factors in young and elderly twins. *Obes Res,* 13, 2139-45.

MAMUN, A. A., CALLAWAY, L. K., O'CALLAGHAN, M. J., WILLIAMS, G. M., NAJMAN, J. M., ALATI, R., CLAVARINO, A. & LAWLOR, D. A. 2011. Associations of maternal pre-pregnancy obesity and excess pregnancy weight gains with adverse pregnancy outcomes and length of hospital stay. *BMC Pregnancy Childbirth,* 11, 62.

MANCOUR, L. V., DAGHESTANI, H. N., DUTTA, S., WESTFIELD, G. H., SCHILLING, J., OLESKIE, A. N., HERBSTMAN, J. F., CHOU, S. Z. & SKINIOTIS, G. 2012. Ligand-Induced Architecture of the Leptin Receptor Signaling Complex. *Mol Cell.*

MANDAL, D., MANDA, S., RAKSHI, A., DEY, R. P., BISWAS, S. C. & BANERJEE, A. 2011. Maternal obesity and pregnancy outcome: a prospective analysis. *J Assoc Physicians India,* 59, 486-9.

MARTINEZ-BARBERA, J. P., RODRIGUEZ, T. A. & BEDDINGTON, R. S. 2000. The homeobox gene Hesx1 is required in the anterior neural ectoderm for normal forebrain formation. *Dev Biol,* 223, 422-30.

MARTYN, C. N., BARKER, D. J. & OSMOND, C. 1996. Mothers' pelvic size, fetal growth, and death from stroke and coronary heart disease in men in the UK. *Lancet,* 348, 1264-8.

MATSUZAKI, H., WANG, P. H., HU, J., RAVA, R. & FU, G. K. 2009. High resolution discovery and confirmation of copy number variants in 90 Yoruba Nigerians. *Genome Biol,* 10, R125.

MATTEVI, V. S., ZEMBRZUSKI, V. M. & HUTZ, M. H. 2002. Association analysis of genes involved in the leptin-signaling pathway with obesity in Brazil. *Int J Obes Relat Metab Disord,* 26, 1179-85.

MAURIEGE, P., MARETTE, A., ATGIE, C., BOUCHARD, C., THERIAULT, G., BUKOWIECKI, L. K., MARCEAU, P., BIRON, S., NADEAU, A. & DESPRES, J. P. 1995. Regional variation in adipose tissue metabolism of severely obese premenopausal women. *J Lipid Res,* 36, 672-84.

MAZEN, I., EL-GAMMAL, M., ABDEL-HAMID, M., FAROOQI, I. S. & AMR, K. 2011. Homozygosity for a novel missense mutation in the leptin receptor gene (P316T) in two Egyptian cousins with severe early onset obesity. *Mol Genet Metab,* 102, 461-4.

MCCARTHY, H. D. 2001. The devlopment of wasit circumference precentiles in British children aged 5.0 - 16.9 years. *European Journal of Clincial Nutrition,* 55, 902 - 907.

MCCARTHY, H. D. & ASHWELL, M. 2006. A study of central fatness using waist-to-height ratios in UK children and adolescents over two decades supports the simple message--'keep your waist circumference to less than half your height'. *Int J Obes (Lond),* 30, 988-92.

MCCARTHY, H. D., COLE, T. J., FRY, T., JEBB, S. A. & PRENTICE, A. M. 2006. Body fat reference curves for children. *Int J Obes (Lond),* 30, 598-602.

MCKEIGUE, P. M., LITHELL, H. O. & LEON, D. A. 1998. Glucose tolerance and resistance to insulin-stimulated glucose uptake in men aged 70 years in relation to size at birth. *Diabetologia,* 41, 1133-8.

MCMILLEN, I. C., MACLAUGHLIN, S. M., MUHLHAUSLER, B. S., GENTILI, S., DUFFIELD, J. L. & MORRISON, J. L. 2008. Developmental origins of adult health and disease: the role of periconceptional and foetal nutrition. *Basic Clin Pharmacol Toxicol,* 102, 82-9.

MCMINN, J., WEI, M., SCHUPF, N., CUSMAI, J., JOHNSON, E. B., SMITH, A. C., WEKSBERG, R., THAKER, H. M. & TYCKO, B. 2006. Unbalanced placental expression of imprinted genes in human intrauterine growth restriction. *Placenta,* 27, 540-9.

MEI, H., CHEN, W., JIANG, F., HE, J., SRINIVASAN, S., SMITH, E. N., SCHORK, N., MURRAY, S. & BERENSON, G. S. 2012. Longitudinal replication studies of GWAS risk SNPs influencing body mass index over the course of childhood and adulthood. *PLoS One,* 7, e31470.

MEI, H., CHEN, W., SRINIVASAN, S. R., JIANG, F., SCHORK, N., MURRAY, S., SMITH, E., SO, J. D. & BERENSON, G. S. 2010. FTO influences on longitudinal BMI over childhood and adulthood and modulation on relationship between birth weight and longitudinal BMI. *Hum Genet,* 128, 589-96.

MEISINGER, C., DORING, A., THORAND, B., HEIER, M. & LOWEL, H. 2006. Body fat distribution and risk of type 2 diabetes in the general population: are there differences between men and women? The MONICA/KORA Augsburg Cohort Study. *American Journal of Clinical Nutrition,* 84, 483-489.

MERGEN, M., MERGEN, H., OZATA, M., ONER, R. & ONER, C. 2001. A novel melanocortin 4 receptor (MC4R) gene mutation associated with morbid obesity. *J Clin Endocrinol Metab,* 86, 3448.

MERICQ, V., ONG, K. K., BAZAES, R., PENA, V., AVILA, A., SALAZAR, T., SOTO, N., INIGUEZ, G. & DUNGER, D. B. 2005. Longitudinal changes in insulin sensitivity and secretion from birth to age three years in small- and appropriate-for-gestational-age children. *Diabetologia,* 48, 2609-14.

MEYRE, D., DELPLANQUE, J., CHEVRE, J. C., LECOEUR, C., LOBBENS, S., GALLINA, S., DURAND, E., VATIN, V., DEGRAEVE, F., PROENCA, C., GAGET, S., KORNER, A., KOVACS, P., KIESS, W., TICHET, J., MARRE, M., HARTIKAINEN, A. L., HORBER, F., POTOCZNA, N., HERCBERG, S., LEVY-MARCHAL, C., PATTOU, F., HEUDE, B., TAUBER, M., MCCARTHY, M. I., BLAKEMORE, A. I., MONTPETIT, A., POLYCHRONAKOS, C., WEILL, J., COIN, L. J., ASHER, J., ELLIOTT, P., JARVELIN, M. R., VISVIKIS-SIEST, S., BALKAU, B., SLADEK, R., BALDING, D., WALLEY, A., DINA, C. & FROGUEL, P. 2009. Genome-wide association study for early-onset and morbid adult obesity identifies three new risk loci in European populations. *Nature Genetics,* 41, 157-9.

MEYRE, D., LECOEUR, C., DELPLANQUE, J., FRANCKE, S., VATIN, V., DURAND, E., WEILL, J., DINA, C. & FROGUEL, P. 2004. A genome-wide scan for childhood obesity-associated traits in French families shows significant linkage on chromosome 6q22.31-q23.2. *Diabetes,* 53, 803-11.

MEYRE, D., PROULX, K., KAWAGOE-TAKAKI, H., VATIN, V., GUTIERREZ-AGUILAR, R., LYON, D., MA, M., CHOQUET, H., HORBER, F., VAN HUL, W., VAN GAAL, L., BALKAU, B., VISVIKIS-SIEST, S., PATTOU, F., FAROOQI, I. S., SAUDEK, V., O'RAHILLY, S., FROGUEL, P., SEDGWICK, B. & YEO, G. S. 2010. Prevalence of loss-of-function FTO mutations in lean and obese individuals. *Diabetes,* 59, 311-8.

MISRA, A. & KHURANA, L. 2009. The metabolic syndrome in South Asians: epidemiology, determinants, and prevention. *Metab Syndr Relat Disord,* 7, 497-514.

MITCHELL, G. F., VERWOERT, G. C., TARASOV, K. V., ISAACS, A., SMITH, A. V., YASMIN, RIETZSCHEL, E. R., TANAKA, T., LIU, Y., PARSA, A., NAJJAR, S. S., O'SHAUGHNESSY, K. M., SIGURDSSON, S., DE BUYZERE, M. L., LARSON, M. G., SIE, M. P., ANDREWS, J. S., POST, W. S., MATTACE-RASO, F. U., MCENIERY, C. M., EIRIKSDOTTIR, G., SEGERS, P., VASAN, R. S., VAN RIJN, M. J., HOWARD, T. D., MCARDLE, P. F., DEHGHAN, A., JEWELL, E. S., NEWHOUSE, S. J., BEKAERT, S., HAMBURG, N. M., NEWMAN, A. B., HOFMAN, A., SCUTERI, A., DE BACQUER, D., ARFAN IKRAM, M., PSATY, B. M., FUCHSBERGER, C., OLDEN, M., WAIN, L. V., ELLIOTT, P., SMITH, N. L., FELIX, J. F., ERDMANN, J., VITA, J. A., SUTTON-TYRRELL, K., SIJBRANDS, E. J., SANNA, S., LAUNER, L. J., DE MEYER, T., JOHNSON, A. D., SCHUT, A. F., HERRINGTON, D. M., RIVADENEIRA, F., UDA, M., WILKINSON, I. B., ASPELUND, T., GILLEBERT, T. C., VAN BORTEL, L., BENJAMIN, E. J., OOSTRA, B. A., DING, J., GIBSON, Q., UITTERLINDEN, A. G., ABECASIS, G. R., COCKCROFT, J. R., GUDNASON, V., DE BACKER, G. G., FERRUCCI, L., HARRIS, T. B., SHULDINER, A. R., VAN DUIJN, C. M., LEVY, D., LAKATTA, E. G. & WITTEMAN, J. C. 2012. Common genetic variation in the 3'-BCL11B gene desert is associated with carotid-femoral pulse wave velocity and excess cardiovascular disease risk: the AortaGen Consortium. *Circ Cardiovasc Genet,* 5**,** 81-90.

MITTELMAN, S. D., VAN CITTERS, G. W., KIRKMAN, E. L. & BERGMAN, R. N. 2002. Extreme insulin resistance of the central adipose depot in vivo. *Diabetes,* 51**,** 755-61.

MODI, N., THOMAS, E. L., UTHAYA, S. N., UMRANIKAR, S., BELL, J. D. & YAJNIK, C. 2009. Whole Body Magnetic Resonance Imaging of Healthy Newborn Infants Demonstrates Increased Central Adiposity in Asian Indians. *Pediatric Research,* 65**,** 584-587.

MONASTA, L., LOBSTEIN, T., COLE, T. J., VIGNEROVA, J. & CATTANEO, A. 2011. Defining overweight and obesity in pre-school children: IOTF reference or WHO standard? *Obesity Reviews,* 12**,** 295-300.

MONTAGUE, C. T., FAROOQI, I. S., WHITEHEAD, J. P., SOOS, M. A., RAU, H., WAREHAM, N. J., SEWTER, C. P., DIGBY, J. E., MOHAMMED, S. N., HURST, J. A., CHEETHAM, C. H., EARLEY, A. R., BARNETT, A. H., PRINS, J. B. & O'RAHILLY, S. 1997. Congenital leptin deficiency is associated with severe early-onset obesity in humans. *Nature,* 387**,** 903-8.

MOORE, V. M., COCKINGTON, R. A., RYAN, P. & ROBINSON, J. S. 1999. The relationship between birth weight and blood pressure amplifies from childhood to adulthood. *J Hypertens,* 17**,** 883-8.

MORITA, M., KUROCHKIN, I. V., MOTOJIMA, K., GOTO, S., TAKANO, T., OKAMURA, S., SATO, R., YOKOTA, S. & IMANAKA, T. 2000. Insulin-degrading enzyme exists inside of rat liver peroxisomes and degrades oxidized proteins. *Cell Struct Funct,* 25**,** 309-15.

MORTON, G. J. & SCHWARTZ, M. W. 2011. Leptin and the central nervous system control of glucose metabolism. *Physiol Rev,* 91, 389-411.

MOTZKAU, M., MEYER, P., MERTENS, P. R. & KLOSE, S. 2012. Monogenic diabetes in a family with 2 unknown HNF-4A gene mutations. *Exp Clin Endocrinol Diabetes,* 120, 89-90.

MUNZBERG, H., FLIER, J. S. & BJORBAEK, C. 2004. Region-specific leptin resistance within the hypothalamus of diet-induced obese mice. *Endocrinology,* 145, 4880-9.

MUST, A., SPADANO, J., COAKLEY, E. H., FIELD, A. E., COLDITZ, G. & DIETZ, W. H. 1999. The disease burden associated with overweight and obesity. *JAMA,* 282, 1523-9.

MYERS, M. G., JR., HEYMSFIELD, S. B., HAFT, C., KAHN, B. B., LAUGHLIN, M., LEIBEL, R. L., TSCHOP, M. H. & YANOVSKI, J. A. 2012. Challenges and opportunities of defining clinical leptin resistance. *Cell Metab,* 15, 150-6.

NAKASHIMA, K., NARAZAKI, M. & TAGA, T. 1997. Leptin receptor (OB-R) oligomerizes with itself but not with its closely related cytokine signal transducer gp130. *FEBS Lett,* 403, 79-82.

NAN, C., GUO, B., WARNER, C., FOWLER, T., BARRETT, T., BOOMSMA, D., NELSON, T., WHITFIELD, K., BEUNEN, G., THOMIS, M., MAES, H. H., DEROM, C., ORDONANA, J., DEEKS, J. & ZEEGERS, M. 2012. Heritability of body mass index in pre-adolescence, young adulthood and late adulthood. *Eur J Epidemiol,* 27, 247-53.

NEEDLEMAN, S. B. & WUNSCH, C. D. 1970. A general method applicable to the search for similarities in the amino acid sequence of two proteins. *J Mol Biol,* 48, 443-53.

NELSON, S. M., MATTHEWS, P. & POSTON, L. 2010. Maternal metabolism and obesity: modifiable determinants of pregnancy outcome. *Hum Reprod Update,* 16, 255-75.

NELSON, T. L., VOGLER, G. P., PEDERSEN, N. L. & MILES, T. P. 1999. Genetic and environmental influences on waist-to-hip ratio and waist circumference in an older Swedish twin population. *Int J Obes Relat Metab Disord,* 23, 449-55.

NEVILL, A. M., STEWART, A. D., OLDS, T. & HOLDER, R. 2006. Relationship between adiposity and body size reveals limitations of BMI. *American Journal of Physical Anthropology,* 129, 151-6.

NEWMAN, C. S., CHIA, F. & KRIEG, P. A. 1997. The XHex homeobox gene is expressed during development of the vascular endothelium: overexpression leads to an increase in vascular endothelial cell number. *Mech Dev,* 66, 83-93.

NG, M. C., HESTER, J. M., WING, M. R., LI, J., XU, J., HICKS, P. J., ROH, B. H., LU, L., DIVERS, J., LANGEFELD, C. D., FREEDMAN, B. I., PALMER, N. D. & BOWDEN, D. W. 2012. Genome-wide association of BMI in African Americans. *Obesity (Silver Spring),* 20, 622-7.

NG, M. C., PARK, K. S., OH, B., TAM, C. H., CHO, Y. M., SHIN, H. D., LAM, V. K., MA, R. C., SO, W. Y., CHO, Y. S., KIM, H. L., LEE, H. K., CHAN, J. C. & CHO, N. H. 2008. Implication of genetic variants near TCF7L2, SLC30A8, HHEX, CDKAL1, CDKN2A/B, IGF2BP2, and FTO in type 2 diabetes and obesity in 6,719 Asians. *Diabetes,* 57, 2226-33.

NISWENDER, K. D. & SCHWARTZ, M. W. 2003. Insulin and leptin revisited: adiposity signals with overlapping physiological and intracellular signaling capabilities. *Front Neuroendocrinol,* 24, 1-10.

NIV-SPECTOR, L., RAVER, N., FRIEDMAN-EINAT, M., GROSCLAUDE, J., GUSSAKOVSKY, E. E., LIVNAH, O. & GERTLER, A. 2005. Mapping leptin-interacting sites in recombinant leptin-binding domain (LBD) subcloned from chicken leptin receptor. *Biochemical Journal,* 390, 475-84.

NOHR, E. A., BECH, B. H., DAVIES, M. J., FRYDENBERG, M., HENRIKSEN, T. B. & OLSEN, J. 2005. Prepregnancy obesity and fetal death: a study within the Danish National Birth Cohort. *Obstet Gynecol,* 106, 250-9.

NYHOLT, D. R. 2004. A simple correction for multiple testing for single-nucleotide polymorphisms in linkage disequilibrium with each other. *Am J Hum Genet,* 74, 765-9.

O'DONOVAN, G., THOMAS, E. L., MCCARTHY, J. P., FITZPATRICK, J., DURIGHEL, G., MEHTA, S., MORIN, S. X., GOLDSTONE, A. P. & BELL, J. D. 2009. Fat distribution in men of different waist girth, fitness level and exercise habit. *Int J Obes (Lond),* 33, 1356-62.

O'RAHILLY, S. & FAROOQI, I. S. 2008. Human obesity: a heritable neurobehavioral disorder that is highly sensitive to environmental conditions. *Diabetes,* 57, 2905-10.

ODDY, W. H. 2012. Infant feeding and obesity risk in the child. *Breastfeed Rev,* 20, 7-12.

ODE, J. J., PIVARNIK, J. M., REEVES, M. J. & KNOUS, J. L. 2007. Body mass index as a predictor of percent fat in college athletes and nonathletes. *Medicine and Science in Sports and Exercise,* 39, 403-409.

OGDEN, C. L., CARROLL, M. D., CURTIN, L. R., MCDOWELL, M. A., TABAK, C. J. & FLEGAL, K. M. 2006. Prevalence of overweight and obesity in the United States, 1999-2004. *JAMA,* 295, 1549-55.

OGDEN, C. L., CARROLL, M. D., KIT, B. K. & FLEGAL, K. M. 2012a. Prevalence of obesity and trends in body mass index among US children and adolescents, 1999-2010. *JAMA,* 307, 483-90.

OGDEN, C. L., CARROLL, M. E., KIT, B. K. & FLEGAL, K. M. 2012b. Prevalence of obesity in the United States, 2009-2010. *NCHS Data Brief,* 1-8.

OGDEN, C. L., FLEGAL, K. M., CARROLL, M. D. & JOHNSON, C. L. 2002. Prevalence and trends in overweight among US children and adolescents, 1999-2000. *JAMA,* 288, 1728-32.

OGURA, K., SAKATA, M., OKAMOTO, Y., YASUI, Y., TADOKORO, C., YOSHIMOTO, Y., YAMAGUCHI, M., KURACHI, H., MAEDA, T. & MURATA, Y. 2000. 8-bromo-cyclicAMP stimulates glucose transporter-1 expression in a human choriocarcinoma cell line. *J Endocrinol,* 164, 171-8.

OKORODUDU, D. O., JUMEAN, M. F., MONTORI, V. M., ROMERO-CORRAL, A., SOMERS, V. K., ERWIN, P. J. & LOPEZ-JIMENEZ, F. 2010. Diagnostic performance of body mass index to identify obesity as defined by body adiposity: a systematic review and meta-analysis. *Int J Obes (Lond),* 34, 791-9.

OKOSUN, I. S., LIAO, Y., ROTIMI, C. N., CHOI, S. & COOPER, R. S. 2000. Predictive values of waist circumference for dyslipidemia, type 2 diabetes and hypertension in overweight White, Black, and Hispanic American adults. *J Clin Epidemiol,* 53, 401-8.

OLIVER, H., JAMESON, K. A., SAYER, A. A., COOPER, C. & DENNISON, E. M. 2007. Growth in early life predicts bone strength in late adulthood: the Hertfordshire Cohort Study. *Bone,* 41, 400-5.

OMORI, S., TANAKA, Y., TAKAHASHI, A., HIROSE, H., KASHIWAGI, A., KAKU, K., KAWAMORI, R., NAKAMURA, Y. & MAEDA, S. 2008. Association of CDKAL1, IGF2BP2, CDKN2A/B, HHEX, SLC30A8, and KCNJ11 with susceptibility to type 2 diabetes in a Japanese population. *Diabetes,* 57, 791-5.

ONG, K. K., AHMED, M. L., EMMETT, P. M., PREECE, M. A. & DUNGER, D. B. 2000. Association between postnatal catch-up growth and obesity in childhood: prospective cohort study. *BMJ,* 320, 967-71.

ONG, K. K. & DUNGER, D. B. 2004. Birth weight, infant growth and insulin resistance. *Eur J Endocrinol,* 151 Suppl 3, U131-9.

ONG, K. K., PETRY, C. J., EMMETT, P. M., SANDHU, M. S., KIESS, W., HALES, C. N., NESS, A. R. & DUNGER, D. B. 2004. Insulin sensitivity and secretion in normal children related to size at birth, postnatal growth, and plasma insulin-like growth factor-I levels. *Diabetologia,* 47, 1064-70.

ONG, K. K., PREECE, M. A., EMMETT, P. M., AHMED, M. L. & DUNGER, D. B. 2002. Size at birth and early childhood growth in relation to maternal smoking, parity and infant breast-feeding: longitudinal birth cohort study and analysis. *Pediatric Research,* 52, 863-7.

OSGERBY, J. C., WATHES, D. C., HOWARD, D. & GADD, T. S. 2004. The effect of maternal undernutrition on the placental growth trajectory and the uterine insulin-like growth factor axis in the pregnant ewe. *J Endocrinol,* 182, 89-103.

OSMOND, C., BARKER, D. J., WINTER, P. D., FALL, C. H. & SIMMONDS, S. J. 1993. Early growth and death from cardiovascular disease in women. *BMJ,* 307, 1519-24.

OSMOND, C., KAJANTIE, E., FORSEN, T. J., ERIKSSON, J. G. & BARKER, D. J. 2007. Infant growth and stroke in adult life: the Helsinki birth cohort study. *Stroke,* 38, 264-70.

OZANNE, S. E. & HALES, C. N. 2004. Lifespan: catch-up growth and obesity in male mice. *Nature,* 427, 411-2.

OZANNE, S. E., JENSEN, C. B., TINGEY, K. J., STORGAARD, H., MADSBAD, S. & VAAG, A. A. 2005. Low birthweight is associated with specific changes in muscle insulin-signalling protein expression. *Diabetologia,* 48, 547-52.

PAINTER, R. C., DE ROOIJ, S. R., BOSSUYT, P. M., SIMMERS, T. A., OSMOND, C., BARKER, D. J., BLEKER, O. P. & ROSEBOOM, T. J. 2006. Early onset of coronary artery disease after prenatal exposure to the Dutch famine. *American Journal of Clinical Nutrition,* 84, 322-7; quiz 466-7.

PARACCHINI, V., PEDOTTI, P. & TAIOLI, E. 2005. Genetics of leptin and obesity: a HuGE review. *Am J Epidemiol,* 162, 101-14.

PARK, H., KIM, J. I., JU, Y. S., GOKCUMEN, O., MILLS, R. E., KIM, S., LEE, S., SUH, D., HONG, D., KANG, H. P., YOO, Y. J., SHIN, J. Y., KIM, H. J., YAVARTANOO, M., CHANG, Y. W., HA, J. S., CHONG, W., HWANG, G. R., DARVISHI, K., KIM, H., YANG, S. J., YANG, K. S., KIM, H., HURLES, M. E., SCHERER, S. W., CARTER, N. P., TYLER-SMITH, C., LEE, C. & SEO, J. S. 2010. Discovery of common Asian copy number variants using integrated high-resolution array CGH and massively parallel DNA sequencing. *Nature Genetics,* 42, 400-5.

PARK, Y. S. & KIM, J. S. 2012. Obesity phenotype and coronary heart disease risk as estimated by the Framingham risk score. *J Korean Med Sci,* 27, 243-9.

PARSONS, T. J., POWER, C. & MANOR, O. 2001. Fetal and early life growth and body mass index from birth to early adulthood in 1958 British cohort: longitudinal study. *BMJ,* 323, 1331-5.

PASCOE, L., TURA, A., PATEL, S. K., IBRAHIM, I. M., FERRANNINI, E., ZEGGINI, E., WEEDON, M. N., MARI, A., HATTERSLEY, A. T., MCCARTHY, M. I., FRAYLING, T. M. & WALKER, M. 2007. Common variants of the novel type 2 diabetes genes CDKAL1 and HHEX/IDE are associated with decreased pancreatic beta-cell function. *Diabetes,* 56, 3101-4.

PATERSON, J. M., SMITH, S. M., HARMAR, A. J. & ANTONI, F. A. 1995. Control of a novel adenylyl cyclase by calcineurin. *Biochem Biophys Res Commun,* 214, 1000-8.

PEELMAN, F., COUTURIER, C., DAM, J., ZABEAU, L., TAVERNIER, J. & JOCKERS, R. 2006a. Techniques: new pharmacological perspectives for the leptin receptor. *Trends Pharmacol Sci,* 27, 218-25.

PEELMAN, F., ISERENTANT, H., DE SMET, A. S., VANDEKERCKHOVE, J., ZABEAU, L. & TAVERNIER, J. 2006b. Mapping of binding site III in the leptin receptor and modeling of a hexameric leptin.leptin receptor complex. *Journal of Biological Chemistry,* 281, 15496-504.

PEELMAN, F., VAN BENEDEN, K., ZABEAU, L., ISERENTANT, H., ULRICHTS, P., DEFEAU, D., VERHEE, A., CATTEEUW, D., ELEWAUT, D. & TAVERNIER, J. 2004. Mapping of the leptin binding sites and design of a leptin antagonist. *Journal of Biological Chemistry,* 279, 41038-46.

PEETERS, A. 2007. Variants in the FTO gene are associated with common obesity in the Belgian population. *Molecular Genetics and Metabolism,* 93, 481 - 484.

PEETERS, M. W., BEUNEN, G. P., MAES, H. H., LOOS, R. J., CLAESSENS, A. L., VLIETINCK, R. & THOMIS, M. A. 2007. Genetic and environmental determination of

tracking in subcutaneous fat distribution during adolescence. *American Journal of Clinical Nutrition,* 86, 652-60.

PEETERS, M. W., THOMIS, M. A., CLAESSENS, A. L., LOOS, R. J., MAES, H. H., LYSENS, R., VANDEN EYNDE, B., VLIETINCK, R. & BEUNEN, G. 2003. Heritability of somatotype components from early adolescence into young adulthood: a multivariate analysis on a longitudinal twin study. *Ann Hum Biol,* 30, 402-18.

PETERS, T., AUSMEIER, K. & RUTHER, U. 1999. Cloning of Fatso (Fto), a novel gene deleted by the Fused toes (Ft) mouse mutation. *Mamm Genome,* 10, 983-6.

PETRY, C. J., SEEAR, R. V., WINGATE, D. L., ACERINI, C. L., ONG, K. K., HUGHES, I. A. & DUNGER, D. B. 2011. Maternally transmitted foetal H19 variants and associations with birth weight. *Hum Genet,* 130, 663-70.

PHILLIPS, D. I., GOULDEN, P., SYDDALL, H. E., AIHIE SAYER, A., DENNISON, E. M., MARTIN, H. & COOPER, C. 2005. Fetal and infant growth and glucose tolerance in the Hertfordshire Cohort Study: a study of men and women born between 1931 and 1939. *Diabetes,* 54 Suppl 2, S145-50.

PICKHARDT, P. J., JEE, Y., O'CONNOR, S. D. & DEL RIO, A. M. 2012. Visceral Adiposity and Hepatic Steatosis at Abdominal CT: Association With the Metabolic Syndrome. *AJR Am J Roentgenol,* 198, 1100-7.

PISCHON, T., BOEING, H., HOFFMANN, K., BERGMANN, M., SCHULZE, M. B., OVERVAD, K., VAN DER SCHOUW, Y. T., SPENCER, E., MOONS, K. G., TJONNELAND, A., HALKJAER, J., JENSEN, M. K., STEGGER, J., CLAVEL-CHAPELON, F., BOUTRON-RUAULT, M. C., CHAJES, V., LINSEISEN, J., KAAKS, R., TRICHOPOULOU, A., TRICHOPOULOS, D., BAMIA, C., SIERI, S., PALLI, D., TUMINO, R., VINEIS, P., PANICO, S., PEETERS, P. H., MAY, A. M., BUENO-DE-MESQUITA, H. B., VAN DUIJNHOVEN, F. J., HALLMANS, G., WEINEHALL, L., MANJER, J., HEDBLAD, B., LUND, E., AGUDO, A., ARRIOLA, L., BARRICARTE, A., NAVARRO, C., MARTINEZ, C., QUIROS, J. R., KEY, T., BINGHAM, S., KHAW, K. T., BOFFETTA, P., JENAB, M., FERRARI, P. & RIBOLI, E. 2008. General and abdominal adiposity and risk of death in Europe. *N Engl J Med,* 359, 2105-20.

POULIOT, M. C., DESPRES, J. P., NADEAU, A., MOORJANI, S., PRUD'HOMME, D., LUPIEN, P. J., TREMBLAY, A. & BOUCHARD, C. 1992. Visceral obesity in men. Associations with glucose tolerance, plasma insulin, and lipoprotein levels. *Diabetes,* 41, 826-34.

PYRZAK, B., WISNIEWSKA, A., KUCHARSKA, A., WASIK, M. & DEMKOW, U. 2009. No association of LEPR Gln223Arg polymorphism with leptin, obesity or metabolic disturbances in children. *Eur J Med Res,* 14 Suppl 4, 201-4.

QU, Y., YANG, Z., JIN, F., SUN, L., FENG, J., TANG, L., ZHANG, C., ZHU, X., SHI, X., SUN, H., WANG, B. & WANG, L. 2008. The haplotype identified in LEPR gene is associated with type 2 diabetes mellitus in Northern Chinese. *Diabetes Res Clin Pract,* 81, 33-7.

QUINTON, N. D., LAIRD, S. M., OKON, M. A., LI, T. C., SMITH, R. F., ROSS, R. J. & BLAKEMORE, A. I. 1999. Serum leptin levels during the menstrual cycle of healthy fertile women. *Br J Biomed Sci,* 56, 16-9.

QUINTON, N. D., LEE, A. J., ROSS, R. J., EASTELL, R. & BLAKEMORE, A. I. 2001. A single nucleotide polymorphism (SNP) in the leptin receptor is associated with BMI, fat mass and leptin levels in postmenopausal Caucasian women. *Hum Genet,* 108, 233-6.

RAGHUPATHY, P., ANTONISAMY, B., GEETHANJALI, F. S., SAPERIA, J., LEARY, S. D., PRIYA, G., RICHARD, J., BARKER, D. J. & FALL, C. H. 2010. Glucose tolerance, insulin resistance and insulin secretion in young south Indian adults: Relationships to parental size, neonatal size and childhood body mass index. *Diabetes Res Clin Pract,* 87, 283-92.

RAGVIN, A., MORO, E., FREDMAN, D., NAVRATILOVA, P., DRIVENES, O., ENGSTROM, P. G., ALONSO, M. E., DE LA CALLE MUSTIENES, E., GOMEZ SKARMETA, J. L., TAVARES, M. J., CASARES, F., MANZANARES, M., VAN HEYNINGEN, V., MOLVEN, A., NJOLSTAD, P. R., ARGENTON, F., LENHARD, B. & BECKER, T. S. 2010. Long-range gene regulation links genomic type 2 diabetes and obesity risk regions to HHEX, SOX4, and IRX3. *Proc Natl Acad Sci U S A,* 107, 775-80.

RALAT, L. A., KALAS, V., ZHENG, Z., GOLDMAN, R. D., SOSNICK, T. R. & TANG, W. J. 2011. Ubiquitin is a novel substrate for human insulin-degrading enzyme. *J Mol Biol,* 406, 454-66.

RAMACHANDRAN, A., SATHYAMURTHY, I., SNEHALATHA, C., SATYAVANI, K., SIVASANKARI, S., MISRA, J., GIRINATH, M. R. & VISWANATHAN, V. 2001a. Risk variables for coronary artery disease in Asian Indians. *Am J Cardiol,* 87, 267-71.

RAMACHANDRAN, A., SNEHALATHA, C., KAPUR, A., VIJAY, V., MOHAN, V., DAS, A. K., RAO, P. V., YAJNIK, C. S., KUMAR, K. M. P., NAIR, J. D. & GRP, D. E. S. 2001b. High prevalence of diabetes and impaired glucose tolerance in India: National Urban Diabetes Survey. *Diabetologia,* 44, 1094-1101.

RAMARAJ, R. & CHELLAPPA, P. 2008. Cardiovascular risk in South Asians. *Postgrad Med J,* 84, 518-23.

RAMSEY, J. J., KEMNITZ, J. W., COLMAN, R. J., CUNNINGHAM, D. & SWICK, A. G. 1998. Different central and peripheral responses to leptin in rhesus monkeys: brain transport may be limited. *J Clin Endocrinol Metab,* 83, 3230-5.

RANADE, S. C., SARFARAZ NAWAZ, M., KUMAR RAMBTLA, P., ROSE, A. J., GRESSENS, P. & MANI, S. 2012. Early protein malnutrition disrupts cerebellar development and impairs motor coordination. *Br J Nutr,* 107, 1167-75.

RAND, L., WINCHESTER, E. C., MILLWOOD, I. Y., PENNY, M. A. & KESSLING, A. M. 2001. Maternal leptin receptor gene variant Gln223Arg is not associated with variation in birth weight or maternal body mass index in UK and South Asian populations. *Int J Obes Relat Metab Disord,* 25, 753-5.

RANKINEN, T., ZUBERI, A., CHAGNON, Y. C., WEISNAGEL, S. J., ARGYROPOULOS, G., WALTS, B., PERUSSE, L. & BOUCHARD, C. 2006. The human obesity gene map: the 2005 update. *Obesity (Silver Spring),* 14, 529-644.

RASMUSSEN, K. M., CATALANO, P. M. & YAKTINE, A. L. 2009. New guidelines for weight gain during pregnancy: what obstetrician/gynecologists should know. *Curr Opin Obstet Gynecol,* 21, 521-6.

RAU, H., REAVES, B. J., O'RAHILLY, S. & WHITEHEAD, J. P. 1999. Truncated human leptin (Delta 133) associated with extreme obesity undergoes proteasomal degradation after defective intracellular transport. *Endocrinology,* 140, 1718-1723.

RAVELLI, A. C., VAN DER MEULEN, J. H., MICHELS, R. P., OSMOND, C., BARKER, D. J., HALES, C. N. & BLEKER, O. P. 1998. Glucose tolerance in adults after prenatal exposure to famine. *Lancet,* 351, 173-7.

RAVELLI, A. C., VAN DER MEULEN, J. H., OSMOND, C., BARKER, D. J. & BLEKER, O. P. 1999. Obesity at the age of 50 y in men and women exposed to famine prenatally. *American Journal of Clinical Nutrition,* 70, 811-6.

RAWLINGS, N. D., MORTON, F. R., KOK, C. Y., KONG, J. & BARRETT, A. J. 2008. MEROPS: the peptidase database. *Nucleic Acids Res,* 36, D320-5.

REES, S. D., HYDRIE, M. Z., SHERA, A. S., KUMAR, S., O'HARE, J. P., BARNETT, A. H., BASIT, A. & KELLY, M. A. 2011. Replication of 13 genome-wide association (GWA)-validated risk variants for type 2 diabetes in Pakistani populations. *Diabetologia,* 54, 1368-74.

REIS, V. M., MACHADO, J. V., FORTES, M. S., FERNANDES, P. R., SILVA, A. J., DANTAS, P. S. & FILHO, J. F. 2007. Evidence for higher heritability of somatotype compared to body mass index in female twins. *J Physiol Anthropol,* 26, 9-14.

RENSTROM, F., PAYNE, F., NORDSTROM, A., BRITO, E. C., ROLANDSSON, O., HALLMANS, G., BARROSO, I., NORDSTROM, P. & FRANKS, P. W. 2009. Replication and extension of genome-wide association study results for obesity in 4923 adults from northern Sweden. *Hum Mol Genet,* 18, 1489-96.

REYNOLDS, K., GU, D., WHELTON, P. K., WU, X., DUAN, X., MO, J. & HE, J. 2007. Prevalence and risk factors of overweight and obesity in China. *Obesity (Silver Spring),* 15, 10-8.

REYNOLDS, R. M., WALKER, B. R., SYDDALL, H. E., ANDREW, R., WOOD, P. J., WHORWOOD, C. B. & PHILLIPS, D. I. 2001. Altered control of cortisol secretion in adult men with low birth weight and cardiovascular risk factors. *J Clin Endocrinol Metab,* 86, 245-50.

RICH-EDWARDS, J. W., COLDITZ, G. A., STAMPFER, M. J., WILLETT, W. C., GILLMAN, M. W., HENNEKENS, C. H., SPEIZER, F. E. & MANSON, J. E. 1999. Birthweight and the risk for type 2 diabetes mellitus in adult women. *Annals of Internal Medicine,* 130, 278-84.

RICH-EDWARDS, J. W., KLEINMAN, K., MICHELS, K. B., STAMPFER, M. J., MANSON, J. E., REXRODE, K. M., HIBERT, E. N. & WILLETT, W. C. 2005. Longitudinal study of birth weight and adult body mass index in predicting risk of coronary heart disease and stroke in women. *BMJ,* 330, 1115.

RICKARD, I. J. & LUMMAA, V. 2007. The predictive adaptive response and metabolic syndrome: challenges for the hypothesis. *Trends Endocrinol Metab,* 18, 94-9.

RILLAMAS-SUN, E., SOWERS, M. R., HARLOW, S. D. & RANDOLPH, J. F., JR. 2012. The relationship of birth weight with longitudinal changes in body composition in adult women. *Obesity (Silver Spring),* 20, 463-5.

RISNES, K. R., NILSEN, T. I., ROMUNDSTAD, P. R. & VATTEN, L. J. 2009. Head size at birth and long-term mortality from coronary heart disease. *Int J Epidemiol,* 38, 955-62.

ROLAND, M. C., FRIIS, C. M., VOLDNER, N., GODANG, K., BOLLERSLEV, J., HAUGEN, G. & HENRIKSEN, T. 2012. Fetal growth versus birthweight: the role of placenta versus other determinants. *PLoS One,* 7, e39324.

ROSE, K. M., NEWMAN, B., MAYER-DAVIS, E. J. & SELBY, J. V. 1998. Genetic and behavioral determinants of waist-hip ratio and waist circumference in women twins. *Obes Res,* 6, 383-92.

ROSEBOOM, T. J., PAINTER, R. C., DE ROOIJ, S. R., VAN ABEELEN, A. F., VEENENDAAL, M. V., OSMOND, C. & BARKER, D. J. 2011. Effects of famine on placental size and efficiency. *Placenta,* 32, 395-9.

ROSEBOOM, T. J., VAN DER MEULEN, J. H., OSMOND, C., BARKER, D. J., RAVELLI, A. C., SCHROEDER-TANKA, J. M., VAN MONTFRANS, G. A., MICHELS, R. P. & BLEKER, O. P. 2000. Coronary heart disease after prenatal exposure to the Dutch famine, 1944-45. *Heart,* 84, 595-8.

ROSEBOOM, T. J., VAN DER MEULEN, J. H., RAVELLI, A. C., OSMOND, C., BARKER, D. J. & BLEKER, O. P. 2001a. Effects of prenatal exposure to the Dutch famine on adult disease in later life: an overview. *Twin Res,* 4, 293-8.

ROSEBOOM, T. J., VAN DER MEULEN, J. H., VAN MONTFRANS, G. A., RAVELLI, A. C., OSMOND, C., BARKER, D. J. & BLEKER, O. P. 2001b. Maternal nutrition during gestation and blood pressure in later life. *J Hypertens,* 19, 29-34.

ROSMOND, R., CHAGNON, M., BOUCHARD, C. & BJORNTORP, P. 2001. A missense mutation in the human melanocortin-4 receptor gene in relation to abdominal obesity and salivary cortisol. *Diabetologia,* 44, 1335-8.

ROSS, R., ARU, J., FREEMAN, J., HUDSON, R. & JANSSEN, I. 2002a. Abdominal adiposity and insulin resistance in obese men. *Am J Physiol Endocrinol Metab,* 282, E657-63.

ROSS, R., FREEMAN, J., HUDSON, R. & JANSSEN, I. 2002b. Abdominal obesity, muscle composition, and insulin resistance in premenopausal women. *J Clin Endocrinol Metab,* 87, 5044-51.

ROSSKOPF, D., BORNHORST, A., RIMMBACH, C., SCHWAHN, C., KAYSER, A., KRUGER, A., TESSMANN, G., GEISSLER, I., KROEMER, H. K. & VOLZKE, H. 2007. Comment on "A common genetic variant is associated with adult and childhood obesity". *Science,* 315, 187; author reply 187.

ROUSKAS, K., MEYRE, D., STUTZMANN, F., PALETAS, K., PAPAZOGLOU, D., VATIN, V., MARCHAND, M., KOUVATSI, A. & FROGUEL, P. 2012. Loss-of-Function Mutations in MC4R Are Very Rare in the Greek Severely Obese Adult Population. *Obesity (Silver Spring).*

ROY, A., KUCUKURAL, A. & ZHANG, Y. 2010. I-TASSER: a unified platform for automated protein structure and function prediction. *Nat Protoc,* 5, 725-38.

SABIN, M. A. & SHIELD, J. P. 2008. Childhood obesity. *Front Horm Res,* 36, 85-96.

SAEED, S., BUTT, T. A., ANWER, M., ARSLAN, M. & FROGUEL, P. 2012. High prevalence of leptin and melanocortin-4 receptor gene mutations in children with severe obesity from Pakistani consanguineous families. *Mol Genet Metab,* 106, 121-6.

SALAS, M., JOHN, R., SAXENA, A., BARTON, S., FRANK, D., FITZPATRICK, G., HIGGINS, M. J. & TYCKO, B. 2004. Placental growth retardation due to loss of imprinting of Phlda2. *Mech Dev,* 121, 1199-210.

SALOPURO, T., PULKKINEN, L., LINDSTROM, J., ERIKSSON, J. G., VALLE, T. T., HAMALAINEN, H., ILANNE-PARIKKA, P., KEINANEN-KIUKAANNIEMI, S., TUOMILEHTO, J., LAAKSO, M. & UUSITUPA, M. 2005. Genetic variation in leptin receptor gene is associated with type 2 diabetes and body weight: The Finnish Diabetes Prevention Study. *Int J Obes (Lond),* 29, 1245-51.

SANCHEZ-PULIDO, L. & ANDRADE-NAVARRO, M. A. 2007. The FTO (fat mass and obesity associated) gene codes for a novel member of the non-heme dioxygenase superfamily. *BMC Biochem,* 8, 23.

SANDOWSKI, Y., RAVER, N., GUSSAKOVSKY, E. E., SHOCHAT, S., DYM, O., LIVNAH, O., RUBINSTEIN, M., KRISHNA, R. & GERTLER, A. 2002. Subcloning, expression, purification, and characterization of recombinant human leptin-binding domain. *Journal of Biological Chemistry,* 277, 46304-9.

SANTORO, N., CIRILLO, G., AMATO, A., LUONGO, C., RAIMONDO, P., D'ANIELLO, A., PERRONE, L. & MIRAGLIA DEL GIUDICE, E. 2006. Insulin gene variable number of tandem repeats (INS VNTR) genotype and metabolic syndrome in childhood obesity. *J Clin Endocrinol Metab,* 91, 4641-4.

SARIC, T., MULLER, D., SEITZ, H. J. & PAVELIC, K. 2003. Non-covalent interaction of ubiquitin with insulin-degrading enzyme. *Mol Cell Endocrinol,* 204, 11-20.

SAULES, K. K., LEVINE, M. D., MARCUS, M. D. & POMERLEAU, C. S. 2007. Differences in smoking patterns among women smokers with childhood versus later onset of weight problems. *Eat Behav,* 8, 418-22.

SAWIN, K. E., LEGUELLEC, K., PHILIPPE, M. & MITCHISON, T. J. 1992. Mitotic spindle organization by a plus-end-directed microtubule motor. *Nature,* 359, 540-3.

SAWIN, K. E. & MITCHISON, T. J. 1995. Mutations in the kinesin-like protein Eg5 disrupting localization to the mitotic spindle. *Proc Natl Acad Sci U S A,* 92, 4289-93.

SAXENA, R., VOIGHT, B. F., LYSSENKO, V., BURTT, N. P., DE BAKKER, P. I., CHEN, H., ROIX, J. J., KATHIRESAN, S., HIRSCHHORN, J. N., DALY, M. J., HUGHES, T. E., GROOP, L., ALTSHULER, D., ALMGREN, P., FLOREZ, J. C., MEYER, J., ARDLIE, K., BENGTSSON BOSTROM, K., ISOMAA, B., LETTRE, G., LINDBLAD, U., LYON, H. N., MELANDER, O., NEWTON-CHEH, C., NILSSON, P., ORHO-MELANDER, M., RASTAM, L., SPELIOTES, E. K., TASKINEN, M. R., TUOMI, T., GUIDUCCI, C., BERGLUND, A., CARLSON, J., GIANNINY, L., HACKETT, R., HALL, L., HOLMKVIST, J., LAURILA, E., SJOGREN, M., STERNER, M., SURTI, A., SVENSSON, M., TEWHEY, R., BLUMENSTIEL, B., PARKIN, M., DEFELICE, M., BARRY, R., BRODEUR, W., CAMARATA, J., CHIA, N., FAVA, M., GIBBONS, J., HANDSAKER, B., HEALY, C., NGUYEN, K., GATES, C., SOUGNEZ, C., GAGE, D., NIZZARI, M., GABRIEL, S. B., CHIRN, G. W., MA, Q., PARIKH, H., RICHARDSON, D., RICKE, D. & PURCELL, S. 2007. Genome-wide association analysis identifies loci for type 2 diabetes and triglyceride levels. *Science,* 316, 1331-6.

SAYER, A. A., SYDDALL, H. E., DENNISON, E. M., GILBODY, H. J., DUGGLEBY, S. L., COOPER, C., BARKER, D. J. & PHILLIPS, D. I. 2004. Birth weight, weight at 1 y of age, and body composition in older men: findings from the Hertfordshire Cohort Study. *American Journal of Clinical Nutrition,* 80, 199-203.

SCHACK-NIELSEN, L., HOLST, C. & SORENSEN, T. I. 2002. Blood pressure in relation to relative weight at birth through childhood and youth in obese and non-obese adult men. *Int J Obes Relat Metab Disord,* 26, 1539-46.

SCHOUSBOE, K., VISSCHER, P. M., ERBAS, B., KYVIK, K. O., HOPPER, J. L., HENRIKSEN, J. E., HEITMANN, B. L. & SORENSEN, T. I. 2004. Twin study of genetic and environmental influences on adult body size, shape, and composition. *Int J Obes Relat Metab Disord,* 28, 39-48.

SCHROEDER, D. G., MARTORELL, R. & FLORES, R. 1999. Infant and child growth and fatness and fat distribution in Guatemalan adults. *Am J Epidemiol,* 149, 177-85.

SCHULZE, M. B., AL-HASANI, H., BOEING, H., FISHER, E., DORING, F. & JOOST, H. G. 2007. Variation in the HHEX-IDE gene region predisposes to type 2 diabetes in the prospective, population-based EPIC-Potsdam cohort. *Diabetologia,* 50, 2405-7.

SCOTT, L. J., MOHLKE, K. L., BONNYCASTLE, L. L., WILLER, C. J., LI, Y., DUREN, W. L., ERDOS, M. R., STRINGHAM, H. M., CHINES, P. S., JACKSON, A. U., PROKUNINA-OLSSON, L., DING, C. J., SWIFT, A. J., NARISU, N., HU, T., PRUIM, R., XIAO, R., LI, X. Y., CONNEELY, K. N., RIEBOW, N. L., SPRAU, A. G., TONG, M., WHITE, P. P., HETRICK, K. N., BARNHART, M. W., BARK, C. W., GOLDSTEIN, J. L., WATKINS, L., XIANG, F., SARAMIES, J., BUCHANAN, T. A., WATANABE, R.

M., VALLE, T. T., KINNUNEN, L., ABECASIS, G. R., PUGH, E. W., DOHENY, K. F., BERGMAN, R. N., TUOMILEHTO, J., COLLINS, F. S. & BOEHNKE, M. 2007. A genome-wide association study of type 2 diabetes in Finns detects multiple susceptibility variants. *Science,* 316, 1341-1345.

SCUTERI, A., SANNA, S., CHEN, W. M., UDA, M., ALBAI, G., STRAIT, J., NAJJAR, S., NAGARAJA, R., ORRU, M., USALA, G., DEI, M., LAI, S., MASCHIO, A., BUSONERO, F., MULAS, A., EHRET, G. B., FINK, A. A., WEDER, A. B., COOPER, R. S., GALAN, P., CHAKRAVARTI, A., SCHLESSINGER, D., CAO, A., LAKATTA, E. & ABECASIS, G. R. 2007. Genome-wide association scan shows genetic variants in the FTO gene are associated with obesity-related traits. *Plos Genetics,* 3, 1200-1210.

SEAL, N., WEAVER, M. & BEST, L. G. 2011. Correlates of the FTO gene variant (rs9939609) and growth of American Indian infants. *Genet Test Mol Biomarkers,* 15, 633-8.

SEBERT, S. P., HYATT, M. A., CHAN, L. L., YIALLOURIDES, M., FAINBERG, H. P., PATEL, N., SHARKEY, D., STEPHENSON, T., RHIND, S. M., BELL, R. C., BUDGE, H., GARDNER, D. S. & SYMONDS, M. E. 2010. Influence of prenatal nutrition and obesity on tissue specific fat mass and obesity-associated (FTO) gene expression. *Reproduction,* 139, 265-74.

SEEBACHER, T., LINDER, J. U. & SCHULTZ, J. E. 2001. An isoform-specific interaction of the membrane anchors affects mammalian adenylyl cyclase type V activity. *Eur J Biochem,* 268, 105-10.

SEGAL, N. L., FENG, R., MCGUIRE, S. A., ALLISON, D. B. & MILLER, S. 2009. Genetic and environmental contributions to body mass index: comparative analysis of monozygotic twins, dizygotic twins and same-age unrelated siblings. *Int J Obes (Lond),* 33, 37-41.

SENG, K. C. & SENG, C. K. 2008. The success of the genome-wide association approach: a brief story of a long struggle. *Eur J Hum Genet,* 16, 554-64.

SETA, K. A. & ROTH, R. A. 1997. Overexpression of insulin degrading enzyme: cellular localization and effects on insulin signaling. *Biochem Biophys Res Commun,* 231, 167-71.

SHAH, N. R. & BRAVERMAN, E. R. 2012. Measuring adiposity in patients: the utility of body mass index (BMI), percent body fat, and leptin. *PLoS One,* 7, e33308.

SHAHKHALILI, Y., MOULIN, J., ZBINDEN, I., APRIKIAN, O. & MACE, K. 2010. Comparison of two models of intrauterine growth restriction for early catch-up growth and later development of glucose intolerance and obesity in rats. *Am J Physiol Regul Integr Comp Physiol,* 298, R141-6.

SHAI, I., JIANG, R., MANSON, J. E., STAMPFER, M. J., WILLETT, W. C., COLDITZ, G. A. & HU, F. B. 2006. Ethnicity, obesity, and risk of type 2 diabetes in women: a 20-year follow-up study. *Diabetes Care,* 29, 1585-90.

SHASTRY, B. S. 2009. SNPs: impact on gene function and phenotype. *Methods Mol Biol,* 578, 3-22.

SHAW, J. E., SICREE, R. A. & ZIMMET, P. Z. 2010. Global estimates of the prevalence of diabetes for 2010 and 2030. *Diabetes Res Clin Pract,* 87, 4-14.

SHOJAEE-MORADIE, F., BAYNES, K. C., PENTECOST, C., BELL, J. D., THOMAS, E. L., JACKSON, N. C., STOLINSKI, M., WHYTE, M., LOVELL, D., BOWES, S. B., GIBNEY, J., JONES, R. H. & UMPLEBY, A. M. 2007. Exercise training reduces fatty acid availability and improves the insulin sensitivity of glucose metabolism. *Diabetologia,* 50, 404-13.

SHULDINER, A. R. 2008. Obesity genes and gene-environment-behavior interactions: recommendations for a way forward. *Obesity (Silver Spring),* 16 Suppl 3, S79-81.

SILVENTOINEN, K., PIETILAINEN, K. H., TYNELIUS, P., SORENSEN, T. I., KAPRIO, J. & RASMUSSEN, F. 2007. Genetic and environmental factors in relative weight from birth to age 18: the Swedish young male twins study. *Int J Obes (Lond),* 31, 615-21.

SILVENTOINEN, K., ROKHOLM, B., KAPRIO, J. & SORENSEN, T. I. 2010. The genetic and environmental influences on childhood obesity: a systematic review of twin and adoption studies. *Int J Obes (Lond),* 34, 29-40.

SILVER, K., WALSTON, J., CHUNG, W. K., YAO, F., PARIKH, V. V., ANDERSEN, R., CHESKIN, L. J., ELAHI, D., MULLER, D., LEIBEL, R. L. & SHULDINER, A. R. 1997. The Gln223Arg and Lys656Asn polymorphisms in the human leptin receptor do not associate with traits related to obesity. *Diabetes,* 46, 1898-900.

SIMPKISS, M. J. 1968. Birth weight, maternal age, and parity among the african population of uganda. *Br J Prev Soc Med,* 22, 234-7.

SINGH, M., BEDI, U. S., SINGH, P. P., ARORA, R. & KHOSLA, S. 2010. Leptin and the clinical cardiovascular risk. *Int J Cardiol,* 140, 266-71.

SINGHAL, A., FEWTRELL, M., COLE, T. J. & LUCAS, A. 2003. Low nutrient intake and early growth for later insulin resistance in adolescents born preterm. *Lancet,* 361, 1089-97.

SINHA, M. K., OPENTANOVA, I., OHANNESIAN, J. P., KOLACZYNSKI, J. W., HEIMAN, M. L., HALE, J., BECKER, G. W., BOWSHER, R. R., STEPHENS, T. W. & CARO, J. F. 1996. Evidence of free and bound leptin in human circulation. Studies in lean and obese subjects and during short-term fasting. *J Clin Invest,* 98, 1277-82.

SKIDMORE, P. M., CASSIDY, A., SWAMINATHAN, R., RICHARDS, J. B., MANGINO, M., SPECTOR, T. D. & MACGREGOR, A. J. 2009. An obesogenic postnatal environment is more important than the fetal environment for the development of adult adiposity: a study of female twins. *American Journal of Clinical Nutrition,* 90, 401-6.

SLADEK, R., ROCHELEAU, G., RUNG, J., DINA, C., SHEN, L., SERRE, D., BOUTIN, P., VINCENT, D., BELISLE, A., HADJADJ, S., BALKAU, B., HEUDE, B., CHARPENTIER, G., HUDSON, T. J., MONTPETIT, A., PSHEZHETSKY, A. V., PRENTKI, M., POSNER, B. I., BALDING, D. J., MEYRE, D., POLYCHRONAKOS, C. & FROGUEL, P. 2007. A genome-wide association study identifies novel risk loci for type 2 diabetes. *Nature,* 445, 881-885.

SMITH, A. J., COOPER, J. A., LI, L. K. & HUMPHRIES, S. E. 2007. INSIG2 gene polymorphism is not associated with obesity in Caucasian, Afro-Caribbean and Indian subjects. *Int J Obes (Lond),* 31, 1753-5.

SMITH, T. F. & WATERMAN, M. S. 1981. Identification of common molecular subsequences. *J Mol Biol,* 147, 195-7.

SNIJDER, M. B. 2003. Associations of hip and thigh circumferences indepedent of waist circumference with the incidence of type 2 diabetes: the Hoorn Study. *American Journal of Clinical Nutrition,* 77, 1192-1197.

SODE-CARLSEN, R., FARHOLT, S., RABBEN, K. F., BOLLERSLEV, J., SANDAHL CHRISTIANSEN, J. & HOYBYE, C. 2009. Assessment of physical function in adults with Prader-Willi syndrome. *Disabil Rehabil,* 31, 1780-4.

SONG, T. M., PERUSSE, L., MALINA, R. M. & BOUCHARD, C. 1994. Twin resemblance in somatotype and comparisons with other twin studies. *Hum Biol,* 66, 453-64.

SORENSEN, H. T., SABROE, S., ROTHMAN, K. J., GILLMAN, M., FISCHER, P. & SORENSEN, T. I. 1997. Relation between weight and length at birth and body mass index in young adulthood: cohort study. *BMJ,* 315, 1137.

SOTO, N., BAZAES, R. A., PENA, V., SALAZAR, T., AVILA, A., INIGUEZ, G., ONG, K. K., DUNGER, D. B. & MERICQ, M. V. 2003. Insulin sensitivity and secretion are related to catch-up growth in small-for-gestational-age infants at age 1 year: results from a prospective cohort. *J Clin Endocrinol Metab,* 88, 3645-50.

SPELIOTES, E. K., WILLER, C. J., BERNDT, S. I., MONDA, K. L., THORLEIFSSON, G., JACKSON, A. U., ALLEN, H. L., LINDGREN, C. M., LUAN, J., MAGI, R., RANDALL, J. C., VEDANTAM, S., WINKLER, T. W., QI, L., WORKALEMAHU, T., HEID, I. M., STEINTHORSDOTTIR, V., STRINGHAM, H. M., WEEDON, M. N., WHEELER, E., WOOD, A. R., FERREIRA, T., WEYANT, R. J., SEGRE, A. V., ESTRADA, K., LIANG, L., NEMESH, J., PARK, J. H., GUSTAFSSON, S., KILPELAINEN, T. O., YANG, J., BOUATIA-NAJI, N., ESKO, T., FEITOSA, M. F., KUTALIK, Z., MANGINO, M., RAYCHAUDHURI, S., SCHERAG, A., SMITH, A. V., WELCH, R., ZHAO, J. H., ABEN, K. K., ABSHER, D. M., AMIN, N., DIXON, A. L., FISHER, E., GLAZER, N. L., GODDARD, M. E., HEARD-COSTA, N. L., HOESEL, V., HOTTENGA, J. J., JOHANSSON, A., JOHNSON, T., KETKAR, S., LAMINA, C., LI, S., MOFFATT, M. F., MYERS, R. H., NARISU, N., PERRY, J. R., PETERS, M. J., PREUSS, M., RIPATTI, S., RIVADENEIRA, F., SANDHOLT, C., SCOTT, L. J., TIMPSON, N. J., TYRER, J. P., VAN WINGERDEN, S., WATANABE, R. M., WHITE, C. C., WIKLUND, F., BARLASSINA, C., CHASMAN, D. I., COOPER, M. N., JANSSON, J. O., LAWRENCE, R. W., PELLIKKA, N., PROKOPENKO, I., SHI, J., THIERING, E., ALAVERE, H., ALIBRANDI, M. T., ALMGREN, P., ARNOLD, A. M., ASPELUND, T., ATWOOD, L. D., BALKAU, B., BALMFORTH, A. J., BENNETT, A. J., BEN-SHLOMO, Y., BERGMAN, R. N., BERGMANN, S., BIEBERMANN, H., BLAKEMORE, A. I., BOES, T., BONNYCASTLE, L. L., BORNSTEIN, S. R., BROWN, M. J., BUCHANAN, T. A., *et al.* 2010. Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index. *Nature Genetics,* 42, 937-48.

SPIEGELMAN, B. M. & FLIER, J. S. 2001. Obesity and the regulation of energy balance. *Cell,* 104, 531-43.

STAIANO, A. E., REEDER, B. A., ELLIOTT, S., JOFFRES, M. R., PAHWA, P., KIRKLAND, S. A., PARADIS, G. & KATZMARZYK, P. T. 2012. Body mass index versus waist circumference as predictors of mortality in Canadian adults. *Int J Obes (Lond).*

STAIGER, H., STANCAKOVA, A., ZILINSKAITE, J., VANTTINEN, M., HANSEN, T., MARINI, M. A., HAMMARSTEDT, A., JANSSON, P. A., SESTI, G., SMITH, U., PEDERSEN, O., LAAKSO, M., STEFAN, N., FRITSCHE, A. & HARING, H. U. 2008. A candidate type 2 diabetes polymorphism near the HHEX locus affects acute glucose-stimulated insulin release in European populations: results from the EUGENE2 study. *Diabetes,* 57, 514-7.

STANNER, S. A., BULMER, K., ANDRES, C., LANTSEVA, O. E., BORODINA, V., POTEEN, V. V. & YUDKIN, J. S. 1997. Does malnutrition in utero determine diabetes and coronary heart disease in adulthood? Results from the Leningrad siege study, a cross sectional study. *BMJ,* 315, 1342-8.

STECK, A. K. & WINTER, W. E. 2011. Review on monogenic diabetes. *Curr Opin Endocrinol Diabetes Obes,* 18, 252-8.

STEIN, A. D., KAHN, H. S., RUNDLE, A., ZYBERT, P. A., VAN DER PAL-DE BRUIN, K. & LUMEY, L. H. 2007. Anthropometric measures in middle age after exposure to famine during gestation: evidence from the Dutch famine. *American Journal of Clinical Nutrition,* 85, 869-76.

STEIN, A. D., ZYBERT, P. A., VAN DER PAL-DE BRUIN, K. & LUMEY, L. H. 2006. Exposure to famine during gestation, size at birth, and blood pressure at age 59 y: evidence from the Dutch Famine. *Eur J Epidemiol,* 21, 759-65.

STEIN, C. E., FALL, C. H., KUMARAN, K., OSMOND, C., COX, V. & BARKER, D. J. 1996. Fetal growth and coronary heart disease in south India. *Lancet,* 348, 1269-73.

STETTLER, N., STALLINGS, V. A., TROXEL, A. B., ZHAO, J., SCHINNAR, R., NELSON, S. E., ZIEGLER, E. E. & STROM, B. L. 2005. Weight gain in the first week of life and overweight in adulthood: a cohort study of European American subjects fed infant formula. *Circulation,* 111, 1897-903.

STEVENS, J., CAI, J., EVENSON, K. R. & THOMAS, R. 2002. Fitness and fatness as predictors of mortality from all causes and from cardiovascular disease in men and women in the lipid research clinics study. *Am J Epidemiol,* 156, 832-41.

STRATIGOPOULOS, G., LEDUC, C. A., MATSUOKA, N., GUTMAN, R., RAUSCH, R., ROBERTSON, S. A., MYERS, M. G., JR., CHUNG, W. K., CHUA, S. C., JR. & LEIBEL, R. L. 2009. Functional consequences of the human leptin receptor (LEPR) Q223R transversion. *Obesity (Silver Spring),* 17, 126-35.

STRATIGOPOULOS, G., PADILLA, S. L., LEDUC, C. A., WATSON, E., HATTERSLEY, A. T., MCCARTHY, M. I., ZELTSER, L. M., CHUNG, W. K. & LEIBEL, R. L. 2008. Regulation of Fto/Ftm gene expression in mice and humans. *Am J Physiol Regul Integr Comp Physiol,* 294, R1185-96.

STROBEL, A., ISSAD, T., CAMOIN, L., OZATA, M. & STROSBERG, A. D. 1998. A leptin missense mutation associated with hypogonadism and morbid obesity. *Nature Genetics,* 18, 213-215.

STUEBE, A. M., LYON, H., HERRING, A. H., GHOSH, J., WISE, A., NORTH, K. E. & SIEGA-RIZ, A. M. 2010. Obesity and diabetes genetic variants associated with gestational weight gain. *Am J Obstet Gynecol,* 203, 283 e1-17.

STUNKARD, A. J., FOCH, T. T. & HRUBEC, Z. 1986a. A twin study of human obesity. *JAMA,* 256, 51-4.

STUNKARD, A. J., SORENSEN, T. I., HANIS, C., TEASDALE, T. W., CHAKRABORTY, R., SCHULL, W. J. & SCHULSINGER, F. 1986b. An adoption study of human obesity. *N Engl J Med,* 314, 193-8.

SUGIHARA, S., SASAKI, N., AMEMIYA, S., KOHNO, H., TANAKA, T. & MATSUURA, N. 2008. Analysis of weight at birth and at diagnosis of childhood-onset type 2 diabetes mellitus in Japan. *Pediatr Diabetes,* 9, 285-90.

SULLIVAN, E. L. & GROVE, K. L. 2010. Metabolic imprinting in obesity. *Forum Nutr,* 63, 186-94.

SUN, Q., CORNELIS, M. C., KRAFT, P., QI, L., VAN DAM, R. M., GIRMAN, C. J., LAURIE, C. C., MIREL, D. B., GONG, H., SHEU, C. C., CHRISTIANI, D. C., HUNTER, D. J., MANTZOROS, C. S. & HU, F. B. 2010. Genome-wide association study identifies polymorphisms in LEPR as determinants of plasma soluble leptin receptor levels. *Hum Mol Genet,* 19, 1846-55.

SUNDHEIM, O., VAGBO, C. B., BJORAS, M., SOUSA, M. M., TALSTAD, V., AAS, P. A., DRABLOS, F., KROKAN, H. E., TAINER, J. A. & SLUPPHAUG, G. 2006. Human ABH3 structure and key residues for oxidative demethylation to reverse DNA/RNA damage. *EMBO J,* 25, 3389-97.

SUNG, J., LEE, K. & SONG, Y. M. 2009. Relationship of eating behavior to long-term weight change and body mass index: the Healthy Twin study. *Eat Weight Disord,* 14, e98-105.

SVENSSON, A. C., PAWITAN, Y., CNATTINGIUS, S., REILLY, M. & LICHTENSTEIN, P. 2006. Familial aggregation of small-for-gestational-age births: the importance of fetal genetic effects. *Am J Obstet Gynecol,* 194, 475-9.

SYDDALL, H. E., SAYER, A. A., SIMMONDS, S. J., OSMOND, C., COX, V., DENNISON, E. M., BARKER, D. J. & COOPER, C. 2005. Birth weight, infant weight gain, and cause-specific mortality: the Hertfordshire Cohort Study. *Am J Epidemiol,* 161, 1074-80.

TAKAHASHI-YASUNO, A., MASUZAKI, H., MIYAWAKI, T., MATSUOKA, N., OGAWA, Y., HAYASHI, T., HOSODA, K., YOSHIMASA, Y., INOUE, G. & NAKAO, K. 2004. Association of Ob-R gene polymorphism and insulin resistance in Japanese men. *Metabolism-Clinical and Experimental,* 53, 650-4.

TANG, W. J. & GILMAN, A. G. 1995. Construction of a soluble adenylyl cyclase activated by Gs alpha and forskolin. *Science,* 268, 1769-72.

TARRY-ADKINS, J. L., CHEN, J. H., SMITH, N. S., JONES, R. H., CHERIF, H. & OZANNE, S. E. 2009. Poor maternal nutrition followed by accelerated postnatal growth leads to telomere shortening and increased markers of cell senescence in rat islets. *Faseb Journal,* 23, 1521-8.

TARTAGLIA, L. A., DEMBSKI, M., WENG, X., DENG, N., CULPEPPER, J., DEVOS, R., RICHARDS, G. J., CAMPFIELD, L. A., CLARK, F. T., DEEDS, J., MUIR, C., SANKER, S., MORIARTY, A., MOORE, K. J., SMUTKO, J. S., MAYS, G. G., WOOL, E. A., MONROE, C. A. & TEPPER, R. I. 1995. Identification and expression cloning of a leptin receptor, OB-R. *Cell,* 83, 1263-71.

THAME, M., OSMOND, C., WILKS, R. J., BENNETT, F. I., MCFARLANE-ANDERSON, N. & FORRESTER, T. E. 2000. Blood pressure is related to placental volume and birth weight. *Hypertension,* 35, 662-7.

THEODORO, M. F., TALEBIZADEH, Z. & BUTLER, M. G. 2006. Body composition and fatness patterns in Prader-Willi syndrome: comparison with simple obesity. *Obesity (Silver Spring),* 14, 1685-90.

THOMAS, E. L., FROST, G., TAYLOR-ROBINSON, S. D. & BELL, J. D. 2012. Excess body fat in obese and normal-weight subjects. *Nutr Res Rev,* 25, 150-61.

THOMAS, E. L., SAEED, N., HAJNAL, J. V., BRYNES, A., GOLDSTONE, A. P., FROST, G. & BELL, J. D. 1998a. Magnetic resonance imaging of total body fat. *J Appl Physiol,* 85, 1778-85.

THOMAS, P. Q., BROWN, A. & BEDDINGTON, R. S. 1998b. Hex: a homeobox gene revealing peri-implantation asymmetry in the mouse embryo and an early transient marker of endothelial cell precursors. *Development,* 125, 85-94.

THORLEIFSSON, G., WALTERS, G. B., GUDBJARTSSON, D. F., STEINTHORSDOTTIR, V., SULEM, P., HELGADOTTIR, A., STYRKARSDOTTIR, U., GRETARSDOTTIR, S., THORLACIUS, S., JONSDOTTIR, I., JONSDOTTIR, T., OLAFSDOTTIR, E. J., OLAFSDOTTIR, G. H., JONSSON, T., JONSSON, F., BORCH-JOHNSEN, K., HANSEN, T., ANDERSEN, G., JORGENSEN, T., LAURITZEN, T., ABEN, K. K., VERBEEK, A. L., ROELEVELD, N., KAMPMAN, E., YANEK, L. R., BECKER, L. C., TRYGGVADOTTIR, L., RAFNAR, T., BECKER, D. M., GULCHER, J., KIEMENEY, L. A., PEDERSEN, O., KONG, A., THORSTEINSDOTTIR, U. & STEFANSSON, K. 2009. Genome-wide association yields new sequence variants at seven loci that associate with measures of obesity. *Nature Genetics,* 41, 18-24.

TIAN, J. Y., CHENG, Q., SONG, X. M., LI, G., JIANG, G. X., GU, Y. Y. & LUO, M. 2006. Birth weight and risk of type 2 diabetes, abdominal obesity and hypertension among Chinese adults. *Eur J Endocrinol,* 155, 601-7.

TIMPSON, N. J., EMMETT, P. M., FRAYLING, T. M., ROGERS, I., HATTERSLEY, A. T., MCCARTHY, M. I. & SMITH, G. D. 2008. The fat mass- and obesity-associated locus and dietary intake in children. *American Journal of Clinical Nutrition,* 88, 971-978.

TINDALL, K. R. & KUNKEL, T. A. 1988. Fidelity of DNA synthesis by the Thermus aquaticus DNA polymerase. *Biochemistry,* 27, 6008-13.

TOBI, E. W., LUMEY, L. H., TALENS, R. P., KREMER, D., PUTTER, H., STEIN, A. D., SLAGBOOM, P. E. & HEIJMANS, B. T. 2009. DNA methylation differences after exposure to prenatal famine are common and timing- and sex-specific. *Hum Mol Genet,* 18, 4046-53.

TORRES, N., BAUTISTA, C. J., TOVAR, A. R., ORDAZ, G., RODRIGUEZ-CRUZ, M., ORTIZ, V., GRANADOS, O., NATHANIELSZ, P. W., LARREA, F. & ZAMBRANO, E. 2010. Protein restriction during pregnancy affects maternal liver lipid metabolism and fetal brain lipid composition in the rat. *Am J Physiol Endocrinol Metab,* 298, E270-7.

TOUVIER, T., CONTE-AURIOL, F., BRIAND, O., CUDEJKO, C., PAUMELLE, R., CARON, S., BAUGE, E., ROUILLE, Y., SALLES, J. P., STAELS, B. & BAILLEUL, B. 2009. LEPROT and LEPROTL1 cooperatively decrease hepatic growth hormone action in mice. *J Clin Invest,* 119, 3830-8.

TRAYHURN, P. & BING, C. 2006. Appetite and energy balance signals from adipocytes. *Philos Trans R Soc Lond B Biol Sci,* 361, 1237-49.

TRAYHURN, P., HOGGARD, N., MERCER, J. G. & RAYNER, D. V. 1999. Leptin: fundamental aspects. *International Journal of Obesity,* 23, 22-28.

TUAN, N. T., BUTTE, N. F. & NICKLAS, T. A. 2012. Body mass index distribution affects discrepancies in weight classifications in children. *Pediatr Int,* 54, 256-65.

TURULA, M., KAPRIO, J., RISSANEN, A. & KOSKENVUO, M. 1990. Body weight in the Finnish Twin Cohort. *Diabetes Res Clin Pract,* 10 Suppl 1, S33-6.

UKKOLA, O., TREMBLAY, A., DESPRES, J. P., CHAGNON, Y. C., CAMPFIELD, L. A. & BOUCHARD, C. 2000. Leptin receptor Gln223Arg variant is associated with a cluster of metabolic abnormalities in response to long-term overfeeding. *J Intern Med,* 248, 435-9.

UUSITUPA, M., LINDI, V., LOUHERANTA, A., SALOPURO, T., LINDSTROM, J. & TUOMILEHTO, J. 2003. Long-term improvement in insulin sensitivity by changing lifestyles of people with impaired glucose tolerance: 4-year results from the Finnish Diabetes Prevention Study. *Diabetes,* 52, 2532-8.

VAISSE, C., CLEMENT, K., GUY-GRAND, B. & FROGUEL, P. 1998. A frameshift mutation in human MC4R is associated with a dominant form of obesity. *Nature Genetics,* 20, 113-4.

VAISSE, C., HALAAS, J. L., HORVATH, C. M., DARNELL, J. E., JR., STOFFEL, M. & FRIEDMAN, J. M. 1996. Leptin activation of Stat3 in the hypothalamus of wild-type and ob/ob mice but not db/db mice. *Nature Genetics,* 14, 95-7.

VALE, R. D. & FLETTERICK, R. J. 1997. The design plan of kinesin motors. *Annu Rev Cell Dev Biol,* 13, 745-77.

VAN DER HOEVEN, F., SCHIMMANG, T., VOLKMANN, A., MATTEI, M. G., KYEWSKI, B. & RUTHER, U. 1994. Programmed cell death is affected in the novel mouse mutant Fused toes (Ft). *Development,* 120, 2601-7.

VAN DOMMELEN, P., DE GUNST, M. C., VAN DER VAART, A. W. & BOOMSMA, D. I. 2004. Genetic study of the height and weight process during infancy. *Twin Res,* 7, 607-16.

VAN HEEK, M., COMPTON, D. S., FRANCE, C. F., TEDESCO, R. P., FAWZI, A. B., GRAZIANO, M. P., SYBERTZ, E. J., STRADER, C. D. & DAVIS, H. R., JR. 1997. Diet-induced obese mice develop peripheral, but not central, resistance to leptin. *J Clin Invest,* 99, 385-90.

VAN ROSSUM, C. T., HOEBEE, B., VAN BAAK, M. A., MARS, M., SARIS, W. H. & SEIDELL, J. C. 2003. Genetic variation in the leptin receptor gene, leptin, and weight gain in young Dutch adults. *Obes Res,* 11, 377-86.

VAN VILET-OSTAPTCHOUK, J. V. 2008. HHEX gene polymorphisms are associated with type 2 diabetes in the Dutch Breda cohort. *European Journal of Human Genetics,* 16, 652 - 656.

VASAN, S. K., NEVILLE, M. J., ANTONISAMY, B., SAMUEL, P., FALL, C. H., GEETHANJALI, F. S., THOMAS, N., RAGHUPATHY, P., BRISMAR, K. & KARPE, F. 2011. Absence of birth-weight lowering effect of ADCY5 and near CCNL, but association of impaired glucose-insulin homeostasis with ADCY5 in Asian Indians. *PLoS One,* 6, e21331.

VEGA, G. L., ADAMS-HUET, B., PESHOCK, R., WILLETT, D., SHAH, B. & GRUNDY, S. M. 2006. Influence of body fat content and distribution on variation in metabolic risk. *Journal of Clinical Endocrinology & Metabolism,* 91, 4459-4466.

VELKOSKA, E., COLE, T. J., DEAN, R. G., BURRELL, L. M. & MORRIS, M. J. 2008. Early undernutrition leads to long-lasting reductions in body weight and adiposity whereas increased intake increases cardiac fibrosis in male rats. *Journal of Nutrition,* 138, 1622-7.

VINER, R. M., COLE, T. J., FRY, T., GUPTA, S., KINRA, S., MCCARTHY, D., SAXENA, S., TAYLOR, S., WELLS, J. C., WHINCUP, P. & ZAMAN, M. J. 2010. Insufficient evidence to support separate BMI definitions for obesity in children and adolescents from south Asian ethnic groups in the UK. *Int J Obes (Lond),* 34, 656-8.

VITALI, A., MURANO, I., ZINGARETTI, M. C., FRONTINI, A., RICQUIER, D. & CINTI, S. 2012. The adipose organ of obesity-prone C57BL/6J mice is composed of mixed white and brown adipocytes. *J Lipid Res,* 53, 619-29.

VLIETINCK, R., DEROM, R., NEALE, M. C., MAES, H., VAN LOON, H., DEROM, C. & THIERY, M. 1989. Genetic and environmental variation in the birth weight of twins. *Behavior Genetics,* 19, 151-61.

VOIGT, M., ZELS, K., GUTHMANN, F., HESSE, V., GORLICH, Y. & STRAUBE, S. 2011. Somatic classification of neonates based on birth weight, length, and head circumference: quantification of the effects of maternal BMI and smoking. *J Perinat Med,* 39, 291-7.

VUGUIN, P. M. 2007. Animal models for small for gestational age and fetal programming of adult disease. *Horm Res,* 68, 113-23.

WADSWORTH, M., BUTTERWORTH, S., MARMOT, M., ECOB, R. & HARDY, R. 2005. Early growth and type 2 diabetes: evidence from the 1946 British birth cohort. *Diabetologia,* 48, 2505-10.

WALCZAK, C. E., VERNOS, I., MITCHISON, T. J., KARSENTI, E. & HEALD, R. 1998. A model for the proposed roles of different microtubule-based motor proteins in establishing spindle bipolarity. *Curr Biol,* 8, 903-13.

WALKER, S. P., GASKIN, P., POWELL, C. A., BENNETT, F. I., FORRESTER, T. E. & GRANTHAM-MCGREGOR, S. 2001. The effects of birth weight and postnatal linear growth retardation on blood pressure at age 11-12 years. *J Epidemiol Community Health,* 55, 394-8.

WALLEY, A. J., ASHER, J. E. & FROGUEL, P. 2009. The genetic contribution to non-syndromic human obesity. *Nature Reviews Genetics,* 10, 431-42.

WANG, C. L., LIANG, L., WANG, H. J., FU, J. F., HEBEBRAND, J. & HINNEY, A. 2006. Several mutations in the melanocortin 4 receptor gene are associated with obesity in Chinese children and adolescents. *J Endocrinol Invest,* 29, 894-8.

WANG, J., THORNTON, J. C., BURASTERO, S., SHEN, J., TANENBAUM, S., HEYMSFIELD, S. B. & PIERSON, R. N., JR. 1996. Comparisons for body mass index and body fat percent among Puerto Ricans, blacks, whites and Asians living in the New York City area. *Obes Res,* 4, 377-84.

WANG, K. S., LIU, X., ZHENG, S., ZENG, M., PAN, Y. & CALLAHAN, K. 2012. A novel locus for body mass index on 5p15.2: a meta-analysis of two genome-wide association studies. *Gene,* 500, 80-4.

WANG, Y., MI, J., SHAN, X. Y., WANG, Q. J. & GE, K. Y. 2007. Is China facing an obesity epidemic and the consequences? The trends in obesity and chronic disease in China. *Int J Obes (Lond),* 31, 177-88.

WANG, Y. F., RIMM, E. B., STAMPFER, M. J., WILLETT, W. C. & HU, F. B. 2005. Comparison of abdominal adiposity and overall obesity in predicting risk of type 2 diabetes among men. *American Journal of Clinical Nutrition,* 81, 555-563.

WARDLE, J., CARNELL, S., HAWORTH, C. M. & PLOMIN, R. 2008. Evidence for a strong genetic influence on childhood adiposity despite the force of the obesogenic environment. *American Journal of Clinical Nutrition,* 87, 398-404.

WARDLE, J., LLEWELLYN, C., SANDERSON, S. & PLOMIN, R. 2009. The FTO gene and measured food intake in children. *International Journal of Obesity,* 33, 42-45.

WAUTERS, M., MERTENS, I., CHAGNON, M., RANKINEN, T., CONSIDINE, R. V., CHAGNON, Y. C., VAN GAAL, L. F. & BOUCHARD, C. 2001a. Polymorphisms in the leptin receptor gene, body composition and fat distribution in overweight and obese women. *Int J Obes Relat Metab Disord,* 25, 714-20.

WAUTERS, M., MERTENS, I., RANKINEN, T., CHAGNON, M., BOUCHARD, C. & VAN GAAL, L. 2001b. Leptin receptor gene polymorphisms are associated with insulin in obese women with impaired glucose tolerance. *J Clin Endocrinol Metab,* 86, 3227-32.

WEEDON, M. N., CLARK, V. J., QIAN, Y., BEN-SHLOMO, Y., TIMPSON, N., EBRAHIM, S., LAWLOR, D. A., PEMBREY, M. E., RING, S., WILKIN, T. J., VOSS, L. D., JEFFERY, A. N., METCALF, B., FERRUCCI, L., CORSI, A. M., MURRAY, A., MELZER, D., KNIGHT, B., SHIELDS, B., SMITH, G. D., HATTERSLEY, A. T., DI RIENZO, A. & FRAYLING, T. M. 2006. A common haplotype of the glucokinase gene alters fasting glucose and birth weight: association in six studies and population-genetics analyses. *Am J Hum Genet,* 79, 991-1001.

WEI, J. N., SUNG, F. C., LI, C. Y., CHANG, C. H., LIN, R. S., LIN, C. C., CHIANG, C. C. & CHUANG, L. M. 2003. Low birth weight and high birth weight infants are both at an increased risk to have type 2 diabetes among schoolchildren in taiwan. *Diabetes Care,* 26, 343-8.

WEISBERG, S. P., MCCANN, D., DESAI, M., ROSENBAUM, M., LEIBEL, R. L. & FERRANTE, A. W., JR. 2003. Obesity is associated with macrophage accumulation in adipose tissue. *J Clin Invest,* 112, 1796-808.

WEISS, R. & CAPRIO, S. 2005. The metabolic consequences of childhood obesity. *Best Pract Res Clin Endocrinol Metab,* 19, 405-19.

WESTENDORP, R. G. 2006. What is healthy aging in the 21st century? *American Journal of Clinical Nutrition,* 83, 404S-409S.

WHITAKER, R. C., WRIGHT, J. A., PEPE, M. S., SEIDEL, K. D. & DIETZ, W. H. 1997. Predicting obesity in young adulthood from childhood and parental obesity. *N Engl J Med,* 337, 869-73.

WHITE, D. W., WANG, D. W., CHUA, S. C., JR., MORGENSTERN, J. P., LEIBEL, R. L., BAUMANN, H. & TARTAGLIA, L. A. 1997. Constitutive and impaired signaling of leptin receptors containing the Gln --> Pro extracellular domain fatty mutation. *Proc Natl Acad Sci U S A,* 94, 10657-62.

WILLER, C. J., SPELIOTES, E. K., LOOS, R. J., LI, S., LINDGREN, C. M., HEID, I. M., BERNDT, S. I., ELLIOTT, A. L., JACKSON, A. U., LAMINA, C., LETTRE, G., LIM, N., LYON, H. N., MCCARROLL, S. A., PAPADAKIS, K., QI, L., RANDALL, J. C., ROCCASECCA, R. M., SANNA, S., SCHEET, P., WEEDON, M. N., WHEELER, E., ZHAO, J. H., JACOBS, L. C., PROKOPENKO, I., SORANZO, N., TANAKA, T., TIMPSON, N. J., ALMGREN, P., BENNETT, A., BERGMAN, R. N., BINGHAM, S. A., BONNYCASTLE, L. L., BROWN, M., BURTT, N. P., CHINES, P., COIN, L., COLLINS, F. S., CONNELL, J. M., COOPER, C., SMITH, G. D., DENNISON, E. M., DEODHAR, P., ELLIOTT, P., ERDOS, M. R., ESTRADA, K., EVANS, D. M., GIANNINY, L., GIEGER, C., GILLSON, C. J., GUIDUCCI, C., HACKETT, R., HADLEY, D., HALL, A. S., HAVULINNA, A. S., HEBEBRAND, J., HOFMAN, A., ISOMAA, B., JACOBS, K. B., JOHNSON, T., JOUSILAHTI, P., JOVANOVIC, Z., KHAW, K. T., KRAFT, P., KUOKKANEN, M., KUUSISTO, J., LAITINEN, J., LAKATTA, E. G., LUAN, J., LUBEN, R. N., MANGINO, M., MCARDLE, W. L., MEITINGER, T., MULAS, A., MUNROE, P. B., NARISU, N., NESS, A. R., NORTHSTONE, K., O'RAHILLY, S., PURMANN, C., REES, M. G., RIDDERSTRALE, M., RING, S. M., RIVADENEIRA, F., RUOKONEN, A., SANDHU, M. S., SARAMIES, J., SCOTT, L. J., SCUTERI, A., SILANDER, K., SIMS, M. A., SONG, K., STEPHENS, J., STEVENS, S., STRINGHAM, H. M., TUNG, Y. C., VALLE, T. T., VAN DUIJN, C. M., VIMALESWARAN, K. S., VOLLENWEIDER, P., *et al.* 2009. Six new loci associated with body mass index highlight a neuronal influence on body weight regulation. *Nature Genetics,* 41, 25-34.

WILSON, P., O'MEARA, S., SUMMERBELL, C. & KELLY, S. 2003. The prevention and treatment of childhood obesity. *Qual Saf Health Care,* 12, 65-74.

WINDHOLZ, J., KOVACS, P., TONJES, A., DITTRICH, K., BLUHER, S., KIESS, W., STUMVOLL, M. & KORNER, A. 2011. Effects of genetic variants in ADCY5, GIPR, GCKR and VPS13C on early impairment of glucose and insulin metabolism in children. *PLoS One,* 6, e22101.

WINKLER, C., BONIFACIO, E., GRALLERT, H., HENNEBERGER, L., ILLIG, T. & ZIEGLER, A. G. 2010. BMI at age 8 years is influenced by the type 2 diabetes susceptibility genes HHEX-IDE and CDKAL1. *Diabetes,* 59, 2063-7.

WITT, K. A. 2005. College athletes with an elevated body mass index often have a higher upper arm muscle mass but not elevated triceps and subscapular skinfolds. *Journal of the American Dietetic Association,* 105, 599-602.

WU, G., BAZER, F. W., CUDD, T. A., MEININGER, C. J. & SPENCER, T. E. 2004. Maternal nutrition and fetal development. *Journal of Nutrition,* 134, 2169-72.

WU, Q., SAUNDERS, R. A., SZKUDLAREK-MIKHO, M., SERNA IDE, L. & CHIN, K. V. 2010. The obesity-associated Fto gene is a transcriptional coactivator. *Biochem Biophys Res Commun,* 401, 390-5.

WU, S., SKOLNICK, J. & ZHANG, Y. 2007. Ab initio modeling of small proteins by iterative TASSER simulations. *BMC Biol,* 5, 17.

WU, Y. 2006. Overweight and obesity in China. *BMJ,* 333, 362-3.

WU, Y., LI, H. X., LOOS, R. J. F., YU, Z. J., YE, X. W., CHEN, L. H., PAN, A., HU, F. B. & LIN, X. 2008. Common variants in CDKAL1, CDKN2A/B, IGF2BP2, SLC30A8, and HHEX/IDE genes are associated with type 2 diabetes and impaired fasting glucose in a Chinese Han population. *Diabetes,* 57, 2834-2842.

XI, B., SHEN, Y., ZHANG, M., LIU, X., ZHAO, X., WU, L., CHENG, H., HOU, D., LINDPAINTNER, K., LIU, L., MI, J. & WANG, X. 2010. The common rs9939609 variant of the fat mass and obesity-associated gene is associated with obesity risk in children and adolescents of Beijing, China. *BMC Med Genet,* 11, 107.

XIAO, X., ZHANG, Z. X., COHEN, H. J., WANG, H., LI, W., WANG, T., XU, T., LIU, A., GAI, M. Y., YING, S., SCHMITZ, O. & YI, Z. 2008. Evidence of a relationship between infant birth weight and later diabetes and impaired glucose regulation in a Chinese population. *Diabetes Care,* 31, 483-7.

YAJNIK, C. 2002. Adiposity and hyperinsulinemai in Indians are present at birth. *Journal of Clinical Endocrinology & Metabolism,* 87, 5575-5580.

YAJNIK, C. S. 2004. Obesity epidemic in India: intrauterine origins? *Proc Nutr Soc,* 63, 387-96.

YAMAGATA, K., FURUTA, H., ODA, N., KAISAKI, P. J., MENZEL, S., COX, N. J., FAJANS, S. S., SIGNORINI, S., STOFFEL, M. & BELL, G. I. 1996. Mutations in the hepatocyte nuclear factor-4alpha gene in maturity-onset diabetes of the young (MODY1). *Nature,* 384, 458-60.

YANG, G. P., PENG, S. H., ZUO, S. Y., WANG, Y. R., PENG, X. N. & ZENG, X. M. 2011. [Meta-analysis on the relationship between leptin receptor Gln223Arg and Pro1019Pro gene polymorphism and obesity in the Chinese population]. *Zhonghua Liu Xing Bing Xue Za Zhi,* 32, 1037-42.

YANG, Q., COGSWELL, M. E., FLANDERS, W. D., HONG, Y., ZHANG, Z., LOUSTALOT, F., GILLESPIE, C., MERRITT, R. & HU, F. B. 2012. Trends in cardiovascular health metrics and associations with all-cause and CVD mortality among US adults. *JAMA,* 307, 1273-83.

YANG, Z., ZHAO, W., ZHANG, X., MU, R., ZHAI, Y., KONG, L. & CHEN, C. 2008. Impact of famine during pregnancy and infancy on health in adulthood. *Obesity Reviews,* 9 Suppl 1, 95-9.

YEO, G. S., FAROOQI, I. S., AMINIAN, S., HALSALL, D. J., STANHOPE, R. G. & O'RAHILLY, S. 1998. A frameshift mutation in MC4R associated with dominantly inherited human obesity. *Nature Genetics,* 20, 111-2.

YFANTI, C., MENGELE, K., GKAZEPIS, A., WEIRICH, G., GIERSIG, C., KUO, W. L., TANG, W. J., ROSNER, M. & SCHMITT, M. 2008. Expression of metalloprotease insulin-degrading enzyme insulysin in normal and malignant human tissues. *Int J Mol Med,* 22, 421-31.

YIANNAKOURIS, N., YANNAKOULIA, M., MELISTAS, L., CHAN, J. L., KLIMIS-ZACAS, D. & MANTZOROS, C. S. 2001. The Q223R polymorphism of the leptin receptor gene is significantly associated with obesity and predicts a small percentage of body weight and body composition variability. *J Clin Endocrinol Metab,* 86, 4434-9.

YLIHARSILA, H., KAJANTIE, E., OSMOND, C., FORSEN, T., BARKER, D. J. & ERIKSSON, J. G. 2007. Birth size, adult body composition and muscle strength in later life. *Int J Obes (Lond),* 31, 1392-9.

YOON, K. H., LEE, J. H., KIM, J. W., CHO, J. H., CHOI, Y. H., KO, S. H., ZIMMET, P. & SON, H. Y. 2006. Epidemic obesity and type 2 diabetes in Asia. *Lancet,* 368, 1681-8.

YU, B., EDSTROM, W. C., BENACH, J., HAMURO, Y., WEBER, P. C., GIBNEY, B. R. & HUNT, J. F. 2006. Crystal structures of catalytic complexes of the oxidative DNA/RNA repair enzyme AlkB. *Nature,* 439, 879-84.

YU, M. X., QIU, X. S., FENG, S. E., MO, Q. P., XIE, X. Y., SHEN, Z. Y. & LIU, Y. Z. 2011. [The effect of birth weight and early growth on body fat composition and insulin sensitivity]. *Zhonghua Yu Fang Yi Xue Za Zhi,* 45, 633-8.

YUDKIN, J. S., YAJNIK, C. S., MOHAMED-ALI, V. & BULMER, K. 1999. High levels of circulating proinflammatory cytokines and leptin in urban, but not rural, Indians. A potential explanation for increased risk of diabetes and coronary heart disease. *Diabetes Care,* 22, 363-4.

ZABANEH, D., CHAMBERS, J. C., ELLIOTT, P., SCOTT, J., BALDING, D. J. & KOONER, J. S. 2009. Heritability and genetic correlations of insulin resistance and component phenotypes in Asian Indian families using a multivariate analysis. *Diabetologia,* 52, 2585-9.

ZABEAU, L., DEFEAU, D., VAN DER HEYDEN, J., ISERENTANT, H., VANDEKERCKHOVE, J. & TAVERNIER, J. 2004. Functional analysis of leptin receptor activation using a Janus kinase/signal transducer and activator of transcription complementation assay. *Mol Endocrinol,* 18, 150-61.

ZABEAU, L., LAVENS, D., PEELMAN, F., EYCKERMAN, S., VANDEKERCKHOVE, J. & TAVERNIER, J. 2003. The ins and outs of leptin receptor activation. *FEBS Lett,* 546, 45-50.

ZANDI-NEJAD, K., LUYCKX, V. A. & BRENNER, B. M. 2006. Adult hypertension and kidney disease: the role of fetal programming. *Hypertension,* 47, 502-8.

ZEGGINI, E., WEEDON, M. N., LINDGREN, C. M., FRAYLING, T. M., ELLIOTT, K. S., LANGO, H., TIMPSON, N. J., PERRY, J. R., RAYNER, N. W., FREATHY, R. M., BARRETT, J. C., SHIELDS, B., MORRIS, A. P., ELLARD, S., GROVES, C. J., HARRIES, L. W., MARCHINI, J. L., OWEN, K. R., KNIGHT, B., CARDON, L. R., WALKER, M., HITMAN, G. A., MORRIS, A. D., DONEY, A. S., MCCARTHY, M. I. & HATTERSLEY, A. T. 2007. Replication of genome-wide association signals in UK samples reveals risk loci for type 2 diabetes. *Science,* 316, 1336-41.

ZHANG, Y. 2008. I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics,* 9, 40.

ZHANG, Y. 2009. I-TASSER: fully automated protein structure prediction in CASP8. *Proteins,* 77 Suppl 9, 100-13.

ZHANG, Y., PROENCA, R., MAFFEI, M., BARONE, M., LEOPOLD, L. & FRIEDMAN, J. M. 1994. Positional cloning of the mouse obese gene and its human homologue. *Nature,* 372, 425-32.

ZHANG, Y. & SKOLNICK, J. 2005. TM-align: a protein structure alignment algorithm based on the TM-score. *Nucleic Acids Res,* 33, 2302-9.

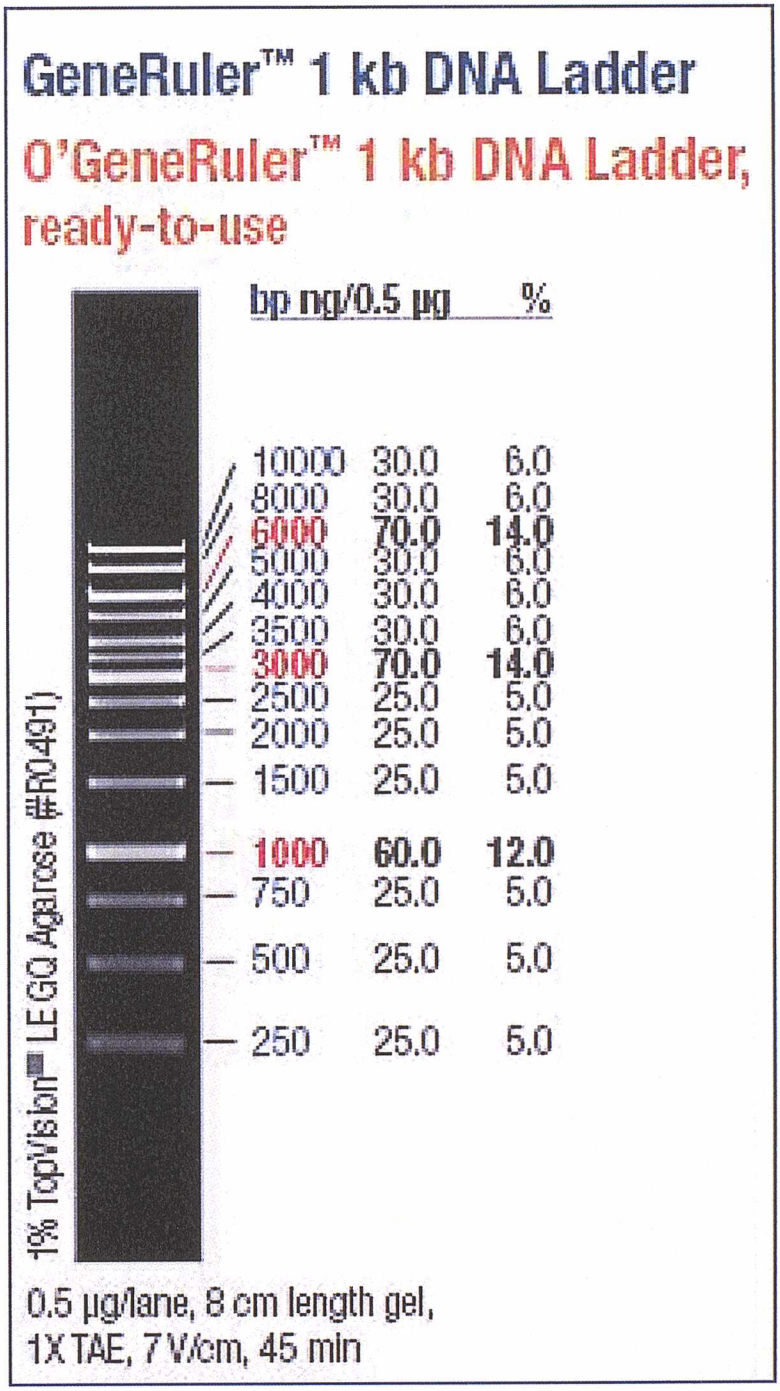
ZHAO, J. 2009. Examination of type 2 diabetes loci implictaes CDKAL1 as a birth weight gene. *Diabetes,* 58, 2414-2418.

ZHAO, J., BRADFIELD, J. P., LI, M., WANG, K., ZHANG, H., KIM, C. E., ANNAIAH, K., GLESSNER, J. T., THOMAS, K., GARRIS, M., FRACKELTON, E. C., OTIENO, F. G., SHANER, J. L., SMITH, R. M., CHIAVACCI, R. M., BERKOWITZ, R. I., HAKONARSON, H. & GRANT, S. F. 2009. The role of obesity-associated loci identified in genome-wide association studies in the determination of pediatric BMI. *Obesity (Silver Spring),* 17, 2254-7.

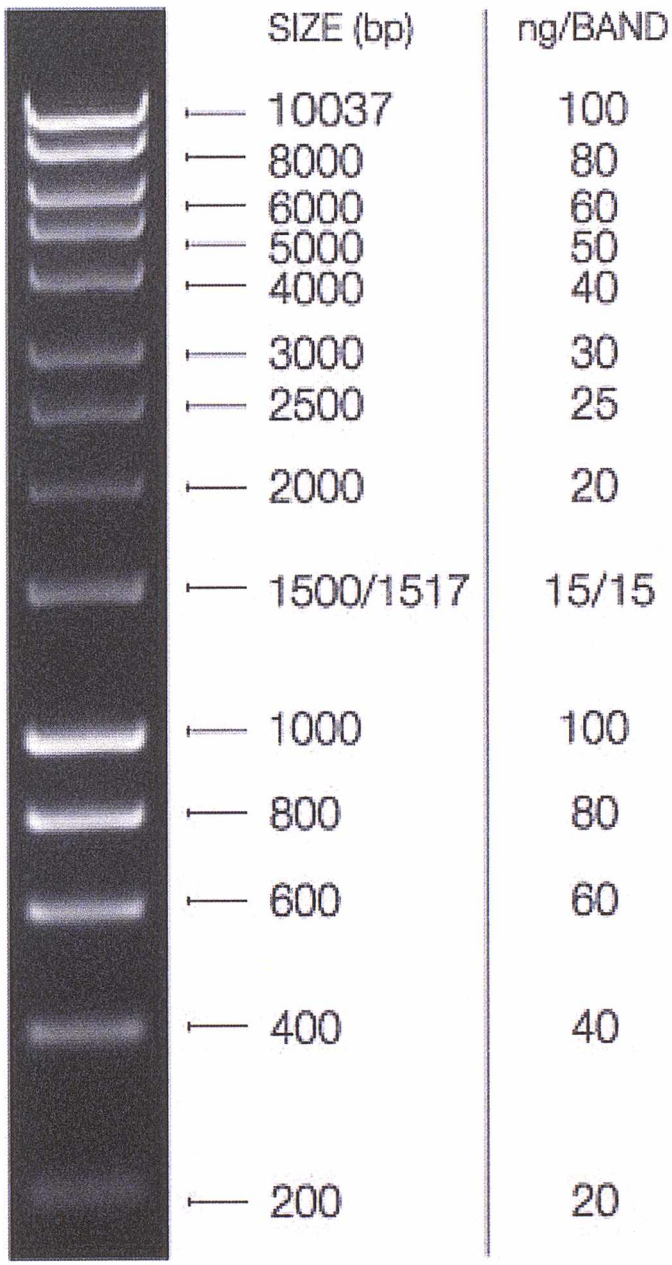
ZHAO, J., BRADFIELD, J. P., ZHANG, H., ANNAIAH, K., WANG, K., KIM, C. E., GLESSNER, J. T., FRACKELTON, E. C., OTIENO, F. G., DORAN, J., THOMAS, K. A., GARRIS, M., HOU, C., CHIAVACCI, R. M., LI, M., BERKOWITZ, R. I., HAKONARSON, H. & GRANT, S. F. 2010. Examination of all type 2 diabetes GWAS loci reveals HHEX-IDE as a locus influencing pediatric BMI. *Diabetes,* 59, 751-5.

ZHAO, Y., WANG, S. F., MU, M. & SHENG, J. 2012. Birth weight and overweight/obesity in adults: a meta-analysis. *Eur J Pediatr.*

GeneRuler (Fermentas),

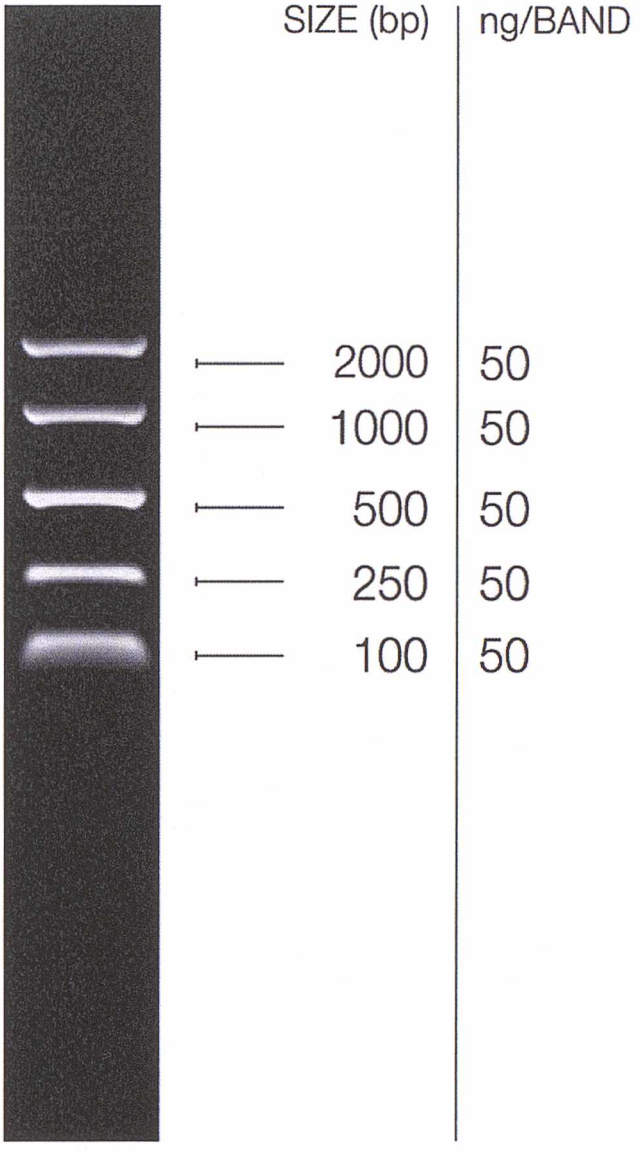


HyperladderI (Bioline)



1% agarose gel
5µl per lane

EasyLadderII (Bioline)



2% agarose gel
5µl per lane

Appendix II: Forward, reverse and extension primer sequences for *HHEX* Sequenom iPLEX assays one and two

HHEX Assay 1 Forward Primer Sequences

Primer Name	Primer Sequence
rs882136F	5' - ACGTTGGATGGAACACTTTGGAGACACCAG - 3'
rs17851141F	5' - ACGTTGGATGAGCTCACCTGTCTCTCGCTG - 3'
rs2497349F	5' - ACGTTGGATGATGAGTAGCCGCATGGTTTC - 3'
rs12262390F	5' - ACGTTGGATGGTTCATTTTCTCAAACCCTG - 3'
rs10509646F	5' - ACGTTGGATGTGGACCTTCATGTCCATAGC - 3'
rs1544210F	5' - ACGTTGGATGGACATGTCCTTCCTTTCAAG - 3'
rs34266926F	5' - ACGTTGGATGTCCCCCAACTCCTCCTTCAC - 3'
rs2488068F	5' - ACGTTGGATGTTCCCAGTTCTGCTGTCTCC - 3'
rs2497311F	5' - ACGTTGGATGCTGAACTGACAACGTGCAAG - 3'
rs17107841F	5' - ACGTTGGATGGTTGAGTCCTCCAAAGTTCC - 3'
rs2497350F	5' - ACGTTGGATGAATAGAGAATCCTGGACGGC - 3'
rs9420592F	5' - ACGTTGGATGCTCAGGACTTGCGAATAGAC - 3'
rs2488067F	5' - ACGTTGGATGGGACATGAATGACACAGGAG - 3'
rs11187146F	5' - ACGTTGGATGGAAAGCCAGGATCCGTAAAG - 3'
rs7081351F	5' - ACGTTGGATGACTAGGTGTAAAGCTGCTCC - 3'
rs4933236F	5' - ACGTTGGATGTATGTGGGAAAAAGGGCTGG - 3'
rs17374868F	5' - ACGTTGGATGTATTTGAAGGTAGATGTGGC - 3'
rs12765131F	5' - ACGTTGGATGCTCATAGCTTCATGGTATGG - 3'
rs10882102F	5' - ACGTTGGATGCAACATTAACCCAACAAGAG - 3'
rs2096177F	5' - ACGTTGGATGAGGTATACTCCACTGAGCTG - 3'

HHEX Assay One Reverse Primer Sequences

Primer Name	Primer Sequence
rs882136R	5' - ACGTTGGATGATCCAATGGATGGGAACCTG - 3'
rs17851141R	5' - ACGTTGGATGGCAGAAATATCTCTCTCCGC - 3'
rs2497349R	5' - ACGTTGGATGTGCACAAGGATTTCACTGTC - 3'
rs12262390R	5' - ACGTTGGATGAGCGAGACTCCATTCTCTAC - 3'
rs10509646R	5' - ACGTTGGATGCTCTGCTTTCAGAGTCTTGC - 3'
rs1544210R	5' - ACGTTGGATGTACTGGGCCTGGGATTTTAC - 3'
rs34266926R	5' - ACGTTGGATGTGCGAGAAGGCTGGATGGAT - 3'
rs2488068R	5' - ACGTTGGATGTAAAGCCAGGAGAGCAAGAG - 3'
rs2497311R	5' - ACGTTGGATGGCTCTTGGCCTTCTTAAATC - 3'
rs17107841R	5' - ACGTTGGATGCTGGTACTGTCCTCTCTCTG - 3'
rs2497350R	5' - ACGTTGGATGGTCCATATACAATAACACCC - 3'
rs9420592R	5' - ACGTTGGATGCAGGCCTGAATCATCCAAAC - 3'
rs2488067R	5' - ACGTTGGATGGATCTTAAAGACTTGCCCAC - 3'
rs11187146R	5' - ACGTTGGATGTCCACCCAATCTCCAGGATG - 3'
rs7081351R	5' - ACGTTGGATGGGCACCATTCTGAATTGCTC - 3'
rs4933236R	5'- ACGTTGGATGAAAATGCCTCCCATGAAGCC - 3'
rs17374868R	5' - ACGTTGGATGGCAGCCAAATGTAGCTGGTC - 3'
rs12765131R	5' - ACGTTGGATGCCCCTCAGACTTTAAAATCC - 3'
rs10882102R	5' - ACGTTGGATGAAGCATGAGCTACCCTTTGG - 3'
rs2096177R	5' - ACGTTGGATGTGCTTTCCAAACCTCTCTGC - 3'

HHEX Assay Two Forward Primer Sequences

Primer Name	Primer Sequence
rs2229328F	5' - ACGTTGGATGAACTCCTCCTTCACCAGCCT - 3'
rs2497304F	5' - ACGTTGGATGTTTCCTTGCCCTCTGTTTCC - 3'
rs2488087F	5' - ACGTTGGATGAGACATGCCTGGGCAATGAA - 3'
rs2488073F	5' - ACGTTGGATGTCTCGAACTCCTGGACTCAA - 3'
rs1418387F	5' - ACGTTGGATGTGTAGTCGTTCACCGTCCG - 3'
rs2497309F	5' - ACGTTGGATGCTAGCAGCCCAAACTAAGAC - 3'
rs11187144F	5' - ACGTTGGATGTGCCTAGGAGTGAAATTGC - 3'
rs1539330F	5' - ACGTTGGATGTCGTGCTCGGTGTGTTCAGT - 3'
rs2275729F	5' - ACGTTGGATGCCTACAGGAAAGAGAGAAAG - 3'
rs11187161F	5' - ACGTTGGATGGATCAAGTGTTCCCAAGAGG - 3'
rs10437472F	5' - ACGTTGGATGTGTAATCCCAGCTACTCAGG - 3'
rs11187158F	5' - ACGTTGGATGGTGACTTGGCCAGATTTGAG - 3'
rs11187152F	5' - ACGTTGGATGATTTCAAGGAGGAAGGATGC - 3'
rs11597458F	5' - ACGTTGGATGACCGGCCCCCCTTTATTTTT - 3'
rs947591F	5' - ACGTTGGATGGGGCTTGTCACCACTATTTC - 3'
rs12780253F	5' - ACGTTGGATGGGAGTCGACAGTTCAATCAG - 3'
rs11187157F	5' - ACGTTGGATGAAAGTAGTGCCCTGTGACTG - 3'

HHEX Assay Two Reverse Primer Sequences

Primer Name	Primer Sequence
rs2229328R	5' - ACGTTGGATGTGCGAGAAGGCTGGATGGAT - 3'
rs2497304R	5' - ACGTTGGATGTCTTGTCCCTTCCTTTGTCC - 3'
rs2488087R	5' - ACGTTGGATGTGCTGCTCTATCAAGTATGG - 3'
rs2488073R	5' - ACGTTGGATGGGTGGTGCAAACCTGTAATC - 3'
rs1418387R	5' - ACGTTGGATGATCCATCCAGCCTTCTCGCA - 3'
rs2497309R	5' - ACGTTGGATGATGAAGCACTTGGCCATTCC - 3'
rs11187144R	5' - ACGTTGGATGAAATGGCACAGTGATGATGG - 3'
rs1539330R	5' - ACGTTGGATGCCAAGAAAGGAAACCTGGAG - 3'
rs2275729R	5' - ACGTTGGATGAGAGAGAGGCGAGGAGCCA - 3'
rs11187161R	5' - ACGTTGGATGTTCACTGTAACTGCCCTTCC - 3'
rs10437472R	5' - ACGTTGGATGAGTGGCACAATCTTGGCTTC - 3'
rs11187158R	5' - ACGTTGGATGATCAAAATAGGCTCCCCACG - 3'
rs11187152R	5' - ACGTTGGATGAGCCATGCAAGTGAAAAGTG - 3'
rs11597458R	5' - ACGTTGGATGCTAGCTGGATTCCCTAATAAG - 3'
rs947591R	5' - ACGTTGGATGAACCTCTATATGCAGCATGG - 3'
rs12780253R	5' - ACGTTGGATGAGCCATATGCATATTGCCCC - 3'
rs11187157R	5' - ACGTTGGATGATCTGGCCCCTGTTTATGTC - 3'

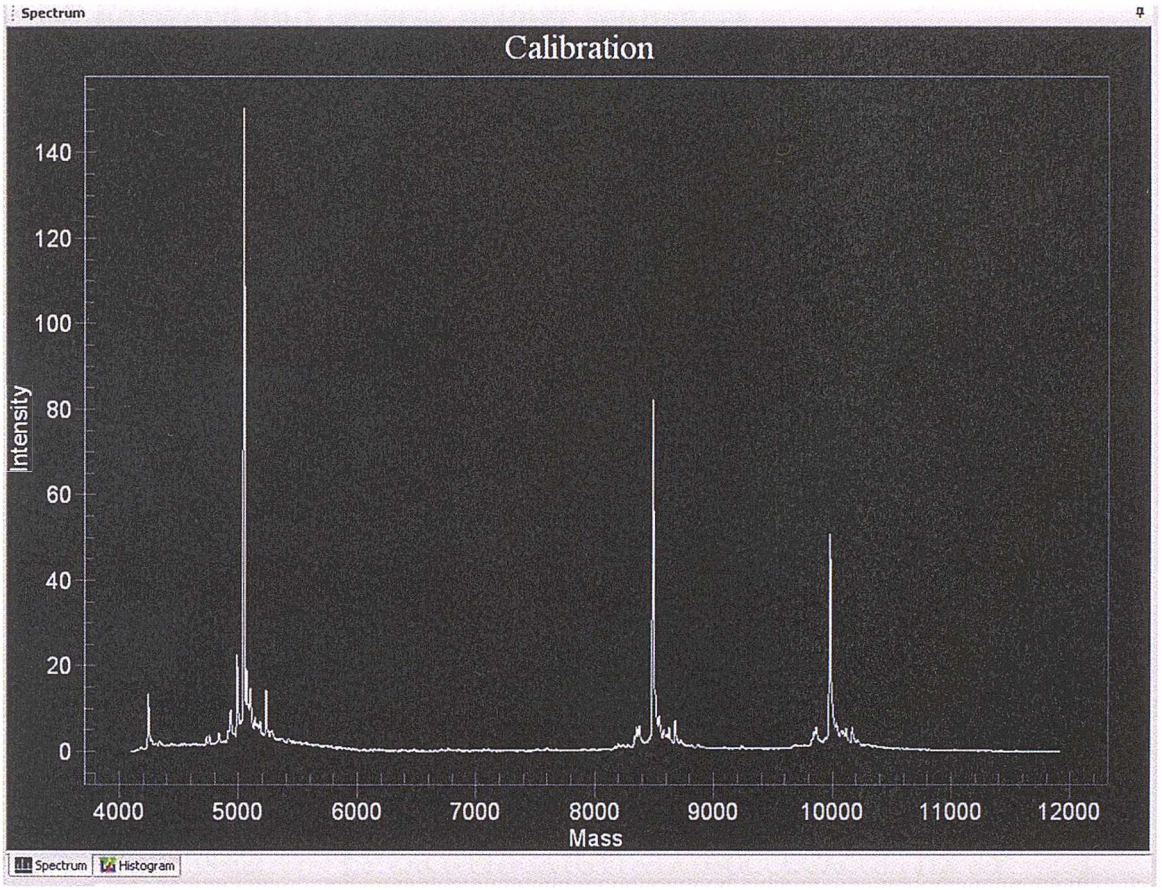
HHEX Assay One Extension Primer Sequences

Primer Name	Primer Sequence
rs882136E	5' - GACACCAGTCACCTC - 3'
rs17851141E	5' - GCTGCAGCATCTTGG - 3'
rs2497349E	5' - ATATGCTCACCCCCTT - 3'
rs12262390E	5' - gAAACCCTGTGCTGCT - 3'
rs10509646E	5' - CCAAAATTGAGTGCCTT - 3'
rs1544210E	5' - TGCAGGAAAACAGTAGC - 3'
rs34266926E	5' - gtgCTCCTTCACCAGCCT - 3'
rs2488068E	5' - cCTCCTCCTTTATCTTCCA - 3'
rs2497311E	5' - tcccACGTGCAAGAACCAT - 3'
rs17107841E	5' - TCCTCCAAAGTTCCATCTAG - 3'
rs2497350E	5' - gggtGACGGCTGTTTTTCAG - 3'
rs9420592E	5' - ccGCGAATAGACTGTTTCTC - 3'
rs2488067E	5' - CCATGTGGTCATAGTATCCAT - 3'
rs11187146E	5' - GAATTGTTGTGGGTATTTCTA - 3'
rs7081351E	5' - tgacAGGTCAACTCCATTTCAA - 3'
rs4933236E	5' - cGCAGAGTGAAGTAATGAATAT - 3'
rs17374868E	5' - TATTTTGTTTCAGCACATAGATA - 3'
rs12765131E	5' - aagcGCGCTTATAAAAAGTGCTC - 3'
rs10882102E	5' - tcGAATTAAGAAATTTTGCCATCT - 3'
rs2096177E	5' - aAACAAAAGCCTGATATATTTGTAT - 3'

HHEX Assay Two Extension Primer Sequences

Primer Name	Primer Sequence
rs2229328E	5' - CACCAGCCTCGTGTC - 3'
rs2497304E	5' - CTACCCTGGAGCAGTA - 3'
rs2488087E	5' - GGGCAATGAATGGAGC - 3'
rs2488073E	5' - CCTGGACTCAAGCAACT - 3'
rs1418387E	5' - gGCCGCCGGGTCCGTAG - 3'
rs2497309E	5' - CCACCCTACAATAAGGAA - 3'
rs11187144E	5' - tGTGAAATTGCTGGGCCA - 3'
rs1539330E	5' - gGTTCAGTGGGAACTCACC - 3'
rs2275729E	5' - AGAGAGAAAGAAAACTTGC - 3'
rs11187161E	5' - TGAGGAGGACAGAAAAGTC - 3'
rs10437472E	5' - ccccgTAAACCCGGCAGATG - 3'
rs11187158E	5' - gtgTGAACTAGGACTGGAGA - 3'
rs11187152E	5' - TGCTTTCATATATTGGGTTTAA - 3'
rs11597458E	5' - TTTATTAACAAAGAGAAAGCAG - 3'
rs947591E	5' - accCTAAATTCTTGTGGGGAAAT - 3'
rs12780253E	5' - gaAGTTCTAGTTTTGAATATGGG - 3'
rs11187157E	5' - gtcAAAAAGAAAGTAGCACAAAAT - 3'

Appendix III: Sequenom iPLEX assay calibrant containing three standard oligonucleotides of known mass: 5100, 8500 and 10,000 mass units.



Appendix IV: Forward, reverse and extension primer sequences for *FTO* Sequenom iPLEX assay

FTO Forward and reverse primer sequences

Primer Name	Primer Sequence
rs3751812F	5' - ACGTTGGATGCTTTTTCGCTGGTAGGATGC - 3'
rs3751812R	5' - ACGTTGGATGAGACCTGAAAATAGGTGAGC - 3'
rs9939973F	5' - ACGTTGGATGTTTCTATCTCTGGCCCTGAC - 3'
rs9939973R	5' - ACGTTGGATGATGACTGGTGTCTGTTCAGC - 3'
rs9940128F	5' - ACGTTGGATGAGTTTTAGGCCTCAGCTTCC - 3'
rs9940128R	5' - ACGTTGGATGTTAGAGACAGTCTGGCGATG - 3'
rs8050136F	5' - ACGTTGGATGAAGGCAAAAACCACAGGCTC - 3'
rs8050136R	5' - ACGTTGGATGTGACAGTGCCAGCTTCATAG - 3'
rs1421085F	5' - ACGTTGGATGGTAGCAGTTCAGGTCCTAAG - 3'
rs1421085R	5' - ACGTTGGATGGGAGACTACCCTACAAATTC - 3'
rs9939609F	5' - ACGTTGGATGTCCCACTCCATTTCTGACTG - 3'
rs9939609R	5' - ACGTTGGATGTTCTAGGTTCCTTGCGACTG - 3'
rs17817449F	5' - ACGTTGGATGCCCTTTGTGTTTCAGCTTGG - 3'
rs17817449R	5' - ACGTTGGATGCTATTAAAGGAGCTGGACTG - 3'
rs7193144F	5' - ACGTTGGATGACTGTCTAGCTTGATGCTTG - 3'
rs7193144R	5' - ACGTTGGATGTCTGAGGCCAGATTTTGTTC - 3'
rs1121980F	5' - ACGTTGGATGCAAGAGTTACAGGTAGGCAG - 3'
rs1121980R	5' - ACGTTGGATGTTATATGGCCCCACCTTTCC - 3'

FTO Extension Primer Sequences

Primer Name	Primer Sequence
rs3751812E	5 - CTCTCCCTGCCAACA - 3'
rs9939973E	5' - cctcCCCTCTCCACAACAG - 3'
rs9940128E	5' - TTTTTCCTTCACCTTTTCC - 3'
rs8050136E	5' - AACCACAGGCTCAGATATT - 3'
rs1421085E	5' - gggCAGGTCCTAAGGCATGA - 3'
rs9939609E	5' - cGAGACTATCCAAGTGCATCAC - 3'
rs17817449E	5' - ggtgCAGCTTGGCACACAGAAAC - 3'
rs7193144E	5' - TGATTCTTATACTTTTTTGTTTAGT - 3'
rs1121980E	5' - ctcttCAGGTGGATCTGAAATCTCA - 3'

Appendix V: Wild type LEPR Protein FASTA Sequence

1020 30405060
MICQKFCVVLLHWEFIYVITAFNLSYPITPWRFKLSCMPPNSTYDYFLLPAGLSKNTSNS

70 8090100 110120
NGHYETAVEP KFNSSGTHFS NLSKTTFHCC FRSEQDRNCS LCADNIEGKT FVSTVNSLVF

130140150 160170180
QQIDANWNIQ CWLKGDLKLF ICYVESLFKN LFRNYNYKVH LLYVLPEVLE DSPLVPQKGS

190 200210220 230240
FQMVHCNCSV HECCECLVPV PTAKLNDTLL MCLKITSGGV IFQSPLMSVQ PINMVKPDPP

250260 270280290300
LGLHMEITDD GNLKISWSSP PLVPFPLQYQ VKYSENSTTV IREADKIVSA TSLLVDSILP

310320330340350360
GSSYEVQVRG KRLDGPGIWS DWSTPRVFTT QDVIYFPPKI LTSVGSNVSF HCIYKKENKI

370380390400410 420
VPSKEIVWWM NLAEKIPQSQ YDVVSDHVSK VTFFNLNETK PRGKFTYDAV YCCNEHECHH

430440450460470 480
RYAELYVIDV NINISCETDG YLTKMTCRWS TSTIQSLAES TLQLRYHRSS LYCSDIPSIH

490500510520530540
PISEPKDCYL QSDGFYECIF QPIFLLSGYT MWIRINHSLG SLDSPPTCVL PDSVVKPLPP

550560570580590600
SSVKAEITIN IGLLKISWEK PVFPENNLQF QIRYGLSGKE VQWKMYEVYD AKSKSVSLPV

610620 630 640650 660
PDLCAVYAVQ VRCKRLDGLG YWSNWSNPAY TVVMDIKVPM RGPEFWRIIN GDTMKKEKNV

670 680 690 700710720
TLLWKPLMKN DSLCSVQRYV INHHTSCNGT WSEDVGNHTK FTFLWTEQAH TVTVLAINSI

730740 750760770780
GASVANFNLT FSWPMSKVNI VQSLSAYPLN SSCVIVSWIL SPSDYKLMYF IIEWKNLNED

790 800810820 830840
GEIKWLRISS SVKKYYIHDH FIPIEKYQFS LYPIFMEGVG KPKIINSFTQ DDIEKHQSDA

850860870880890 900
GLYVIVPVII SSSILLLGTL LISHQRMKKL FWEDVPNPKN CSWAQGLNFQ KPETFEHLFI

910 920 930940950960
KHTASVTCGP LLLEPETISE DISVDTSWKN KDEMMPTTVV SLLSTTDLEK GSVCISDQFN

970980 990100010101020
SVNFSEAEGT EVTYEDESQR QPFVKYATLI SNSKPSETGE EQGLINSSVT KCFSSKNSPL

1030104010501060107010801090
KDSFSNSSWE IEAQAFFILS DQHPNIISPH LTFSEGLDEL LKLEGNFPEE NNDKKSIYYL

1090 1100 1110 1120 1130 1140
GVTSIKKRES GVLLTDKSRV SCPFPAPCLF TDIRVLQDSC SHFVENNINL GTSSKKTFAS

1150 1160
YMPQFQTCST QTHKIMENKM CDLTV

Highlighted in red are the positions of the three polymorphisms Lys^{109} (K), Gln^{223} (Q) and Lys^{656} (K).

Appendix VI: LEPR mRNA Sequence

1 ccggtctggc ttgggcaggc tgcccgggcc gtggcaggaa gccggaagca gccgcggccc
61 cagttcggga gacatggcgg gcgttaaagc tctcgtggca ttatccttca gtggggctat
121 tggactgact tttcttatgc tgggatgtgc cttagaggat tatgggtgta cttctctgaa
181 gtaagatgat ttgtcaaaaa ttctgtgtgg ttttgttaca ttgggaattt atttatgtga
241 taactgcgtt taacttgtca tatccaatta ctccttggag atttaagttg tcttgcatgc
301 caccaaattc aacctatgac tacttccttt tgcctgctgg actctcaaag aatacttcaa
361 attcgaatgg acattatgag acagctgttg aacctaagtt taattcaagt ggtactcact
421 tttctaactt atccaaaaca actttccact gttgctttcg gagtgagcaa gatagaaact
481 gctccttatg tgcagacaac attgaaggaa agacatttgt ttcaacagta aattctttag
541 tttttcaaca aatagatgca aactggaaca tacagtgctg gctaaaagga gacttaaaat
601 tattcatctg ttatgtggag tcattattta agaatctatt caggaattat aactataagg
661 tccatctttt atatgttctg cctgaagtgt tagaagattc acctctggtt ccccaaaaag
721 gcagttttca gatggttcac tgcaattgca gtgttcatga atgttgtgaa tgtcttgtgc
781 ctgtgccaac agccaaactc aacgacactc tccttatgtg tttgaaaatc acatctggtg
841 gagtaatttt ccagtcacct ctaatgtcag ttcagcccat aaatatggtg aagcctgatc
901 caccattagg tttgcatatg gaaatcacag atgatggtaa tttaaagatt tcttggtcca
961 gcccaccatt ggtaccattt ccacttcaat atcaagtgaa atattcagag aattctacaa
1021 cagttatcag agaagctgac aagattgtct cagctacatc cctgctagta gacagtatac
1081 ttcctgggtc ttcgtatgag gttcaggtga ggggcaagag actggatggc ccaggaatct
1141 ggagtgactg gagtactcct cgtgtcttta ccacacaaga tgtcatatac tttccaccta
1201 aaattctgac aagtgttggg tctaatgttt cttttcactg catctataag aaggaaaaca
1261 agattgttcc ctcaaaagag attgtttggt ggatgaattt agctgagaaa attcctcaaa
1321 gccagtatga tgttgtgagt gatcatgtta gcaaagttac ttttttcaat ctgaatgaaa
1381 ccaaacctcg aggaaagttt acctatgatg cagtgtactg ctgcaatgaa catgaatgcc
1441 atcatcgcta tgctgaatta tatgtgattg atgtcaatat caatatctca tgtgaaactg
1501 atgggtactt aactaaaatg acttgcagat ggtcaaccag tacaatccag tcacttgcgg
1561 aaagcacttt gcaattgagg tatcatagga gcagccttta ctgttctgat attccatcta
1621 ttcatcccat atctgagccc aaagattgct atttgcagag tgatggtttt tatgaatgca
1681 ttttccagcc aatcttccta ttatctggct acacaatgtg gattaggatc aatcactctc
1741 taggttcact tgactctcca ccaacatgtg tccttcctga ttctgtggtg aagccactgc
1801 ctccatccag tgtgaaagca gaaattacta taaacattgg attattgaaa atatcttggg
1861 aaaagccagt ctttccagag aataaccttc aattccagat tcgctatggt ttaagtggaa
1921 aagaagtaca atggaagatg tatgaggttt atgatgcaaa atcaaaatct gtcagtctcc
1981 cagttccaga cttgtgtgca gtctatgctg ttcaggtgcg ctgtaagagg ctagatggac
2041 tgggatattg gagtaattgg agcaatccag cctacacagt tgtcatggat ataaaagttc
2101 ctatgagagg acctgaattt tggagaataa ttaatggaga tactatgaaa aaggagaaaa
2161 atgtcacttt actttggaag cccctgatga aaaatgactc attgtgcagt gttcagagat
2221 atgtgataaa ccatcatact tcctgcaatg gaacatggtc agaagatgtg ggaaatcaca
2281 cgaaattcac tttcctgtgg acagagcaag cacatactgt tacggttctg gccatcaatt
2341 caattggtgc ttctgttgca aattttaatt taacctttc atggcctatg agcaaagtaa
2401 atatcgtgca gtcactcagt gcttatcctt taaacagcag ttgtgtgatt gtttcctgga
2461 tactatcacc cagtgattac aagctaatgt attttattat tgagtggaaa aatcttaatg
2521 aagatggtga aataaaatgg cttagaatct cttcatctgt taagaagtat tatatccatg
2581 atcattttat ccccattgag aagtaccagt tcagtcttta cccaatattt atggaaggag
2641 tgggaaaacc aaagataatt aatagtttca ctcaagatga tattgaaaaa caccagagtg
2701 atgcaggttt atatgtaatt gtgccagtaa ttatttcctc ttccatctta ttgcttggaa
2761 cattattaat atcacaccaa agaatgaaaa agctattttg ggaagatgtt ccgaaccccа
2821 agaattgttc ctgggcacaa ggacttaatt ttcagaagcc agaaacgttt gagcatcttt

2881 ttatcaagca tacagcatca gtgacatgtg gtcctcttct tttggagcct gaaacaattt
2941 cagaagatat cagtgttgat acatcatgga aaaataaaga tgagatgatg ccaacaactg
3001 tggtctctct actttcaaca acagatcttg aaaagggttc tgtttgtatt agtgaccagt
3061 tcaacagtgt taacttctct gaggctgagg gtactgaggt aacctatgag gacgaaagcc
3121 agagacaacc ctttgttaaa tacgccacgc tgatcagcaa ctctaaacca agtgaaactg
3181 gtgaagaaca agggcttata aatagttcag tcaccaagtg cttctctagc aaaaattctc
3241 cgttgaagga ttctttctct aatagctcat gggagataga ggcccaggca ttttttatat
3301 tatcagatca gcatcccaac ataatttcac cacacctcac attctcagaa ggattggatg
3361 aacttttgaa attggaggga aatttccctg aagaaaataa tgataaaaag tctatctatt
3421 atttaggggt cacctcaatc aaaaagagag agagtggtgt gcttttgact gacaagtcaa
3481 gggtatcgtg cccattccca gccccctgtt tattcacgga catcagagtt ctccaggaca
3541 gttgctcaca ctttgtagaa aataatatca acttaggaac ttctagtaag aagacttttg
3601 catcttacat gcctcaattc caaacttgtt ctactcagac tcataagatc atggaaaaca
3661 agatgtgtga cctaactgtg taatttcact gaagaaacct tcagatttgt gttataatgg
3721 gtaatataaa gtgtaataga ttatagttgt gggtgggaga gagaaaagaa accagagtca
3781 aatttgaaaa taattgttcc aaatgaatgt tgtctgtttg ttctctctta gtaacataga
3841 caaaaaattt gagaaagcct tcataagcct accaatgtag acacgctctt ctattttatt
3901 cccaagctct agtgggaagg tcccttgttt ccagctagaa ataagcccaa cagacaccat
3961 cttttgtgag atgtaattgt tttttcagag ggcgtgttgt tttacctcaa gttttttgttt
4021 tgtaccaaca cacacacaca cacacattct taacacatgt ccttgtgtgt tttgagagta
4081 tattatgtat ttatattttg tgctatcaga ctgtaggatt tgaagtagga ctttcctaaa
4141 tgtttaagat aaacagaatt c

LEPR and nested LEPR primers are located within the mRNA sequence. LEPR forward primer is highlighted in red and LEPR reverse primer is highlighted in blue. LEPR nested forward primer is highlighted in green and LEPR nested reverse primer is highlighted in orange.

Appendix VII: LEPR Protein Haplotype Sequences

Haplotype 2 Protein sequence (109Arg Polymorphism)

```
10          20          30          40          50          60
MICQKFCVVL LHWEFIYVIT AFNLSYPITP WRFKLSCMPP NSTYDYFLLP AGLSKNTSNS

          70          80          90          100         110         120
NGHYETAVEP KFNSSGTHFS NLSKTTFHCC FRSEQDRNCS LCADNIEGRT FVSTVNSLVF

          130         140         150         160         170         180
QQIDANWNIQ CWLKGDLKLF ICYVESLFKN LFRNYNYKVH LLYVLPEVLE DSPLVPQKGS

          190         200         210         220         230         240
FQMVHCNCSV HECCECLVPV PTAKLNDTLL MCLKITSGGV IFQSPLMSVQ PINMVKPDPP

          250         260         270         280         290         300
LGLHMEITDD GNLKISWSSP PLVPFPLQYQ VKYSENSTTV IREADKIVSA TSLLVDSILP

          310         320         330         340         350         360
GSSYEVQVRG KRLDGPGIWS DWSTPRVFTT QDVIYFPPKI LTSVGSNVSF HCIYKKENKI

          370         380         390         400         410         420
VPSKEIVWWM NLAEKIPQSQ YDVVSDHVSK VTFFNLNETK PRGKFTYDAV YCCNEHECHH

          430         440         450         460         470         480
RYAELYVIDV NINISCETDG YLTKMTCRWS TSTIQSLAES TLQLRYHRSS LYCSDIPSIH

          490         500         510         520         530         540
PISEPKDCYL QSDGFYECIF QPIFLLSGYT MWIRINHSLG SLDSPPTCVL PDSVVKPLPP

          550         560         570         580         590         600
SSVKAEITIN IGLLKISWEK PVFPENNLQF QIRYGLSGKE VQWKMYEVYD AKSKSVSLPV

          610         620         630         640         650         660
PDLCAVYAVQ VRCKRLDGLG YWSNWSNPAY TVVMDIKVPM RGPEFWRIIN GDTMKKEKNV

          670         680         690         700         710         720
TLLWKPLMKN DSLCSVQRYV INHHTSCNGT WSEDVGNHTK FTFLWTEQAH TVTVLAINSI

          730         740         750         760         770         780
GASVANFNLT FSWPMSKVNI VQSLSAYPLN SSCVIVSWIL SPSDYKLMYF IIEWKNLNED

          790         800         810         820         830         840
GEIKWLRISS SVKKYYIHDH FIPIEKYQFS LYPIFMEGVG KPKIINSFTQ DDIEKHQSDA

          850         860         870         880         890         900
GLYVIVPVII SSSILLLGTL LISHQRMKKL FWEDVPNPKN CSWAQGLNFQ KPETFEHLFI

          910         920         930         940         950         960
KHTASVTCGP LLLEPETISE DISVDTSWKN KDEMMPTTVV SLLSTTDLEK GSVCISDQFN

          970         980         990         1000        1010        1020
SVNFSEAEGT EVTYEDESQR QPFVKYATLI SNSKPSETGE EQGLINSSVT KCFSSKNSPL

          1030        1040        1050        1060        1070        1080
KDSFSNSSWE IEAQAFFILS DQHPNIISPH LTFSEGLDEL LKLEGNFPEE NNDKKSIYYL
```

1090 1100 1110 1120 1130 1140
GVTSIKKRES GVLLTDKSRV SCPFPAPCLF TDIRVLQDSC SHFVENNINL GTSSKKTFAS

1150 1160
YMPQFQTCST QTHKIMENKM CDLTV

Haplotype 3 Protein Sequence (223Arg Polymorphism)

```
10          20          30          40          50          60
MICQKFCVVL LHWEFIYVIT AFNLSYPITP WRFKLSCMPP NSTYDYFLLP AGLSKNTSNS

          70          80          90          100         110         120
NGHYETAVEP KFNSSGTHFS NLSKTTFHCC FRSEQDRNCS LCADNIEGKT FVSTVNSLVF

         130         140         150         160         170         180
QQIDANWNIQ CWLKGDLKLF ICYVESLFKN LFRNYNYKVH LLYVLPEVLE DSPLVPQKGS

         190         200         210         220         230         240
FQMVHCNCSV HECCECLVPV PTAKLNDTLL MCLKITSGGV IFRSPLMSVQ PINMVKPDPP

         250         260         270         280         290         300
LGLHMEITDD GNLKISWSSP PLVPFPLQYQ VKYSENSTTV IREADKIVSA TSLLVDSILP

         310         320         330         340         350         360
GSSYEVQVRG KRLDGPGIWS DWSTPRVFTT QDVIYFPPKI LTSVGSNVSF HCIYKKENKI

         370         380         390         400         410         420
VPSKEIVWWM NLAEKIPQSQ YDVVSDHVSK VTFFNLNETK PRGKFTYDAV YCCNEHECHH

         430         440         450         460         470         480
RYAELYVIDV NINISCETDG YLTKMTCRWS TSTIQSLAES TLQLRYHRSS LYCSDIPSIH

         490         500         510         520         530         540
PISEPKDCYL QSDGFYECIF QPIFLLSGYT MWIRINHSLG SLDSPPTCVL PDSVVKPLPP

         550         560         570         580         590         600
SSVKAEITIN IGLLKISWEK PVFPENNLQF QIRYGLSGKE VQWKMYEVYD AKSKSVSLPV

         610         620         630         640         650         660
PDLCAVYAVQ VRCKRLDGLG YWSNWSNPAY TVVMDIKVPM RGPEFWRIIN GDTMKKEKNV

         670         680         690         700         710         720
TLLWKPLMKN DSLCSVQRYV INHHTSCNGT WSEDVGNHTK FTFLWTEQAH TVTVLAINSI

         730         740         750         760         770         780
GASVANFNLT FSWPMSKVNI VQSLSAYPLN SSCVIVSWIL SPSDYKLMYF IIEWKNLNED

         790         800         810         820         830         840
GEIKWLRISS SVKKYYIHDH FIPIEKYQFS LYPIFMEGVG KPKIINSFTQ DDIEKHQSDA

         850         860         870         880         890         900
GLYVIVPVII SSSILLLGTL LISHQRMKKL FWEDVPNPKN CSWAQGLNFQ KPETFEHLFI

         910         920         930         940         950         960
KHTASVTCGP LLLEPETISE DISVDTSWKN KDEMMPTTVV SLLSTTDLEK GSVCISDQFN

         970         980         990         1000        1010        1020
SVNFSEAEGT EVTYEDESQR QPFVKYATLI SNSKPSETGE EQGLINSSVT KCFSSKNSPL

         1030        1040        1050        1060        1070        1080
KDSFSNSSWE IEAQAFFILS DQHPNIISPH LTFSEGLDEL LKLEGNFPEE NNDKKSIYYL
```

1090 1100 1110 1120 1130 1140
GVTSIKKRES GVLLTDKSRV SCPFPAPCLF TDIRVLQDSC SHFVENNINL GTSSKKTFAS

1150 1160
YMPQFQTCST QTHKIMENKM CDLTV

Haplotype 4 Protein Sequence (656Asn Polymorphism)

```
10          20          30          40          50          60
MICQKFCVVL LHWEFIYVIT AFNLSYPITP WRFKLSCMPP NSTYDYFLLP AGLSKNTSNS

          70          80          90          100         110         120
NGHYETAVEP KFNSSGTHFS NLSKTTFHCC FRSEQDRNCS LCADNIEGKT FVSTVNSLVF

         130         140         150         160         170         180
QQIDANWNIQ CWLKGDLKLF ICYVESLFKN LFRNYNYKVH LLYVLPEVLE DSPLVPQKGS

         190         200         210         220         230         240
FQMVHCNCSV HECCECLVPV PTAKLNDTLL MCLKITSGGV IFQSPLMSVQ PINMVKPDPP

         250         260         270         280         290         300
LGLHMEITDD GNLKISWSSP PLVPFPLQYQ VKYSENSTTV IREADKIVSA TSLLVDSILP

         310         320         330         340         350         360
GSSYEVQVRG KRLDGPGIWS DWSTPRVFTT QDVIYFPPKI LTSVGSNVSF HCIYKKENKI

         370         380         390         400         410         420
VPSKEIVWWM NLAEKIPQSQ YDVVSDHVSK VTFFNLNETK PRGKFTYDAV YCCNEHECHH

         430         440         450         460         470         480
RYAELYVIDV NINISCETDG YLTKMTCRWS TSTIQSLAES TLQLRYHRSS LYCSDIPSIH

         490         500         510         520         530         540
PISEPKDCYL QSDGFYECIF QPIFLLSGYT MWIRINHSLG SLDSPPTCVL PDSVVKPLPP

         550         560         570         580         590         600
SSVKAEITIN IGLLKISWEK PVFPENNLQF QIRYGLSGKE VQWKMYEVYD AKSKSVSLPV

         610         620         630         640         650         660
PDLCAVYAVQ VRCKRLDGLG YWSNWSNPAY TVVMDIKVPM RGPEFWRIIN GDTMKNEKNV

         670         680         690         700         710         720
TLLWKPLMKN DSLCSVQRYV INHHTSCNGT WSEDVGNHTK FTFLWTEQAH TVTVLAINSI

         730         740         750         760         770         780
GASVANFNLT FSWPMSKVNI VQSLSAYPLN SSCVIVSWIL SPSDYKLMYF IIEWKNLNED

         790         800         810         820         830         840
GEIKWLRISS SVKKYYIHDH FIPIEKYQFS LYPIFMEGVG KPKIINSFTQ DDIEKHQSDA

         850         860         870         880         890         900
GLYVIVPVII SSSILLLGTL LISHQRMKKL FWEDVPNPKN CSWAQGLNFQ KPETFEHLFI

         910         920         930         940         950         960
KHTASVTCGP LLLEPETISE DISVDTSWKN KDEMMPTTVV SLLSTTDLEK GSVCISDQFN

         970         980         990         1000        1010        1020
SVNFSEAEGT EVTYEDESQR QPFVKYATLI SNSKPSETGE EQGLINSSVT KCFSSKNSPL

         1030        1040        1050        1060        1070        1080
KDSFSNSSWE IEAQAFFILS DQHPNIISPH LTFSEGLDEL LKLEGNFPEE NNDKKSIYYL
```

1090 1100 1110 1120 1130 1140
GVTSIKKRES GVLLTDKSRV SCPFPAPCLF TDIRVLQDSC SHFVENNINL GTSSKKTFAS

1150 1160
YMPQFQTCST QTHKIMENKM CDLTV