UNIVERSITY OF SOUTHAMPTON

FACULTY OF MEDICINE, HEALTH AND LIFE SCIENCES

Infection, Inflammation and Immunity Division

Molecular Basis of Gene-Environment Interactions in the Pathogenesis of Asthma and COPD

by

Matthew JJ Rose-Zerilli

This thesis is submitted to the University of Southampton for the degree of Doctor of Philosophy

Abstract

UNIVERSITY OF SOUTHAMPTON

FACULTY OF MEDICINE, HEALTH AND LIFE SCIENCES
INFECTION, INFLAMMATION AND IMMUNITY DIVISION

Doctor of Philosophy

MOLECULAR BASIS OF GENE-ENVIRONMENT INTERACTIONS IN THE PATHOGENESIS OF ASTHMA AND COPD

by Matthew JJ Rose-Zerilli

The origins of respiratory disease, such as asthma in childhood and COPD in later life are unclear. Maternal smoking during pregnancy and low birth weight is associated with increased risk of asthma, poor lung function in adults and COPD in old age. Exposure to oxidative stress and poor nutrition *in utero* is thought to cause damage to the lung and alter the normal course of lung development.

Glutathione S-transferases (GST) are potent antioxidants. In this work, genetic polymorphisms that alter GST enzyme activity were genotyped in a family-based childhood asthma cohort (341 families, n = 1508) and analysed to investigate whether they alter the risk of developing asthma when individuals are exposed to environmental tobacco smoke. Real-time PCR based copy number variation methodology was developed to genotype the common gene deletion polymorphism of *GSTT1* and *GSTM1* genes, for other GST genes (*GSTP1* and *GSTO2*) SNP haplotypes were constructed. A rare *GSTO2* haplotype was negatively associated with asthma susceptibility, atopy severity, and FEV_1 values. Asthmatic children with a *GSTT1* gene deletion, or a common *GSTP1* haplotype, developed more severe asthma compared to individuals with a *GSTT1* gene or non-carriers of the *GSTP1* haplotype. Total IgE levels were increased in *GSTT1**0 individuals when exposed to tobacco smoke in early life, suggesting a gene-environment interaction. *GSTO2* may be a shared susceptibility locus for asthma in childhood and COPD in later life.

Animal models of maternal protein-restriction during pregnancy can induce hypertension, diabetes and endothelial dysfunction in offspring and in some of these models alterations to lung gene expression and lung architecture have been reported. This work established that a rat model of maternal dietary protein-restriction during pregnancy known to induce hypertension in the offspring, results in persistent alterations to the expression of genes in the lungs of adult offspring (120 days), including genes involved in glucocorticoid action (*Hsd11b2*), growth (*Igf1* & 2 and *Pcdh1*) and alveolar development (*Tp53*). Lung microRNA expression profiles were also altered in response to exposure to protein restriction *in utero*. These findings suggest a role for nutritional programming in respiratory disease susceptibility in later life and a role for microRNAs in the study of the developmental origins of health and disease in general. Further work will include the investigation of epigenetic mechanisms that control nutritional programming in lungs of animals exposed to protein-restriction *in utero*.

This work has demonstrated that GST polymorphism is a risk factor for childhood asthma and certain genotypes can offer some protection against the development of severe asthma. There was little evidence to suggest that GST polymorphism modulates the effects of smoke exposure in early life. In addition, we have demonstrated that maternal diets that are poor in nutrition could predispose her offspring to respiratory disease in later life by altering the course of normal lung development in early life or response to environmental stimuli in later life.

Table of Contents

Abstract ii
Table of Contents iii
List of Figures v
List of Tables vii
List of Tables vii
List of Equations xii
List of Equations xii
Declaration of Authorship xiii
Declaration of Authorship xiii
Acknowledgements xiv
Abbreviations xv

1 Chapter 1: Introduction 1
1.1 Asthma 1
1.2 Pathogenesis of Asthma 5
1.3 Aetiology of Asthma 12
1.4 Genetic Epidemiology 23
1.5 Role of Glutathione S-Transferases in Protection Against Oxidative Stress 31
1.6 Developmental Origins of Health and Disease 38
1.7 Aims of Research 46

2 Chapter 2: General Materials and Methods 47
2.1 Materials 47
2.2 Methods 53

3 Chapter 3: Optimisation and Validation of Glutathione S-transferase Copy Number Variation Assays 67
3.1 Introduction 67
3.2 Aims 70
3.3 Methods 71
3.4 Results 81
3.5 Discussion 87
3.6 Concluding Remarks and Future Work 88

4 Chapter 4: Environmental Tobacco Smoke, Glutathione S-Transferase Gene Polymorphisms and Childhood Asthma 89
4.1 Introduction 89

4.2 Aims 96
4.3 Methods 97
4.4 Results 111
4.5 Discussion 130
4.6 Concluding Remarks and Future Work 136

5 Chapter 5: Protein-Restricted Rat Lung: Gene Expression Analysis 138
5.1 Introduction 138
5.2 Aims 150
5.3 Methods 151
5.4 Results 163
5.5 Discussion 180
5.6 Concluding Remarks and Future Work 190

6 Protein-Restricted Rat Lung: micro RNA Expression Analysis 192
6.1 Introduction 192
6.2 Aims 199
6.3 Methods 200
6.4 Results 206
6.5 Discussion 215
6.6 Conclusion 222
6.7 Future Work 223

7 General Discussion 225
7.1 Gene-Environment Interactions in the Early Life Origins of Asthma 227
7.2 Modelling the Developmental Origins of Respiratory Disease Hypothesis 230
7.3 Conclusion 233
7.4 Future Directions 234

8 References 235
9 Website Bibliography 266
10 Appendices 267
10.1 Reagent Preparations for Total RNA, DNA and Protein Isolation Using TRIzol™ Reagent: 268
10.2 Robotic Liquid Handling Protocols using the Biomek 3000 Platform 269
10.3 Use of the Lightcycler480™ Real-Time PCR Machine 270
10.4 Real-Time PCR Data Analysis 273
10.5 Full FBAT Result Outputs 277
10.6 Publications Arising From This Thesis 281

List of Figures

Figure 1-1: Structural Changes in Asthmatic Airways.....5

Figure 1-2: Lung Remodelling in Asthmatic Airways 11

Figure 1-3: Glutathione Metabolism Downstream of Glutathione S-Transferases. 34

Figure 1-4: DNA Methylation.....45

Figure 2-1: Example 384 Well Reaction Plate Layout. 64

Figure 3-1: Determining Gene Dosage by Real-Time PCR. 69

Figure 3-2: TaqMan PCR Thermalcycling Reaction. 72

Figure 3-3: TaqMan Probe Designs. 73

Figure 3-4: Copy Number Variation Ranked XY Scattergraph. 76

Figure 3-5: Zoomed Copy Number Variation Ranked XY Scattergraph. 77

Figure 3-6: Full Screenshot of the CNV Analysis Spreadsheet User Interface. 78

Figure 3-7: Colour Coded 384 & 96 Well Plate Layout Arrays for Data Quality Control. ..80

Figure 3-8: *GSTM1* Real-Time Amplification Plots in Reactions with a Decreasing Probe Concentration.....81

Figure 3-9: *GSTM1* Standard Curve Analysis. 82

Figure 3-10: *GSTT1* Standard Curve. 83

Figure 3-11: Matrix Histograms of the Assay Stability Trial. 85

Figure 3-12: *GSTT1* CNV Scattergraph from the Roche Lightcycler 480 Machine. 86

Figure 4-1: Mechanism of Cigarette Smoke Mediated Lung Injury by Glutathione Depletion 92

Figure 4-2: Principles of SNP Genotyping by High Resolution Melting Curve Analysis ..102

Figure 4-3: Normalisation and Temperature-Shifting of HRM Data. 105

Figure 5-1: Protein Restriction Study Design. 152

Figure 5-2: Lung Fixation Apparatus.....154

Figure 5-3: TRIzol™ Workflow Diagram. 155

Figure 5-4: Relative Abundance of Target Genes. 160

Figure 5-5: Boxplots of F1 225 Day Old Lung Measurements. 163

Figure 5-6: Validation of mRNA Stabilisation Methods Amplification Plots. 164

Figure 5-7: Validation of mRNA Stabilisation Methods Relative Expression Chart. 165

Figure 5-8: SYBRgreen Melting Curve Analysis. 166

Figure 5-9: *Gstp1* Standard Curve Analysis. 167

Figure 5-10: 120 Day F_1 Generation Rat Lung qPCR mRNA Expression Results. 171

Figure 5-11: 120 Day F_1 Generation Rat Lung qPCR mRNA Differential Expression Results. 172

Figure 5-12: Fold Difference in Median Relative Quantification Values. 173

Figure 5-13: IUGR mRNA Expression in 225 Day F_1 Generation Rat Lung 174

Figure 5-14: GST mRNA Expression in 120 Day F1 Generation Rat Lung. 175

Figure 5-15: F_2 Generation Rat Lung qPCR mRNA Expression Results. 177

Figure 5-16: Ratio of PR Versus Control mRNA Levels from Pooled 120 Day Lung and Liver cDNA Templates. 178

Figure 6-1: microRNA Biogenesis. 193

Figure 6-2: Principles of miScript miRNA Real-Time PCR System. 202

Figure 6-3: Principles of miRcury miRNA Real-Time PCR System. 203

Figure 6-4: Double-Log Scatter Plot of Re-Ratio Values from the Microarray Experiment. 206

Figure 6-5: miScript miRNA Real-Time PCR Validation Using Trizol™ Total RNA Isolates Enriched for miRNAs. 208

Figure 6-6: miScript miRNA Real-Time PCR Validation Using Miltenyi Biotec RNA Isolates. 209

Figure 6-7: miRcury miRNA Real-Time PCR Validation Using Miltenyi Biotec RNA Isolates. 210

Figure 6-8: miRNA Gene Expression by Chromosome. 212

Figure 10-1: Biomek 3000 Robot in the Asthma Genetics Laboratory 269

Figure 10-2: Thermalcycling Conditions for TaqMan and UPL Based qPCR Assays 270

Figure 10-3: Thermalcycling Conditions for SYBRgreen Based qPCR Assays 270

Figure 10-4: SYBRgreen Thermal Melt Program 272

Figure 10-5: Standard Curve PCR Efficiency (E) Estimation. 274

Figure 10-6: LinRegPCR Program Screenshot. 274

List of Tables

Table 1-1: Objective Measurements for Asthma Diagnosis....2

Table 1-2: Static and Synamic Pulmonary Function Tests Used to Diagnose and Monitor Asthma....3

Table 1-3: Immune Cell Types Involved in Asthma. 8

Table 1-4: Key Cytokine Signalling Molecules Involved in Asthma Pathogenesis. 9

Table 1-5: Table of Positionally Cloned Asthma Genes 22

Table 1-6: Two-SNP Haplotype Combinations. 26

Table 1-7: Glutathione S-Transferase Gene Polymorphisms 36

Table 1-8: Windows of Developmental Plasticity. 40

Table 2-1: Reagents 47

Table 2-2: Enzymes and Buffers....47

Table 2-3: TaqMan Probe Assays for qPCR....48

Table 2-4: UPL Probe or miRNA Assays for qPCR 48

Table 2-5: SYBRgreen Assays for qPCR 49

Table 2-6: DNA Cohorts and Human Random Control DNA Samples 49

Table 2-7: Protein Restriction Rat Lung Tissue Samples. 50

Table 2-8: Commercial Laboratory Kits 50

Table 2-9: Plasticware 50

Table 2-10: Apparatus 51

Table 2-11: Software....51

Table 2-12: Miscellaneous....52

Table 2-13: RT Primer Reaction Setup....59

Table 2-14: RT Master Mix Setup 59

Table 2-15: Primer Concentrations Utilised in Optimisation Reactions 60

Table 2-16: Master Mix Setup Tables for Primer Titration Reactions. 61

Table 2-17: Final Probe Concentrations Used in Probe Titration Optimisation. 62

Table 2-18: Master Mix Setup Tables for Probe Titration Reactions 62

Table 2-19: Two-Fold Serial Dilution of Pooled gDNA Template (ECACC HRC1 DNA Panel) for the GST Copy Number Variation Assays. 63

Table 2-20: Four-Fold Serial Dilution of cDNA Template for mRNA or miRNA Assays..... 63

Table 2-21: Ten-Fold Serial Dilution of cDNA Template of the mRNA or miRNA Assays. 63

Table 2-22: ECACC Genomic Control Sample Names Used as Positive Controls for the GST CNV Assays 64

Table 2-23: CNV Master Mix Setup. 65

Table 3-1: GST CNV Genotype Frequencies in the Southampton Asthma Cohort. 84

Table 4-1: Antioxidant Defences in the Lung. 90

Table 4-2: Advantages and Disadvantages of SNP Genotyping by High-Resolution Melting Curve Analysis 102

Table 4-3: *GSTO2* and *GSTP1* HRM Primers 103

Table 4-4: HRM-PCR Reagent Master Mix Setup 103

Table 4-5: FBAT Regression Analysis Outcomes..... 107

Table 4-6: PBAT Power Calculation Settings 110

Table 4-7: GST Copy Number Variation Genotype Frequencies in the Final Data Set ... 111

Table 4-8: *GSTO2* and *GSTP1* SNP Allele Frequencies..... 111

Table 4-9: *GSTO2* FBAT Haplotype Frequencies 112

Table 4-10: *GSTP1* FBAT Haplotype Frequencies..... 112

Table 4-11: FBAT Results from Association Analysis Involving *GSTM1*, *GSTT1* CNV Genotypes and the Diagnosis of Asthma..... 113

Table 4-12: FBAT Results from Association Analysis Involving *GSTO2* and *GSTP1* Haplotypes and the Diagnosis of Asthma. 113

Table 4-13: FBAT Results from Association Analysis Involving *GSTM1*, *GSTT1* CNV Genotypes and the Log Age-Corrected Total IgE Levels in Asthmatic Siblings. 114

Table 4-14: FBAT Results from Association Analysis Involving *GSTO2* and *GSTP1* Haplotypes and the Log Age-Corrected Total IgE Levels in Asthmatic Siblings 114

Table 4-15: FBAT Results from Association Analysis Involving *GSTM1*, *GSTT1* CNV Genotypes and the Measure of BHR on the Asthmatic Siblings. 115

Table 4-16: FBAT Results from Association Analysis Involving *GSTO2* and *GSTP1* Haplotypes and the Measure of BHR on the Asthmatic Siblings. 115

Table 4-17: FBAT Results from Association Analysis Involving *GSTM1*, *GSTT1* CNV Genotypes and Percentage Predicted FEV1 Values on Asthmatic Siblings. 116

Table 4-18: FBAT Results from Association Analysis Involving *GSTO2* and *GSTP1* Haplotypes and Percentage Predicted FEV_1 Values on Asthmatic Siblings.....116

Table 4-19: FBAT Results from Association Analysis Involving *GSTM1*, *GSTT1* CNV Genotypes and Atopy Severity Scores on Asthmatic Siblings.....117

Table 4-20: FBAT Results from Association Analysis Involving *GSTO2* and *GSTP1* Haplotypes and Atopy Severity Scores on Asthmatic Siblings. 117

Table 4-21: FBAT Results from Association Analysis Involving *GSTM1*, *GSTT1* CNV Genotypes and Asthma Severity Scores on Asthmatic Siblings.....118

Table 4-22: FBAT Results from Association Analysis involving *GSTO2* and *GSTP1* Haplotypes and Asthma Severity Scores on Asthmatic Siblings. 118

Table 4-23: FBAT Results from Association Analysis Involving *GSTM1*, *GSTT1* CNV Genotypes and the Diagnosis of Asthma on Smoke Exposed Individuals. 119

Table 4-24: FBAT Results from Association Analysis Involving *GSTT1* CNV Genotypes and the Log Age-Corrected Total IgE Levels on Smoke Exposed Individuals. 119

Table 4-25: FBAT Results from Association Analysis Involving *GSTT1* CNV Genotypes and Asthma Severity Scores on Smoke Exposed Individuals. 120

Table 4-26: FBAT Results from Association Analysis Involving *GSTM1*, *GSTT1* CNV Genotypes and the Diagnosis of Asthma on Individuals Unexposed to Smoke. 120

Table 4-27: FBAT Results from Association Analysis Involving *GSTO2* and *GSTP1* Haplotypes and the Diagnosis of Asthma on Individuals Unexposed to Smoke. 121

Table 4-28: FBAT Results from Association Analysis Involving *GSTO2* and *GSTP1* Haplotypes and Log Age-Corrected IgE Levels in Individuals Unexposed to Smoke. 121

Table 4-29: FBAT Results from Association Analysis Involving *GSTO2* and *GSTP1* Haplotypes and Percentage Predicted FEV_1 in Individuals Unexposed to Smoke.....122

Table 4-30: FBAT Results from Association Analysis Involving *GSTO2* and *GSTP1* Haplotypes and Atopy Severity Score in Individuals Unexposed to Smoke. 122

Table 4-31: FBAT Results from Association Analysis Involving *GSTT1* CNV Genotypes and the Asthma Severity Scores on Individuals Unexposed to Smoke. 123

Table 4-32: FBAT Results from Association Analysis Involving *GSTM1*, *GSTT1* CNV Genotypes and the Log Age-Corrected Total IgE Levels on Asthmatic Siblings Using Smoke Exposure as a Covariate. 123

Table 4-33: FBAT Results from Association Analysis Involving *GSTO2* and *GSTP1* Haplotypes and the Log Age-Corrected Total IgE Levels on Asthmatic Siblings Using Smoke Exposure as a Covariate. 124

Table 4-34: FBAT Results from Association Analysis Involving *GSTM1*, *GSTT1* CNV Genotypes and the Measure of BHR on Asthmatic Siblings Using Smoke Exposure as a Covariate....124

Table 4-35: FBAT Results from Association Analysis Involving *GSTO2* and *GSTP1* Haplotypes and the Measure of BHR on Asthmatic Siblings Using Smoke Exposure as a Covariate....125

Table 4-36: FBAT Results from Association Analysis Involving *GSTM1*, *GSTT1* CNV Genotypes and Percent Predicted FEV_1 Values on Asthmatic Siblings Using Smoke Exposure as a Covariate....125

Table 4-37: FBAT Results from Association Analysis Involving *GSTO2* and *GSTP1* Haplotypes and Percent Predicted FEV_1 Values on Asthmatic Siblings Using Smoke Exposure as a Covariate....126

Table 4-38: FBAT Results from Association Analysis Involving *GSTM1*, *GSTT1* CNV Genotypes and Atopy Severity Scores on Asthmatic Siblings Using Smoke Exposure as a Covariate....126

Table 4-39: FBAT Results from Association Analysis Involving *GSTO2* and *GSTP1* Haplotypes and Atopy Severity Scores on Asthmatic Siblings Using Smoke Exposure as a Covariate....127

Table 4-40: FBAT Results from Association Analysis Involving *GSTM1*, *GSTT1* CNV Genotypes and Asthma Severity Scores on Asthmatic Siblings Using Smoke Exposure as a Covariate....127

Table 4-41: FBAT Results from Association Analysis Involving *GSTO2* and *GSTP1* Haplotypes and Asthma Severity Scores on Asthmatic Siblings Using Smoke Exposure as a Covariate....128

Table 4-42: Summary of the FBAT Association and Regression Analyses....129

Table 5-1: Milestones in Human Lung Development. 143

Table 6-1: Rattus Norvegicus miRNA Gene Clusters. 194

Table 6-2: Rat Lung Developmental Stages and miRNA Expression. 198

Table 6-3: miScript MicroRNA PCR Reaction Master Mix. 202

Table 6-4: miRcury MicroRNA Real-Time PCR Reaction Master Mixes. 204

Table 6-5: Putative Up or Down Regulated miRNA Species in Protein-Restricted Offspring Lung Tissue. 207

Table 6-6: miRNA Species Found Only in Protein-Restricted or Control Diet Offspring Lung Tissue. 207

Table 6-7: miRNA Species Validated by Real-Time PCR.....210

Table 6-8: miRNAs in Protein Coding Genes 211

Table 6-9: Predicted Target Genes for rno-mir-126, 186 and 365.....214

Table 10-1: Biomek 3000 Plasticware for DNA Arrays 269

Table 10-2: Advantages and Disadvantages of the Different PCR Efficiency Estimation Methods 275

List of Equations

Equation 1: Percent Predicted FEV_1 Formula....4
Equation 2: Hardy-Weinberg Equilibrium Calculation....23
Equation 3: Glutathione Redox Reactions.....31
Equation 4: Glutathione S-Transferases....33
Equation 5: Calculation of BHR Variable....99
Equation 6: Total Lung Volume Extrapolation....153
Equation 7: Lung Weight Index....153
Equation 8: Lung Volume Index....153
Equation 9: Formula for Calculating the Minimum Detectable Difference in Expression. 162
Equation 10: Concentration Ratio Formula....271
Equation 11: Normalised Ratio Formula....271
Equation 12: PCR Reaction Calculation.....273
Equation 13: Comparative Ct Method Standard Deviation.....276
Equation 14: Relative Standard Curve Standard Deviation.....276

Declaration of Authorship

I, Matthew John Jerome Rose-Zerilli, declare that the thesis entitled, 'Molecular Basis of Gene-Environment Interactions in the pathogenesis of Asthma and COPD' and the work presented in it is my own. I confirm that:

The work was done wholly or mainly while in candidature for a research degree at this University;

- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
- Where I have consulted the published work of others, this is always clearly attributed;
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
- I have acknowledged all main sources of help;
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
- Part of this work has been published as:

Rose-Zerilli MJ, Barton SJ, Henderson AJ, Shaheen SO, Holloway JW.Copy-Number Variation Genotyping of GSTT1 and GSTM1 Gene Deletions by Real-Time PCR. Clin Chem. 2009 Sep;55(9):1680-5.

Minelli C & Granell R, Newson R, **Rose-Zerilli MJ**, Torrent M, Ring SM, Holloway JW, Shaheen SO, Henderson AJ. Glutathione-S-transferase genes and asthma phenotypes: a HuGE systematic review and meta-analysis including unpublished data. Int J Epidemiol. 2010 Apr;39(2):539-62.

Signed:

Date:

Acknowledgements

Where do I start to begin to thank everyone that's helped me over the last three years?

Firstly, I would like to thank Dr John Holloway for all of the support and great supervision. I would also like to extend my gratitude to my co-supervisor Dr. Christopher Torrens; Dr. Jude Holloway for showing me how to perform RNA extractions properly and Ms Sheila Barton for her help with statistics. Secondly, a big shout out to all my colleagues in the Asthma Genetics Group, Human Genetics Division and the molecular pathology guys for the 'occasional' tea and biccy breaks.

All my love to my girlfriend/long suffering common law partner Frances 'FrannyPants' Monk for the all the support, understanding and yummy food! A special thanks to my parents and Brenda and Steve Monk for their support. I mustn't forget to say cheers to my close friends, Tom, Chris, Kristen, Fraser, Jen and James for the good times, one more year and I will be able to buy a round!

Finally, I would like to thank the School of Medicine for funding my studentship.

Abbreviations

Abbreviation	Full name
1 x TBE	1 x Tris-Borate EDTA
Abs	Absorbance
APCs	Antigen Presenting Cells
ASM	Airway Smooth Muscle
ATP	Adenosine Tri-Phosphate
BHR	Bronchial Hyper-Responsiveness
bp	basepair
BTS	British Thoracic Society
CD14	Monocyte differentiation antigen
CNV	Copy Number Variation
Cp	Crossing point value
Ct	Crossing threshold
DEPC	Diethyl Pyrocarbonate
dsDNA	Double-stranded DNA
EAR	Early phase Allergic Response
EBC	Exhaled Breath Condensate
ECACC	European Collection of Cell Cultures
EDTA	EthyleneDiamineTetraacetic Acid
ETS	Environmental Tobacco Smoke
FEV_1	Forced Expiratory Volume in 1 second
GCS	Glutamylcysteine
GLP	Good Laboratory Practice
GM-CSF	Granulocyte Monocyte Colony Stimulating Factor
GR	Glucocorticoid Receptor
grm	gram
GSH	Glutathione
GSSG	Glutathione disulphide
GST	Glutathione S-Transferase
H_2O_2	Hydrogen Peroxide
HDM	House dust Mite
htSNP	Haplotype tagging Single Nucleotide Polymorphism
HWE	Hardy-Weinberg Equilibrium
IgE	Immunoglobulin E
IL-13	Interleukin-13
IL-4	Interleukin-4
IL-5	Interleukin-5
IL-9	Interleukin-9
IUGR	Intrauterine Growth Restriction
LAR	Late phase Allergic Response
LNA	Locked Nucleic Acids
LPS	Lipopolysaccharide
LTs	Leukotrienes
Mb	Megabase (DNA)
miRNA	microRNA
ml	millilitre
NAPDH	Nicotinamide Adenine Dinucleotide Phosphate
NFKB	Nuclear Factor kappa B
nM	Nano Molar
NO	Nitric Oxide
NOS	Reactive Nitrogen Species

Abbreviation	Full name
PEF	Peak Expiratory Flow
PGs	Prostaglandins
pM	Pico Molar
PR	maternal Protein-Restriction
qPCR	quantitative PCR or Real-Time PCR
RISC	RNA Induced Silencing Complex
ROS	Reactive Oxygen Species
RTLF	Respiratory Tract Lining Fluid
RT-PCR	Reverse Transcription PCR
RT-ve	Reverse Transcriptase Negative Control

Abbreviation	Full name
SNP	Single Nucleotide Polymorphism
TGFβ	Transforming Growth Factor beta
T_H1	T helper Cell Type 1 response
T_H2	T helper Cell Type 2 response
Tm	Melting Temperature
TNF	Tissue Necrosis Factor
TNFα	Tissue Necrosis Factor alpha
μM	Micro Molar

1 Chapter 1: Introduction

1.1 Asthma

1.1.1 Incidence

The current estimate of asthma incidence worldwide by the World Health Organisation was 300 million in 2005 and is now the most common chronic disease in children.[1] In the UK there are 5.2 million asthmatics (1.1 million of which are children) and the healthcare burden of asthma is around £996 million in the UK, with a total economic cost of £2.3 billion per annum.(www.asthmauk.org.uk) It is important to remember that this disease can be a killer with just over 1,300 deaths in the UK and 255,000 individuals dying in 2005 from fatal asthma exacerbations worldwide.[1]

1.1.2 Prevalence

Worldwide prevalence of asthma symptoms (current wheeze) in 6-7 year old children was ~11% and ~13% in 13-14 year olds in 2006 (mean increase of 0.13% and 0.06% per year).[2, 3] The International Study of Asthma and Allergies in Children (ISAAC) determined that asthma prevalence had decreased slightly in English speaking countries and Western Europe (-0.51% and -0.09%; for 13-14 and 6-7 year olds respectively) in the last 10 years, while asthma had increased in those countries where prevalence was previously low.[3] However, the percentage of children reported to have had asthma at some time in their lives increased by 0.28% per year in the 13-14 year age group and by 0.18% per year in the 6-7 year age group.[3] The increases in asthma symptom prevalence in Africa, Latin America and parts of Asia indicate that the global burden of asthma is continuing to rise, but global prevalence differences are lessening.[3]

1.1.3 Clinical Symptoms

Asthma is a chronic respiratory disease involving episodic narrowing of the airways by bronchoconstriction and inflammation. Acute asthma exacerbations (commonly known as 'asthma attacks') can be triggered by various aero-allergens (pollen, house dust mite, chemicals and animal allergens), respiratory virus infections, exercise, cold air, tobacco smoke and pollution (e.g. diesel exhaust particles). Asthma triggers vary person to person and can worsen the asthma symptoms of wheeze, shortness of breath, chest tightness and coughing without always causing an asthma attack.

1.1.4 Diagnosis

Asthma is simply defined as a reversible airways obstruction. The British Thorax Society (BTS) published updated asthma management guidelines in 2007, detailing asthma diagnosis by combining symptoms with a family history of asthma or atopy and several objective measurements of lung function (updated guidelines are available via www.brit-thoracic.org.uk; originally published in Thorax in 2003[4]). As symptoms are shared with other respiratory disorders is it important for the diagnosing practitioner to have clear guidelines for diagnosis and be aware of the co-morbidity of the symptoms to avoid misdiagnosis. In the UK, practitioners examine adult patients for any signs of wheeze (diffuse, bilateral or expiratory), or tachypnea (rapid breathing) it is common for patients to present with no signs. Records of personal or family history of asthma or atopy (eczema, allergic rhinitis), adverse effects after aspirin/NSAID ingestion or use of β-blockers, recognised triggers, and pattern and severity of symptoms or exacerbations aid in diagnosis. They also perform at least one of the following objective spirometry measurements to measure airway function in order to confirm airway obstruction (see table 1.1).

Table 1-1: Objective Measurements for Asthma Diagnosis

Objective Measurements for Asthma Diagnosis
>20% diurnal variation on ≥3 days in a week for two weeks on Peak Expiratory flow (PEF) diary
FEV_1 ≥ 15% (and 200ml) increase after short acting β-agonists (eg salbutamol)
FEV_1 ≥ 15% (and 200ml) increase after trial of steroid tablets (prednisolone 30mg/day for two weeks)
FEV_1 ≥ 15% decrease after 6 minutes of exercise (running)
Histamine or methacholine challenge in difficult cases

If the diagnosis is unclear or in doubt the patient is referred to a specialist. In children the diagnosis of asthma can be more complicated, as it is not often possible to measure airway function, so the BTS provide alternative guidelines. These guidelines recommend identifying a key feature of asthma, monitor the child's response to trials of asthma treatments and perform repeated assessment of the child to detect if management is ineffective.[4]

1.1.4.1 Lung function testing

Pulmonary function tests objectively measure the ability of the respiratory system to perform gas exchange by assessing its ventilation, diffusion and mechanical properties. Spirometry and measurements of lung volumes determine the presence and severity of obstructive and restrictive pulmonary dysfunction. Asthma causes obstructive pulmonary dysfunction and sufferers generally have less than predicted airflow rates as shown in table 1.2.

Table 1-2: Static and Dynamic Pulmonary Function Tests Used to Diagnose and Monitor Asthma

Static Tests	Dynamic Tests
Vital capacity (**VC**): Change in volume of gas in the lung from complete inspiration to complete expiration	Forced Expiratory Volume in first second (**FEV_1**): The volume of air expelled in the first second of a forced expiration starting from full inspiration. **FEV_1%**: This is the FEV_1 expressed as a percentage of the total volume (FEV_1 Ratio or FEV_1/VC% or FEV_1/FVC%)
Forced Vital Capacity (**FVC**): The maximum volume of air in litres that can be forcibly and rapidly exhaled following a maximum inspiration. FVC is the basic manoeuvre in spirometry tests	Peak Expiratory Flow Rate (**PEFR**): The greatest flow that can be sustained for 10 milliseconds on forced expiration from full inflation, this is measured with a peak flow meter in Litres per minute. Forced Expiratory Flow Rate (**FEFR 25%-75%**): Average expiratory flow rate at the middle part of the FVC manoeuvre. Expressed in litres per second it gives an indication of the function of the lower airways and can act as an early warning sign of disease.

Bronchial hyperresponsiveness and Percent Predicted Forced Expiratory Volume in first second (%FEV_1 predicted) are two commonly used pulmonary function measurements used to monitor asthma severity and airways obstruction.

1.1.4.1.1 Bronchial Hyperresponsiveness: Methacholine Challenge

Bronchial Hyperresponsiveness (BHR) is a hallmark of asthma and is classed as 'specific', when the airways narrow in response to allergen (e.g. cat dander) or 'non-specific', when the airways narrow in response to agonists (e.g. ozone, histamine). Non-specific responsiveness can occur in both asthmatics and non-asthmatics. In response to a challenge the smooth muscle around airways contract and narrow the airway, reducing lung function. A histamine or methacholine challenge is used as an objective measurement for diagnosis and in determining asthma severity. BHR can be induced by inhaling increasing concentrations (0.06 - 16 mg/ml) of nebulised methalcholine or histamine; the concentration required to induce a 20% reduction in FEV_1 is termed as the 'Provocative Concentration' or PC_{20} (mg/ml). Asthmatic airways in general, require a smaller dose of stimulant to induce bronchoconstriction than normal airways (this is defined as airways sensitivity).

The FEV_1 values can be plotted against stimulant concentration and the reciprocal of the linear regression slope is used as measure of the individuals airway reactivity. Thus, the slope will be in general steeper in asthmatics as the airways are more reactive (i.e. airway narrowing per unit of agonist is greater) and the maximum reduction in FEV_1 will be larger than normal airways, this is referred to as 'maximal response'.

1.1.4.1.2 Percent Predicted FEV1

FEV_1 is the volume of air expelled in the first second of a forced expiration starting from full inspiration; this dynamic test of pulmonary function is the most widely used index of assessing airways obstruction. Percent predicted FEV_1 is an individual's baseline FEV_1 expressed as a percentage of the predicted FEV_1 value according to their height and age using this formula (see equation 1).

Equation 1: Percent Predicted FEV_1 Formula

$$\%\,\text{FEV1 Predicted} = \frac{\text{Obtained Baseline FEV1 Value}}{\text{Predicted Baseline FEV1 Value}}$$

Predicted FEV_1 values are determined from large population based epidemiological studies of lung function such as the standardised lung volume and forced ventilatory flow tables of the European Community for steel and coal (ESCS) workers.[5]

1.2 Pathogenesis of Asthma

Asthma is an inflammatory disease of the airways characterised by airways hyper-activity, chronic eosinophilic inflammation, mucus hyper-secretion and reversible bronchoconstriction (see figure 1.1).[6] Various cellular pathophysiological pathways are therefore involved in asthma, and sufferers can be categorised as atopic, non-atopic or aspirin sensitive asthmatics. The majority of asthma (especially in childhood) is mediated by allergic inflammation.[7] Childhood asthma tends to have T_H2 driven responses to allergens such as house dust mite and therefore has significant co-morbidity with atopic IgE mediated diseases like allergic rhinitis and eczema.[8] Adult onset asthma on the other hand tends to be more severe and is often not mediated by IgE (non-atopic).[9] Finally, some asthmatics are sensitive to aspirin (and other cyclooxygenase inhibitors) and exhibit symptoms when exposed to these agents.[10] This form of asthma is also considered to be non-atopic and sufferers can be desensitised to aspirin.[10] Therefore, not all asthmatics are atopic, non-atopic asthma is particularly prevalent in adult-onset asthma but is also found in children and may result from reactions to cold air, respiratory virus infection and environmental air pollution.[11]

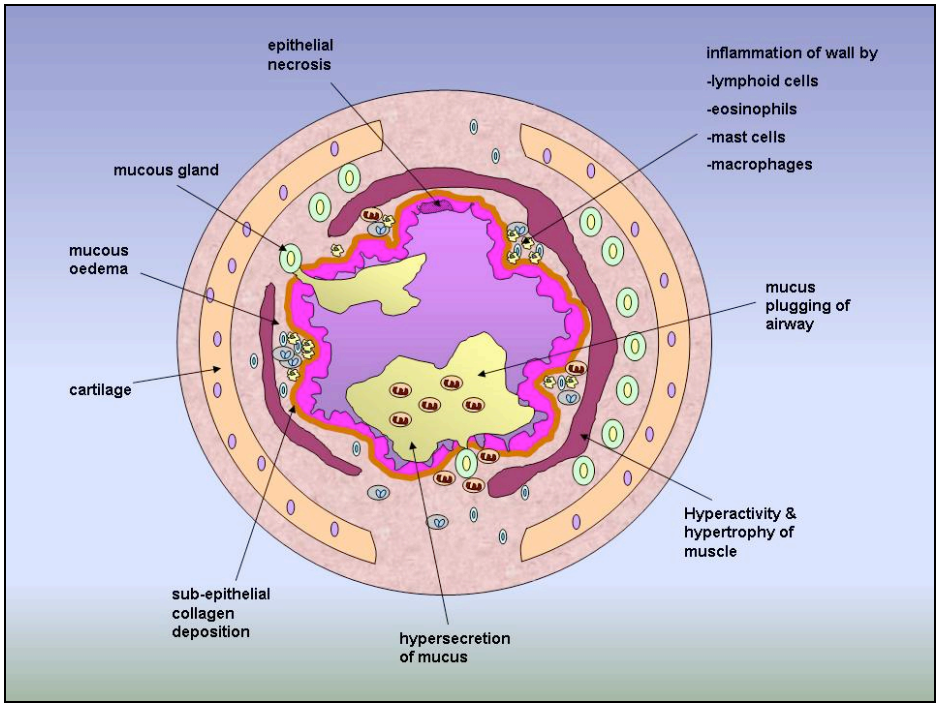


Figure 1-1: Structural Changes in Asthmatic Airways

Asthmatic airways are generally characterised by the influx of inflammatory cells, hypersecretion of mucus (yellow), thickening of the airway smooth muscle (purple) and collagen deposition (brown layer).

1.2.1 Atopic Asthma

Atopic Asthma is also referred to as extrinsic asthma; this type of asthma is triggered by identifiable external allergens. Patients with this form of asthma have IgE mediated allergic responses to common allergens. Atopic asthma is also known as early-onset asthma and is the type commonly seen in children. Numerous population based studies have shown that prevalence of atopy can range from 4 - 64% and asthma from 1 - 38%.[11] Therefore, atopy is not always associated with asthma, and not all asthmatics are atopic.

1.2.2 Non-Atopic Asthma

In non-atopic asthma (intrinsic asthma or late onset asthma) the agents or conditions responsible for attacks are generally unknown but can include pollution (diesel exhaust particulate matter and ozone), cold air and infection with virus. The sufferers tend to be adults that are not atopic.

1.2.3 Categories of Asthma

1.2.3.1 Seasonal Asthma

Asthma may be exacerbated by seasonal increases in specific aeroallergens (pollens, e.g. birch pollen and ragweed) in certain sensitised individuals. Seasonal asthma is usually associated with allergic rhinitis and can occur intermittently with the patient being completely asymptomatic between seasons.

1.2.3.2 Occupational Asthma

Occupational asthma is asthma acquired in the workplace due to exposure to chemicals by inhalation. A wide range of chemicals are known to induce asthma, from chemicals such as isocyanates to animal and plant products. Occupational asthma can be misdiagnosed as chronic bronchitis or COPD; absence of asthma before employment and reduction in symptoms when absent from the workplace are key signs for diagnosis.

1.2.3.3 Cough Variant Asthma

This type of asthma is characterized by recurrent cough particularly at night. These patients seldom have wheezing episodes and consequently display normal signs under examination; administration of bronchodilators can be diagnostic.

1.2.4 Allergic Inflammation

The allergic cascade is crucial in the development of airway inflammation in allergic asthma.[7] Inhaled allergens that escape mucociliary clearance are taken up by antigen presenting cells (APCs) that are found throughout the respiratory tract (nasal mucosa to lung pleura). After migrating to the draining lymph nodes, the processed antigen is presented to allergen specific T and B cells. Activation of T helper (T_h) cells by APCs leads to the production of cytokines (IL-4, IL-5, IL-9, IL-13) that regulate the iso-type switching of B-cells to produce immunoglobulin E (IgE), mast cell development, mucus overproduction, airway hyperresponsiveness (IL-4, IL-13 & IL-9) and eosinophil accumulation (IL-5). Released IgE antibodies circulate in the blood before binding to the high affinity IgE receptor FcεR1 on mast cells in tissue or on peripheral blood basophils.

After re-exposure, the allergen cross-links to mast cell bound specific IgE (activating membrane and cytosolic pathways) triggering the release of preformed pro-inflammatory mediators such as, histamine, tryptase, chymase, and newly synthesised prostaglandins (PGs), leukotrienes (LTs), and cytokines are transcribed (*TNFα* & *VEGF*). The release of these mediators cause the early-phase asthmatic reaction (EAR) which is characterised by mucus production, vascular leakage, constriction of airway smooth muscle (ASM) cells and enhanced airway hyperresponsiveness. The EAR is immediate, lasting up to one hour and is followed by the late-phase asthmatic reaction (LAR) between 4-6 hours later. Table 1.3 and 1.4 details the role of immune cells in asthma and key cytokine molecules that are essential for maintaining inflammation.

The late phase asthmatic reaction consists of persistent and excessive inflammation driven by recruitment of inflammatory cells such as macrophages/monocytes and eosinophils, that in the long-term results in structural changes to the airways tissue. This is known as tissue remodelling and is discussed in detail in section 1.2.5.

Table 1-3: Immune Cell Types Involved in Asthma.

This table provides brief overview of the role of different immune cells involved in the disease.

Lineage	Cell type	Role in inflammation	Role in asthma
Lymphocytes	B cell	Matured in bone marrow. Synthesis and display of immunoglobulin (antibody) molecules that bind antigen. Clonal differentiation into specific plasma cells (effector cell) or memory cells. Interactions with T cells and macrophages by presenting processed antigen bound on Class II MHC molecules	Induction of the asthmatic reaction
	T cell	Matured in the thymus. Binds processed antigen though T-cell receptor. Recognises only antigen bound to MHC molecules on APC or virus infected cells, cancer cells or grafts.	Induction of the asthmatic reaction, but important during all of the allergic response especially T_H2 lymphocytes.
Mononuclear	Dendritic cell	Express MHC class II on cell surface. Antigen presenting cell. Presents antigen to blood T lymphocytes.	Involved in late phase asthmatic reaction
	Monocyte	Circulates in the bloodstream then differentiates into tissue specific Macrophage cells	See macrophage.
	Macrophage	Phagocytic cell. Antimicrobial and cytoxic activity.Antigen processing and presentation (to T_H cells). Secretion of cytokine factors. Promotes pro-inflammatory response.	Involved in late phase asthmatic reaction Airways remodelling
Granulocytes	Basophils	Circulating non-phagocytic cells. Release of pharmacologically active substances such as histamine and MBP.	Involved in early phase asthmatic reaction.
	Eosinophil	Motile phagocytic cells. Cytotoxic defence against parasitic organisms. Secretes major basic protein (MBP) and Eosinophilic cationic protein.	Involved in late phase asthmatic reaction. Airways inflammation and remodelling Shedding of respiratory epithelium eosinophilia in the surrounding lung tissue is a characteristic of severe asthma exacebations
	Neutrophil	Accumulates at the site of infection. Phagocytic cell. Antimicrobial defence via production of ROS and RNS.	Involved in late phase asthmatic reaction Neutrophilia a marker of severe asthma
	Mast cells	Releases Histamine, prostaglandins and leukotrienes, TNFα, VEGF and other active substances. Pro-inflammatory.	Involved in early phase asthmatic reaction

Table 1-4: Key Cytokine Signalling Molecules Involved in Asthma Pathogenesis.

This table details only a few examples of the many cytokines involved in asthma.

Type	Cytokine	Role in Inflammation	Effect in asthma
Lymphokine	IL-4	T cell differentiation (T_H2 cells ↑ & T_H1 cells ↓) isotype switching of B cells cell adhesion eosinphil growth ↑	Induces allergic airways inflammation associated with increased BHR
	IL-5	eosinphil maturation apoptosis ↓ T_H2 cells ↑	Chemoattraction of eosinophils increased BHR
	IL-9	Mast cell protease expression Goblet cell hyperplasia	increased mucus production
	IL-13	similar to IL-4, except it does not drive T_H2 differentiation activates eosinophils IgE ↑	Airway eosinophilia increased BHR
Growth factor	TGFβ	Cell proliferation, differentiation and angiogenesis	Lung remodelling
Pro-inflammatory	GM-CSF	Granulocyte macrophage colony stimulating factor Regulates growth of several airways cells, macrophages, eosinophils, T cells, fibroblasts, endothelial, smooth muscle and epithelial cells	Chronic eosinophila and airways remodelling
	TNFα	Induced by IL-1, IL-13, GM-SCF, IFNγ Pro-inflammatory Induction of ICAM1 Up regulation of IL-1, 8 & MHC which attracts neutrophils	Amplifies asthmatic inflammation and may initiate chronic inflammation Increased neutrophils in sputum
Anti-inflammatory	IFNγ	Produced by T_H1 cells. Inhibits T_H2 cells. Amplifies release of THFα from macrophages	Promotes cell cytotoxic responses while Inhibiting allergic inflammation and IgE synthesis

1.2.5 Remodelling of Asthmatic Lung Tissue

Research into the pathogenesis of asthma has mainly focused on the immunological component of disease; however histological examination of lung tissue from asthmatics reveals extensive tissue remodelling and this is now recognised as an important aspect of this condition. Changes in lung tissue involve a range of processes. Goblet cell hyperplasia and metaplasia; these cells are responsible for hypersecretion of mucus during exacerbation. Myofibroblast cells increase in number and thicken the reticular basement membrane by collagen deposition resulting in oedema, increased airways obstruction and is correlated with airway responsiveness.[12] Remodelling in the smooth muscle layer (smooth muscle cell hyperplasia) and proliferation of airway blood vessels and nerves results in increased hypersensitivity of the airways and airway narrowing.[13]

Long-standing inflammation is considered to cause remodelling. But signs of remodelling have been detected in tissue biopsies from children up to 4 years before the onset of asthma symptoms.[14] It has been suggested that rather than being a consequence of airway inflammation, airway remodelling may run in parallel or is required for the establishment of persistent inflammation.[15]

Tissue remodelling is required for the normal development of the human lung in early life and starts *in utero.*[16] In asthmatic lungs it appears that an aberrant epithelium repair response may lead to the reactivation of these developmental pathways. A number of studies have provided evidence for an abnormal response to inflammation by the structural cells of the airways in asthmatics. Epithelial cells from asthmatics are more susceptible to damage from oxidative stress[17] and environmental oxidative stress agents (diesel particulates) may trigger epithelial activation in response to injury.[18] Increased clusters of sloughed epithelial cells (creola bodies) are observed in the sputum, especially during exacerbations and BAL fluids of asthmatics contain more epithelial cells.[12]

Injury to the epithelium causes the release of fibrogenic growth factors, such as TGF-β, which promotes the proliferation of underlying fibroblasts and differentiation into myofibroblasts that secrete interstitial collagens and growth factors (Endothelin 1 and VEGF). These growth factors are mitogens for smooth muscle and endothelial cells.[19]

Signals from the repairing epithelium are amplified deep into the submucosa layer; this communication between the epithelium and mesenchymal cells has been termed the Epithelium-Mesenchymal Trophic Unit (EMTU) and is similar to the remodelling process in embryonic lung development.[16] It is thought that the EMTU remains active or is re-activated in asthma, driving the remodelling process of asthmatic airways (see figure 1.2).[20]

The lung remodelling model suggests that in certain people lung epithelial cells are more susceptible to injury from inhaled environmental exposures and/or have impaired repair mechanisms that induce airways inflammation and remodelling, leading to the development of asthma.

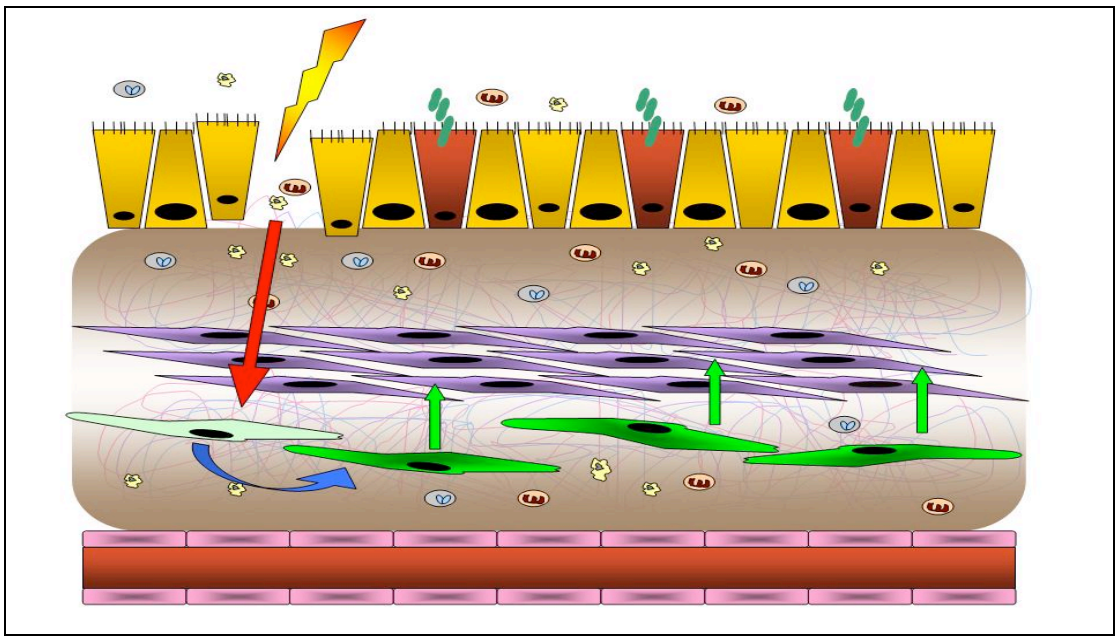


Figure 1-2: Lung Remodelling in Asthmatic Airways

Injury to the respiratory epithelium causes the release of fibrogenic growth factors (red arrow) and the signal is amplified into the deep submucosa, promoting proliferation of fibroblasts and differentiation into myofibroblasts (blue arrow). The myofibroblasts secret growth factors and collagen into the interstitial space (green arrow)

1.3 Aetiology of Asthma

The exact cause of asthma remains uncertain; several factors are known to influence the disease, genetic factors (e.g. family history of asthma and atopy), environmental factors (e.g. allergen exposure, viral infection, air pollution, occupational exposure) and life style choices (e.g. smoking and nutrition).

Asthma is a complex disease formed from the interaction of multiple genetic and environmental factors; these interactions are still not fully understood even though in recent years there have been remarkable advances in understanding the disease pathogenesis and the availability of improved treatments. Genetic and environmental factors can also influence the severity of the disease and the patient's response to medication. Certain factors only affect asthma when they form a gene-environment interaction, for example genetic factors may only have disease relevance in a particular environment. This makes elucidating the genetic component of asthma extremely challenging. Differences in asthma prevalence between populations provide evidence for certain environmental exposures (including life style choices) contributing to the development of this disease. Several theories have been proposed for the main environmental causes of asthma and are outlined below:

1.3.1 Environment

1.3.1.1 Hygiene Hypothesis: Role of Infection

Early childhood exposures to infection inhibit the tendency to develop allergic disease. Westernised lifestyles are protective against early life infections common in the developing world, but increase the risk of developing allergic disease. Numerous exposures have been found to be protective against asthma and atopy, living on a farm,[21,22] drinking unpasteurised milk,[21] living with animals[23,24] and having several siblings.[23,25,26] Abundant microorganism exposure in childhood seems to provide protection by switching the child's immunological memory to a T_H1 phenotype so that T_H2 cell responses do not dominate when exposed to allergens. The hygiene hypothesis cannot account for the 67-75% of adult asthmatics who are not atopic; therefore other causes of this disease must exist.[27]

1.3.1.2 Antibiotics

Antibiotic use in early life has been associated with the development of asthma;[28-31] although this is not conclusive as several studies have refuted those findings.[32-35] A recent prospective study of Canadian children (n = 13,116) found a 1.86 fold increased risk of developing asthma by age seven, when using broad spectrum antibiotics (cephlasporin) to treat non-respiratory tract infections.[28] Treatment with antibiotics can be thought of as an extension of the hygiene hypothesis, where exposure to microbial endotoxin in early life can be protective against the development of allergy and asthma. The microflora hypothesis proposes that external microbial exposure is not as important as the presence of commensal microbial flora in the gastro-intestinal (GI) tract, for the maturation of the mucosal immune system during infancy.[36] Several epidemiological studies have linked variation in GI flora and probiotic administration with less allergy and asthma.[37-39] Murine models of antibiotic administration reports altered flora, impaired barrier function, diminished T_H1 responses and allergic disease.[40, 41] It is still not clear whether antibiotics are responsible for the development of asthma and allergy as it is difficult to control for differences in antibiotics, environmental exposures, and population variability. Reverse causality in any association with risk of disease is difficult to exclude.

1.3.1.3 Vaccination

Early life vaccination programmes in westernised countries has been suggested as a cause for the increased prevalence of allergic disease in these nations by limiting the induction of a T_H1 immune response in infants by protecting against common childhood infections.[42] There is no recent evidence to support an increased risk of atopy or asthma from vaccination.[43, 44] In fact, a recent cross-sectional study that found a positive association between pertussis infection and atopic disorders in pertussis vaccinated children could not conclude whether this was a causal association or resulted from reverse causality.[45] Moderate to severe asthmatic individuals are prone to exacerbations caused by respiratory tract infections and influenza vaccination is considered as a preventative measure; even though there is uncertainty to the level of protection it affords.[46]

There is inconclusive evidence to suggest there is an increased risk of allergy or asthma with early life vaccination. But from an evidenced based viewpoint there is also no evidence to suggest that vaccination has absolutely no effect on susceptibility to atopy or asthma and may even be protective against respiratory tract infection induced exacerbations in later life.

1.3.1.4 Air Pollution

Ozone and particulate matter (diesel exhaust) have been associated with increased asthma exacerbations with increased levels of both pollutants in Mexico City[47] and cities in the US.[48, 49] There is a close correlation between daily asthma symptom scores and exposure to ozone and particulates;[50] in addition exposure to diesel exhaust increases the airways response to allergens in experimental airway challenges.[51] *In vitro* studies have shown that small particulate matter (≤10 μM in diameter; PM10) is responsible for the production and release of inflammatory cytokines by the respiratory tract epithelium as well as activating the $NF_{\kappa}B$ transcription factor.[52-53] It is hypothesised that oxidative stress may be a component of the mechanism of this response because many of the absorbed materials on the particles surfaces are direct oxidants (metals & quinones) that indirectly produce reactive oxidant species.[52]

1.3.1.5 Environmental Tobacco Smoke

Environmental Tobacco smoke exposure (ETS) has been found to increase both the prevalence and severity of asthma;[54] evidenced by the increased frequency of attack and resulting hospitalisations.[55,56] *In utero* smoke exposure (maternal smoking during pregnancy) is associated with wheezing and reduced lung function in children[57,58] and elevated levels of IgE in cord blood.[59] Passive smoking has been shown to induce oxidative damage *in vivo*[60] and markers of oxidative stress are elevated in children exposed to passive smoking.[61, 62] A small prospective asthma exacerbation study has found that levels of malondialdehyde (MDA) a biomarker for lipid peroxidation are elevated during asthma attacks and correlate with urinary cotinine levels.[63]

1.3.1.6 Maternal Paracetamol Intake

Shaheen et al.[64] proposed that as paracetamol can cross the placenta and the foetus has the ability to metabolise paracetamol (acetaminophen), the production of the toxic paracetamol metabolite, N acetyl-p-benzoquinoneimine (NAPQI), could deplete foetal antioxidant glutathione molecules resulting in oxidative stress and as a consequence altered lung development and increased susceptibility to asthma.[64] Furthermore, NAPQI clearance by conjugation to glutathione may be reduced in late gestation because of a decline in expression of foetal Glutathione S-Transferases (GSTs) enzymes (Fryer[65] and Cossar[66] cited in Shaheen et al, 2005[64]). This could cause damage to the foetal airway epithelium and increase the vulnerability of the airways to more profound oxidant damage after birth.[64] High doses of paracetamol can deplete pulmonary glutathione and cause necrosis of bronchiolar epithelium in mice;[67, 68] it is not known in humans whether therapeutic doses have the same effect. Oxidative stress can promote T_H2 responses in animals and these changes to immune cell differentiation could lead to increased production of IgE after birth.[64] Epidemiological studies have shown that paracetamol use in late pregnancy (20-32 weeks) is associated with an increased risk of early wheeze in the offspring; this exposure would account for 1% of the population prevalence of wheezing in early childhood.[69] At age 7, maternal intake of paracetamol in late pregnancy was positively associated with asthma, wheeze and IgE levels; the proportion of asthma attributable to paracetamol use in late pregnancy was 7% assuming a causal relationship.[64] Frequent use of paracetamol in adults may also contribute to asthma morbidity and rhinitis.[70] Subsequent epidemiological studies in several different countries also support an association with paracetamol intake by children and increased asthma risk.[71-75]

1.3.1.7 Oxidant/Antioxidant Balance

Changes in the oxidant/antioxidant balance are evident during an asthma attack, as antioxidant levels (Glutathione peroxidise) measured in saliva and serum are lower compared to levels after recovery.[76, 77] Reactive oxygen species (superoxide anion generation by leukocytes, lipid peroxidation products, total nitrates and nitrites, total protein carbonyls, total protein sulfhydrils) levels in blood are elevated in asthmatics compared to non-asthmatics.[78]

Elevated exhaled NO levels and the presence of inflammatory biomarkers in Exhaled Breath Condensate (EBC) can be used to predict the presence, control and severity of asthma.[79]

Epidemiological studies of diets low in antioxidants (Selenium, Vitamin A, C & E) have reported an increased risk for asthma and allergy,[80, 81] asthma severity,[82] deficits in lung function[83] and supplementation can reduce ozone mitigated lung function decline.[84]

A maternal diet deficient of vitamin D in combination with lack of exposure to sunlight has been postulated as the cause of the increase of asthma, allergies and other autoimmune diseases.[85] In recent years, changes in diet and lifestyle (time spent indoors) have ran in parallel with increased prevalence of T_H2 diseases (asthma, allergic rhinitis, eczema and food allergy) and also T_H1 type diseases in industrial countries at higher longitudes, that cannot be explained by the hygiene hypothesis alone. Vitamin D is a micro-nutrient that is not normally sourced from the diet, but is synthesised via exposure to sunlight. Vitamin D insufficiency leads to dysfunctional T cell regulation by a lack of down regulation of T_H1 and T_H2 inflammation;[85] and can potentially explain all of the epidemiological aspects of recent increase in allergic and autoimmune disease.

Selenium supplementation of a small group of tourism workers exposed to ETS showed a reduction in lipid peroxidation and levels of antioxidant enzyme levels (GSH peroxidise) after 60 days of supplementation.[86] Subsequent randomised placebo-controlled trials of selenium and vitamin E have shown no benefit to supplementation in adult asthma.[87, 88] Most intervention studies have been performed on a small sample size and the benefit of supplementation is neither conclusive nor corroborated by placebo control trials performed to date. In consideration of the supporting evidence for oxidative stress in asthma, therapeutic intervention (dietary or pharmacological) with antioxidants is a rational cost effective approach to combat the deleterious effects of excessive reactive oxygen species in lung tissue but nevertheless it remains controversial.

1.3.2 Genetics of Asthma

1.3.2.1 Heritability

It has been recognised that asthma results from both inherited (familial factors) and environmental exposures for over three hundred and fifty years (by Sennartus in 1650, cited in Wiener, 1936.[89] and Salter, 1868.[90, 91]). The existence of a genetic component to asthma and BHR has been confirmed in a plethora of more recent twin and family studies with current estimates of asthma heritability ranging from 50-90%.[92-95] The large range in asthma heritability estimates between studies are due to factors such as differences between the populations studies, age of onset of disease, the type of environment exposure and study design. In several twin studies, the consistent finding is that the non-shared environment between twin pairs is a more important risk factor than the shared environment and that the heritable asthma component may be higher in males (74%) than females (58%).[96, 97]

1.3.2.2 Candidate Asthma Genes

Candidate gene analysis has been extensively used in identifying the genetic contribution to asthma or atopy phenotypes. Association tests are used to determine the presence of a relationship between genetic variation in genes thought to be involved in disease pathogenesis and asthma phenotypes.

It is possible to use categorical phenotypes in the study of asthma genetics, but because of the heterogeneity of the disease, interpretation of an association can be problematic. Therefore, quantitative intermediate phenotypes such as BHR and FEV_1 are an attractive alternative as 'affected' individuals can be defined as those exceeding a certain threshold. The design of candidate gene studies will not be mentioned here as the concepts are discussed at length in section 1.6.2.2.

An extensive review by Ober & Hoffjan in 2006 found over 500 papers on asthma or atopy that had identified over 100 genes with an association to one or more phenotypes.[98] In total, 54 genes were associated in two to five independent studies and a further 25 genes were replicated in six or more populations. The remaining studies have failed or are as yet to be replicated.

Reassuringly, a number of positionally cloned asthma genes such as *GPRA, PHF11, HLA-G, DPP10 and ADAM33* (see section 1.3.2.4) have shown subsequent association in candidate gene studies performed in different populations.[98]

It has been reported that in 2006-2007 alone an extra 53 candidate asthma genes have been discovered.[99] The acceleration in discovery of new asthma genes in recent years can be accredited to the development of high throughput genotyping and advances in technology have facilitated the first genome-wide association studies of asthma.[100] A review of the literature identifies a selection of elite candidate asthma genes that have consistently shown association with asthma phenotypes in numerous populations. Although, the exact type of association may not be replicated in every study because of differences in phenotype definitions, environmental exposures or ethnicity. Even so, it is clear that these genes play a major role in the genetic contribution to asthma pathogenesis. This elite group contains the genes for *IL-4, IL-4RA, IL-13, ADRB2, TNFα, HLA-DRB1 & DQB1, FCER1B, CD14* and *ADAM33*.[98] Some of the elite genes will be examined in more detail in the next part of this section, along with a couple of more recently discovered candidate asthma genes.

IL-4 and IL-13 were chosen as candidate genes because they are key effector T_H2 cytokine molecules (see table 1.4) expressed by T-cells and their genetic location on 5q31-33 (see section 1.3.2.4). Several polymorphisms in the promoter (-1111C/T) and coding region (R130Q) of IL-13 are associated with asthma and/or atopy and related phenotypes.[101, 102] Resequencing of the gene in different ethnic groups revealed that R130Q is the only non-synonymous substitution present in all populations[103] and that this variant may act synergistically with other functional variants in the IL-4 receptor gene (IL-4RA).[104-106] In functional assays, IL-13 130Q showed more activity than the wild type, therefore it has been concluded that this polymorphism can enhance the pathways of allergic inflammation.[107]

Polymorphism in CD14 has shown strong interactions and risk of atopy and asthma, with environmental exposures (House dust mite, farms, animals, ETS and LPS).[108-111] CD14 is part of a membrane bound pattern recognition receptor on macrophage cells that binds bacterial ligands and activates innate antimicrobial defences; early life infection is thought to be protective against the development of atopy and asthma (section 1.3.1.1). Polymorphism in this gene that alters expression or function of the protein could be expected to module responses to microbial exposure. Initial reports of association in asthma atopy were inconsistent,[112] indicating that the level of environmental exposure may alter the effect of *CD14* polymorphism.[113] Further studies of the CD14 -159 C/T polymorphism provide evidence for an 'endotoxin switch' where there is a dose-dependant response to endotoxin for each genotype.[114-116]

The recent groundbreaking discovery of the role of the Filaggrin gene in maintaining epithelial barrier function in atopic dermatitis[117] has highlighted that skin defence systems are important in allergic pathways. Two stop mutations (R501X & 2282 del4) that result in complete loss of peptide production were consistently associated with atopic dermatitis,[117] and with asthma and asthma severity, but only in the presence of atopic dermatitis.[118, 119] A defective epidermal barrier allows allergens easier entry through the skin to initiate local and systemic allergy and predisposes the individual to allergic disease.

Findings to date confirm the efficacy of candidate gene association strategies for identifying common gene variation that can modulate either the risk of developing asthma and subsequent severity of disease; or identify gene-environmental interactions that alter with type, time, length and strength of exposure. Functional studies of the mechanism of genetic variation are essential to fully understand interactions in this disease.

1.3.2.3 Problem of Non-replication of Candidate Gene Associations in Asthma

As described earlier (section 1.3.2.2), hundreds of candidate gene association studies have been performed in asthma alone; with tens of thousands of studies performed in other diseases. A common problem in all genetic association studies and especially those in asthma is that of replicating an association. In the early years of genetic association studies many findings were hard to replicate because of poor study design (see section 1.4.2.2).

It has become apparent with improvements in study design that the genetic nature of asthma is difficult to study as many components of the disease have shared genetic loci with atopy and other allergic diseases. Asthmatic individuals from different populations could have a shared genetic component but with different associations of the same genetic marker because this is dependant on differences in the type, time and level of environmental exposures that interact with asthma. Therefore, it is prudent for asthma researchers to qualify replication primarily as the repeated association of the same gene region with an asthma phenotype in different populations. Rather than discounting a new association that could potentially replicate a previous association just because the opposite allele is associated when compared to the first study.

1.3.2.4 Positionally Cloned Asthma Genes

Positional cloning uses genetic linkage in affected families to search for susceptibility loci in an hypothesis independent manner; through this technique regions of the genome can be identified which contain genes with previously unknown function in the pathogenesis of disease. The principles of positional cloning are described in detail in section 1.4.2.1.

This technique has been highly successful in indentifying loci involved in single gene diseases (such as Huntingdon's and Cystic Fibrosis) because of the strong genetic signal transmitted to the next generation of affected individuals. However, positional cloning does not have the same sensitivity in complex diseases as many genes of moderate effect are likely to be involved and their role can be modified by environmental exposures, as in the case of asthma. Nevertheless, several regions of the genome are repeatedly identified in different populations as having an important contribution to the development of asthma. For example, atopy, and by extension asthma, have been consistently showing linkage to Chromosome 5q33, this locus contains the inflammatory interleukin gene cluster (IL-4, IL-5, IL9, IL-13) and to Chromosome 6p21 which contains HLA, TNFα and LTα genes. Conversely, regions have been linked to atopy only, such as the STAT6 transcription factor on Chromosome 12q13 with gene polymorphisms associated with increased IgE levels and eosinophil cell count. Therefore, this locus is likely to contribute to eosinophilia and changes in IgE levels seen in asthmatics rather than susceptibility.[120]

As atopy is a component of a majority of asthma it can be difficult to separate out the contribution to the linkage signal for the different phenotypes at shared genetic regions. There has been some success in identifying novel regions using intermediate asthma phenotypes such as bronchial hyperresponsiveness; where addition of BHR produces a stronger linkage signal than an asthma/atopy phenotype (i.e. asthma + IgE levels), as in the discovery of the *ADAM33* BHR susceptibility gene.[121] In theory, positional cloning is less sensitive for complex diseases, but in practice several important discoveries have still been made that further the understanding of asthma and highlight novel biological mechanisms.

Several novel asthma and BHR susceptibility genes such as *ADAM33, DPP10, PHF11, HLA-G* and *IRAK-M*[121-124] have been discovered to date and their probable functional roles are shown in table 1.5. It now remains for further candidate polymorphism, immunohistochemical and gene function analyses of these positional cloned asthma genes to elucidate the actual role they play in the susceptibility and subsequent progression of asthma

Table 1-5: Table of Positionally Cloned Asthma Genes

Gene Name (GeneID)	Chr	Associated Phenotypes	*Gene Product:* possible functional role in asthma	Associated Variation	Size of study	Population	Reference of Initial Study	Replication of association
DPP10 (57628)	2q14.1	Asthma Atopy (SPT) Asthma severity	*Dipeptidyl peptidase:* Potassium channel regulator with no detectable protease activity.Involved in Cytokine processing (especially in T-cells)	D2S308*3/*5 Multiple SNPs Haplotypes	244 families	Australian UK German	Allen et al[122]	YES[122]
CYFIP2 (26999)	5q33.3	Atopic Asthma Childhood Asthma	*Cytoplasmic fragile X mental retardation (FMRP) interaction protein 2:* May be involved in differentiation of T-cells	Multiple SNPs Haplotypes	155 families	Korean	Noguchi et al[125]	NO
HLA-G (3135)	6p21.3	Asthma BHR Atopy	*Class I, Histo-compatibility antigen-G:* Inhibits Th1-mediated inflammation and only the soluble form (HLA-G5) is expressed in asthmatic bronchial epithelial cells.	D6S1281 MOGc Multiple SNPs	129 families	Caucasian Hutterite Dutch	Nicolae et al[124]	YES[124]
GPRA (387129)	7q14.3	Asthma Total IgE BHR Specific IgE Childhood asthma	*G-Protein coupled receptor:* Bronchial epithelial and smooth muscle surface receptor. May modulate asthma by increasing expression levels in tissues and potential inhibitory effect of GPRA-A on cell growth[126]	Haplotypes	86 families & 103 trios	Finnish Canadian German Italian Chinese	Laitinen et al[127]	YES[127-131]
IRAK-M (11213)	12q14.3	Early onset persistent asthma	*Interleukin-1 receptor associated kinase 3:* Negative regulator of TOLL-like receptor/IL-1R pathways. Master regulator of NF-Kappaβ and inflammation.	Haplotypes	100 families	Sardinian Italian	Balaci et al[123]	YES[123]
PHF11 (51131)	13q14.3	Total IgE Asthma Asthma severity	*Zinc Finger transcription factor:* Possibly involved in Chromatin mediated transcription regulation. B-cell clonal expansion and regulation of immunoglobulin expression may operate through shared mechanisms at this locus.	Multiple SNPs Haplotypes	230 families	Australian UK European	Zhang et al[132]	YES[132, 133]
ORMDL3 (94103)	17q12-21.1	Child onset asthma	*ORM1 (Saccharomyces cerevisiae)-like 3:* Third member of a novel class of genes of unknown function that encode transmembrane proteins anchored in the endoplasmic reticulum. Involved in protein folding.	Multiple SNPs Increased gene expression in asthmatics	Genome wide Association & combined gene expression 994 cases 1243 controls	British German	Moffat et al[100]	YES[100]
ADAM33 (80332)	20p13	Asthma+BHR	*Metalloproteinase:* Involved in Airway remodelling by fibroblasts and smooth muscle hyperactivity	D20S482 Multiple SNPs Haplotypes	460 families	Caucasian (UK & US)	Van Eerdewegh et al[121]	YES[134-141].

1.4 Genetic Epidemiology

1.4.1 Population Genetics

Population genetics is the study of the genetic structure of populations and changes to the genetic composition of populations because of influence from natural selection, genetic drift, mutation, and gene migration. Population geneticists study the allele frequencies in a population by using mathematical models that can monitor any changes in genotype frequency of a population over time and question why this is occurring. This discipline linked the most important theories of evolution and genetics, Darwin's theory of natural selection and Mendelian inheritance together. Natural selection provides a mechanism for evolution to occur, but did not explain how 'fitness' was inherited. Mendel's theory of genetic inheritance provided the mechanism and the independent work of G.H. Hardy and W. Weinberg worked out the mathematical basis of population genetics in 1908.[142] Their formula predicts the expected genotype frequencies using allele frequencies in a diploid Mendelian population; this formula is known as Hardy-Weinberg Equilibrium (HWE).

1.4.1.1 Hardy-Weinberg Equilibrium

Hardy and Weinberg showed that if the population size is very large and random mating takes place, allele frequencies (of a biallelic locus) will remain the same and the genotype frequencies will follow equation 2:

Equation 2: Hardy-Weinberg Equilibrium Calculation

HWE: $p^2 + 2pq + q^2 = 1$

Genotypes: AA Aa aa

If the expected genotype frequencies match the observed frequencies then that population is said to be in Hardy-Weinberg Equilibrium. Should the population genotypes deviate from HWE then some selective pressure has been applied to that population and the following assumptions do not hold:

1. Genotypes have equal fitness (no selection for a particular genotype)
2. No mutation (germline) and migration (loss of one allele from population) affecting allele frequencies in population
3. Males and Females have similar genotype frequencies
4. Random mating
5. Infinite population size

In the laboratory, HWE calculations are routinely used to validate that genotyping techniques are not biasing genotype calls. A genotyping methodology that is inaccurate in assigning genotypes will increase or decrease the observed frequency of a particular genotype and the genotype frequencies will rapidly deviate from HWE. This basic quality control procedure is applied to each data set to ensure that a genetic association is not due to experimental error.

1.4.1.2 Genetic Association

Genetic association is the association of a particular genotype with a non-genetic trait (phenotype). A phenotype is an observable trait in an organism. Some phenotypes are purely genetic, while the majority are formed from interactions of the genotype with the environment. Phenotypic plasticity is the ability of an organism with a particular genotype to change its phenotype to adapt to an environmental change. This concept of genetic association and phenotypic plasticity is the fundamental principle for genetic studies into disease pathogenesis. Genetic polymorphisms are screened against disease phenotypes in an affected population. The association of certain genotypes with altered disease susceptibility, severity or the presence of disease modifying gene-environment interactions, indentifies at risk individuals and can further understanding into the pathogenesis of that disease.

1.4.1.3 Genetic Linkage

Genetic Linkage occurs when genetic markers are inherited together on a region of a chromosome, linkage between markers can be severed by recombination at meiosis. This term is also used to describe the linkage of a marker to an unknown disease contributing gene. Randomly spaced genetic markers are used to search for regions within or near to disease contributing genes in an affected family or population.

1.4.1.4 Linkage Disequilibrium

Linkage disequilibrium is the non-random association of adjacent polymorphisms on a single strand of DNA in a population; the allele of one polymorphism in an LD block can predict the allele of adjacent polymorphisms.[143] The size of the LD blocks depends on the recombination rate in that region and the number of generations since the first variant arose in an ancestral individual in that population. LD can be measured in a number of ways and common measures of LD used are D' and r^2.[144] D' is the deviation of observed haplotypes from the expected haplptypes, that has been normalised, by dividing it with the theoretical maximum for the observed allele frequencies.[143] r^2 is the squared correlation coefficient of linkage disquilibrium between pairs of loci.[145]

1.4.1.5 Haplotypes

A haplotype is a set of closely linked genetic markers present on one chromosome which tend to be inherited together in a population. Haplotypes can be as few as two genetic markers in length and can extend for thousands of basepairs until a recombination event occurs between two loci and breaks the link; thereby forming a new haplotype block. Table 1.6 displays all possible haplotype combinations that could occur in a two marker haplotype. When an individual is heterozygous at two or more loci it is impossible to 'phase' the haplotype on each chromatid.

Table 1-6: Two-SNP Haplotype Combinations[a].

	AA	**AT**	**TT**
GG	AG/AG	AG/TG	TG/TG
GC	AG/AC	AG/TC or AC/TG	TG/TC
CC	AC/AC	AC/TC	TC/TC

[a]*There are nine possible haplotype combinations from two biallelic markers. When an individual is heterozygous for both markers it is impossible to phase the haplotypes (i.e. allele 1 could be AG or AC) with out sequencing the region on one chromatid*

There are several mathematical techniques that can predict haplotypes by using the frequencies of genotypes in the studied population. A widely used technique is the Expectation-Maximisation (EM) algorithm that predicts the most probable haplotypes for an individual based on the likelihood of that haplotype occurring in the population.[146]

The International HapMap project's goal was to determine the common pattern of DNA sequence variation in the human genome in four different populations (Yoruba in Ibadan, Nigeria; Japanese in Tokyo, Japan; Han Chinese in Beijing, China and CEPH which are Utah residents with ancestry from northern and western Europe).[147] Researchers can freely access this information and select a panel of SNPs that act as markers (Haplotype tagging or htSNPs) for all known genetic variation in a region of linkage disequilibrium and test for association with disease, as an allele in a haplotype can mark for the unknown disease-causing functional variant.[147] Selecting htSNPs for genetic association studies is a statistically more powerful technique than using single SNPs, as htSNPs provide information on the entire genetic region in that particular linkage disequilibrium block; thereby reducing the number of statistical comparisons required to test for association.[148]

1.4.2 Genetic Association Techniques

1.4.2.1 Positional Cloning

Positional cloning involves family based co-inheritance analysis of genetic markers spread evenly throughout a genetic region of interest. The region of interest can be genome–wide or a region of a single chromosome. At the whole genome-level, positional cloning is a hypothesis-independent approach; it has the potential to identify genes and biological pathways that were not previously implicated in the pathogenesis of the disease. Micro-satellite polymorphisms are the marker of choice for positional cloning because they are highly polymorphic in the general population. All members of the family group are genotyped and the increased transmission of a particular allele to disease affected individuals indicates genetic linkage with the unknown disease gene. Linkage is broken down by recombination, therefore in a family group the resulting linkage region can be relatively large (approximately 1Mb in length) as only one recombination event is likely to occur between generations over this distance.[149]

A region of the genome identified as being linked to a disease phenotype will contain several to hundreds of potential disease contributing genes. Traditionally, DNA sequencing of the linkage region in individuals will identify new polymorphisms or mutations that may be closer to the gene of interest. These can be used as further markers to narrow the linkage region in subsequent analyses.

In the era of the Human Genome Project researchers do not have to sequence large regions of a chromosome to find extra markers as a simple database search will provide a summary of all known polymorphism previously identified in that region. The positional cloning technique as been successful in identifying high penetrance, high risk genes responsible for Mendelian inherited disorders, such as Huntingtons,[150] and cystic fibrosis.[151] This technique does not have the same statistical power to identify low penetrance, low-medium risk genes involved in complex diseases.[152, 153] Linkage analysis has less power than an equivalently sized case-control study as association is only tested in the probands. In complex diseases it is often difficult to define an appropriate disease phenotype.[154] There is variability in the severity of the disease in individuals and the age of disease onset may vary leading to individuals being inaccurately identified as unaffected at recruitment.

Also, diagnosed individuals may have apparently identical symptoms but they result from different aetiologies involving various biological pathways. Researchers may also have difficulty in choosing the best population to study and studies may be confounded by population stratification.

1.4.2.2 Candidate Gene Analysis

This technique is a hypothesis dependant approach because rather than utilising a random selection of evenly spaced genetic markers, genes are chosen on the basis of *a priori* hypothesis about their role in a disease. The selection of a candidate gene is based on the involvement of the gene product in biological processes relevant to the disease in question. Evidence for candidate gene selection can be drawn from a broad range of areas, for example: involvement in other diseases with phenotypic overlap, biological function, differential expression, affected tissues, cell type(s) involved and findings from animal models. Candidate gene analysis has been extensively utilised in the study of complex diseases states. For example, more than 500 studies have now examined the association of genetic variants in over 200 genes with atopic and allergic diseases alone.[98, 155, 156]

Case-Control studies are commonly used in the candidate approach. They are essentially a population based sample of affected and unaffected individuals. Any significant differences in genotype frequency found between the two groups are potentially associated with the disease or susceptibility phenotype. Findings should be interpreted with caution as genetic variants showing association with a disease are not necessary causal because of the phenomenon of linkage disequilibrium (LD). Case-control association studies have greater statistical power to detect genes of small effect than linkage based approaches.[154] This is highly relevant as it is assumed that polymorphisms of milder functional effect in multiple genes in the general population play a role in susceptibility to complex genetic disorders. The candidate gene approach has been criticised for non-replication of findings which may be due to a number of factors including poor study design, population stratification and different LD patterns between individuals of different ethnicities. Unfortunately, the genetic association approach can also be limited by under-powered studies and loose phenotype definitions.[157] Therefore it is important that due consideration is given to all aspects of genetic epidemiological study design to ensure that relevant conclusions can be drawn.

1.4.2.3 Combination Techniques

A hybrid of the two approaches described above is the selection of candidate genes based on their function and on their position within a genetic region previously linked to the disease. A good example of the 'positional candidate' approach is the identification of the *SOCS5* gene as a potential candidate gene responsible for linkage to BHR susceptibility on Chromosome 2p in 364 asthmatic families.[158] The region of linkage was 12.2Mb in size and contained approximately 75 genes; rather than continue with comprehensive analysis of the region, *SOCS5* (suppressor of cytokine signalling 5; GeneID: 9655, MIM: 607094) was selected as an interesting candidate because SOCS proteins are implicated in the control of the balance of Th1 and Th2 cells[159] and *SOCS5* is a specific inhibitor of IL-4 signalling.[160] The confirmation of *SOCS5* as an asthma susceptibility gene awaits further studies.

While this linkage study design is sound and a reasonable candidate gene choice has been made to reduce the considerable workload required to narrow a large linkage region of DNA to a single gene to test for association with the disease. It clearly identifies the difficulties facing researchers in identifying the causal gene(s) under a peak of linkage. A disadvantage of this technique is that without further high-resolution association analysis to narrow the genetic region carrying the risk allele, researchers can only pursue potential candidate genes based on limited current knowledge rather than directly identifying the causal gene by hypothesis independent approaches. With the advent of the first whole genome association studies in asthma this combined approach may become less prevalent.[100]

Another combination approach is to select potential candidate genes on the basis of their differential expression in diseased tissue.[161-163] An example of this is the gene encoding tenascin-C (*TNC*), located in a region of the genome previously linked with asthma[164] and whose mRNA expression was found to be up-regulated in bronchial epithelial cells exposed to a T_H2 cytokine environment.[165] A subsequent case-control study by Matsuda et al.[166] has shown association of *TNC* polymorphisms with asthma susceptibility. It will not be uncommon in the near future to combine whole genome expression micro-array and whole genome association data sets in order to elucidate the roles of specific polymorphisms on gene expression in complex disease states.[100, 167]

1.4.2.4 Whole Genome Association Studies

The advances brought about by the Human Genome Project and the International HapMap Project[168] in cataloguing and mapping the extent of human genetic variation and the availability of new genotyping methodologies providing high throughput with low cost per genotype call, has given rise to the possibility of *genome-wide association studies* in complex genetic diseases. In these studies of cases and controls, array-based technology (over 1 million genetic markers in the latest chips) is used to genotype single nucleotide polymorphisms (SNPs) or copy number variations (CNV) across the genome. Recent results from large disease genetics consortia have provided exciting gene discovery results for genetic susceptibility to chronic diseases.

For example, the Wellcome Trust Case Control Consortium (WTCC) studied ~2,000 cases with 7 chronic diseases and a shared set of ~3,000 controls, and discovered genes associated with bipolar disorder, coronary artery disease, Crohn's disease, rheumatoid arthritis, and type I and type II diabetes mellitus, with p values $< 5 \times 10^{-7}$.[169] High density SNP arrays scale up genetic association studies to the whole genome level and combine the advantages of association studies (greater statistical power for a given number of subjects and easier cohort recruitment) with the hypothesis independent whole genome approach. The examples provided by the WTCC study together with recent publications for obesity,[170] diabetes,[171, 172] and breast cancer[173, 174] demonstrate that this approach can be applied successfully to the identification of complex disease genes. The first whole genome association study of asthma has already identified a novel gene of unknown function; several SNPs were found to regulate ORMDL3 expression and contribute to the risk of childhood asthma.[100]

With the advent of genome on a chip technology, researchers can rapidly investigate family or case-control studies for up to millions of SNPs spread over the genome for any association. The high density arrays provide greater genome coverage ensuring that the maximum amount of linkage information can be retrieved and reduce the size of the critical linkage region for further analysis. By using cohorts with environmental exposure variables it may possible to define how genes interact with the environment to cause disease.

1.5 Role of Glutathione S-Transferases in Protection Against Oxidative Stress

Free radicals, reactive oxygen species (ROS) and reactive nitrogen species (NOS) are a by-product of normal metabolic pathways and are present in such abundance in the environment that they can cause damage to cellular molecules such as, lipid membranes, protein, amino acids and DNA. Therefore, cells have evolved mechanisms to limit and repair damage caused to molecules. Glutathione molecules and the action of Glutathione S-Transferase enzymes are one mechanism that provides protection from oxidative stress.

1.5.1 Role of Glutathione

Glutathione (GSH) is an abundant low molecular weight thiol synthesised from the three amino acids: glutamate, cysteine and glycine in a two step process by γ-glutamylcysteine (GCS) and GSH synthetase. 85-90% of all GSH molcules are found in the cytosol with the remaining in organelles. GSH is a potent scavenger of free radicals (such as Superoxide anion $O_2^{\bullet}$), ROS and NOS. Scavenging of electrophilic compounds by GSH can occur directly or via an enzymatic reaction such as the reduction of H_2O_2 by glutathione peroxidase by oxidising GSH as seen in equation 3.

Equation 3: Glutathione Redox Reactions.

Reduction of hydrogen peroxide by glutathione peroxidase

$$H_2O_2 + 2GSH \text{ ---> } GSSG + 2H_2O \quad (1)$$

$$GSSH + NADPH + H^+ \text{ ---> } 2GSH + NADP^+ \quad (2)$$

Presence of a cysteine amino acid allows for the GSH molecule to be readily oxidised by many electrophilic compounds into the pro-radical, $GS^{\bullet}$ which in turn binds to another $GS^{\bullet}$ forming Glutathione disulphide (GSSG). As cysteine is only found at low concentrations in intracellular pools it is considered to be the rate limiting amino acid for the synthesis of more GSH molecules a cell or tissue may require. It is known that increasing cysteine uptake or providing methionine supplementation (a precursor of cysteine) will increase GSH synthesis and reduce oxidative stress.

GSH and GSSG molecules form a redox couple which allows for the homeostasis of GSH levels in cells. Normal (0.5-10 mmol/L) cellular levels of GSH can be maintained by reduction of the oxidised glutathione back to GSH by the NAPDH dependant glutathione reductase.

If the normal ratio of [GSH]:[GSSH] (>10 under normal physiological conditions) cannot be maintained the demand for NADPH will increase and this will necessitate an increase in glucose metabolism via the pentose cycle. If NADPH cannot be replenished the scavenging rate of GSH will be less than the production of free radicals and oxidative damage will occur in those cells or tissues. The presence of GSH molecules at sufficient concentrations in the cytosol negatively regulates the production of more GSH by GCS. Shifting the GSH:GSSG redox couple to the oxidising state activates key signalling pathways (PK-B, NF-κB and MAPK) that reduces cell proliferation and increases cell apoptosis. Therefore oxidative stress plays a key role in the pathogenesis of many diseases such as cancer, inflammation, myocardial infarction and diabetes.

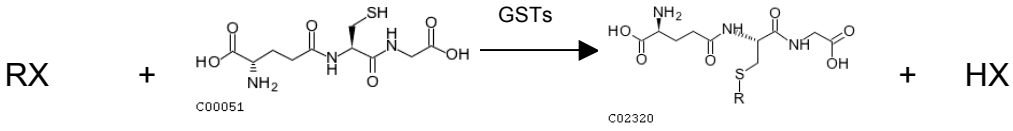
1.5.2 Role of Glutathione S-Transferases

Glutathione-S-transferases (GSTs) (EC 2.5.1.18) are phase II detoxification enzymes that are freely found in the cytosol and catalyse the conjugation of electrophilic compounds such as xenobiotics (activated by Phase I detoxification enzymes, such as Cytochrome P450) to Glutathione (GSH), UDP-glucuronic acid or glycine; producing stable more soluble compounds that can be excreted or compartmentalised by phase III enzymes.[175] The hydrophilic GSH molecule stops the conjugated hydrophobic compound from freely diffusing across plasma membranes and causing cell damage via lipid peroxidation; the GSH-conjugate has to be actively pumped out of the cell by a transmembrane ATPase such as the GS-X pump.[176]

GSTs catalyse the conjugation of glutathione to many substrates including free radicals, xenobiotics and physiological metabolites producing mercapturates (ultimately Mercapturic acid) that can be excreted from the cell and tissues (see equation 4 and figure 1.3). This mechanism produces a net loss of GSH as it is not recycled as in the NAPDH-dependent redox couple catalysed by GSH reductase.

Equation 4: Glutathione S-Transferases

GSTs catalyse the conjugation of reduced glutathione to free radicals, xenobiotic and physiological metabolites. Utilising the thiol group from the GSH molecule GSTs produce a variety of thiolesters. GSTs perform the initial step in the synthesis mercapturic acid. Key: R = aliphatic, aromatic or heterocyclic group; X = sulphate, nitrile or halide group and HX = Halo acid.



Glutathione R-S-Glutathione

Chemical Formulas from KEGG database

If a tissue is exposed to a high concentration or mixtures of xenobiotics or eletrophillic compounds, such as cigarette smoke, this can result in an imbalance in the [GSH]:[GSSG] ratio. The action of GST enzymes could place the cell in greater risk from oxidative stress by free radicals and other eletrophilic compounds if the GSH sink becomes depleted. GSH homeostasis plays an important role in the maintenance of the epithelial barrier forming the lung airspace (cited in Rahman & MacNee[177]). Decreased levels of GSH in lung tissue have been shown to lead to the loss of this barrier and increased permeability into the surrounding tissue.[178]

GSTs have other functions besides GSH conjugation. The Alpha class GST molecules can isomerise steroids and provide direct protection against oxidative stress by its selenium-independent GSH peroxidase activity.[179] The Pi class molecules provide further defence by inhibiting Jun N-terminal Kinase, via its ligandin function and therefore prevent hydrogen peroxide induced cell death.[180]

All GSTs are able to non-catalytically bind a wide range of endogenous and exogenous ligands. The theta class GST have been shown to be distinct from the Alpha/Mu and Pi classes; they only have an overall 7% sequence identity and have a catalytically essential serine rather than a tyrosine in the N-terminus.[181] The theta GST enzymes are catalytically more efficient than the other classes who appear to have evolved to remove electrophilic compounds by product retention over catalytic action.[182]

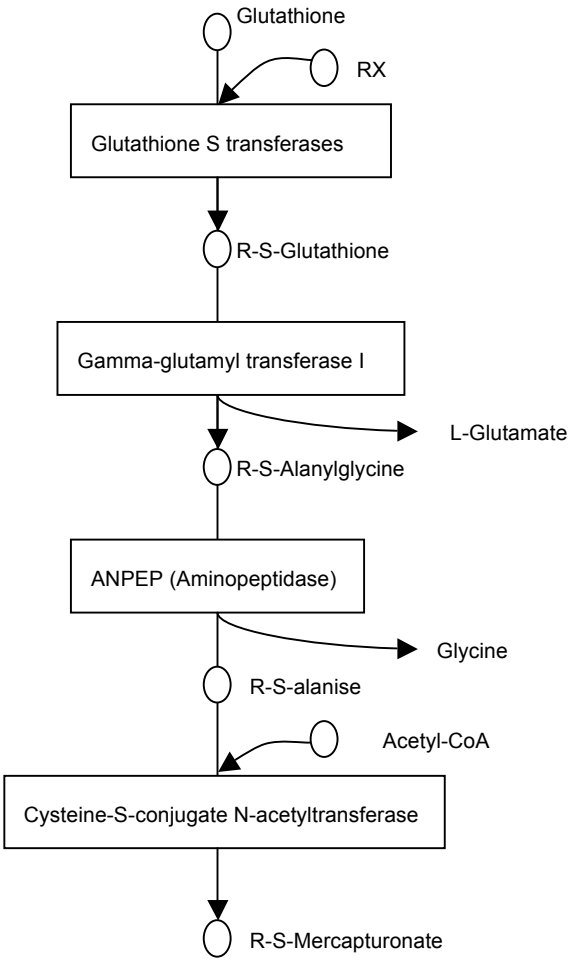


Figure 1-3: Glutathione Metabolism Downstream of Glutathione S-Transferases.

The actions of GSTs can produce a net loss of glutathione as it is excreted from the cell as mercapturic acid

1.5.3 Genetics of GSTs

1.5.3.1 Human GST Superfamily

The Glutathione-S-Transferase (GST) gene family in humans is grouped into the 'canonical' classes: Alpha, Mu, Pi, Sigma, Theta, Zeta and Omega based on the similarity of primary and tertiary protein structures, substrate/inhibitor specificity and immunological identity. The four most prevalent soluble dimeric cytosolic protein classes Alpha (GSA), Mu (GSTM) Pi (GSTP) and Theta (GSTT) have been widely studied. Kappa class GSTs are also found in humans and localised to mitochondria and peroxisomes;[183] but their physiological substrates are not known. Forty percent amino-acid sequence similarity is required to become a member of a class of GST enzyme[184] and the Alpha, Mu and Pi protein structure alters at two sites. The sub unit interface is conserved within each class, this allows for the formation of heterodimers only within classes. The second difference is at the active site, Alpha and Mu enzymes have partially occluded sites compared to Pi[185] and the electrophile binding is always hydrophobic. These structural differences determine substrate specificity between the classes of GSTs. Theta GSTs on the other hand have the active site tyrosine replaced with a serine residue that can still activate the glutathione molecule. Theta GST enzymes have a higher K_m value for glutathione than the other classes and act more like a catalyst than the other classes (Alpha, Mu, Pi) which lose catalytic efficiency but act as a sink for their own products.[186]

1.5.3.2 Expression of GST Enzymes

GST enzymes are widely expressed in most human tissues. *GSTP1* is highly expressed in the lung, placenta, brain and oesophagus.[187-189] In embryonic and early fetal tissues glutathione and GSTs are present in high amounts with the *GSTP1* isoform the most predominantly expressed in all organs, with highest concentrations in the small intestine, kidney and lung.[188] GSTA is the predominant isoform of GST in liver and GSTM and GSTT are present in relatively low concentrations in numerous organs including the lung.[188, 190]

1.5.3.3 Polymorphism of GST Genes

The human GST gene superfamily consists of 21 genes that are highly polymorphic; hundreds of polymorphic sites have been indentified in the genome,

the majority of which are of unknown function.[191] The table below details common functional polymorphisms that alter enzyme activity for *GSTM1*, *GSTO2*, *GSTP1* and *GSTT1* (see table 1.7). Also, there are numerous promoter polymorphisms that could affect the regulation of gene expression though altering transcription factor binding sites, enhancer and repressor sequence motifs. However, these polymorphisms have not been well characterised and the effect of each of these polymorphisms on gene function is not fully understood.[192, 193]

Common copy number variation (full gene deletions are present at ~50% for *GSTM1* and ~20% for *GSTT1* in Caucasian populations) have a dose effect on *GSTM1* and *GSTT1* protein levels, affecting that individual's ability to efficiently detoxify compounds and provide protection from oxidative stress.[194, 195]

Table 1-7: Glutathione S-Transferase Gene Polymorphisms

Gene (Gene ID)	Chromosome	Alleles	Nucleotide Change	Amino Acid Change	Functional Consequence
GSTM1 (2944)	1p13.3	M1*A	519G	Lys173	Reference[196]
		M1*B	519C	Asn173	No change[197]
		M1*0	Gene Deletion	No protein	No Activity[198]
		M1*Ax2	Gene Duplication	Overexpression	Increased Activity[199]
GSTO2 (119391)	10q25.1	O2*A†	1052A	Asn142	Reference[200]
		O2*B†	1052G	Asp142	No change[200]
		O2*C†	del at residue 155	loss of residue 155	loss of heat stability & increased activity.[200]
GSTP1 (2950)	11q13	P1*A	313A,341C, 555C	Ile105,Ala114,Ser185	Reference[201, 202]
		P1*B	313G,341C, 555T	Val105,Ala114,Ser185	Substrate dependant[201, 202]
		P1*C	313G,341T, 555T	Val105,Val114,Ser185	Substrate dependant[201, 202]
		P1*D	313A, 341T	Ile105,Val114	No change[201, 202]
GSTT1 (2952)	22q11.23	T1*A	310A	Thr104	Reference[203]
		T1*B	310C	Pro104	Decreased Activity[203]
		T1*0	Gene Deletion	No Protein	No Activity[203]

† These are provisional allele designations as *GSTO2 has not been fully characterised.*

1.5.3.4 Glutathione S-Transferases and Disease

Due to the pleiotropic nature of the GST gene family there have been many diseases with varying aetiologies associated with polymorphic variants in the canonical GST genes. Cancers of the head, neck, lung and bowel have been associated with *GSTM1*, *GSTP1*, and *GSTT1* polymorphism.[191] *GSTM1* deficiency may be a risk factor for cancer by providing increased sensitivity to particular chemical carcinogens.[204] Single or combined GST gene deletions increase susceptibility to lung cancer in smokers,[205] and in non-smokers exposed to household environmental tobacco smoke (ETS).[206] A recent meta-analysis of lung cancer using data from 130 studies, involving a total of 23,452 lung cancer cases and 30,397 controls has shown a weak positive association (Relative Risk (RR) = 1.18; 95% CI = 1.14-1.23) and (RR = 1.09; 95% CI = 1.02-1.16) with *GSTT1* and *GSTM1* null alleles.[207]

There is also an increased cardiovascular disease risk in smokers who a have GST deficiency.[208] *GSTP1* polymorphism is a risk factor for increased susceptibility to Parkinson's disease when exposed to pesticide or herbicides.[209] The age of onset of type 1 diabetes is associated with the presence of a *GSTM1* gene, although the mechanism is not clear it could involve cellular GSH depletion or a decrease in inflammatory response.[209, 210] Polymorphism in *GSTO2* have been associated with ovarian and gastric cancer and some studies have examined the age of onset of Parkinson's disease.[211-214]

1.5.3.5 Glutathione S-Transferases and Drug Response

GSTM1 and *GSTT1* gene deletion polymorphisms have consequences for the detoxification of xenobiotics.[215] This impacts the efficacy of drug treatments in the general population and outlines the clinical relevance of a personalised medicine based approach in the post genome era, namely the introduction of phamacogenetics.[216] As an example, GST polymorphisms have been shown to be indicative of response to treatment and event-free survival in acute paediatric acute lymphoblastic leukaemia and in ovarian cancer.[217] It is thought that patients with GST gene deletions respond better because of their inability to metabolise and inactivate anti-cancer treatment.[218]

1.6 Developmental Origins of Health and Disease

Epidemiological studies worldwide have established a link between low birth weight and increased risk from chronic diseases in adulthood.[219, 220] Low birth weight is not the cause, but rather it is a surrogate indicator for impaired growth during pregnancy.[221] The associations are not down to genetics alone, but rather the interaction between the early-life environment and genetic influences that predispose individuals to susceptibility to disease in adult life. These observations have lead to the proposal of several hypotheses:

1.6.1 Thrifty Phenotype or Barker's Fetal Origins of Disease Hypothesis

This hypothesis proposes that *in utero* under-nourishment results in permanent detrimental changes in the foetus leading to the development of chronic disease in later life.[222] The Dutch hunger winter famine of 1944-45 provided the initial supporting evidence for this hypothesis in humans.[223-225] The enforced famine conditions in Nazi controlled Holland during World War II provided epidemiological evidence that under-nutrition in early pregnancy resulted in an increased incidence coronary heart disease, body mass index, and glucose intolerance in the offspring in later life. Famine exposure in early or mid-gestation was also associated with decreased adult lung function (FEV_1, FVC, FEV_1/FVC) and increased risk of COPD.[226] Subsequent studies have shown that peri-natal nutritional deficiencies predispose adult offspring to chronic diseases such as, metabolic syndrome, obesity, cardio-vascular disease (CVD), hypertension and type II diabetes.(reviewed in McMillen et al, 2005)[227]

1.6.2 Predictive Adaptive Response (PAR) Hypothesis

This hypothesis was proposed in 2004[228] as an extension to the Thrifty phenotype, and theorises that the foetus makes adaptations based on the predicted postnatal environment. Prenatal nutrition is believed to be the primary predictor of the post-natal environment available to the foetus.

When the PAR is appropriate the offspring's phenotype is normal, when there is a 'mis-match' (either a high or low pre-natal nutritional environment) during critical developmental periods and the adult environment disease will develop in the offspring. The timings of these windows of developmental plasticity will vary between organs, so there is the potential for the plasticity phase to bridge across in to early post-natal development and environments. This hypothesis extends the Barker hypothesis[222] to include the effect of over-nutrition on subsequent susceptibility to adult disease.

1.6.3 Nutritional Programming

Nutritional programming *in utero* can affect the incidence and severity of disease in adults. The hypothesis states that:

'A relationship exists between the nutritional environment during critical windows of developmental plasticity and offspring disease in adult life'.

This hypothesis is derived from the two prior hypotheses:

- Thrifty phenotype or Barker foetal origins of disease.[229]
- Predicted Adaptive Response (PAR) hypothesis.[228]

This hypothesis seeks to establish the link between nutritional environment during critical windows of developmental plasticity and disease in the offspring in later life. Understanding nutritional programming during critical developmental periods will have enormous potential for improving the health of future generations.

1.6.4 Developmental Plasticity

The definition of developmental plasticity is the ability of a single genotype to produce more than one alternative form of structure, or behaviour in response to environmental conditions.[230] Programming mechanisms can act at several critical windows of developmental plasticity that occur *in utero* and post-natal life to adapt the offspring's phenotype (see Table 1.8). The nature of the adaptation, is in part, dependant on the type of the environmental exposure.

Adaptations made to the offspring's phenotype during these windows of development may have a life long impact such as aiding survival to reproductive age at the expense of a healthy old age. The duration of the windows of developmental plasticity is time-limited, because when organogenesis is completed, plasticity is no longer possible.

Table 1-8: Windows of Developmental Plasticity.

Adapted from Mcmillen et al, 2005.[227]

Critical windows of developmental plasticity	Programming mechanisms
Primordial germ cell ↓ Mature germ cell ↓ Ovulation ↓ Fertilisation: Zygote	← Epigenetic regulation of gene expression
↓ Blastocyst: Inner cell mass & Trophectoderm ↓ Differentiation	← Cell cycle regulation: proliferation & apoptosis
↓ Organogenesis ↓ Fetal growth & development of integrated system responses	← Adaptive cardiovascular, metabolic, neuroendocrine responses: hormonal programming
↓ Postnatal growth & development of integrated system responses ↓ Puberty ↓ Pregnancy	← Placental development

1.6.5 Modelling the DOHAD Hypothesis

All of the hypotheses in the developmental origins of adult disease concept have been established on the basis of the outcomes of epidemiological studies or *in utero* undernourishment in humans. In order to test these hypotheses experimentally and establish a greater understanding of the mechanisms involved it is necessary to use animal models of nutritional restriction. Numerous models have been developed in several species; all involve some form of change to the offspring's nutritional intake during and/or after gestation.

1.6.5.1 Animal Models

An extensive review of the literature on the different animal models utilised in the study of the DoHaD hypothesis is beyond the scope of this introduction. Nevertheless, this section aims to provide an overview of the contribution of animal models in furthering the understanding the developmental origins of cardiovascular disease, hypertension, obesity and type 2 diabetes.

1.6.5.1.1 Cardiovascular Disease and Hypertension

One of the first models to show experimental evidence for the DoHaD hypothesis was a study using guinea pigs. In this model, one uterine horn was ligated in pregnancy to restrict fetal growth; resulting in severely restricted fetal growth and hypertension in the offspring.[231] Since then nearly all animal models (mainly rat and sheep) of intrauterine growth restriction have shown similar programming of hypertension.[232-234] Low protein diets have been extensively used in study of hypertension, these models do not seem to display an effect on cardiac function (i.e. cardiac hypertrophy), but do appear to affect endothelial-dependant vasodilation and kidney development (changes to nephron, glomeruli number & kidney/body weight ratios).[232, 235-237] It is thought that protein under-nutrition is not directly responsible for programming of cardiovascular function *per se*.[238] Rather, that the imbalance of amino acids and other nutrients in this diet is disrupting the correct pattern of DNA methylation, through increased homocyteinemia, during critical periods of organogenesis and vasculogenesis.[238]

Meanwhile, chronic hypoxia (10.5% O_2) of pregnant animals can induce cardiomyocyte hyperplasia, hypertrophy and apoptosis in rat and sheep offspring.[239] Prenatal exposure to dexamethasone (synthetic glucocorticoid

steroid) increases the cardiac output of sheep at 40 months; these animals also show ventricular hypertrophy, reduced cardiac reserve and increased collagen content in adult life.[240, 241]

1.6.5.1.2 Obesity

Severely restricted maternal diets during pregnancy in animals have resulted in growth restriction at birth with later hyperphagia (increase in appetite and food consumption) that can be amplified by hypercaloric diets post weaning.[242] Maternal under-nutrition has also been associated with decreased voluntary locomotor activity in offspring and changes to the programming of appetite and behaviour that could acerbate any effects of metabolic programming in later adult life.[243, 244] In rodents, if there is a reduction in litter size early in life, the remaining pups show an increased growth rate and remain over-weight.[245] Cross-fostering experiments have shown that 'catch up' pups (experienced maternal protein restriction and born small) gain weight more than the control pups; whereas normal birth weight animals grow slowly and show no additional weight gain on a highly palatable diet compared to control animals on a standard chow diet.[246] Leptin is an adipose derived hormone that controls appetite and metabolism.[247] Administration of leptin to pregnant rats late in gestation makes their male offspring less susceptible to high-fat diet induced weight gain and insulin resistance.[248] It is thought that increased maternal leptin programs the ability of the offspring to respond better to a high fat diet in the post-natal environment. This is an example of a predictive adaptive response (see section 1.6.2).[227]

1.6.5.1.3 Type 2 Diabetes

Numerous animal models of intra-uterine growth restriction (IUGR) have been extensively used in the study of pancreatic development. These models have utilised uterine artery ligation, restricted maternal nutrition or a low protein diet.[249-251] All of the different IUGR challenges have resulted in changes to foetal pancreatic development and interaction with the pattern of pancreatic development in post-natal life.(Reviewed in McMillen et al, 2005[227]) The low protein model results in normal foetal glycemia, but diabetic phenotypes can be induced through a decrease in specific amino acids (taurine), impacting the proliferation/apoptosis of β-cells and insulin response.[252] Global energy reduction predominantly affects pancreatic cellular neogenesis.[249]

Overall, as reviewed in McMillen et al., 2005[227] IUGR results in a reduction of β-cell mass and secretory capacity in the pancreas; decreased glucose uptake and increased gluconeogenesis in the liver. In skeletal muscle, IUGR has been found to increased lipid oxidation and decrease insulin sensitivity and in adipocytes, lipolysis in inhibited by insulin. These effects of IUGR result in decreased pancreatic functional capacity, glucose intolerance and insulin resistance. Thus, IUGR can program all of the hallmark pathways underlying type 2 diabetes.

1.6.6 Mechanisms of Developmental Plasticity

In order for the offspring to make a predictive adaptive response for the expected post-natal environment a dynamic mechanism must exist to 'program' the offspring's phenotype during windows of plasticity in gestation. It is proposed that the offspring achieves this via epigenetic regulation of cellular processes. The next section discusses the mechanisms of epigenetic regulation in more detail.

1.6.6.1 Epigenetic Regulation

The word 'Epigenetic' literately means 'on top of' genetics (Greek 'Epi' prefix). So 'epigenetic regulation' describes regulatory mechanisms that control the expression of genes without altering the DNA sequence. As a general rule, organisms (or cells) are exposed to environments that change all the time, in order to survive they must adapt to the changing environment. The common misconception (especially in the post genome era) is to think cells as fixed (in relation to gene expression) in their metabolism and cellular processes. However, every cell has to differentiate during embryogenesis, respond to injury and changes in the micro-environment. Through switching genes on and off cells can survive or undergo programmed cell death. The switching of genes on or off is regulated by epigenetic processes.

In humans, there are epigenetic mechanisms that control how the DNA is packaged in the nucleus of the cell and which genes are accessible to transcription machinery and post-transcription there is a mechanism for inhibiting translation of the mRNA molecules into protein, microRNAs. Epigenetic regulation mechanisms are described in detail in the following sections:

1.6.6.1.1 Histone Modification

In order to form chromatin 146 basepairs of DNA is wrapped twice around an octamer of proteins containing two copies of H2A, H2B, H3 and H4 histones, this arrangement forms a nucleosome and is repeated along the DNA in the cell nucleus. Linker histones (H1) and other proteins interact with the nucleosomes and package them in to higher-order chromatin. The N-terminal tail of histones can carry post-translational modifications that alter DNA-histone interactions within and between nucleosomes. These modifications alter higher–order chromatin structures by collaborating with cellular processes such as transcription, DNA repair and replication and cell cycle progression.[253] The types of histone modification are described below:

- Histone Acetylation: Acetylation of lysine residues on the four core histones by HATs (Histone acetyltransferases) change histone-DNA interactions and create a more open chromatin structure, facilitating activation of gene expression as transcriptional machinery has access to genes at that site. Open chromatin structures can be closed by histone deacetylation which is performed by HDACs (histone deacetyltransferases) enzymes. H3 acetylation is associated with genome stability (DNA repair) and H4 acetylation with transcription.[254, 255]
- Histone Methylation: Histones H3 and H4 can be methylated on lysine and arginine residues by histone methyltransferases creating an open or closed chromatin structure.[256] Histone methyltransferases appear to be more specific about their targets than HATs and it is likely that each enzyme can regulate different genes or cellular processes.
- Histone Phosphorylation: Histones are often phosphorylated at specific sites during cell division and can be induced by DNA-damage signalling pathways. Phosphorylation is catalysed by several kinase enzymes and appears to have a conserved role for compaction of chromatin during germ cell development.(reviewed by Wendt et al, 2006)[257]
- Histone Ubiquitination: Udiquitination of histone residues (H2A & H2B) correlates with active transcription. This relationship is more complex than a straight positive correlation with transcription as histone ubiqutination of

histone H2B can signal for the methylation of other histones through a "histone crosstalk" pathway.[258]

1.6.6.1.2 DNA Methylation

DNA methylation negatively regulates gene transcription by stopping transcription factor binding or by the modification of chromatin structure. DNA methylation is important in normal development and cellular differentiation by altering the gene expression pattern. DNA methylation is the addition of a methyl group to the 5th carbon atom of a cytosine pyrimidine ring (see figure 1.4). Methylated Cytosine bases are grouped together forming CpG islands in the promoter of methylated genes. Methylation patterns in imprinted genes are inherited from either the maternal or paternal germ line.[259]

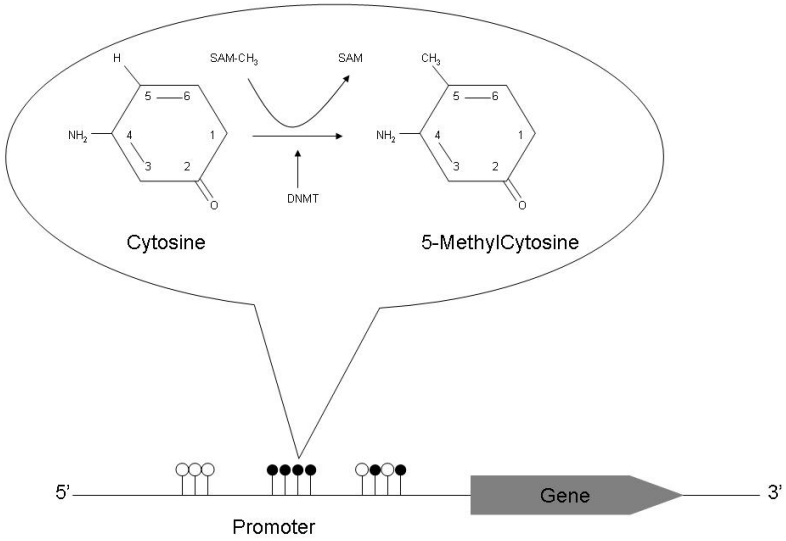


Figure 1-4: DNA Methylation.

The maintenance of methylated cytosine bases is due to the action of DNA methyltransferases (DNMT) which add the universal methyl donor S-adensyl-L-methionine (SAM-CH_3) to cytosine.

1.7 Aims of Research

The pathogenesis of asthma and COPD is complex and heterogeneous in nature. Much of the research into asthma has focussed on the immunological component of the disease. While this research had lead to a greater understanding and generation of effective treatment regimes, there has been no cure and the origins of the disease is still not fully understood. Recent evidence from studies of the developmental origins of chronic adulthood disease (e.g. cardiovascular disease) has shown that exposures during *in utero* development and early life are important factors in the manifestations of disease in offspring.[228] Exposure to excessive levels of oxidative stress and impaired nutritional status during pregnancy may be detrimental to or alter the trajectory of lung development predisposing offspring to a greater risk of respiratory disease, namely asthma in childhood and the development of COPD in later life.[260, 261]

The first part of this thesis examines the effect of genetic polymorphism in important antioxidant genes in asthmatic children when exposed to parental tobacco smoke in early life (Chapter 4). For the genetic association study, a copy number variation real-time PCR genotyping methodology was developed and is described in chapter 3.

The second part of this thesis investigates whether an established animal model of maternal protein-restriction (Southampton Protein-Restriction Rat Model[262]) developed for the study of the developmental origins of hypertension is suitable for the study of the developmental origins of respiratory disease. Gene expression analysis was performed on rat lung tissue to examine whether maternal protein restriction during pregnancy results in persistent changes to the expression profile of key genes involved in metabolism, lung growth, lung development and antioxidant protection (Chapter 5). MicroRNAs have an important role in lung development by modulating developmental timing, cell fate determination, apoptosis and growth but little is known about microRNAs expression in response to protein-restriction *in utero*. Chapter 6 investigates whether rat lung microRNA expression is altered by maternal protein-restriction during pregnancy.

2 Chapter 2: General Materials and Methods

2.1 Materials

All reagents were stored at room temperature, unless otherwise stated. All chemicals were molecular Biology or Biotechnology grade.

Table 2-1: Reagents

Reagent	Cat no:	Supplier
Absolute Ethanol	51976	Sigma
DEPC treated water	N/A	N/A
2xQPCR Mastermix (PD)	Mastermix-R	Primerdesign ltd
2xQPCR Mastermix for probes (E)	RT-QP2X-03	Eurogentec, SA
PCR grade water	04 909 631 001	Roche diagnostics, UK
Iso-propanol	I-0398	Sigma
Sodium dodecyl sulphate	S/P530NC/48	Fisher Biochemicals
Sodium Hydroxide	S-8045	Sigma
Guanidine hydrochloride	G-3272	Sigma
Sodium Citrate	S-4641	Sigma
Chloroform	BP1145-1	Fisher BioChemicals
TRIzol Reagent	15596-026	Invitrogen
Agarose	EP-0010-05	Eurogentec, SA
Smartladder	MW-1700-10	Eurogentec, SA
50bp Ladder	N3236	New England Biolabs (NEB)
Blue-Orange Loading Buffer	G109A	Promega, UK
Ethidium Bromide	E-1510	Sigma
Diethyl Pyrocarbonate	D-5758	Sigma
10 x TBE Buffer	EC-860	National Diagnostics

Table 2-2: Enzymes and Buffers

Enzyme	Cat no:	Supplier
AmpliTaq Gold	4311816	Applied Biosystems, Foster City, CA, USA

Table 2-3: TaqMan Probe Assays for qPCR

Gene	Ref sequence	Forward Primer sequence (5'-3')	Reverse Primer sequence (5'-3')	Probe sequence (5'-3')	Oligo Conc[n]
ALBUMIN	M12523	CTGTCATCTCTTGTGGGCTGTAA	GGCATGACAGGTTTTGCAATATT	VIC-CATCGTCTAGGCTTAAGAG-MGB-NFQ	*GSTM1*: 900 / 600 / 150 nM *GSTT1*: 450 / 300 / 150 nM
GSTM1	X68676	GACTCTTGCATCCTGCACACA	GGAAAGCACTTGGAGGATGAAT	6FAM-TGGTCTTAAGTCCCTGGTAC-MGB-NFQ	600 / 900 / 150 nM
GSTT1	af240786	CAGGTGAACCCCCTCAAGAA	GTCCCAGTTACCTCTCCGTCAA	6FAM-CAGCCTTGAAGGACGG-MGB-NFQ	675 / 450 / 150 nM
Nr3c1 (*GR*)	NM_012576	TGCATGTATGACCAATGTAAACACA	CCTTCCTTAGGAACTGAGGAGAGAA	TCCTCTGAATTACAAAGATTGCAGGTATCC TATGAAGA	900 / 900 / 200 nM
Actb	NM_031144	CGTGAAAAGATGACCCAGATCA	CACAGCCTGGATGGCTACGT	TTTGAGACCTTCAACACCCCAGCCAT	900 / 900 / 200 nM
B2m	NM_012512	ACTCTGAAGGAGCCCAAAACC	TCCAGATGATTCAGAGCTCCATAG	CACCTGGGACCGAGACATGTAATCAAGC	900 / 900 / 200 nM
Ubc	NM_017314	CGTACCTTTCTCACCACAGTATCTAGA	GAAAACTAAGACACCTCCCCATCA	AGAGCCCTTCTTGTGCTTGTTCTTGGGT	900 / 900 / 200 nM

Table 2-4: UPL Probe or miRNA Assays for qPCR

Gene	Ref sequence	Forward Primer sequence (5'-3')	Reverse Primer sequence (5'-3')	Allele 1 Probe sequence (5'-3') or part no.	Oligo Conc[n]
Adam33	XM_001079402.1	TAGTGGCCCTGCACAGTCT	TGGTAGTAGCACCAGGCTAGG	UPL #74 CAT NO: 04688970001	900 / 900 / 200 nM
Pcdh1	NM_225997.4 & NM_001066229.1	CAAATACCCCAGCAAGCAGT	CTGTGCTGTGATGGATCCTG	UPL #49 CAT NO: 04688104001	900 / 900 / 200 nM
5S rRNA	NA	Unknown	Unknown	Exiqon: 201509 Qiagen: MS00007574	manufacturers recommended conditions
*rno-miR-22**	miRBase v13	Unknown	Unknown	Exiqon: 202827	manufacturers recommended conditions
rno-miR-126	miRBase v13	Unknown	Unknown	Exiqon: 202141 Qiagen: MS00005607	manufacturers recommended conditions
rno-miR-140	miRBase v13	Unknown	Unknown	Qiagen: MS00005635	manufacturers recommended conditions
rno-miR-153	miRBase v13	Unknown	Unknown	Exiqon: 202160	manufacturers recommended conditions
*rno-miR-181a**	miRBase v13	Unknown	Unknown	Exiqon: 204110	manufacturers recommended conditions
rno-miR-186	miRBase v13	Unknown	Unknown	Exiqon: 202167 Qiagen: MS0000511	manufacturers recommended conditions
rno-miR-365	miRBase v13	Unknown	Unknown	Exiqon: 204622	manufacturers recommended conditions
rno-miR-674	miRBase v13	Unknown	Unknown	Qiagen: MS000013797	manufacturers recommended conditions

Table 2-5: SYBRgreen Assays for qPCR

Gene	Ref sequence (Position; bp)	Forward Primer sequence (5'-3')	Reverse Primer sequence (5'-3')	Oligo Conc^{n}
Atp1a1	NM_012504.1	TGGATCAATGATGTGGAGGA	CAATATTCCTCTGCTCGTAGG	600 / 600 nM
Atp1a2	NM_012505.1	GACCTCATCATTTGCAAGACC	CTAGCAGCCCAAAAATCAGG	600 / 600 nM
Hsd11b1	NM_017080.2	TGTCCTCGGCTTCATAGACA	GCGCAGAACTGTGCCTTT	600 / 600 nM
Hsd11b2	NM_017081.1	GGGGTATCAAGGTCAGCATC	TCCCAGAGGTTCACATTAGTCA	600 / 600 nM
Gstm1	NM_017014.1	TGTTACAACCCCGACTTTGA	TCTTCTCAGGGATGGTCTTCA	600 / 600 nM
Gstp1	NM_012577.1	TGGTACCCTCATCTACACTAACTATGA	CAGCAGGGTCTCAAAAGGTT	600 / 600 nM
Gstt1	NM_053293.2	CCTGGCTGACGTGGTAGC	GACGCCCTTCAAAGACTGG	600 / 600 nM
Igf1	NM_178866.2	CGGCCTCATAATACCCACTC	AAGACGACATGATGTGTATCTTTATTG	600 / 600 nM
Igf2	NM_031511.1	CGCTTCAGTTTGTCTGTTCG	GCAGCAACTCTTCCACGATG	600 / 600 nM
Nr3c2 (MR)	NM_013131.1	CCTGGCAGCGAAACAGAT	TCCTCGAGAGGCAAGTTTTT	600 / 600 nM
Tp53	NM_030989.1	GAGAGCACTGCCCACCAG	AACATCTCGAAGCGCTCAC	600 / 600 nM
Ucp2	NM_019354.1	GACTCTGTAAAGCAGTTCTACACCAA	GGGCACCTGTGGTGCTAC	600 / 600 nM

Table 2-6: DNA Cohorts and Human Random Control DNA Samples

Study Name	Study type	No. of individuals	Concentration	Notes:
Southampton asthma study	Family-based	1508	1, 5, 10 or 100 ng/µl	Originally described in Van Eerdewegh et al. Nature 2002[121] Refer to chapter 4.
ECACC	Human Random Control Panel	96	10 ng/ µl	HRC panel 1 (www.ecacc.org.uk)
Pooled control gDNA	Human Random Control Panel	96	1 or 10 ng/ µl	Pooled gDNA template from the ECACC dna samples

Table 2-7: Protein Restriction Rat Lung Tissue Samples.

Frozen archived lung tissue was provided by Dr. Christopher Torrens.

Study	Generation	Treatment group	No. of samples	Sex	Age (days)	Reference
DohaD Rat PR study [a]	F_1	PR	11	Male	120	*Torrens et al.*[263]
	F_1	Control	7	Male	120	
	F_1	PR	6	Male	225	*Torrens et al.*[263]
	F_1	Control	6	Male	225	
	F_2	PR	13	Male	80	*Burdge et al.*[264]
	F_2	Control	13	Male	80	

[a] *Liver RNA isolates were also available for the same 120 day old F_1 generation offspring.*

Table 2-8: Commercial Laboratory Kits

Kit	Cat no:	Supplier
ImPromII™ - Reverse Transcription Kit	A3800	Promega, Southampton, UK
SYBRgreen™ Core qPCR Mastermix Kit	RT-SN10-05	Eurogentec, SA
DNAfree Kit	AM1906	Ambion, USA

Table 2-9: Plasticware

Plastic ware	Cat no:	Supplier
ABgene 96 well V-Bottomed storage plates	AB-1058	ABgene, UK
ABgene Thermo-Fast 384 well plates	TF-0384	ABgene, UK
Universal tubes	CW3900	Alpha Laboratories
Bijoux Tubes	-	Alpha Laboratories
15ml, 25ml & 50ml Falcon Centrifuge tubes	LW1120/LW3078	Alpha Laboratories
1.5ml Tubes	616201	Greiner Bio-one
0.5ml Tubes	LT2172	Alpha Laboratories
0.2ml Tubes	LW2130	Alpha Laboratories
0.2ml Tube strips	LW2500	Alpha Laboratories
10ul Pipette tips	771288	Greiner Bio-one
100, 200, 1000ul Pipette tips	772288/739288/740288	Greiner Bio-one
10ml & 25 ml Serological Pipettes	607180/760160	Greiner Bio-one
Petri dishes	633102	Greiner Bio-one
2ml Screw cap tubes	CP5320WL	Alpha Laboratories

Table 2-10: Apparatus

Apparatus	Serial no:	Location	Supplier
Lightcycler480™ real-time PCR machine	1230	Asthma Genetics Lab	Roche diagnostics, UK.
Nano-drop absorbance spectrophotometer	unknown	Lab 18, Human Genetics	Nano-drop technologies, USA
7900ht ABI Prism Sequence detection System	500722	Molecular Pathology lab,	Applied Biosystems, Foster City, CA, USA
MJ Research DNA engine Tetrad Thermal cycler	TD002380	Asthma Genetics Lab	BioRad, USA
Labofuge 400 Plate Centrifuge	40340180	Asthma Genetics Lab	Heraeus, Germany
Biomek™ 3000 Robotic Liquid Handling Platform	9861200140	Asthma Genetics Lab	Beckman-Coulter, UK.
ABgene Plate Heat-sealer	951/2563	Asthma Genetics Lab	ABgene, UK.
BioPulverizer	59014H (cat no)	Asthma Genetics Lab	BioSpec Inc, US
DigiDoc-it UV trans-illuminator	-	Lab 18, Human Genetics	UVP, LLC. Upland, CA, US
RiboLyser	-	Brook Lab	Thermo Fisher Scientific Inc
UV PCR Cabinet	MC 0239	Asthma Genetics Lab	Lab-Caire, UK
M-240R bench-top centrifuge	0000138-02-00	Asthma Genetics Lab	BOECO, Germany
AB54-S Balance	1120133255	Asthma Genetics Lab	Mettler-Toledo, USA

Table 2-11: Software

Program	Version	Supplier	Web-link
SPSS	14.0	SPSS Inc	www.spss.com
LinRegPCR[265]	7.5	N/A	Requested by email
Roche Lightcycler480 Relative Expression Module	1.2	Roche Diagnosics, UK	www.roche-applied-science.com
Roche Lightcycler Basic Software	1.2	Roche Diagnosics, UK	www.roche-applied-science.com
FBAT and PBAT	1.4 & 1.0	N/A	http://www.biostat.harvard.edu/~fbat/default.html
SDS (ABI 7900ht Real-time PCR machine software)	2.1	Applied Biosystems	www.appliedbiosystems.com

Table 2-12: Miscellaneous

Description	Supplier
Foil	N/A
ABgene clear seal Diamond plate seals	ABgene, UK
ABgene clear seal Strong	ABgene, UK
Nalgene glass bottles	Fisher Scientific
Parafilm	Fisher Scientific
Scalpel & blades	-
Liquid Nitrogen Dewar	-
Metal spatulas	-
Sterile tweezers	-
RNAzap	Ambion, USA
RNAzap wipes	Ambion, USA

2.2 Methods

2.2.1 Gel electrophoresis of PCR Products

PCR products were electrophoresised on either a 2 or 3% w/v Agarose gel. A 3% w/v agarose gel was used for PCR products less than 100 bp in length. 2 or 3 grams of Agarose (Eurogentec, SA) was weighed and dissolved in 100 mls of 1 x TBE buffer by heating in the microwave. The agarose solution was allowed to cool to ~40°C before pouring into the mini-gel former and left to set. 5 µl of the PCR product was mixed with 1 µl of the blue-orange loading buffer (x 6 Concentration; Promega, UK) and loaded on to the gel. 5 µl of either the Smart-ladder (200 bp -10 kb; Eurogentec, SA) or 50 bp Ladder (Promega, UK) was loaded to the first and last lane of the gel to size the products. The PCR products were electrophoresed for 30-40 mins dependant of the size of the PCR products. The gel was post-stained with Ethidium Bromide (EtBr) by submerging the gel in 100 mls of 1 x TBE buffer containing EtBr (0.5 µg/ml) for 15 mins and de-stained in 100 mls of water for 15 minutes before visualisation on the DigiDoc-it™ system (UVP, LLC. Upland, CA, US).

2.2.2 Nucleic Acid Quantification

2.2.2.1 Nano-drop Absorbance Spectrophotometer

RNA and DNA quantification was performed on a Nanodrop D-1000 (Nandrop Technologies, US) absorbance spectrophotometer; only 1-2 µl of an undiluted nucleic acid solution aliquot is required to accurately quantify nucleic acids over a large linear range (2 to 3700 ng/µl of dsDNA). This methodology conserves precious samples by only using a small quantity of sample that is recoverable and is faster than cuvette spectrophotometers.

2.2.2.2 Nano-drop DNA/RNA Quantification Protocol

1.5 µl of water was loaded onto the pedestal and used as a reference to initialise the spectrophotometer; after a successful initialisation the pedestal was cleaned by wiping the bottom and top sample pedestal 5 times with a clean soft tissue. The machine is 'blanked' before the 1st and after the 10th nucleic acid sample using 1.5 µl of water to check the baseline and sample carry-over.

1.5 µl of DNA or RNA is also used for each sample reading and between samples the pedestal was cleaned as before. DNA and RNA samples were quantified in ng/µl and the 260/230, 260/280 ratios recorded for determining DNA/RNA quality. RNA samples were considered to be 'pure' when the 260/230 (Abs of DNA/Abs of Salts, Solvents & Protein) and 260/280 ratios (Abs of DNA/Abs of Proteins containing tryptophan) were between 1.8 - 2.2 and ~2.0, respectively. For DNA samples the ranges were between 1.8 - 2.2 and ~1.8. RNA samples were then normalised (0.5 µg/µl) and frozen at ⁻80°C, ready of use as RT-PCR templates.

2.2.3 Total RNA, DNA and Protein Isolation using TRIzol™ Reagent

2.2.3.1 RNA Work GLP

To maintain the integrity of the RNA isolate it is necessary to work on ice at all times and refrigerate the samples during centrifugation. Also, to protect the RNA samples from RNase degradation all plasticware, glassware, work-surfaces and equipment was either certified as RNase-free (in the case of plastic-ware) or treated/wiped down with RNAzap spray or wipes (Ambion, US) and gloves were worn at all times. Where applicable, all isolation steps were performed in the UV PCR extraction hood (Lab-Caire Systems Ltd, UK); filter tips and a designated set of RNA pipettes were used at all times. Appendix 10.1 details the preparation of reagents used in the TRIzol™ protocol.

2.2.3.2 RNA Stabilisation Treatments

In order to maintain RNA integrity during the lengthy tissue processing steps, it was necessary to evaluate the effectiveness of the two most commonly used RNA stabilisation treatments in combination with a tissue homogenisation protocol:

2.2.3.2.1 Liquid Nitrogen-Bio-Pulverizer Method

The mortar of the Bio-Pulverizer (BioSpec Products Inc, US) and one of the spatulas were placed in to a liquid nitrogen dewar. Once the liquid nitrogen had ceased to bubble the mortar was removed and placed on to the bench-top work surface. A 0.5 cm^3 section of lung tissue was cut using a sterile scalpel blade on a foil covered Petri dish top, weighed and quickly transferred to the liquid nitrogen remaining in the mortar.

While the liquid nitrogen was freezing the tissue and evaporating away, the hammer plate was introduced into the liquid nitrogen to cool. After the liquid nitrogen had evaporated the hammer plate was centred over the mortar, pressed down into position and the spring-loaded hammer trigger was depressed three times reducing the tissue section into a fine powder. Using the cold spatula the powder was scraped into a Lysing Matrix-D tube and 1.0 ml of TRIzol was added. The capped tube was mixed by inversion and stored in the ⁻20°C freezer ready for immediate homogenisation in the Ribolyser machine. The mortar, hammer plate and spatula were cleaned using UHQ water and RNaseZap wipes before repeating the process for more tissue samples.

2.2.3.2.2 RNALater-ICE RNA Stabilisation Treatment

In advance, 10 volumes (ie: 1 ml for 100 mg tissue) of RNALater-ICE regent (Cat no: 7030, Ambion) were pre-chilled to ⁻80°C for each lung sample. A piece of lung tissue (0.5 cm^3) was sectioned on a foil covered petri dish lid using a sterile scalpel blade and the weight was recorded on an accurate balance. The tissue sample had to be >5 cm in one dimension to ensure uniform permeation of the RNA stabilising reagent. The reagent contains a blue dye to identify treated samples and to see how far the reagent has permeated into the tissue. Working on dry-ice, the tissue section was place into 1 ml of the pre-chilled RNALater-ICE in a 1.5 ml tube. The tube was capped and inverted several times to mix. The tissue was soaked in RNALater-ICE at ⁻20°C overnight; ready for tissue homogenisation in the Ribolyser machine. At this temperature the tissue sample can be stored for up to 6 months. The next day, the treated lung section was finely chopped on an RNase-free petri dish lid using a sterile scalpel and transferred to a labelled Lysing matrix-D tube containing 1ml of TRIzol reagent. The samples were then stored on ice ready for tissue homogenisation in the Ribolyser machine (see table 2.10).

2.2.3.3 Tissue Homogenisation Protocol

The Ribolyser machine (Thermo Fisher Scientific Inc.) was programmed to run at speed 4.0 for 40 seconds. The lysing matrix-D tubes containing the lung sample in 1 ml of TRIzol were transferred to the Ribolyser machine, balanced and locked into place. After the each of the two runs the tubes were removed from the Ribolyser and placed on ice to cool for 2 mins. After the second incubation on ice the samples was checked for complete tissue homogenisation; if required the Ribolyser run program was repeated until samples were homogenous. The homogenised samples were incubated for 5mins at room temperature; then transferred to a ¯20°C freezer for storage until the RNA isolation step. The tissue homogenate in TRIzol can be stored at ¯70°C for up to one month protected from RNase degradation.

2.2.3.4 RNA Isolation

The homogenate was carefully transferred from the Lysing Matrix-D tubes into clean 1.5 ml micro-centrifuge tubes. The lysing Matrix tubes were retained at ¯80°C. 0.2 ml of chloroform was added to the 1.5 ml tube and shaken vigorously by hand for 15 seconds; this was then incubated at room temperature for 3 mins and then centrifuged at 12,000 x g (11,100 RPM) for 15 mins at 4°C. The upper aqueous layer that contains the RNA was carefully transferred to a fresh 1.5 ml tube. The remaining interphase and organic layer for DNA and Protein precipitation was saved and stored on ice (this preparation can be stored overnight at 4°C). The RNA was precipitated by mixing in 0.5 ml of chilled iso-propanol. The RNA precipitate in iso-propanol can be stored at 4°C for up to one week or ¯20°C for up to one year. The RNA precipitates were incubated at room temperature for 10 mins; then samples were centrifuged at 12,000 x g for 10 mins at 4°C, with the tube hinges facing out from the centre of the rotor. The RNA pellet forms on the hinge side and bottom of the tube. The supernatant was carefully removed from the tube (in 200 µl aliquots) without disturbing the pellet. The pellet was washed once with 1.3 mls of 75% ethanol. The pellet was mixed by vortex and then centrifuged at no more than 7,500 x g for 5 mins at 4°C. Most of the ethanol above the pellet was carefully removed and the pellet was air-dried for 5 - 10 mins to remove the excess ethanol ready for DNase treatment.

2.2.3.5 DNase Treatment of RNA Isolates and Storage

20 µl of the Ambion DNA-free™ master mix, containing 2 µl of 10 x DNase I Buffer 1ul rDNase I (2 U/µl), 17 µl RNase-free water was added to each tube and RNA pellet was dissolved by gentle aspiration. The samples were incubated at 37°C for 1 hour on the DNA engine tetrad thermalcycler (MJ Research, US). 2 µl of the DNase inactivation reagent was then added, mixed by vortex and then incubated at room temperature for 2 mins with occasional mixing. The sample was then centrifuged at 10,000 x g for 1.5 mins at 4°C to pellet the inactivation reagent. The RNA solution was transferred to a fresh 0.2 ml RNase-free tube and 30 µl of RNase-free water was added and mixed gently by pipette aspiration. 25 µl of the RNA solution was then aliquoted into another 0.2 ml RNase-free tube and both aliquots stored at ˉ80°C.

2.2.3.6 DNA Isolation and Storage

Any remaining aqueous phase left over the interphase (retained from the RNA isolation step) was carefully removed. 0.3 ml of 100% ethanol was added to the top of the interphase and the samples were mixed by inversion. The samples were then incubated at room temperature for 3 mins and then centrifuged at 2000 x g for 5 mins at 4°C to pellet the DNA. The phenol-ethanol supernatant was removed and transferred to a clean 1.5 ml tube and left on ice ready for the protein isolation step. The DNA pellet was washed twice in 1 ml of 0.1 M Sodium Citrate in 10% Ethanol. After the addition of each wash, the sample was incubated 30 mins at room temperature with periodic mixing and centrifuged at 2,000 x g for 5 mins at 4°C. The pellet was suspended in 1.5 ml of 75% ethanol and incubated at room temperature for 15 mins with periodic mixing and then centrifuged at 2,000 x g for 5 mins at 4°C. The 75% ethanol was carefully removed from above the pellet and air dried for 5-15 minutes in the open tube. The pellet was dissolved in 100-300 µl of TE solution (for 2 hours or overnight at room temperature). Any insoluble material was removed by centrifugation at 12,000 x g for 10 mins and the clear DNA supernatant was transferred to a new tube.

2.2.3.7 Protein Isolation and Storage

The phenol-ethanol supernatant retained at the DNA isolation stage was aliquoted in equal volumes into two pre-labelled 2.0 ml screw-cap tubes and treated in parallel for the rest of the protein isolation procedure as if 0.5 ml of TRIzol reagent was used initially. The proteins were precipitated from the phenol-ethanol supernatant with 750 µl of chilled iso-propanol. The precipitates were stored for 10 mins at room temperature and sedimented by centrifugation at 12,000 x g for 10 mins at 4°C. The supernatant was removed and the pellet was washed 3 times in 1 ml of 0.3 M Guanidine hydrochloride in 95% ethanol preparation. During each wash, the samples were incubated for 20 minutes at room temperature and centrifuged at 7,500 x g for 5 mins at 4°C. Following addition of the third wash, the pellet was snap-frozen in liquid nitrogen and stored at $^{-}$80°C.

2.2.4 Reverse Transcription

2.2.4.1 RNA Work GLP

Refer to methods section 2.2.3.1

2.2.4.2 Reverse Transcription Primer Reaction Setup

Reverse Transcription (RT) was performed on 1 µg of total RNA isolate using the Improm™II kit (Promega, UK). All RT reaction set-ups contained positive (Kanamycin RNA), Reverse Transcription negative (RT-ve) and water controls (No-template control) to control for reagent failure, internal and external gDNA contamination respectively. Firstly, the RNA template and primers were mixed as in table 2.13 in 0.2 ml PCR tubes and denatured at 70°C for 5 mins on a thermalcycler to remove secondary RNA structures, facilitating random/Oligo dT primer annealing. At the end of the denaturing step the sample tubes were immediately placed on ice.

Table 2-13: RT Primer Reaction Setup

Reagents	Positive control Reaction	Negative (No-template) Control Reaction	Experimental/ RT-ve reaction
1.2kb Kanamycin Positive control RNA (1 µg)	2 µl	-	-
Experimental RNA (1 µg/reaction)	-	-	2 µl
Oligo(dT)$_{15}$ Primer (0.5 µg/reaction)	1 µl	-	-
Random Primer (0.5 ng/reaction)	-	1 µl	1 µl
Nuclease-Free Water	2 µl	4 µl	2 µl
Final Volume	5 µl	5 µl	5 µl

2.2.4.3 RT Master Mix Setup

15 µl RT master mixes were made up as follows (table 2.14) and added to the 5 µl Primer/template mixes; mixed by flicking the tubes and spun down in a bench-top mini-centrifuge.

Table 2-14: RT Master Mix Setup

Reagents	Positive Control Reaction	No RT Control Reaction (RT-ve)	Experimental/ No-template Control Reaction
Nuclease-Free water (to a final volume of 15 µl)	4.2 µl	8.8 µl	7.3 µl
5 x ImPromII™ Reaction buffer	4.0 µl	4.0 µl	4.0 µl
$MgCl_2$ (Final conc 1.5 mM)	4.8 µl	1.2 µl	1.2 µl
dNTP mix	1.0 µl	1.0 µl	1.0 µl
RNasin Ribonuclease Inhibitor (20 U)	0.0 µl	0.0 µl	0.5 µl
ImPromII™ RTase	1.0 µl	0.0 µl	1.0 µl
Final Volume	15 µl	15 µl	15 µl

The tubes were transferred to the themalcycler and heated to 25°C for 5 mins, 42°C for 1 hr, 70°C for 15 mins and then finally cooled and held at 20°C. The cDNA samples were then stored at -20°C ready for use as templates in real-time PCR reactions.

2.2.5 Quantitative PCR:

2.2.5.1 Assay Design

PCR primers and TaqMan Probes were designed using PrimerExpress (version 2.0; Applied Biosystems) software and online Universal Probe Library assay design centre (www.roche-applied-science.com/sis/rtpcr/upl/index.jsp) for UPL probe assays.

Where possible, mRNA assays were designed over a splice site common to all known transcripts, where possible. The CNV assay sequences were BLAST SNP searched (www.ncbi.nlm.nih.gov/SNP/snp_blastByOrg.cgi) and all known polymorphism was avoided. To ascertain assay specificity the amplicon, the primer and probe sequences were also BLAST searched online. (http://www.ncbi.nlm.nih.gov/blast/Blast.cgi)

2.2.5.2 Assay Optimisation

2.2.5.2.1 Primer Concentration Titrations

Asymmetrical primer concentration titrations were performed for all assays and the concentration combination with the lowest Crossing threshold (Ct) or Crossing point (Cp) value was chosen (see table 2.15).

Table 2-15: Primer Concentrations Utilised in Optimisation Reactions

Forward Primer [FINAL] \ Reverse Primer [FINAL]	300 nM	600 nM	900 nM
300 nM	Reaction Set 1	Reaction Set 2	Reaction Set 3
600 nM	Reaction Set 4	Reaction Set 5	Reaction Set 6
900 nM	Reaction Set 7	Reaction Set 8	Reaction Set 9

The quantitative PCR (qPCR) master mix cocktail was made as follows (see table 2.16); each primer combination was run in triplicate wells using the thermalcycling programs in (see appendix 10.3). An absolute quantification analysis was performed on the real-time data and the concentration with the lowest average Ct or Cp value was considered to be the optimal primer concentrations.

Table 2-16: Master Mix Setup Tables for Primer Titration Reactions.

The left tables were used for probe based assays and the right tables for SYBRgreen assays. Primers diluted in 1 x TE buffer to stock concentrations of 3.33, 6.66 & 10 µM gave final reaction concentrations of 300, 600 & 900 nM respectively.

Master Mix Probes	x1	x37
2 x QPCR Mix	5.0	185
10uM probe [100 nM]	0.1	3.7
DNA (1 ng/µl)	2.0	74
dH2O	1.1	40.7
Total Volume (µl)	8.2	303.4

Master Mix SYBRgreen	x1	x37
2 x QPCR Mix	5.0	185
SYBRgreen dye	0.3	11.1
DNA (1 ng/µl)	2.0	74
dH2O	0.9	33.3
Total Volume (µl)	8.2	303.4

Primer Reaction Sets 1-9	x1	x4
Master Mix Probes	8.2	32.8
Forward Primer [3.33, 6.66 or 10 µM]	0.9	3.6
Reverse Primer [3.33, 6.66 or 10 µM]	0.9	3.6
Total Volume (µl)	10	40

Primer Reaction Sets 1-9	x1	x4
Master Mix SYBR green	8.2	32.8
Forward Primer [3.33, 6.66 or 10 µM]	0.9	3.6
Reverse Primer [3.33, 6.66 or 10 µM]	0.9	3.6
Total Volume (µl)	10	40

2.2.5.2.2 Melting Curve Analysis

Assay specificity was determined by post PCR melting curve analysis (see appendix 10.3). SYBRgreen assays are melted immediately after the PCR run while the probe based assays were spiked with SYBRgreen dye to a final 3% v/v concentration before being analysed.

2.2.5.2.3 Probe Concentration Titrations

Using the optimised primer concentrations a subsequent probe concentration titration experiment revealed the lowest probe concentration to use without sacrificing assay sensitivity (see table 2.17).

Table 2-17: Final Probe Concentrations Used in Probe Titration Optimisation.

Probes diluted in 1 x TE buffer to stock concentrations of 2, 4, 8 & 10 µM gave final reaction concentrations of 50, 100, 150 & 200 nM respectively.

Probe Reaction Set	1	2	3	4
Probe [Final]	200 nM	150 nM	100 nM	50 nM
Probe [Stock]	8 µM	6 µM	4 µM	2 µM
Volume (µl)/10µl reaction	0.25	0.25	0.25	0.25

The qPCR mastermix cocktail was made as follows (see table 2.18); each probe concentration was run in triplicate wells using the required thermalcycling program in appendix 10.3. An absolute quantification analysis was performed on the real-time data and the optimal concentration was determined by choosing the lowest probe concentration that produced similar real-time amplification curves to the highest concentration reactions.

Table 2-18: Master Mix Setup Tables for Probe Titration Reactions

Master Mix Probes	x1	X4
2 x QPCR Mix	5.0	20
Forward Primer [3.33, 6.66 or 10 µM]	0.9	3.6
Forward Primer [3.33, 6.66 or 10 µM]	0.9	3.6
2 µM - 8 µM Probe	0.25	1.0
DNA (1 ng/µl)	2.0	8.0
dH2O	0.95	3.8
Total Volume (µl)	8.2	40

2.2.5.3 Assay Validation

Standard curves were performed to test the linear range, sensitivity and determine the PCR efficiency of the qPCR assays (see tables 2.19 – 2.21). Each step of the serial dilution was run in triplicate wells on a 384 well PCR reaction plates using reaction volumes of 5 µl (CNV assays) or 10 µl (mRNA assays). Serial dilutions were made up on the day of the qPCR thermalcycling run in 0.2 ml PCR tubes (mixing by gentle vortex at each dilution step) and scaled up in volume if testing multiple assays on the same plate. Data was analysed according to appendix 10.3.

Table 2-19: Two-Fold Serial Dilution of Pooled gDNA Template (ECACC HRC1 DNA Panel) for the GST Copy Number Variation Assays.

Serial dilution[a] (DNA/reaction)	20 ng	10 ng	5 ng	2.5 ng	1.25 ng	0.625 ng	0.3125 ng	0.156 ng	No DNA
Pooled DNA (10 ng/2µl)	20	10	10	10	10	10	10	10	0
H20	0	10	10	10	10	10	10	10	10

[a] 1 µl of each standard per reaction well

Table 2-20: Four-Fold Serial Dilution of cDNA Template for mRNA or miRNA Assays.

Serial dilution[a] (cDNA/reaction)	4^0	4^{-1}	4^{-2}	4^{-3}	4^{-4}	4^{-5}	4^{-6}	4^{-7}	No cDNA
Pooled cDNA (~25 ng/µl)	18.0	4.5	4.5	4.5	4.5	4.5	4.5	4.5	0
H20	0.0	13.5	13.5	13.5	13.5	13.5	13.5	13.5	13.5

[a] 1 µl of each standard per reaction well

Table 2-21: Ten-Fold Serial Dilution of cDNA Template of the mRNA or miRNA Assays.

Serial dilution[a] (cDNA/reaction)	10^0	10^{-1}	10^{-2}	10^{-3}	10^{-4}	10^{-5}	No cDNA
Pooled cDNA (~25 ng/µl)	14	1	1	1	1	1	0
H20	0	13	13	13	13	13	13

[a] 1 µl of each standard per reaction well

2.2.5.4 Reaction Plate Setup and Analysis

2.2.5.4.1 Copy Number Variation Reaction Plate Setup

2.0 µl of gDNA (1 ng/µl) was transferred to duplicate wells on a 384 well reaction plate (see figure 2.1) using an electronic multi-channel pipette (Biohit, 0.5-10 µl) or the Biomek robot (appendix 10.2).

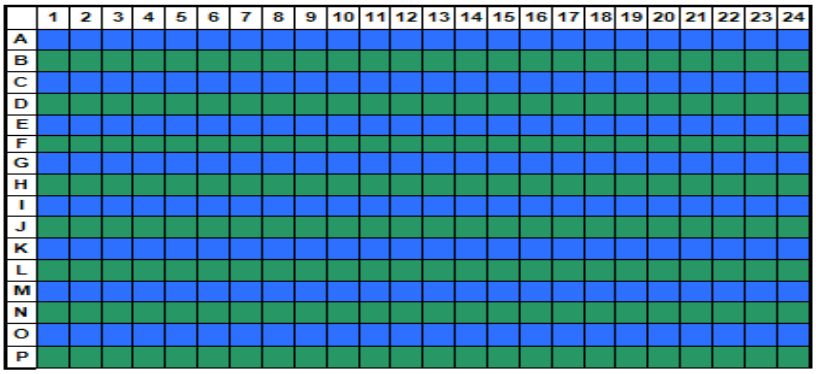


Figure 2-1: Example 384 Well Reaction Plate Layout.

The first 96 well DNA was duplicated on to the blue rows (i.e. first sample occupies wells A1 and A2). The second 96 well DNA plate was duplicated on to the green rows (i.e. first sample occupies wells B1 and B2)

Each 384 well plate contained negative control wells known as No Template Controls (NTCs) and positive control DNA (see table 2.22).

Table 2-22: ECACC Genomic Control Sample Names Used as Positive Controls for the GST CNV Assays

Copy number	*GSTT1*	*GSTM1*
2	C0858 & C0901	C0145 & C0150
1	C0906 & C0722	C0027 & C0088
0	C0007 & C0960	C0073 & C0202

The DNA template was dried down by heating the DNA solution to 80°C for 10 minutes in the DNA engine tetrad thermalcycler (Genetic Research Inc, US). To each sample well on the plate 5 µl of CNV master mix (see table 2.23) was added using a multichannel electronic Pipette (Matrix Technologies, US). The plates were sealed, then spun down in a plate centrifuge for 1 min at 3000 rpm and incubated at 4°C for 1 hr before PCR thermalcycling (see section 2.2.5.4.1.1).

Table 2-23: CNV Master Mix Setup.

Primer/Probe premixes were made up in large batches in advance to minimise batch to catch variability. CNV Master Mixes were made up in excess for pipetting (420 reaction volumes of CNV master mix were required for each 384 well plate)

***GSTM1* CNV Master Mix**	x1	x420
2xQPCR Mix	2.5	1050
Primer/Probe Mix per 5 µl rxn	0.165	69.3
H20	2.335	980.7
Total Volume (µl)	5	2100

***GSTM1* Primer/Probe Premix**	x1	x420	[FINAL]
100 µM *ALB* MGB Forward Primer	0.045	18.9	900 nM
100 µM *ALB* MGB Reverse Primer	0.03	12.6	600 nM
100 µM *GSTM1* MGB Forward Primer	0.03	12.6	600 nM
100 µM *GSTM1* MGB Reverse Primer	0.045	18.9	900 nM
100 µM *ALB* MGB Probe	0.0075	3.15	150 nM
100 µM *GSTM1* MGB Probe	0.0075	3.15	150 nM
Total Volume (µl)	0.165	69.3	

***GSTT1* CNV Master Mix**	x1	x420
2xQPCR Mix	2.5	1050
Primer/Probe Mix per 5 µl rxn	0.064	26.8
H20	2.436	1023
Total Volume (µl)	5	2100

***GSTT1* Primer/Probe Premix**	x1	x420	[FINAL]
100 µM *ALB* MGB Forward Primer	0.0113	4.73	225 nM
100 µM *ALB* MGB Reverse Primer	0.0075	3.15	150 nM
100 µM *GSTT1* MGB Forward Primer	0.0225	9.45	450 nM
100 µM *GSTT1* MGB Reverse Primer	0.015	6.3	300 nM
100 µM *ALB* MGB Probe	0.0038	1.58	75 nM
100 µM *GSTT1* MGB Probe	0.0038	1.58	75 nM
Total Volume (ul)	0.06375	26.8	

2.2.5.4.1.1 Thermalcycling Reaction Profile

PCR amplification and real-time fluorescent data collection was performed on the ABI Prism 7900HT Sequence Detection System (Applied Biosystems). Thermal cycling conditions were 95°C for 10 minutes then 40 cycles of 95°C for 15 seconds and 60°C for 60 seconds without 9600 emulated temperature-ramping. Fluorescent data was collected at the 60°C annealing/extension cycles and Ct values were calculated using the automatic Ct analysis settings on an absolute quantification run (SDS v2.1, Applied Biosystems).

2.2.5.4.1.2 Copy Number Variation Analysis

See section 3.3.7 in chapter 3.

2.2.5.4.2 mRNA Relative Expression Reaction Plate Setup and Analysis

9 µl of mRNA expression master mix was transferred to duplicate wells in 384 well reaction plates, standard curves and NTCs required triplicate wells. Each qPCR assay grouping on the plate also includes extra control wells for RT-PCR negatives (RT -ve), extraction negatives (Ex -ve) and calibrator cDNA template. The plate was temporally sealed and spun down in the plate centrifuge. Working on ice, 1 µl of cDNA template, calibrator, RT -ve, Ex -ve or standard curve serial dilution cDNA was added to each well using an electronic multi-channel pipette (using clean tips for each assay grouping). The plate was sealed, mixed, spun down and then run on the Lightcycler 480 real-time PCR machine according to appendix 10.3.

Note: qPCR assays were grouped by type (Probe or SYBRgreen dye) and ran on the same 384 well reaction plates.

2.2.5.4.2.1 Thermalcycling Reaction Profile

The profile used depended on the type of qPCR assay used (i.e. TaqMan™, UPL probes or SYBRgreen dye); see appendix 10.3 for a detailed description of each thermalcycling profile.

2.2.5.4.2.2 PCR Efficiency Determination by LinRegPCR Analysis

PCR efficiency estimates were performed on amplification plot data text files exported from the LightCycler480 system software. LinRegPCR analyses were performed in line with LinRegPCR user's manual;[265] fitting the regression line with 4-6 data points and best correlation. All regression lines were visually checked and the fit was corrected as necessary to ensure accurate estimates of PCR efficiency. See appendix 10.4 for a detailed description of the LinRegPCR program.

2.2.5.4.2.3 Relative Expression Using Multiple Reference Genes Analysis Protocol

Relative expression values were calculated for each target/reference gene combination in the Lightcycler480 relative expression module (appendices 10.3 and 10.4). When the standard curve range did not include all sample concentrations, LinRegPCR efficiency values were determined and relative expression values were calculated in Microsoft® Excel spreadsheets using the Roche E-method.[266] The geometric mean of values from each target/reference gene combination was then used for all further data analysis.

3 Chapter 3: Optimisation and Validation of Glutathione S-Transferase Copy Number Variation Assays

3.1 Introduction

The advent of microarray-based comparative genomic hybridisation (array CGH) technologies and whole genome sequencing has enabled the detection of unexpectedly heterogeneous structural variation (deletions, duplications, inversions and translocations) in the human genome.[267-269] These variations may have a greater effect on disease susceptibility than previously thought; therefore it is important for any genetic study to consider the presence of structural variations in the region of interest and have suitable technologies to detect them.[270] Current technologies such as array-CGH and MLPA (Multiplex ligation-dependant probe amplification; HRC Holland) facilitate the detection of structural variation across the genome or a smaller chromosomal region respectively.[271, 272]

In order to examine glutathione s-transferase gene polymorphism in the Southampton asthma cohort (table 2.6) it was necessary to develop a high-throughput genotyping technology capable of providing gene dosage information. Previous work in the genetics of asthma has shown that deletions of these genes increase the risk of developing asthma and interact with environmental oxidative stress exposures (section 4.1.4).

GSTT1 and *GSTM1* gross gene deletions were created by separate equal or un-equal recombination cross-over events between two highly homologous repeat regions flanking each gene.[194, 195] The separate recombination events result in deletion junction regions spanning up to several kilobases with very high homology (>98%) to the flanking repeat regions on chromosomes containing the genes. This phenomenon severely restricts unique priming sequences for standard PCR methods to identify the presence of a deletion.

Sprenger et al. described a multiplex PCR method using unique priming sites in the *GSTT1* deletion junction.[194] This method appears to have limited suitability in large scale genetic epidemiological studies as it involves two-steps with PCR amplification followed by agarose gel electrophoresis being required to visualise the 1.46 kb deletion junction and 466 bp gene specific PCR products. Unbiased, co-amplification of products differing this much in size can be problematic (especially when DNA quality is variable) as preferential amplification of the product with the higher PCR efficiency (most likely the smaller product) may occur.[273]

Brasch-Andersen et al.[274] have shown that it is possible to utilise the increased sensitivity of real-time PCR assays to provide dosage of *GSTT1* and *GSTM1*.[274] This method can distinguish between individuals with two, one or zero copies of a gene. As *GSTT1* or *GSTM1* copy number variations (also known as, 'gene dosage') are correlated with altered enzyme activity, [194, 199] analysis in a dose-dependant manner would best describe any disease outcome association.

The principle of applying real-time PCR to copy number variation genotyping is relatively simple, each copy number variation will require two sets of primers and fluorescent probes, the first set being specific to the gene of interest (i.e. variable copy number gene) and the second set specific for a stable reference gene (i.e. single copy number gene). The fluorescent signal generated by PCR amplification of the variable copy number gene can be normalised to the fluorescent signal from the stable reference PCR reaction. When these PCR reactions are combined in a single well ('multiplexed'), the reference gene PCR assay controls for the starting amount of template DNA in each PCR reaction and permits the calculation of copy number (i.e. no gene present, 1 copy, 2 copies, 3 copies…etc) expressed either as a concentration ratio of or difference from the reference gene Ct/Cp value (see figure 3.1).

This Chapter describes the real-time PCR assay design, optimisation, validation and novel data analysis software required to generate a high-throughput genotyping technology for assaying common GST gene deletion polymorphisms.

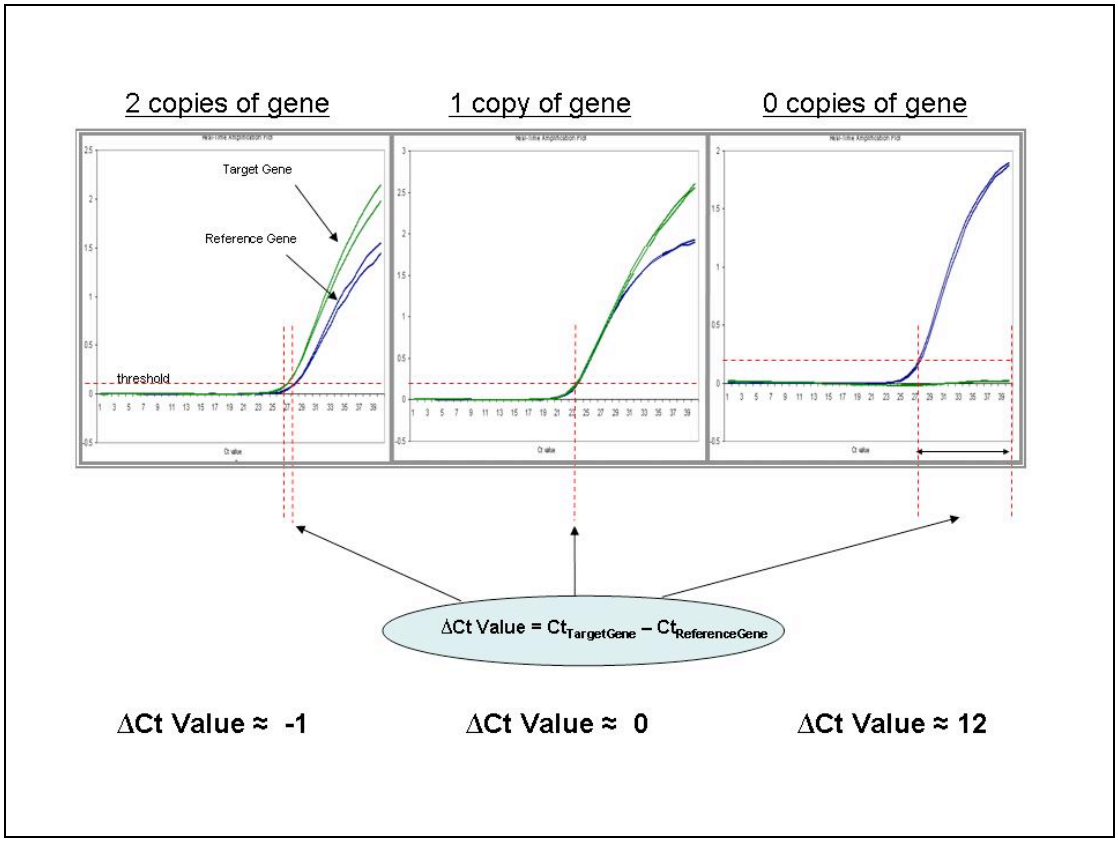


Figure 3-1: Determining Gene Dosage by Real-Time PCR.

The level of target gene amplification signal is normalised for starting DNA template concentration by the reference gene amplification signal. Ct values are derived from the point at which the fluorescent signal crosses the threshold. The resulting ΔCt values are used to determine the gene copy number in that DNA sample. In this example, samples with no target gene have large ΔCt values (i.e. no amplification of target gene, therefore the reaction end point of 40 is substituted into the equation); in samples with one copy of the target gene the ΔCt value is determined to be approximately zero. Samples with two copies of the target gene have a ΔCt value of about -1 as the target gene amplification signal has crossed the threshold before the signal from the reference gene.

3.2 Aims

The aim of this work was to:

- Design real-time PCR based copy number variation assays for the *GSTM1* and *GSTT1* genes.
- Optimise and validate the use of multiplex real-time PCR assays for high-throughput Copy Number Variation genotyping of the *GSTM1* and *GSTT1* genes.
- Design a custom analysis program for processing high-throughput real-time Copy Number Variation data.

3.3 Methods

3.3.1 Principles of the TaqMan Probe Based Real-Time PCR Assay

TaqMan Probe based real-time PCR detection is achieved via the 5' nuclease activity of Taq polymerase.[275] Hybridisation of a sequence specific fluorescent oligonucleotide probe to template DNA or PCR product is achieved using stringent PCR conditions that favour cleavage of a correctly matched probe during extension by the Taq polymerase. Upon cleavage the fluorochrome molecule is separated from the quencher molecule and releases light that can be detected by the real-time machine.

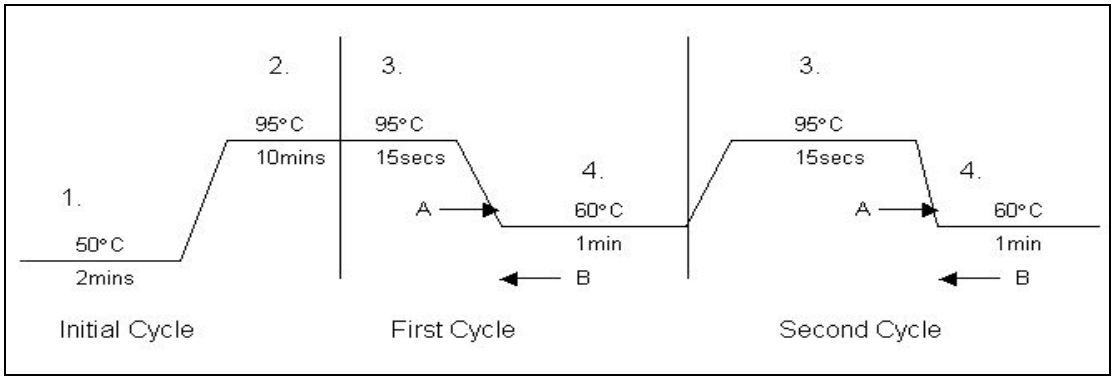
3.3.1.1 TaqMan PCR Reaction

At the start of the PCR reaction (step 1 in figure 3.2), Uracil-N-Glycosylase (UNG) is activated at 50°C for two minutes; this enzyme degrades any DNA containing the Uracil base and therefore stops the amplification of any PCR product that may be contaminating the reaction well. Step 2, activates the hot-start Taq polymerase enzyme by breaking the synthetic peptide bond covering the active site. Activation of the Taq polymerase enzyme at a high temperature ensures that no non-specific amplification can occur and interfere with the PCR reaction. Step 3, denatures the DNA template into single strands and breaks any secondary structure in the probes and primers ready for the first cycle of PCR. Step 4, cools the reaction from 95°C to 60°C; this cooling step allows the annealing of the correctly matched probe to the sequence of interest and then the primers anneal to their respective priming sites. The Taq polymerase activity extends the 3' end of the primer and cleaves the probe in the 5' to 3' direction releasing the fluorescent dye from the close proximity of the quencher molecule. This results in the break down of Förster-type resonance energy transfer (FRET) between the two molecules allowing for the reporter fluorescence to be visible when excited by the excitation light source.[276] Steps 3 and 4, are repeated up to 44 times generating a two-fold increase in PCR product at each cycle and the amplified reporter fluorescence signal is detected at the end of each cycle of PCR. A hybridised mis-matched probe will be displaced by the Taq polymerase rather than cleaved by the 5' nuclease activity.

The probe Melting temperatures are designed so that the annealing temperature of the reaction allows for the binding and cleaving of only the correctly matched probe (point A in figure 3.2). The melting temperature (Tm) of a miss-matched probe is lower than the annealing temperature of the reaction (Point B), so any partial hybridisation of a mis-matched probe will lead to displacement of the probe strand by the Taq polymerase, with no release of reporter fluorescence.

Figure 3-2: TaqMan PCR Thermalcycling Reaction.

The diagram demonstrates that through stringent assay design and constrained thermalcycling parameters the hybridisation of a mis-matched probe is virtually impossible as the melting temperature of a miss-matched probe/template duplex is lower than the annealing temperature of the thermalcycling reaction. Therefore, the fluorescent probe signal is highly specific to the target region only.



Key

1. UNG activation
2. AmpliTaq Gold Activation
3. Denaturing of DNA
4. Annealing and extension Stage

} Up to 45 Cycles

A. Melting Temperature (Tm) of a hybridised match probe
B. Melting Temperature (Tm) of a hybridised miss-matched probe

3.3.1.2 TaqMan® Probe Designs

There are two designs of TaqMan® probes. A Tamra™ probe (See figure 3.3) consists of a reporter dye and the Tamra™ quencher (Tamra™ is a fluorescence dye which can quench FAM or VIC reporter dye signals by absorbing the released energy and emitting its light at a higher wavelength). The Tm of a Tamra™ probe is totally dependant on the melting temperature of the sequence. The MGB probe is similar to the Tamra™ probe except for the addition of a DNA Minor Groove Binding (MGB) moiety that increases the Tm of the probe and the use of a non-fluorescent quencher (releases energy as heat). The stabilisation effect of the minor groove binder allows for shorter probes to be designed (increases signal strength), and the probing of AT rich sequence regions (low Tm regions) in the genome.

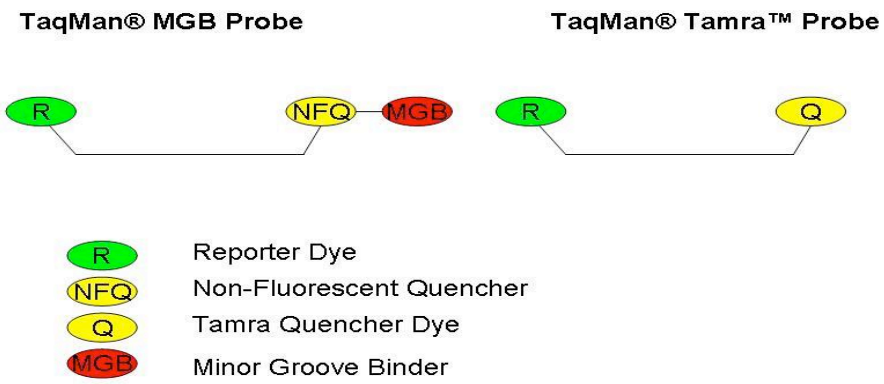


Figure 3-3: TaqMan Probe Designs.

There are two forms of the TaqMan hydrolysis probe; both probes have a reporter dye on a linker arm attached to the 5' end of the oligonucleotide probe and a quencher molecule on the 3' end of the probe. The MGB moiety is a crescent shaped molecule that fits into the minor groove of the DNA, whereby stabilising the shorter TaqMan MGB probe:DNA duplex.

3.3.2 Copy Number Variation qPCR Assay Design

Real-time PCR TaqMan® MGB probe assays were designed as previously described in the general methods section 2.2.7.1 for the *GSTM1*, *GSTT1* and ALBUMIN (*ALB*) gene which acts a reference to normalise out differences in the starting DNA template concentrations between samples. As the PrimerExpress v2.0 software designs up to 200 assays for each gene; the optimum assay designs were chosen as follows:

- Lowest Penalty score (automatically generated by software, assays are ranked by closeness to optimum TaqMan® assay design parameters)
- Matched amplicon size to *ALB* assay
- Primers lie over an exon-intron boundary
- Primers do not form excessive dimers under multiplex PCR conditions (Performed using the primer test analysis in PrimerExpress software v2.0)

The *GSTM1* and *GSTT1* TaqMan® MGB probes were 5' labelled with a 6-FAM reporter dye and the *ALB* probe with VIC™ (refer to table 2.3).

3.3.3 Optimisation of Copy Number Variation Assays

Using a pooled ECACC Human Random Control (HRC1) DNA samples as template (10 ng/reaction), asymmetric primer concentration PCR reactions were performed for all three assays as detailed in the general method section 2.2.5.2.1 to optimise priming conditions. Probe concentration titration reaction cocktails (section 2.2.5.2.3) were then made up using the primer concentrations determined from the previous experiment. Once the optimal concentrations of primers and probe were determined for each assay extra PCR product was generated and spiked post-PCR with SYBRgreen dye and a meltcurve analysis was performed according to the general methods section (section 2.2.5.2.2.) to test the primer specificity for each gene target. Single (e.g. *GSTT1*, *GSTM1* or *ALB*) and multiplex (e.g. *GSTT1* & *ALB* or *GSTM1* & *ALB*) two-fold serial dilution standard curves were constructed according to general methods section 2.2.5.3 and used to determined the PCR efficiencies of each probe assay.

3.3.4 Validation of Copy Number Variation Assays

ECACC Human Random Control (HRC1) DNA samples (www.ecacc.org.uk) were used to optimise reaction conditions, determine reaction plate storage stability and validated subsequent experiments as positive controls. Table 2.22 details the ECACC HRC1 DNA samples used as genomic controls after their genotypes were confirmed using the Brasch-Andersen *GSTT1/M1* TaqMan assays.[274]

All GST copy number variation reaction mixture cocktails were made up, aliquotted to 384 well reaction plates and thermalcycled according to the general methods section 2.2.5.4.1. Inter-plate and intra-plate duplicate reactions and Inheritance checks (using FBAT v1.4[277]) were performed on the Southampton family-based asthma cohort (table 2.6) to determine genotyping accuracy.

3.3.4.1 Hardy-Weinberg Equilibrium Calculations

Hardy-Weinberg calculations were performed on the parents genotypes using the online version of the Definetti program[278] (http://ihg.gsf.de/cgi-bin/hw/hwa1.pl) A Pearson's p value of less than 0.05 was considered to be a significant deviation of genotype frequency from Hardy-Weinberg Equilibrium.

3.3.5 Copy Number Variation Stability Tests

Assay plate stability trials were performed at 0 hr, 12 hr, 24 hr and 48 hr time intervals using the ECACC HRC1 DNA panel (table 2.6) as template for PCR. The four identical plates were set up at the same time. The first plate was run immediately on the ABI Prism 7900ht sequence detection system, the remaining three plates were stored at room temperature (21°C) and ran at 12, 24 and 48 hour time points.

3.3.6 Performance of GST CNV Assays on Other Real-Time PCR Machines

One reaction plate using the ECACC HRC1 DNA panel (see table 2.6) as template was set up according to the general methods section 2.2.5.4.1 and was thermalcycled on the Roche Lightcycler 480 real-time PCR machine (see apparatus table 2.10 and appendix 10.3). The reaction plate was thermalcycled according to section 2.2.5.4.1.1, a colour comparison was performed and the data was analysed according to appendix 10.3.

3.3.7 Design of Copy Number Variation Analysis Spreadsheet

A macro-based Microsoft® Excel spreadsheet was designed for the analysis of gene dosage real-time PCR data. The results and clipped data files generated by the SDS v2.1 software (Applied Biosystems) for each plate run can imported into the spreadsheet and each duplicate data point can be visualised in a ranked XY scatter graph (ΔCt versus ΔCt ranked sample position; see figure 3.4).

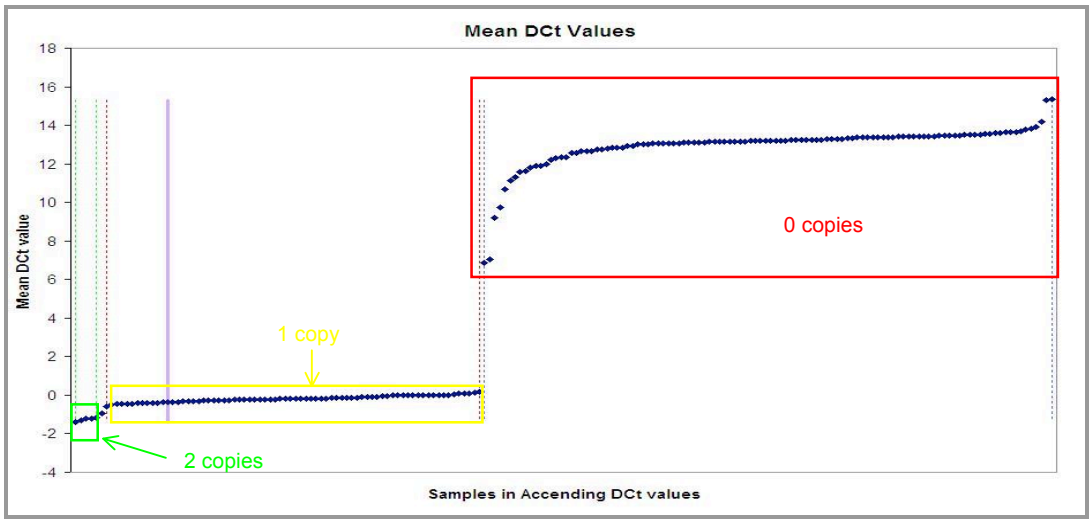


Figure 3-4: Copy Number Variation Ranked XY Scattergraph.

This is a screenshot of the ranked XY scattergraph from the analysis spreadsheet. A blue diamond data point corresponds to the mean ΔCt value for one sample. The vertical dashed bars can be moved by the user to determine the start and end ΔCt values for each genotype group. At this zoom level it is relatively easy to distinguish between 0 copy and gene present (1 or 2 copies) samples.

The ΔCt value for each duplicate can be compared directly to the real-time amplification curve plots generated by each sample well grouping (see figure 3.5). The interactive sliding rules allow the researcher to semi-automatically assign genotype groups and outliers on the basis of ΔCt value grouping on the ranked XY scatter graph.

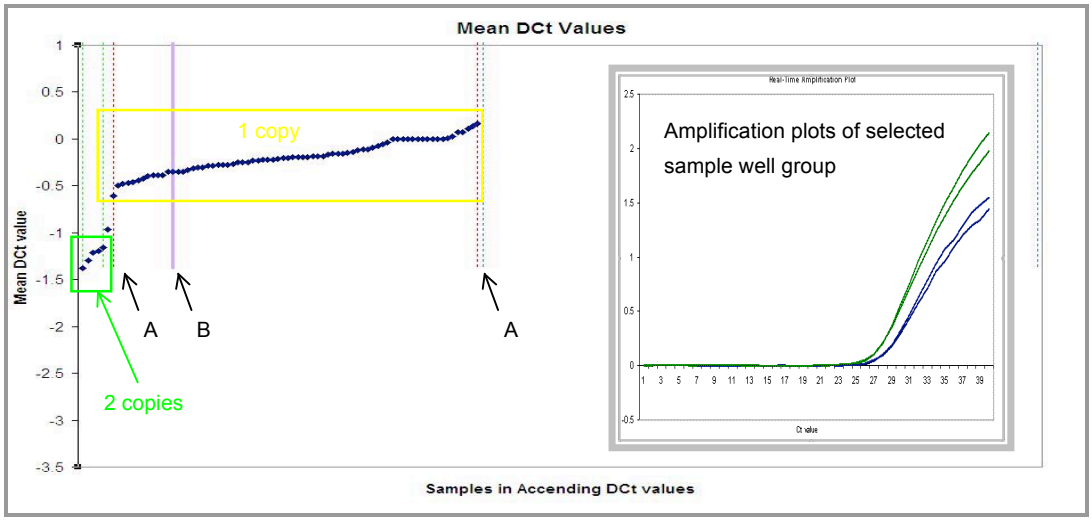


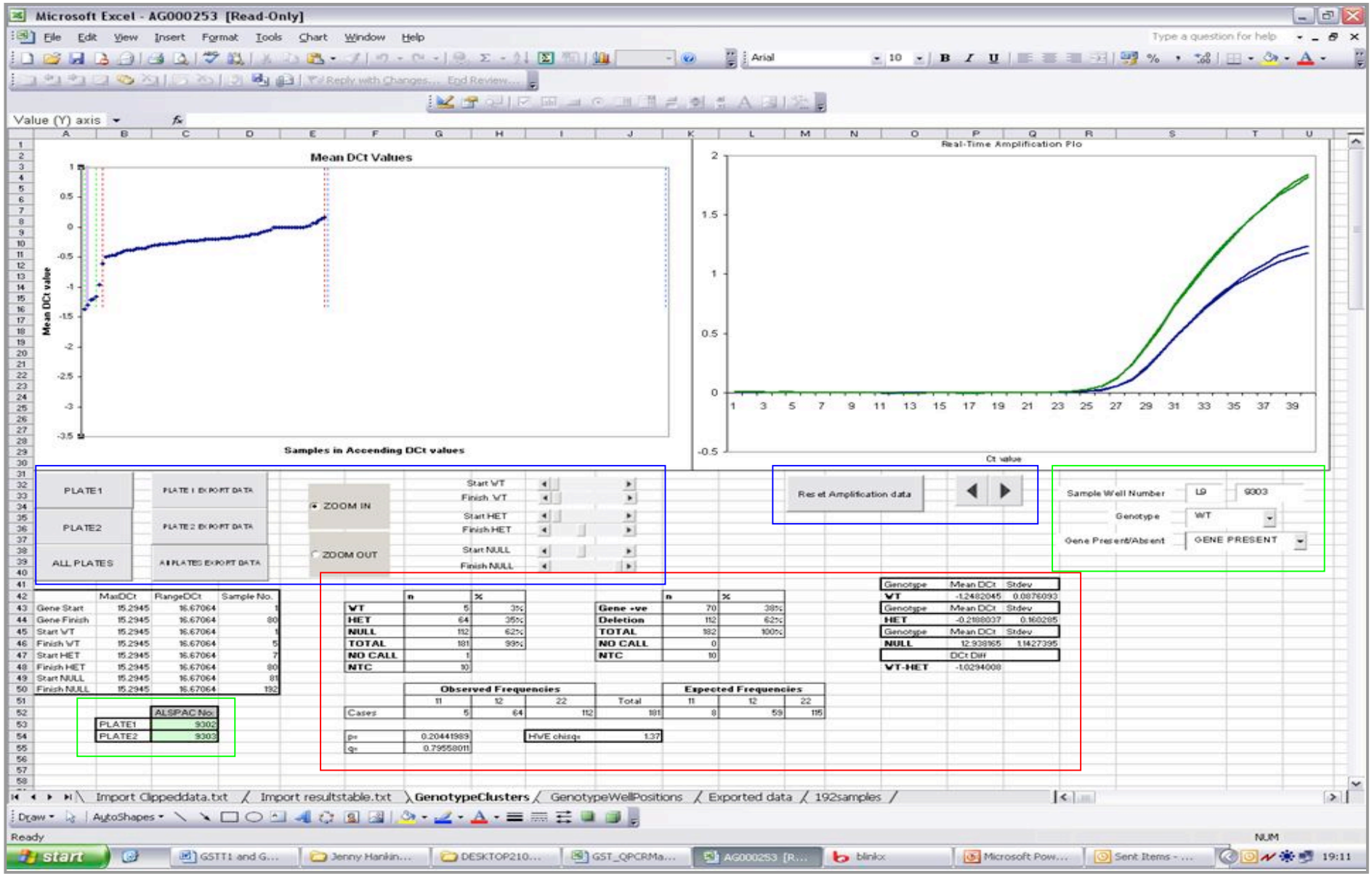
Figure 3-5: Zoomed Copy Number Variation Ranked XY Scattergraph.

This is a screenshot of the ranked XY scattergraph zoomed into focus on gene present data points only. A blue diamond data point corresponds to the mean ΔCt value for one sample. The vertical dashed bars (labelled 'A') can be moved by the user to determine the start and end ΔCt values for each genotype group. The purple solid bar (labelled 'B') can also be controlled by the user; this bar displays the real-time amplification data used to calculate the ΔCt data point. At this zoom level it is relatively easy to distinguish between 1 copy, 2 copy samples and abnormal data points that can be manually excluded from genotype assignment by examining the amplification plots directly.

Importantly, interrogation of the real-time amplification curve plots allows the researcher to rapidly determine the validity of any genotype call and correct manually. The spreadsheet automatically calculates copy number variation and gene present/absent frequencies, deviation from Hardy-Weinberg equilibrium and several quality control parameters (i.e. mean ΔCt value and SD of each copy number group). Figure 3.6 displays a full screenshot of the analysis spreadsheet user interface.

Figure 3-6: Full Screenshot of the Copy Number Variation Analysis Spreadsheet User Interface.

Data importing is controlled by three mouse reactive control buttons, one to import all plate data, and two to import the first or second 96 well DNA array data (1st button column in the left blue box); at this point DNA array names can be typed in (left green box). Users can assign genotype groups on the ranked XY scattergraph (top left chart) using sliding control buttons for each genotype group (4th button column in the left blue box), zoom into the gene present results (3rd button column in the left blue box) and simultaneously examine amplification plots (top right chart) using scrolling buttons (right blue box). After examining the amplification plot data, any manual genotype calls or exclusions are performed using drop down list boxes (right green box). Genotype frequencies, Hardy-Weinberg calculations and ∆Ct value statistics are also displayed in this interface (red box). Analysed copy number variation data is then exported in a database format using the three export data buttons (2nd button column in the left blue box).


Microsoft Excel - AG000253 [Read-Only]
Mean DCt Values
Mean DCt value
Samples in Accending DCt values
Real-Time Amplification Plo
Ct value
PLATE1
PLATE2
ALL PLATES
ZOOM IN
ZOOM OUT
Start WT
Finish WT
Start HET
Finish HET
Start NULL
Finish NULL
Reset Amplification data
Sample Well Number
Genotype
Gene Present/Absent
GENE PRESENT
Import Clippeddata.txt
Import resultstable.txt
GenotypeClusters
GenotypeWellPositions
Exported data
192samples
Ready
NUM

In addition, the interactive and colour-coded 384 and 96 well plate layout arrays generated by the macro allow the well position of any genotype call (with corresponding %CV of target and reference Ct values between each sample pair), water control or failure to be visualised post-analysis; this utility enables the researcher to monitor any problematic samples, experimental controls and identify any possible plate-edge effects on data quality (see figure 3.7).

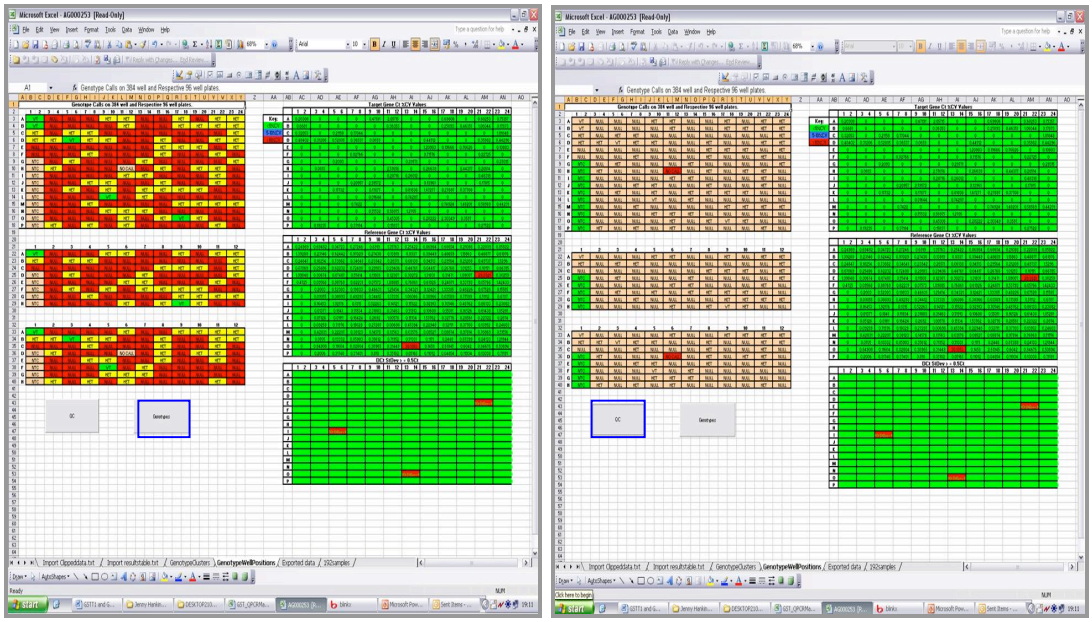


Figure 3-7: Colour Coded 384 & 96 Well Plate Layout Arrays for Data Quality Control.

The well position of any genotype call (Left panel: left array column; 0 copies in red, 1 copy in yellow and 2 copies in green) can be visualised by pressing genotype control button (blue box in left panel). Water controls and genotyping failures (Right panel: left array column) are highlighted in green and red respectively by pressing the QC control button (blue box in right panel). %CV of Ct values for each sample well duplicate is displayed on the right array column in both panels.

Finally, the macro exports all analysed data to a new sheet in the spreadsheet in a database compatible format (Microsoft® Access).

3.4 Results

3.4.1 Assay Optimisation

The sensitivity of the qPCR assays were substantially improved by utilising the asymmetric primer concentrations listed in table 2.3; these conditions reduced Ct values by up to 1.38 Ct values. Probe concentration titrations (figure 3.8) revealed that a 150 nM or 75 nM final concentration of probe for each assay (*GSTM1* or *GSTT1* respectively) could be used in a reaction without any significant decrease in signal strength (i.e. lower signal to noise ratio) or sensitivity (increase in Ct value). SYBRgreen melt curve analysis revealed single melting peaks for all assays.

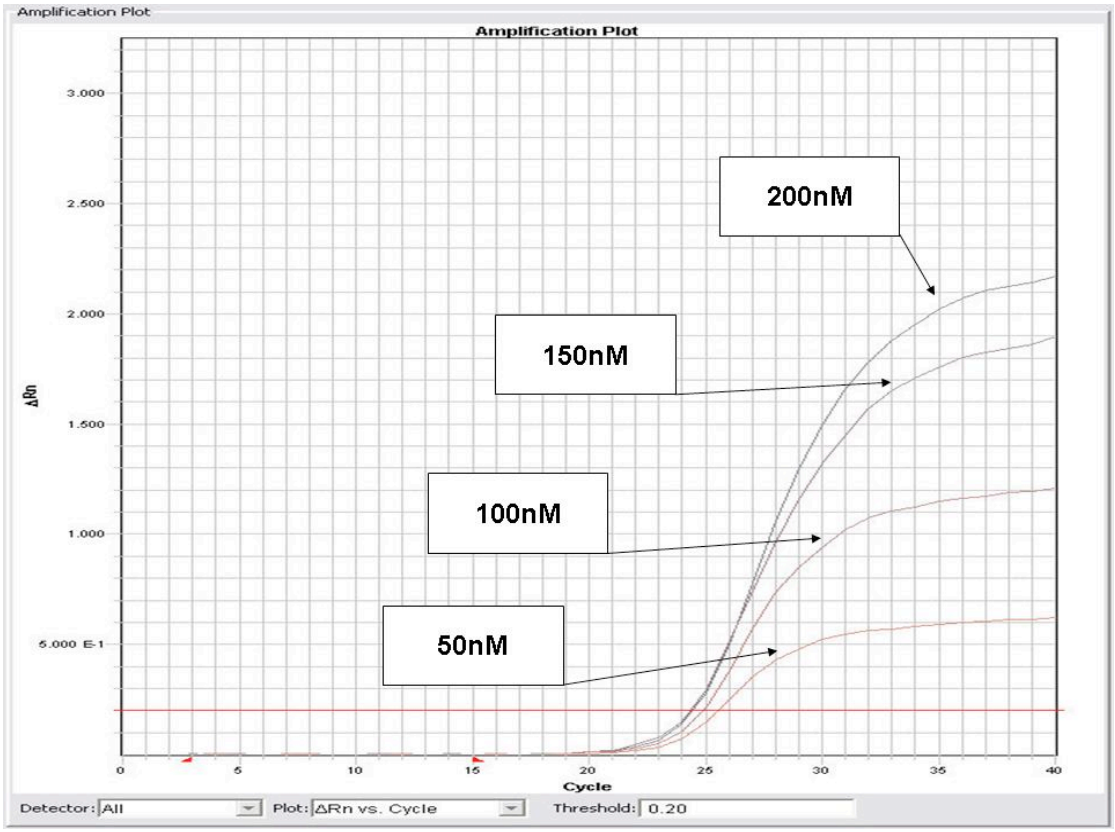


Figure 3-8: *GSTM1* Real-Time Amplification Plots in Reactions with a Decreasing Probe Concentration.

It is clear from the amplification plot data that 150 nM of probe can be used without any change in Ct value or significant decrease in fluorescence signal magnitude compared to the highest probe concentration (200 nM); thus the use of 150 nM of probe in reactions is a probe cost saving of 25%.

3.4.2 Assay Validation

3.4.2.1 Standard Curve Analysis

Under multiplex conditions all assays performed well using the optimised asymmetric primer and probe concentrations described above. The standard curves in figures 3.9 and 3.10 demonstrate that the multiplex reactions are capable of linear amplification over the full range of DNA template concentrations. Over the 20 – 0.156 ng DNA template dilution range all PCR reaction efficiency values ($E = 10^{1/s} - 1$; where s is the slope of the standard curve) were between 1.95 and 1.99; which are close to the maximum theoretical value (E = 2.0) where a 1 Ct value is equal to a doubling of PCR product at each cycle of PCR. The inter-assay PCR efficiency variation was less than 1% in both multiplexes (E_{GSTM1} - $E_{ALBUMIN}$ = 0.65%; E_{GSTT1} & $E_{ALBUMIN}$ = 0.01%) confirming that both PCR products in the multiplex are amplifying at almost the same rate.

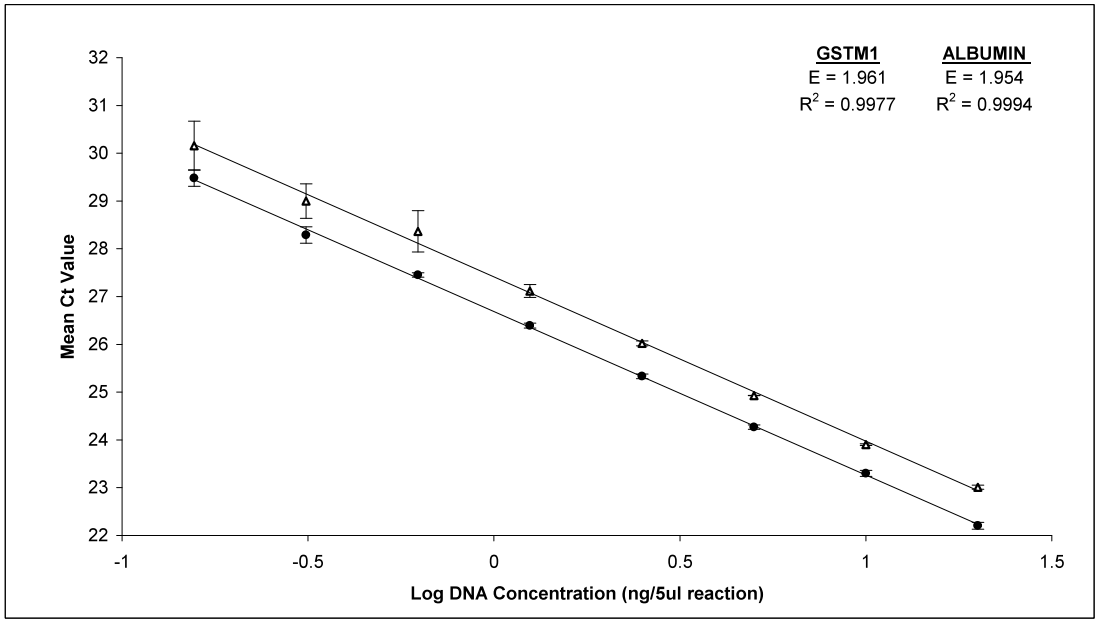


Figure 3-9: *GSTM1* Standard Curve Analysis.

Standard curves generated using pooled ECACC DNA template in a GSTM1 (●) + ALBUMIN (∆) multiplex reaction. Each data point is the mean of three duplicate wells, errors bars = 95% CI. DNA concentration range 0.156 – 20 ng/well. The PCR efficiencies (E) are almost the same in both PCR reactions and the linear regression coefficients are > 0.997.

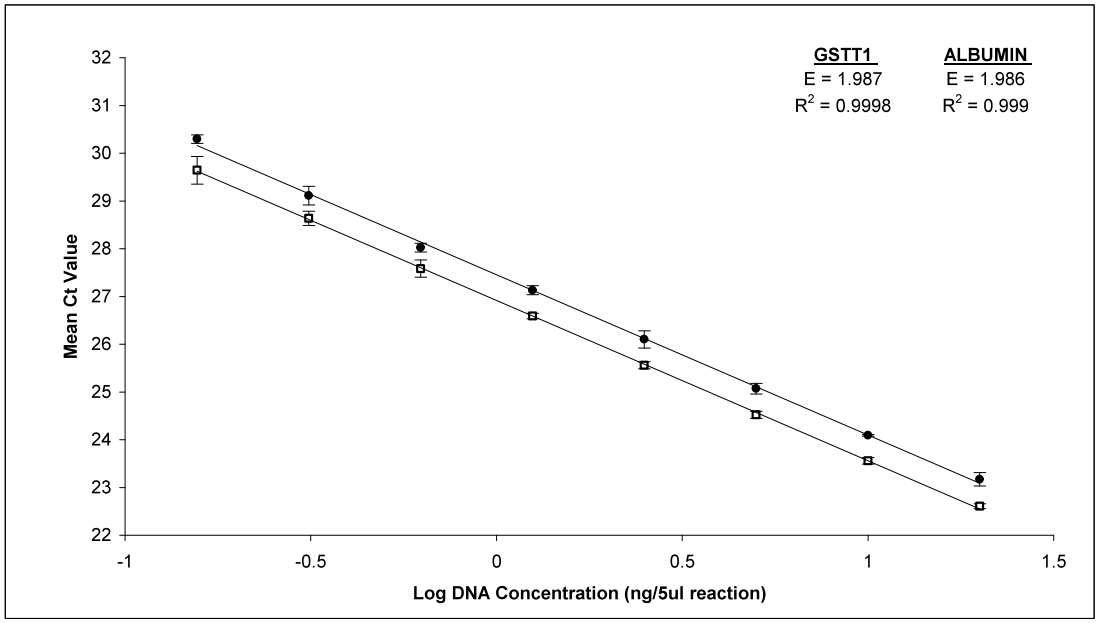


Figure 3-10: *GSTT1* Standard Curve.

Standard curves generated using pooled ECACC DNA template in a GSTT1 (●) + ALBUMIN (□) multiplex reaction. Each data point is the mean of three duplicate wells, errors bars = 95% CI. DNA concentration range 0.156 – 20 ng/well. The PCR efficiencies (E) are virtually identical in both PCR reactions and the linear regression coefficients are > 0.999.

Any experimental variation in the ΔCT (ΔCT = $CT_{target} – CT_{reference}$) value used to determine the gene of interest copy number can be estimated by the %CV (Percentage Coefficient of Variation) of the mean of the ΔCt values determined at each DNA dilution point along the standard Curve. Over the full dilution range (20-0.156 ng DNA) there is 13.34 %CV and 13.92 %CV in ΔCT values for the *GSTM1* and *GSTT1* assays respectively.

3.4.2.2 ΔCt Variation by Genotype

Initial genotyping of 192 DNA samples for both genes demonstrated that the mean ∆Ct values (± 1 SD) for 0 copy individuals were 12.9 ± 1.14 for *GSTM1*, 14.15 ± 1.77 for *GSTT1*; 2, and 1 copy individuals for *GSTM1* were -1.25 ± 0.09 and -0.22 ± 0.16 respectively. For *GSTT1*, individuals with 2 copies have mean ∆Ct values of -1.83 ± 0.21 and -0.74 ± 0.18 for 1 copy of the gene.

The $\Delta Ct_{2copies} - \Delta Ct_{1copy}$ difference was initially determined to be 1.03 and 1.08 for *GSTM1* and *GSTT1*; this value is very close to the theoretical 1 Ct value associated with a doubling in concentration (i.e. one extra copy of the gene). In practise, this value varies slightly between plate runs (0.85-1.20); this variation does not impact genotyping accuracy as data from each plate is analysed independently in the analysis spreadsheet.

3.4.2.3 Genoytyping Success and Accuracy

Table 3.1 shows the full cohort and parent genotype frequencies for both gene copy number variations and non-significant Pearson test results for deviation from Hardy-Weinberg Equilibrium. Undetermined copy number variation call rates (e.g. 2, 1, and 0 copies) were 4 and 9% for *GSTM1* and *GSTT1* genes respectively; and 2% and 3% for the gene absent/present call.

Genotyping accuracy of the *GSTM1* (99.87%) and *GSTT1* (99.46%) assays was determined by subtracting the number of non-Mendelian inheritances detected by the FBAT program from the number of child genotypes (*GSTM1* n = 770; *GSTT1* n = 735). All of the non-Mendelian inheritances (*GSTM1* n = 1; *GSTT1* n = 4) resulted from child '2 copy' genotype calls when the parental genotype combination permitted only 0 or 1 copy offspring. Duplicate genotyping of 6% of the cohort demonstrated 100% concordance of genotype calls between plate runs and analyses.

Table 3-1: GST CNV Genotype Frequencies in the Southampton Asthma Cohort.

Hardy-Weinberg calculations were only performed to the parent samples to avoid any potential confounding by related individuals.

Copy Number Variation Frequencies n (%)					
Group	**Gene**	**2 copies**	**1 copy**	**O copies**	**HWE P value**
All cohort	*GSTT1*	415 (31)	690 (51)	247 (18)	N/A
	GSTM1	98 (7)	534 (38)	786 (55)	N/A
Parents only	*GSTT1*	193 (31)	321 (52)	103 (17)	0.117
	GSTM1	42 (6)	250 (39)	356 (55)	0.831

3.4.3 Stability Trials

All plates were analysed using the copy number variation analysis spreadsheet and genotypes were assigned to the three later time point plates based on the genotype results from the first plate (time = 0). The reported Ct values for each genotype were then compared to determine the maximum length of storage without loss of data quality (figure 3.11). The assay plate stability trial results determined that sealed reaction plates can be stored protected from light at room temperature for up to 24 hours before the real-time PCR thermalcycling run without any degradation in the real-time data quality.

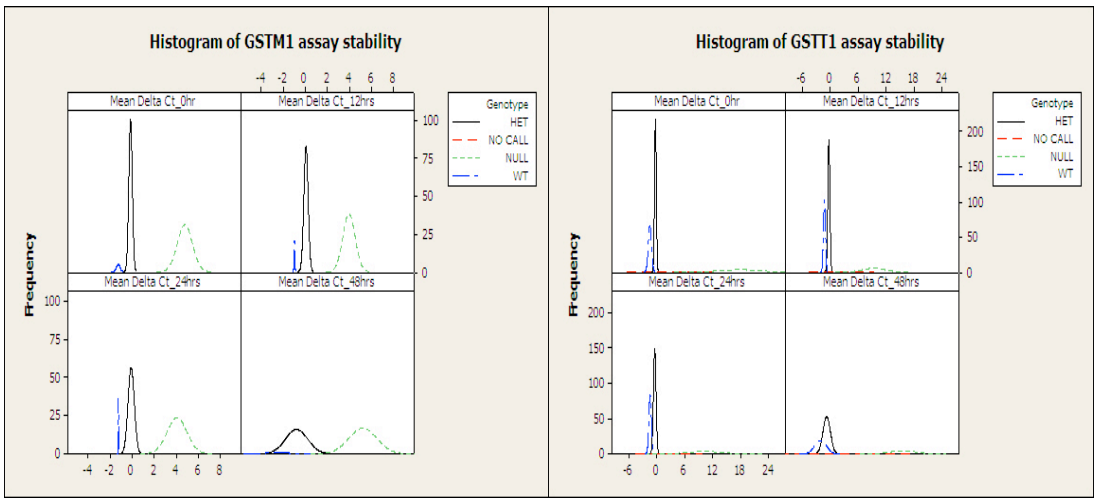


Figure 3-11: Matrix Histograms of the Assay Stability Trial.

Left panel: histogram matrix of ΔCt values by GSTM1 genotype (WT = 2 copies, HET = 1 copy, NULL = 0 copies) at four time points (0, 12, 24, 28 hours). Right panel: histogram matrix of ΔCt values by GSTT1 genotype at the four time points (0, 12, 24, 28 hours). It is apparent from the distinct distribution of ΔCt values for each genotype that the master mix remains stable up to 24 hours after plating out.

3.4.4 Performance of GST CNV Assays on Other Real-Time PCR Machines

The analysed GST CNV results from the Lightcycler 480 real-time PCR machine show that the assays perform well on other real-time PCR platforms without any further optimisation or validation. All of the genotype assigned to the ECACC HRC1 DNA panel on the Lightcycler 480 machine matched the genotypes determined previously on the ABI Prism 7900HT sequence detection system. The *GSTT1* XY scatter graph generated in excel directly from the relative quantification concentration ratios shows that sample data points can be discriminated into individuals with 0, 1 and 2 copies of the *GSTT1* genes (see figure 3.12). The data from the *GSTM1* XY scatter graph was very similar and clustered into each of the three possible genotypes (data not shown).

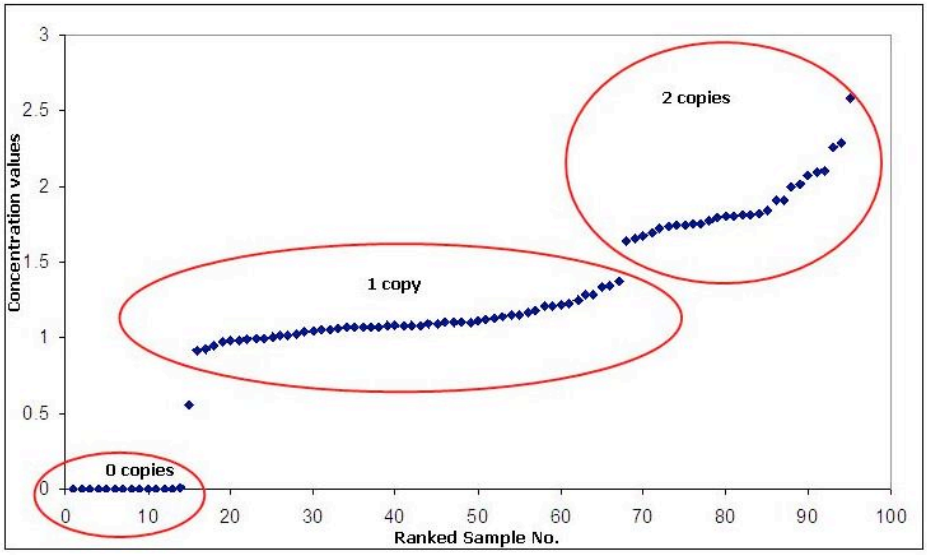


Figure 3-12: *GSTT1* CNV Scatter graph from the Roche Lightcycler 480 Machine.

As the Lightcycler relative quantification software module calculates the difference between the target and reference gene using a ratio ($Conc_{Target}/Conc_{Reference}$); a 0 copy genotype has a concentration value of 0 and 1 & 2 copy genotypes have respective concentration values of about 1 & 2. It is clear from the distribution of the data that this assay performs well on other real-time PCR platforms without further optimisation.

3.5 Discussion

This new methodology provides a significant advance in enabling the study of gene copy number variation in large sample numbers required by genetic association studies. This is not the first application of real-time PCR to assay *GSTM1* and *GSTT1* copy number variation; several other protocols have been described in the literature.[274, 279] Unfortunately, these methods do not have true high-throughput capability because of low sample density on reaction plates due to triplicate wells and large reaction volumes (≥ 5 µl) resulting in high costs per sample. With this method 96 samples can be genotyped in duplicate wells for 2 copy number variations on 1 reaction plate (or 192 samples for 1 copy number variation) in under two hours with a proportional cost saving of 66% in 2 x qPCR master mix and 25% in probe costs. The protocol presented in this thesis has true high-throughput capability as it is compatible with standard laboratory liquid handling robotic platforms for DNA aliquoting, reaction plate setup and real-time PCR machine plate loading. Assay stability trials have shown that 12 x 384 well reaction plates can be processed with in a 24 hour period with out any degradation of real-time PCR data quality.

The analysis methodology described in the Brasch-Anderson paper[274] requires substantial post-run data processing and analysis by Probit/Rankit plots. Whereas, the macro-driven analysis spreadsheet described in this chapter provides automation to the data analysis procedure and allows for the examination of real-time amplification plots for every well on each reaction plate while assigning genotypes to each ΔCt value. The analysis spreadsheet provides non-experienced and experienced real-time PCR users alike a complete analysis, quality control, troubleshooting and data handling solution to copy number variation genotyping on the ABI 7900ht sequence detection machine. This Microsoft Excel® spreadsheet removes the considerable bottleneck of processing real-time PCR data and can be applied to real-time data from other copy number variations. Initial experimentation on the Lightcycler 480® Real-time PCR machine (Roche Diagnostics, UK) demonstrated the portability of these assays without further optimisation of reaction or run conditions. This system is readily adaptable to study other copy number variations in the human genome.

3.6 Concluding Remarks and Future Work

Two multiplex copy number variation real-time PCR assays have been successfully designed and tested. These assays provide full gene dosage information for common gene deletion polymorphisms in the *GSTM1* and *GSTT1* genes. The assays have been optimised down to small volumes (2 x 5 µl per DNA sample) enabling true high-throughput capability by providing substantial economic savings 66% of master mix and 25% of probe cost) and returning 96 - 99% of results on the first round of genotyping. In order to speed up data analysis, an excel macro-driven analysis package has been designed to process the numerous data handling and calculation steps required to generate copy number genotypes from the raw real-time PCR data. Streamlining the process in this way, with a semi-automatic system has removed the considerable bottleneck of data analysis.

Further developments would include the generation of a standalone software application to analyse real-time PCR copy number variation data, with support for other real-time PCR platforms currently available on the commercial market. This technology could also be applied to other copy number variation polymorphism and the assay design, optimisation and validation protocol described in this chapter should aid in this. In the case of further study into *GSTM1* polymorphism it would be interesting to determine if the assay has the capabilities to detect the GSTM1 gene duplication (> 2 copies) that had previously been described in a Arabic population.[199] Doing so, would require fine tuning of the genotype cut-off windows in the analysis spreadsheet as analysis would be performed on the extreme end of the data range, where inaccuracy and imprecision is at its maximum. This would provide the opportunity to test and implement a procedure for automatic genotype assignment, either by a linear regression or statistical method. These improvements could also increase the overall successful genotype call rate on first round of genotyping.

4 Chapter 4: Environmental Tobacco Smoke, Glutathione S-Transferase Gene Polymorphisms and Childhood Asthma

4.1 Introduction

Asthma is mediated by inflammation in lung tissue, reactive oxygen species (ROS) can be formed as a result of inflammatory processes and from environmental exposures. Glutathione is a tripeptide that protects cells from ROS damage through conjugation with reactive substances (section 1.5.1). Glutathione S-transferases catalyse the conjugation of glutathione to reactive substances (see section 1.5.2). Therefore, glutathione S-transferase enzymes are potent antioxidants and play an important role protecting against cellular damage caused by oxidative stress. GST genes are highly polymorphic and polymorphisms that reduce enzyme activity have been associated with increased risk from numerous cancers,[204, 280, 281] Reduced GST activity can increase cancer risk by increasing cellular sensitivity to environmental toxins and carcinogens and by modifying the efficacy and toxicity of certain anti-cancer treatments.[282, 283]

4.1.1 Oxidative Stress in the Lung

The potential for oxidative stress mediated damage in the lung is high as a consequence of breathing air. The pathology of respiratory diseases such as cystic fibrosis, COPD and asthma are characterised by an imbalance between reactive oxygen species and antioxidant defences (reviewed in Kelly, 1999).[284] In the lung, ROS arise from endogenous sources such the influx of inflammatory cells and exogenous sources such as air pollution and cigarette smoke.[284] Chronic inflammation, a characteristic of asthma, can be induced in the lung through a continuing imbalance between increased ROS levels and/or inadequate antioxidant defence mechanisms.[285] The lung has a range of antioxidant defences to protect tissue from damage by oxidative stress, employing either enzymatic or non-enzymatic mechanisms (See table 4.1).

Table 4-1: Antioxidant Defences in the Lung.

Adapted from Kirkham et al, 2006.[286]

Enzymatic antioxidants	Location in lung
Catalase	Fibroblasts, macrophages, pneumocytes
Cu,ZnSOD	Bronchial & alveolar epithelium, macrophages, neutrophils, vascular walls, pneumocytes
EC-SOD	Bronchial epithelium, macrophages, neutrophils, vascular walls, pneumocytes
Glutamate cysteine ligase	Alveolar, bronchical epithelium, macrophages
Glutathione Peroxidase	ELF cells, epithelium, macrophages and other lung cells
Glutathione S-transferases	Bronchial & alveolar epithelium, clara cells, avelolar macrophages, pneumocytes[287]
Heme Oxygenase 1	Alveolar, bronchical epithelium, macrophages, inflammatory cells of the lung
Mn,SOD	Bronchial epithelium, macrophages, neutrophils, vascular walls, pneumocytes
Thioredoxin	Bronchial epithelium, macrophages

Non-enzymatic antioxidants	Concentration in lung epithelium lining fluid (μM)
Albumin-SH	70
Ascorbic acid	100
β-Carotene	-
Glutathione	100
α-Tocopherol	2.5
Uric Acid	90

4.1.1.1 Antioxidant Vitamins in the Lung

Vitamin C (L-ascorbic acid) is an essential nutrient for humans and must be obtained from the diet, it is a potent antioxidant that is recycled back into its reduced state by glutathione; vitamin C is also required for the synthesis of collagen (reviewed by Linster et al, 2007).[288] Deficiency of this vitamin in humans causes scurvy. Vitamin C is present in the extracellular fluid lining of the lung and low intake of this vitamin has been associated with lower lung function and respiratory symptons.[289] Vitamin E is the generic term for eight related tocopherols and tocotrienols, that are antioxidant lipid soluble vitamins. These molecules are thought to protect membranes from oxidation by reacting with lipid radicals produced by lipid peroxidation (reviewed by Wang et al, 1999).[290] In the lung, vitamin E is actively secreted with surfactant lipids by type II pneumocytes and is considered to be the most important lipophilic antioxidant molecule protecting the lung surfactant lipid layer.[291]

4.1.1.2 Measuring Oxidative Stress

In cells, reactive oxygen species degrade lipids forming Malondialdehyde (MDA), this molecule is highly reactive and causes toxic stress in cells by forming protein adducts and reacts with bases in DNA to form mutagenic DNA adducts.[292] MDA is widely used as a biomarker to measure the level of oxidative stress in an organism or disease state.(Reviewed by Rio et al)[293] The level of MDA a biological sample can be determined by a colorimetric assay using thiobarbituric acid.[294]

4.1.1.3 Oxidative Stress from Tobacco Smoke

Cigarette smoke contains thousands of reactive chemical species that have damaging effects on lung tissue. (reviewed by Yoshida et al, 2007)[295] Nitric oxide (NO) is highly abundant in the gas phase of tobacco smoke and the tar phase contains reactive oxygen and nitrogen species (ROS & RNS), quinones and phenol. At physiological levels, nitric oxide acts as a key signalling molecule in virtually every cell and organ in the body, controlling signal transduction and regulation of gene expression.[296] At higher levels, nitric oxide is a highly deleterious and cytotoxic free radical.[297] Nicotine components from tobacco smoke have been shown to activate NF-κB signalling in several cell lines[298] and long-term exposure to nicotine leads to thickening of the airway wall in chronic bronchitis because of an imbalance between apoptosis and cell proliferation.[299]

4.1.1.4 Glutathione in the Lung

The antioxidant, glutathione is present in the intracellular, and vasculature systems and in respiratory tract lining fluid (RTLF).[300] Glutathione (GSH) is a key intracellular antioxidant in the lung, that is maintained at high concentrations in RTLF and levels are higher in smokers and asthmatics, and lower in people with CF, HIV and idiopathic pulmonary fibrosis (reviewed by Kelly, 1999).[284] It is clear that low levels of GSH in the RTLF will leave the surface of the lung more susceptible to damage from exogenous and endogenous sources (see figure 4.1).

However, is not known if GSH status is the cause or a consequence in these conditions.[284] It is thought that reduced GSH levels in the RTLF could lead to inactivation of the antiproteolytic screen in the airways, leaving the lung epithelium more susceptible to proteolytic attack.[284]

The GSH redox status of the airways plays a key role in the regulation of the inflammatory response in the lung to stimuli via NF-κB signalling, inducing transcription of pro-inflammatory genes (reviewed by Rahman et al, 2006).[301]

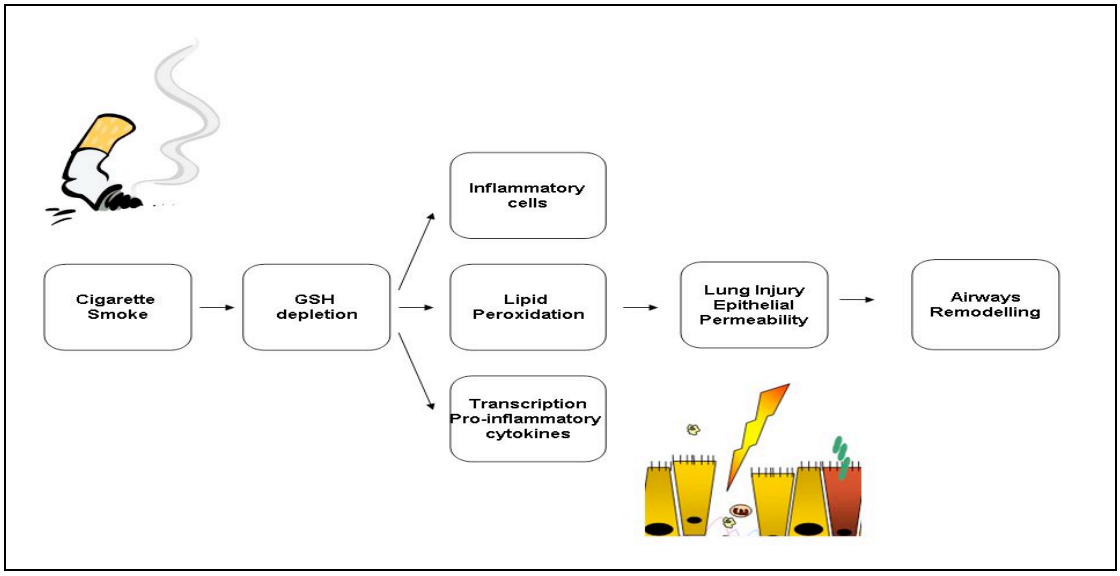


Figure 4-1: Mechanism of Cigarette Smoke Mediated Lung Injury by Glutathione Depletion

Reactive components of cigarette smoke are thought to overwhelm antioxidant defences, namely by depleting glutathione levels, whereby unbound ROS are free to damage cells and initiate airway remodelling mechanisms.

4.1.2 Reactive Oxygen Species in Asthma

ROS generation can induce many of the pathological features of asthma such as, increased arachidonic acid release, airways smooth muscle contraction, airways reactivity, secretions, vascular permeability, and synthesis of chemoattractants (reviewed in Kirkman et al, 2006).[286] Triggers for asthma exacerbations include ozone (O_3), environmental tobacco smoke and viral infections (see section 1.2). Malondialdehyde, NO and H_2O_2 can be used as markers of oxidative stress in the lung and are found to be increased in breath condensate from asthmatics.[302] ROS can stimulate histamine release from mast cells and increases mucus production by airways epithelial cells.[303] Inflammation has been shown to be driven by oxidative stress in asthmatics.[304] Inflammatory and immune cells such as, macrophages, neutrophils and eosinophils release increased amounts of ROS in asthmatic individuals.[305] Finally, ROS can directly damage epithelial cells and cause cell shedding.[306]

4.1.3 Environmental Tobacco Smoke Exposure in Early Life

Environmental tobacco smoke is a major risk factor for respiratory disease in children.[307, 308] Exposure from maternal smoking during pregnancy and the first years of life in the home has consistently been found to affect the respiratory system and is associated with symptoms such as wheezing, cough, bronchitis and development of asthma.[58, 309-311] The impact of ETS appears to be strongest during foetal development and in the first years of life, affecting intrauterine growth, lung development and lung function growth.[57, 312-314] The results from studies examining the effect of ETS exposure on atopy are not as consistent as those focussing on wheeze and asthma, indicating that non-immune related mechanisms may be more important in determining susceptibility.[315, 316] The increasing prevalence of respiratory symptoms in recent years is thought to be due to a genetic predisposition in individuals exposed to detrimental environmental exposures. Maternal smoking during pregnancy is the largest avoidable exposure and can be tackled through awareness programs aimed at expectant mothers.

4.1.4 GST Polymorphism and Risk of Asthma

The majority of genetic epidemiological studies of GST polymorphism (*GSTP1*, *M1* and *T1*) provide supporting evidence for an association with increased asthma susceptibility in children and adults. In addition, these polymorphisms, as might be expected from their function, interact with environmental oxidative stress exposures such as ETS and air pollution.

In Danish atopic families the presence of *GSTM1* and *GSTT1* gene deletions are a risk factor for the development of asthma in a dose response manner.[274] Other studies of asthma have also confirmed *GSTM/T1* gene deletion as a risk factor for asthma development.[279] *GSTM1*/*T1* polymorphism and ETS interactions have also been shown to contribute to the development of childhood asthma[317] and may leave asthmatic children more susceptible to the deleterious effects of ozone.[84] Gilliland et al.[318] showed that *GSTM1* and *GSTP1* polymorphism appears to modify the adjuvant effect of diesel exhaust particles on allergic inflammation,[318] and children with *GSTM1* deletions have also been shown to have decreased lung function growth.[319]

A recent meta-analysis of GST polymorphism in adult asthma involving 14 studies (2292 asthma patients and 5718 controls) reported an overall asthma risk of 1.20 (95% CI: .08-1.35) with the *GSTM1* null genotype.[320] When stratified for age and

smoking status this risk increased to 1.56 (95% CI: 1.25-1.94) and 1.95 (95% CI: 1.21-3.13) in non-smokers. The *GSTT1* null genotype was also associated with asthma risk in non-smoking adults (OR = 2.06, 95% CI: 1.21-3.71). Further analysis with combined GST genotypes revealed that individuals with both *GSTM1* and *GSTT1* null genotypes were at a significantly higher risk of developing asthma (2.15, 95% CI: 1.39-3.33). The authors concluded that *GSTM1* and *GSTT1* genes do not protect against the development of asthma in adults with a positive history of smoking, because chronic smoking carries such a high dose of toxins into the body that it overloads the capacity of either *GSTM1* or *GSTT1* detoxification system.[320]

Findings from *GSTP1* association studies are not so consistent. The variant *GSTP1* Val105 allele, that is associated with decreased glutathione activity,[321] has been shown to be protective against asthma,[202] atopy[322] and reduced airway responsiveness.[323] Carroll et al.[324] observed increased lung function in children carrying the *GSTP1* Val105 allele.[324] Furthermore, the same group also observed that mothers who are homozygous for *GSTP1* Val105 have children who are more likely to have increased lung function, decreased atopy and airways hyperresponsiveness.[325,326] An eleven year follow up study in the SAPALDIA cohort,[327] found there was an increased risk of progression of BHR to adult onset asthma in the Swiss general population who are homozygous for *GSTP1* Ile105.[328] The *GSTP1* Val105 genotype has been shown to interact with a high EPHX1 (Microsomal epoxide hydrolase) phenotype, increasing the risk of lifetime asthma, in children who lived within 75 metres of a major road.[329] In a relatively small study of isocyanate-induced occupational asthma in 131 individuals exposed for over 10 years, a protective effect of homozygosity for the *GSTP1* Val105 allele was observed; Val105 carriers experienced milder airways responsiveness and the protective effect increased in proportion to the duration of exposure to toluene diisocyanate.[330] However, in the published literature there are several findings that contrast with these reports; the Val105 allele has also been associated with risk of asthma,[279] increased environmental tobacco smoke effects on lung function[331] and ozone induced breathing difficulties.[332] Furthermore, several publications have also found no association between *GSTP1* Val105 genotype and asthma.[274,333]

In the last year, studies have tried to address these contrasting findings by performing comprehensive *GSTP1* haplotyping studies examining gene-environmental interactions with risk factors. *GSTP1* haplotypes have been associated with respiratory illness-related absence, decreased risk of exercise associated asthma in high ozone and particular haplotypes have been shown to modulate the association of *in utero* maternal smoking on asthma/wheezing outcomes in shoolchildren.[334-336]

Very few genetic epidemiological studies have been performed on *GSTO2* polymorphism, and the ones that have been done, generally focus on arsenic exposure and cancer risk.[337] To date there have been no studies performed that examine the role of *GSTO2* polymorphism in asthma susceptibility. However, the results from a recent large genome-wide association study of population based lung function found two *GSTO2* single nucleotide polymorphisms (rs156697 and rs156699) to be strongly associated with mean FVC and FEV_1 values.[338] A recent study of susceptibility to chronic obstructive pulmonary disease (COPD) failed to replicate the lung function findings from the genome-wide association study, but did find that the *GSTO2* polymorphism, rs156697 was associated with increased risk of developing COPD.[339]

4.1.5 Hypotheses

It is proposed that:

'*Glutathione S-transferase gene polymorphisms modulate the susceptibility to, and subsequent severity of, childhood asthma when exposed to environmental tobacco smoke in pregnancy and/or early life'.*

The work presented in this chapter attempts to test the above hypothesis by utilising the novel real-time PCR based copy number variation technique presented in Chapter 3 and further GST SNP genotyping described in this chapter. *GSTM1*, *GSTT1* CNV genotyping and *GSTO2* and *GSTP1* SNP genotyping was performed on DNA from a large UK family-based study of childhood asthma. Family based association tests were used to investigate whether GST polymorphism modulated the risk or severity of asthma and if a gene-environment interaction exists between GST genotype and tobacco smoke exposure in early life.

4.2 Aims

The aim of this work was to:

- Perform gene-only association analysis of GST gene polymorphism with asthma phenotypes in the Southampton asthma family cohort.
- Perform gene-environment interaction association analysis of GST gene polymorphism and asthma phenotypes from asthmatic siblings stratified by parent smoking status during pregnancy or early childhood.

4.3 Methods

4.3.1 Study Population

Recruitment to the Southampton asthma study began in 1997 by the Asthma Genetics group at the University of Southampton in collaboration with Genome Therapeutics Corporation (Waltham, MA, USA). In the UK, families were recruited into the study from the Southampton area, with multi-centre sites in Southampton, Portsmouth, Bournemouth and the Isle of Wight. Potential candidates were identified by general practitioners and the families were contacted in writing. Families were entered into the study if there were two siblings (5 to 21 years old) diagnosed with asthma and currently using asthma medication. Families were excluded if both parents were affected with asthma. Data was collected from 341 families, comprising 1,508 individuals. In 2002, a genome-wide linkage analysis of this study population identified a novel asthma susceptibility gene called *ADAM33*.[121]

4.3.2 Study Questionnaire

The case report form included questions on demographics, medical history including medication, a health survey on the incidence and frequency of asthma, wheezing, eczema, hay fever, nasal problems, smoking and home environment.[121]

4.3.3 Preparation of DNA samples

DNA was previously isolated from peripheral blood samples collected from the participants at recruitment and archived in the laboratory for future analysis.

1 ng/µl working concentration DNA arrays were aliquoted from the archived 100 ng/ul stock DNA arrays stored at ⁻80°C using the Biomek 3000 liquid handling system. (See appendix 10.2)

4.3.4 Asthma Phenotypes

All of the phenotype measurements were performed at recruitment. Numerous measurements were made but only six phenotypes were used in the GST CNV FBAT analysis:

- Asthma diagnosis
- Log age corrected total IgE levels
- BHR
- % Predicted FEV_1
- Positive atopy severity score
- Positive asthma severity score

4.3.4.1 Asthma Diagnosis

The criteria for being affected with asthma in this study required a physician's diagnosis of asthma and current asthma medication use.

4.3.4.2 Atopy Measurement

Total and specific IgE levels (house dust mite, cat, dog, grass) were measured from the blood samples by the ImmunoCAP™ system (Pharmacia Diagnostics). Asthmatic individuals were dichotomised using an age-specific cut-off for elevated total IgE levels (age 5–9 yr, >52 kilounits (kU) l^{-1}; age 10–14, >63 kUl^{-1}; age 15–18, >75 kUl^{-1}; age ≥ 19, >81 kUl^{-1}). An individual was assigned a positive specific IgE value if his/her level was positive (to grass or tree) or elevated (≥0.35 $kUl^{-1}$for cat, dog, *Dermatophagoides pteronyssinus, D. farinae, Alternaria*, ragweed) for at least one such measure.[121] Skin prick tests using the same allergens were also performed.[121]

4.3.4.3 Bronchial Hyperresponsiveness

Bronchial hyperresponsiveness was measured using a methacholine challenge (section 1.1.4.1.1). Bronchial responsiveness was only measured in participants with a baseline forced expired volume in 1 second (FEV_1) of ≥70% predicted. Doubling methacholine concentrations (0.06 – 16.0 mg/ml) were administered in continuous 5 minute durations until a 20% reduction in FEV_1 or the highest concentration has been reached. FEV_1 values were recorded 30 seconds after each dose and continued at 1 minute intervals thereafter. A linear model was then fitted to the recorded FEV_1 values and corresponding cumulative dose received. The least squares slope value could then be derived and used in the following equation, to provide a measure of each participant's bronchial responsiveness.

Equation 5: Calculation of BHR Variable

$$\left[\frac{1}{(\text{Least Squares slope} + 30)}\right] \times 1000$$

The reciprocal transformation was used to improve uniformity of variance, while the constant value of 30 was added to avoid negative values. Furthermore, multiplying by the constant 1000 avoided the need for three decimal places. This BHR variable has a negative correlation with BHR (i.e. larger value = less BHR).

4.3.4.4 Lung Function Measurements

A baseline FEV_1 measurement was established by monitoring the participant's performance in pulmonary function tests at 60 second intervals. Percentage predicted FEV_1 were calculated using the formula and the population-based sample used to calculate the Predicted FEV_1 values is described in section 1.1.4.1.2.

4.3.4.5 Atopy Severity Score

Each patient's atopy score was derived via an equation involving the following components:

1. The 1st principal component (PC) for mean wheal diameter from the skin prick test. This represents the magnitude of the response.
2. The numbers of positive skin prick responses representing the range of the response. If the mean wheal diameter was greater than or equal to 3mm then that particular test was considered positive.
3. The level of specific IgE response as measured by the 1st PC of specific IgE's. This represents the magnitude of the response.
4. The number of positive responses to specific IgE. This represented the range of the response.

An appropriate constant value was added to the severity score to ensure all values were positive.

4.3.4.6 Asthma Severity Score

Each patient's asthma score was calculated via an equation involving three components:

1. $$\left[\frac{1}{(\text{Least Squares slope} + 30)}\right] \times 1000$$
2. A treatment score, representing a value assigned to each patient according to the British Thoracic Society's guidelines. This was dependent on the type of medication they were receiving. A high treatment score indicates the patient is being treated for severe asthma.
3. A symptom score, derived from ten questions regarding asthma on the health survey questionnaire. The score was calculated by assigning a weight indicating the degree of importance to each question. These weights were based on the opinion of experienced professionals, namely GP's, chest physicians, paediatricians with a respiratory interest and respiratory nurse specialists.

A high asthma severity score implied the patient had severe asthma. A constant value was added to all scores to ensure all values were positive.

4.3.5 Environmental Tobacco Smoke Exposure

Children were classified as smoke exposed when the mother reported that one or more of the following situations occurred:

- the mother had smoked during pregnancy
- the mother had smoked during early childhood
- the father had smoked during early childhood

4.3.6 Genotyping Methodologies

4.3.6.1 Copy Number Variation Genotyping

GSTT1 and *GSTM1* CNV genotyping of the Southampton asthma cohort is described section 3.3.4.

4.3.6.2 SNP Genotyping by High Resolution Melting Curves

The *GSTP1* (rs1695 & rs1138272) and *GSTO2* (rs156697 & rs156699) SNPs were genotyped by High Resolution Melting Curve Analysis[340] (HRM-PCR) after PCR amplification on the Lightcycler 480 real-time PCR machine.

4.3.6.2.1 Principles of high-resolution melting curve analysis

This SNP genotyping technique incorporates a saturating dsDNA intercalating florescent dye (SYTO9, Invitrogen) into a PCR reaction (see table 4.2). Once bound to the PCR product it fluoresces greatly compared to the unbound form of the dye molecule. During the post-PCR dissociation stage (when the thermalcycler ramps slowly from 60 to 95ºC) the double stranded PCR product begins to separate and the dye dissociates losing fluorescence. The rate of decline in fluorescence is dependant on the length of the PCR product and base composition, and can be detected by the real-time PCR machine optics (melting curve analysis). A single base pair change in the product can result in a change in the rate of decline of fluorescence that can be detected by the real-time PCR instrument. The presence of a SNP in the PCR product will form a melting curve shape unique to that genotype (see figure 4.2).

Table 4-2: Advantages and Disadvantages of SNP Genotyping by High-Resolution Melting Curve Analysis

Advantages of HRM-PCR	Disadvantages of HRM-PCR
Cheap. Only requires primers and fluorescence dye. **No** expensive probes. Can detect multiple or unknown mutations in a PCR amplicon.	Resolution of the assay is affected by the type of basepair transition. i.e. for A>T changes the difference in melting temperature can be indistinguishable. However, this can be resolved by creating artificial heterozygotes by spiking PCR reactions with wild type (AA) DNA template. AA samples will generate AA melting curves and TT samples will generate heteroduplexes and therefore will present as AT melting curves.

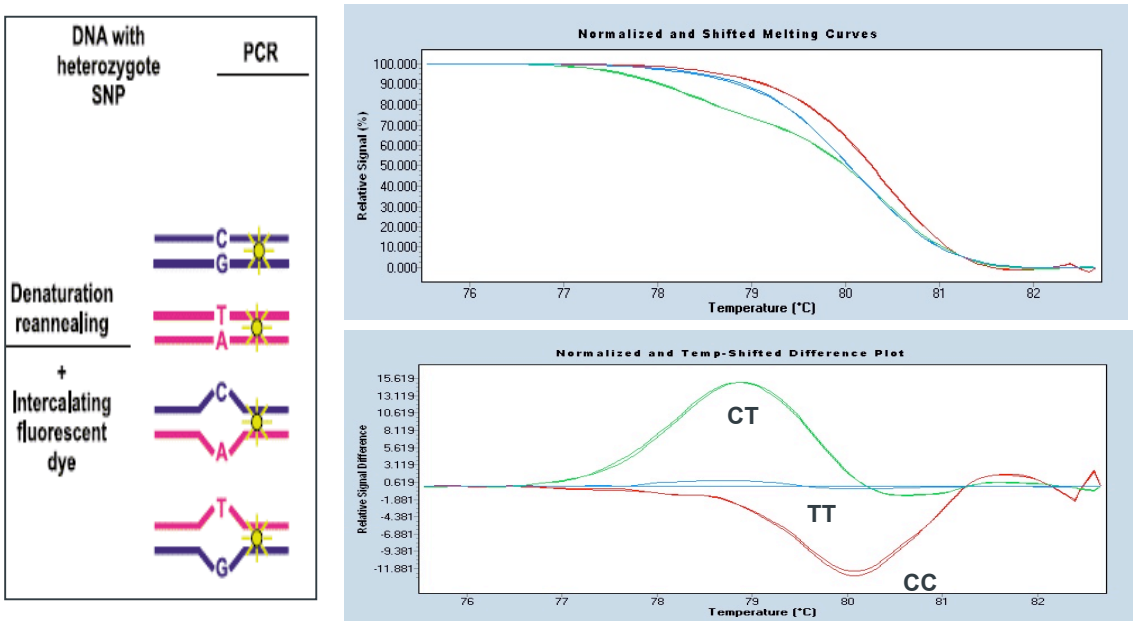


Figure 4-2: Principles of SNP Genotyping by High Resolution Melting Curve Analysis

The presence of a different single nucleotide in the short PCR amplicon (left panel) can have a profound effect on the melting temperature (top right panel). DNA with a heterozygote SNP form heteroduplex products containing a basepair mismatch that decreases the melting temperature further, allowing all three genotypes to be discriminated based on shape and temperature of melting curves (bottom right panel).

4.3.6.2.2 High-Resolution Melting Curve Genotyping Assay Design and Validation

GSTO2 and *GSTP1* SNP assays were designed using the Primer3 website (http://frodo.wi.mit.edu/primer3/) setting the design parameters for a PCR product size of less than 100 basepairs, primer length ≤ 27 basepairs with a melting temperature of 60°C. Primer sequences are shown in table 4.3. Assays were validated by comparing HRM-derived genotypes of commercially available human control DNA samples (www.coriell.org/) to publically available individual SNP genotype datasets (http://www.ncbi.nlm.nih.gov/projects/SNP/).

Where there was no individual SNP genotype for the human control DNA samples in the database; all three genotypes were confirmed by sequence validation.

Table 4-3: *GSTO2* and *GSTP1* HRM Primers

Gene	Ref sequence (Position; bp)	Forward Primer sequence (5'-3')	Reverse Primer sequence (5'-3')	Oligo Conc^{n}
GSTP1	rs1695	AGCCCTGGTGGACATGGTGA	GCCCAACCCTGGTGCAGAT	600 / 600 nM
GSTP1	rs1138272	TGGGAGGGATGAGAGTAGGA	GGGCAGTGCCTTCACATAGT	600 / 600 nM
GSTO2	rs156697	ACCTCTTCCAGGTTGCTGAA	GCCTGGTAGCGTTGAGATGT	600 / 600 nM
GSTO2	rs156699	CCGGTTTGCACTGGAACTAA	TTGAATGACACAGGGTTGACA	600 / 600 nM

4.3.6.2.3 HRM Genotyping Experiment Setup

1.5 µl of DNA template was pipeted into each well on 384 well reaction plates and dried down according to section 2.2.5.4.1. Each reaction plate included a minimum of three negative and 3 genomic DNA controls (Corriell samples; one of each SNP genotype). The HRM-PCR reaction mixture was made up in bulk volumes (see table 4.4 below), and 10 µl was transferred to each well using a multi-channel electronic pipette. The plates were sealed, spun down and incubated at room temperature for 20 minutes with occasional mixing.

Table 4-4: HRM-PCR Reagent Master Mix Setup

HRM mastermix	X1
2xQPCR mix (Eurogentec)	5.0
Forward Primer (10 µM)	0.6
Reverse Primer (10 µM)	0.6
SYTO9 dye (Invitrogen)	1.0
DNA (10 ng/ul; dried)	--
dH20	2.8
Total	10 µl

4.3.6.2.4 HRM-PCR Thermalcycling Reaction

The Lightcycler 480 real-time PCR machine thermalcycling conditions for HRM-PCR were as follows: Taq Activation: 95°C for 10 mins. 45 cycles of PCR: 95°C for 15 secs, 62°C for 30 secs and 72°C for 30 secs. Directly after the PCR reaction the post-PCR dissociation profile was performed: 95°C for 1 min, 55°C for 1 min, 60 °C for 1 sec, ramping from 60 to 95°C with 25 data acquisitions per °C/s.

4.3.6.2.5 HRM-PCR Data Analysis

Analysis of the data was performed using the Lightcycler 480 gene scanning software module v1.5 (Roche Diagnostics). The SYTO9 dye fluorescent signal intensity generated by each reaction well is normalised (vertical adjustment) and temperature-shifted (horizontal adjustment) to generate groups of melting curve shapes (see figure 4.3). Each group of melting curve data corresponds to a SNP genotype and is automatically compared to the in-run genomic DNA control samples by the analysis software. At this stage, dubious genotype calls or abnormal melting curve data can be identified and excluded.

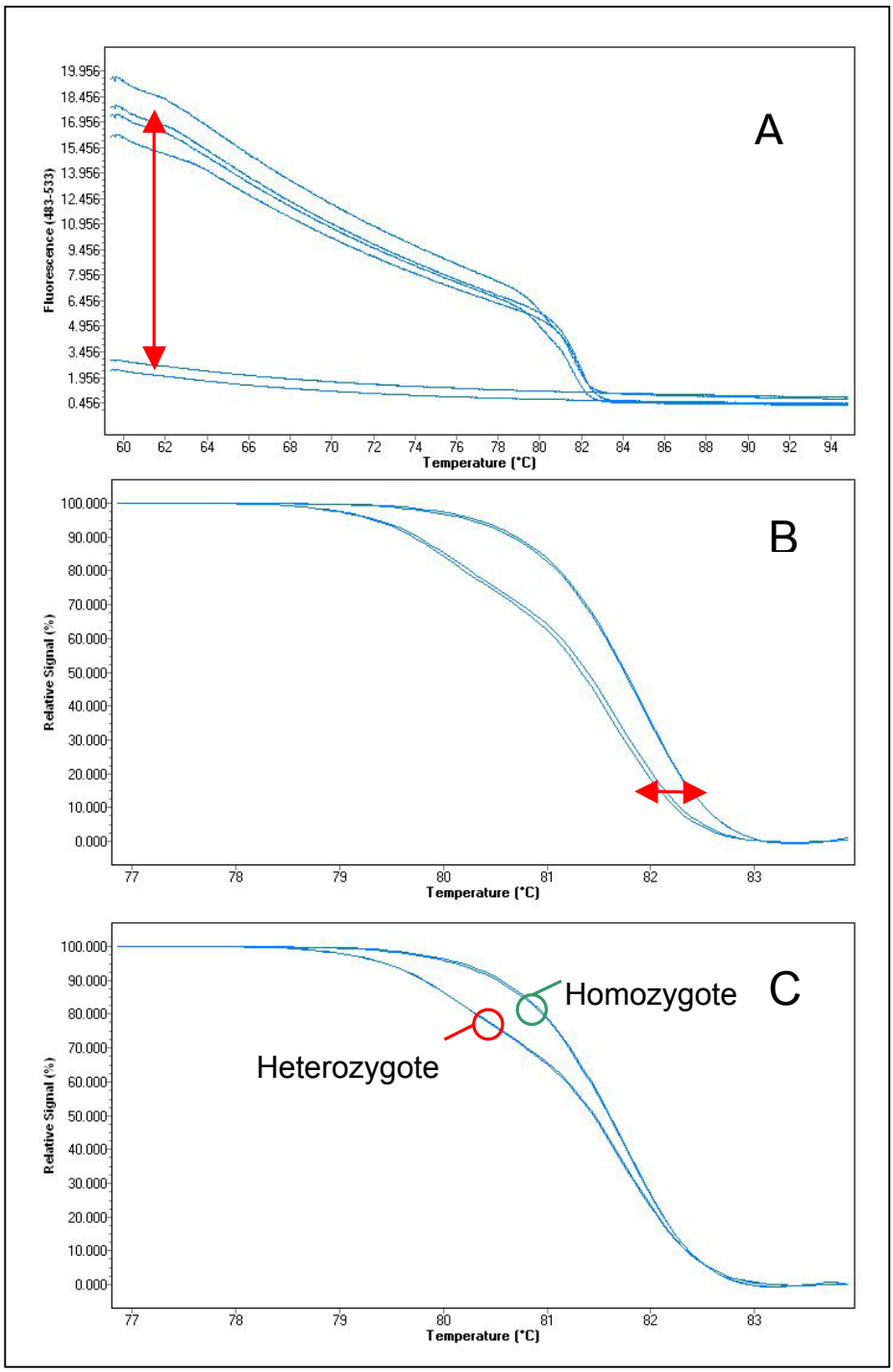


Figure 4-3: Normalisation and Temperature-Shifting of HRM Data.

Before genotype assignment can be performed the fluorescent data signals from all reactions in an experiment requires normalisation to the same intensity (Panel A). Normalised data (Panel B) is then temperature-shifted, this allows the software to distinguish any differences in the shape of the melting curves (Panel C) and assign the correct genotype.

4.3.7 Preliminary Analysis

4.3.7.1 Data Processing, Genotype Frequency and Hardy-Weinberg Equilibrium

The genotype results were imported into a database directly from the GST CNV analysis spreadsheet and duplicate sample and control DNA results were filtered from the final data set. An exclusion variable was then applied to the data set to remove excluded individuals (e.g. identical twins or those withdrawn from the study) from any further analysis. The genotypes were then recoded in to the FBAT program pedigree file format.[277, 341] The genotype frequencies for each gene were counted and Hardy-Weinberg Equilibrium was tested according to section 3.3.4.1 in the parents only.

4.3.7.2 Inheritance Checks

Mendelian inheritance checks are automatically performed by the FBAT program when the pedigree file is loaded into the software. Families with any inheritance errors are immediately recoded by the program and are excluded from any further FBAT analysis.

4.3.8 Family Based Association Tests

This section details the Family Based Association Test (FBAT) method that was used to analyse GST copy number variation genotypes and the asthma phenotypes described in the last section. The FBAT software also facilitates the exploration of any relationship between the smoking status of the parents, GST polymorphism and asthma in their children.

The FBAT association and regression analyses described in this chapter were performed with the assistance of Mr. Jake Kashani and Ms. Sheila Barton.

4.3.8.1 Principles of Transmission Disequilibrium Test

A transmission disequilibrium test (TDT) uses genotype data from families with at least one affected child and evaluates the transmission of the associated marker allele from an informative parent to an affected offspring. TDT tests can detect linkage between the marker locus and the disease locus only if association (due to linkage disequilibrium) is present.

As the test compares the number of times a heterozygous parent transmits the associated marker to an affected offspring to the number of times they transmit the alternate marker allele within the family pedigree, this test is not affected by the presence of population stratification.[342]

4.3.8.2 The FBAT Software

FBAT version 1.4 (table 2.11 for details) is a computer program designed by Rabinowitz and Laird,[277, 341] that builds on the TDT method[342] and integrates tests of different genetic models, sample designs, disease phenotypes, missing parents and tests of different null hypotheses.

4.3.8.2.1 Regression Analysis Using the FBAT Software

FBAT can be modified to include environmental covariates through a covariate-adjusted phenotype defined as the continuous or dichotomous covariate subtracted from the dichotomous indicator of affection status.[343] It is possible to develop a susceptibility score using residuals derived from a logistic regression model of affection status as a function of only known environmental predictors (see table 4.5).

Table 4-5: FBAT Regression Analysis Outcomes

Residual values	Inference
Highly positive	**Affected with low environmental risk** = *Strong Genetic predisposition*
Close to zero	**Environmental risk predicts the outcome**. *Therefore it is unlikely that genotype information would provide additional information to the association test.*
Highly negative	**Unaffected with high environmental risk** = *No genetic predisposition*

In this analysis, instead of modelling affection status the residuals were derived from a linear regression model of the phenotype variables as functions of exposure to smoke. Pearson/Standardised residuals were calculated from the asthma phenotype measures (dependant variables) and exposure to smoke was expressed as a dichotomy (independent variable).

The advantage of using the regression approach is that the power of the statistical tests is often improved because the number of informative families present at each stage of the analysis remained relatively large, whilst taking account of the differences in exposure to cigarette smoke.

4.3.8.3 Haplotype Based Association Tests (HBAT) Using the FBAT Software

An extension of the FBAT analysis program is the ability to test for linkage and association between a haplotype and a disease susceptibility locus.[277] In the presence of multiple tightly linked SNP markers a FBAT test using haplotypes may be more beneficial than using the markers one by one as in the standard FBAT test procedure as less statistical tests are performed.

The *GSTP1* and *GSTO2* SNPs are close together on their respective genes and form bi-allelic haplotypes. Therefore, these SNPs were analysed under the HBAT protocol.[344] *GSTP1* and *GSTO2* haplotypes are constructed by the FBAT program using an EM algorithm.(See Section 1.4.1.5 regarding EM algorithms)

4.3.8.4 FBAT Association Analysis Plan

This study aims to determine whether there is an association between the asthma phenotypes described in section 4.3.4 and GST genotype status, stratified by whether the parents smoked during pregnancy or early childhood.

The analysis plan is broken down as follows:

1. Perform gene only association tests under a null hypothesis of no linkage and no association. Association analysis for the diagnosis of asthma was performed on asthmatic and non-asthmatic individuals (siblings); while the remaining asthma phenotypes were tested for association in asthmatic siblings only.
2. *Gene-environment analysis*: Stratify the genotype and phenotype data for the asthmatic siblings by parental smoking status during pregnancy and early childhood and perform and compare the association analysis results for the smoke exposed and unexposed asthmatic siblings. Note: some families may have appeared in each group as the parents smoking status may have changed during different pregnancies.
3. Finally, perform regression analysis using FBAT using the smoking variable as a covariate in the series of regression models for each of the five asthma phenotypes (Log age corrected total IgE levels, BHR, %Predicted FEV_1, Positive atopy severity score, Positive asthma severity score) using *SPLUS version 6.2 Professional* (www.insightful.com/products/splus/default).

FBAT test hypotheses were as follows:

H_1 = association between the presence or absence of the gene and the asthma phenotype

H_0 = 'no association and no linkage' between the presence or absence of the gene and the asthma phenotype

As there is more than one affected sibling in some families the null hypothesis was set to 'no association and no linkage' so that the family pedigree can be broken down into nuclear families that can be treated as independent.

4.3.9 Post-Hoc PBAT Power Analysis

PBAT is an integrated FBAT and study design software package that has conditional power calculation functions for virtually any type of family based study design.[345, 346] The program can handle different types of traits (dichotomous/binary or continuous), family structures and genetic models and all power calculation results can be verified by Monte-Carlo simulations. As this study is a retrospective study of asthma susceptibility and examines a gene-environment interaction, two estimates of power are required. The first calculation estimates the power of the study to detect a gene only association with the primary outcome (asthma diagnosis) and the second estimates the power to detect an association with asthma diagnosis in smoke exposed individuals. PBAT power calculation settings were as follows:

Table 4-6: PBAT Power Calculation Settings

Parameters	Gene only	Smoke Exposed (30 % of families)
Study Design:		
no. of siblings per family =	2	2
no. of missing parents =	0	0
no. of families =	341	102
ascertainment condition for proband 1 =	1	1
ascertainment condition for proband 1 =	1	1
Genetic Model:		
Allele freq. of disease gene =	0.53/0.17[†]	0.53/0.17[†]
Increment in allele freq. Per iteration =	0	0
Mode Of Inheritance =	0 (Additive)	0 (Additive)
Odds Ratio of disease allele =	1.20/1.36[a]	1.6/unknown[a]
Is disease locus = Marker locus	1 (Yes)	1 (Yes)
Statistical Parameters:		
Significance level =	0.05	0.05
Offset value =	0 (default)	0 (default)
Sigificance level =	0.05	0.05
Power Calculation option:		
1 = Numerical integration;	2	2
2 = Approximation; 3= Simulation (Monte carlo)		

[†](GSTM1/GSTT1). [a]Odds ratios from a published meta-analysis of GST polymorphism, smoke exposure and asthma. (Saadat et al 2007)[347]

4.4 Results

4.4.1 Preliminary Genotype Analysis Results

Table 4.7 shows the CNV genotype frequencies of the final data set of 1456 individuals available for FBAT analysis after merging the pedigree and phenotype data files.

Table 4-7: GST Copy Number Variation Genotype Frequencies in the Final Data Set

Gene	2 copies of gene n (%)	1 copy of gene n (%)	0 copies of gene n (%)	Total number of genotypes n (%)	Undetermined genotypes n (%)
GSTM1	97 (6.7)	528 (36.3)	773 (53.1)	1398 (96.0)	58 (4.0)
GSTT1	407 (30.0)	683 (47.0)	242 (16.6)	1332 (91.5)	128 (8.5)

The FBAT program identified a number of non-mendelian inheritances within families in the final data set. The number and type of non-mendelian inheritances present are described in section 3.4.2.3; in order to proceed with the FBAT analysis all samples from the affected family were excluded from the data set. Hardy-Weinberg equilibrium calculations were performed on the parent subset as described in section 3.3.4.1 and both GST CNV genotype frequencies did not deviate from the expected distribution.

Table 4.8 shows the genotype frequencies of the final data set for the *GSTO2* and *GSTP1* SNPs, the number of undetermined alleles were less than 5%. As with the CNV genotypes, a number of non-mendelian inheritances were present within some families. Data from these families were excluded prior to FBAT analysis.

Table 4-8: *GSTO2* and *GSTP1* SNP Allele Frequencies

Gene	SNP (allele1>allele2)	Allele 1 n (%)	Allele 2 n (%)	Total number of Alleles n (%)	Undetermined alleles n (%)
GST02	*rs156697 (T>C)*	1945 (68)	905 (32)	2850 (96.5)	52 (3.5)
	rs156699 (T>C)	2048 (72)	796 (28)	2844 (96.3)	55 (3.7)
GSTP1	*rs1695 (A>G)*	2030 (70)	880 (30)	2910 (98.5)	22 (1.5)
	rs1138272 (C>T)	2606 (90)	274 (10)	2880 (97.5)	37 (2.5)

4.4.1.1 *GSTO2* and *GSTP1* Haplotype Frequencies

Haplotype frequencies for the *GSTO2* SNPs are shown in table 4.9 and were different to previously published frequencies from the HapMap database.[147]

Table 4-9: *GSTO2* FBAT Haplotype Frequencies

GSTO2 Haplotype	Alleles (rs156697 rs156699)	Frequency	Published Freq.
h1	T T	0.65	0.59
h2	C C	0.23	0.36
h3	C T	0.07	0.04
h4	T C	0.04	--

Haplotype frequencies for the *GSTP1* SNPs are shown in table 4.10 and were similar to previously published frequencies from population-based case-control studies conducted in Sweden, southeast England, Denmark, and Finland.[348]

Table 4-10: *GSTP1* FBAT Haplotype Frequencies

GSTP1 Haplotype	Alleles (rs1695 rs1138272)	Frequency	Published Freq
h1	A C	0.67	0.65
h2	G C	0.24	0.27
h3	G T	0.08	0.08
h4	A T	0.01	--

4.4.2 Environmental Tobacco Smoke Exposure

Out of a total of 781 siblings who participated in the study 93 were exposed to smoke during pregnancy, 192 were exposed to their mother smoking during childhood, and 124 were exposed to their father smoking during childhood. In total of 240 siblings (30.7%) were defined as being exposed to cigarette smoke in early life.

4.4.3 FBAT Analysis Results

4.4.3.1 Gene Only Association Analyses

4.4.3.1.1 Diagnosis of Asthma

There is no significant association at the 5% level between diagnosis of asthma and the GST CNV genotypes (see table 4.11). Comparison of the test statistic (S) against the expected test statistic (E(S)) for both genes reveals that there are more null alleles (0 copy) present than expected under the null hypothesis.

Table 4-11: FBAT Results from Association Analysis Involving *GSTM1*, *GSTT1* CNV Genotypes and the Diagnosis of Asthma.

Marker	No. of Copies	Allelic Frequency	No. of Informative Families	S[a]	E(S)	Var(S)	Z	P
GSTM1	1	0.262	199	252	260.8	110.8	-0.839	0.401
	0	0.738	199	**508**	**499.2**	110.8	0.839	0.401
GSTT1	1	0.573	224	450	468.7	146.2	-1.544	0.123
	0	0.427	224	**414**	**395.3**	146.2	1.544	0.123

[a]*Where S is the test statistic; E(S) and Var(S) are the expected value and variance of the test statistic under H_0; Z and P are the Z statistic (S normalized using E(S) and Var(S)) and its corresponding p-value.*

There is a rare *GSTO2* haplotype (h4; freq = 0.04) that is significantly associated (uncorrected P-value = 0.046) with the diagnosis of asthma phenotype (see table 4.12). Comparison of the test statistic (S) against the expected test statistic (E(S)) for the *GSTO2* h4 haplotype shows that there are less h4 haplotypes present than expected under the null hypothesis.

Table 4-12: FBAT Results from Association Analysis Involving *GSTO2* and *GSTP1* Haplotypes and the Diagnosis of Asthma.

Marker	Haplotype	Frequency	No. of Informative Families	S	E(S)	Var(S)	Z	P
GSTP1	h1	0.669	146	381.990	362.309	156.901	1.571	0.116
	h2	0.239	127	157.010	167.858	141.421	-0.912	0.361
	h3	0.081	60	61.990	70.076	59.303	-1.050	0.293
	h4	0.012	16	14.010	14.758	9.372	-0.244	0.807
GSTO2	h1	0.652	115	322.701	312.567	110.296	0.965	0.334
	h2	0.230	101	122.701	125.174	97.507	-0.250	0.802
	h3	0.075	43	52.299	49.524	43.059	0.423	0.672
	h4	0.043	26	**18.299**	**28.735**	27.426	-1.993	**0.046**

4.4.3.1.2 Log Age-Corrected Total IgE Levels

There is no significant association at the 5% between log age-corrected total IgE levels in siblings diagnosed with asthma. Inspection of the test statistic (see table 4.13) indicates that individuals with at least one copy of a GST gene have lower total IgE levels compared to the expected values under the null hypothesis.

Table 4-13: FBAT Results from Association Analysis Involving *GSTM1*, *GSTT1* CNV Genotypes and the Log Age-Corrected Total IgE Levels in Asthmatic Siblings.

Marker	No. of Copies	Allelic Frequency	No. of Informative Families	S	E(S)	Var(S)	Z	P
GSTM1	1	0.262	208	**496.8**	**527.8**	545.4	-1.327	0.185
	0	0.738	208	981.1	950.1	545.4	1.327	0.185
GSTT1	1	0.571	229	**870.1**	**901.7**	770.7	-1.136	0.256
	0	0.429	229	817.5	785.9	770.7	1.136	0.256

There is a rare *GSTO2* haplotype (h4; freq = 0.04) that is significantly associated (uncorrected P-value = 0.01) with log age-corrected total IgE levels in siblings diagnosed with asthma (see table 4.14). Inspection of the FBAT test statistic indicates that individuals with a *GSTO2* h4 haplotype have lower total IgE levels compared to the expected values under the null hypothesis.

Table 4-14: FBAT Results from Association Analysis Involving *GSTO2* and *GSTP1* Haplotypes and the Log Age-Corrected Total IgE Levels in Asthmatic Siblings

Marker	Haplotype	Frequency	No. of Informative Families	S	E(S)	Var(S)	Z	P
GSTP1	h1	0.669	213	1128.968	1080.413	945.247	1.579	0.114
	h2	0.239	197	438.045	474.181	805.373	-1.273	0.202
	h3	0.081	96	194.263	205.237	344.714	-0.591	0.554
	h4	0.012	20	21.788	23.233	34.389	-0.246	0.805
GSTO2	h1	0.652	182	1034.040	1008.784	620.825	1.014	0.311
	h2	0.23	168	386.747	400.462	543.412	-0.588	0.556
	h3	0.075	61	185.570	162.303	279.048	1.393	0.163
	h4	0.043	39	**61.306**	**96.113**	180.381	-2.592	**0.010**

4.4.3.1.3 Bronchial Hyperresponsiveness

Association analysis of the bronchial hyperresponsiveness variable was non-significant at the 5% level. Examination of the test statistic revealed no discernable trend compared to the expected test statistic under the null hypothesis (see table 4.15 & 4.16).

Table 4-15: FBAT Results from Association Analysis Involving *GSTM1*, *GSTT1* CNV Genotypes and the Measure of BHR on the Asthmatic Siblings.

Marker	No. of Copies	Allelic Frequency	No. of Informative Families	S	E(S)	Var(S)	Z	P
GSTM1	1	0.262	206	3726.1	3643.7	31267.2	0.467	0.641
	0	0.738	206	6099.2	6181.7	31267.2	-0.467	0.641
GSTT1	1	0.571	227	5693.0	6021.7	41501.4	-1.614	0.107
	0	0.429	227	5195.2	4866.4	41501.4	1.614	0.107

Table 4-16: FBAT Results from Association Analysis Involving *GSTO2* and *GSTP1* Haplotypes and the Measure of BHR on the Asthmatic Siblings.

Marker	Haplotype	Frequency	No. of Informative Families	S	E(S)	Var(S)	Z	P
GSTP1	h1	0.669	214	9043.048	8676.365	52782.440	1.596	0.110
	h2	0.239	197	3936.955	4221.142	48037.139	-1.297	0.194
	h3	0.081	96	1566.882	1658.404	20521.335	-0.639	0.522
	h4	0.012	21	265.092	256.067	3471.073	0.153	0.878
GSTO2	h1	0.652	183	8267.687	7996.126	47388.469	1.247	0.212
	h2	0.23	170	3438.961	3634.717	46778.193	-0.905	0.365
	h3	0.075	61	1236.262	1211.921	20504.296	0.170	0.865
	h4	0.043	40	663.595	763.741	6530.893	-1.239	0.215

4.4.3.1.4 Percentage Predicted FEV_1 Values

Association analysis using the percent predicted FEV_1 phenotype shows a strong non-significant trend (P = 0.068) towards the presence of a copy of the *GSTT1* gene being associated with a lower percentage predicted FEV_1 value (worse lung function) than expected under the null hypothesis (see table 4.17; smaller S value than expected E(S)).

Table 4-17: FBAT Results from Association Analysis Involving *GSTM1*, *GSTT1* CNV Genotypes and Percentage Predicted FEV1 Values on Asthmatic Siblings.

Marker	No. of Copies	Allelic Frequency	No. of Informative Families	S	E(S)	Var(S)	Z	P
GSTM1	1	0.262	209	25933	26876.5	1003691.8	-0.942	0.346
	0	0.738	209	50851	49907.5	1003691.8	0.942	0.346
GSTT1	1	0.571	230	**44261**	**46361.0**	1325268.0	-1.824	**0.068**
	0	0.429	230	41147	39047.0	1325268.0	1.824	**0.068**

Association analysis using the percent predicted FEV_1 phenotype shows a significant association (uncorrected P value = 0.038) towards a *GSTO2* h4 haplotype being associated with a lower percentage predicted FEV_1 value (worse lung function) than expected under the null hypothesis (see table 4.18; smaller S value than expected E(S)).

Table 4-18: FBAT Results from Association Analysis Involving *GSTO2* and *GSTP1* Haplotypes and Percentage Predicted FEV_1 Values on Asthmatic Siblings.

Marker	Haplotype	Frequency	No. of Informative Families	S	E(S)	Var(S)	Z	P
GSTP1	h1	0.669	214	60930.179	59356.212	1670589.513	1.218	0.223
	h2	0.239	196	26647.821	27474.455	1541932.247	-0.666	0.505
	h3	0.081	97	10868.179	11601.078	600675.041	-0.946	0.344
	h4	0.012	21	1499.821	1514.255	85322.529	-0.049	0.960
GSTO2	h1	0.652	183	56179.845	54491.765	1233602.038	1.520	0.128
	h2	0.23	169	22361.845	23518.537	1112849.537	-1.096	0.272
	h3	0.075	61	8671.155	8065.866	557730.342	0.810	0.417
	h4	0.043	40	**3866.155**	**5002.831**	301345.059	-2.071	**0.038**

4.4.3.1.5 Atopy Severity Score

There is no significant association at the 5% level between the atopy severity score phenotype and GST polymorphism (see tables 4.19 & 4.20) in the gene-only analyses.

Table 4-19: FBAT Results from Association Analysis Involving *GSTM1*, *GSTT1* CNV Genotypes and Atopy Severity Scores on Asthmatic Siblings.

Marker	No. of Copies	Allelic Frequency	No. of Informative Families	S	E(S)	Var(S)	Z	P
GSTM1	1	0.262	187	317.0	336.4	272.1	-1.180	0.238
	0	0.738	187	617.6	598.1	272.1	1.180	0.238
GSTT1	1	0.571	207	600.6	620.1	420.8	-0.951	0.341
	0	0.429	207	531.9	512.4	420.8	0.951	0.341

Table 4-20: FBAT Results from Association Analysis Involving *GSTO2* and *GSTP1* Haplotypes and Atopy Severity Scores on Asthmatic Siblings.

Marker	Haplotype	Frequency	No. of Informative Families	S	E(S)	Var(S)	Z	P
GSTP1	h1	0.669	199	762.698	730.210	461.643	1.512	0.130
	h2	0.239	181	277.697	293.311	370.212	-0.811	0.417
	h3	0.081	89	103.076	120.376	177.551	-1.298	0.194
	h4	0.012	20	14.886	14.461	14.687	0.111	0.911
GSTO2	h1	0.652	169	608.658	613.204	272.360	-0.275	0.782
	h2	0.23	158	250.980	243.633	231.726	0.483	0.629
	h3	0.075	55	105.709	93.379	115.299	1.148	0.250
	h4	0.043	39	41.959	57.090	78.175	-1.711	0.087

4.4.3.1.6 Asthma Severity Score

A highly significant association exists between the *GSTT1* CNV genotype and the asthma severity score phenotype (P = 0.015; table 4.21). It appears from the test statistic that the asthma severity score variable is higher in asthmatic siblings with no copies of the *GSTT1* gene (0 copy genotype) than expected under the null hypothesis (S = 1631.7 vs E(S) = 1513.4). Therefore, it would appear that inheritance of one or more copies of the *GSTT1* gene would protect against the development of more severe asthma in this cohort.

Table 4-21: FBAT Results from Association Analysis Involving *GSTM1*, *GSTT1* CNV Genotypes and Asthma Severity Scores on Asthmatic Siblings.

Marker	No. of Copies	Allelic Frequency	No. of Informative Families	S	E(S)	Var(S)	Z	P
GSTM1	1	0.262	201	1027.4	1076.8	1873.3	-1.142	0.254
	0	0.738	201	2059.1	2009.7	1873.3	1.142	0.254
GSTT1	1	0.571	217	1705.1	1823.5	2351.4	-2.441	**0.015**
	0	0.429	217	**1631.7**	**1513.4**	2351.4	2.441	**0.015**

A significant association exists between the *GSTP1* h1 haplotype and the asthma severity score phenotype (P = 0.047; see table 4.22). It appears from the test statistic that the asthma severity score variable is higher in asthmatic siblings with a *GSTP1* h1 haplotype than expected under the null hypothesis Therefore, it would appear that inheritance of a *GSTP1* h1 haplotype would protect against the development of more severe asthma in this cohort.

Table 4-22: FBAT Results from Association Analysis involving *GSTO2* and *GSTP1* Haplotypes and Asthma Severity Scores on Asthmatic Siblings.

Marker	Haplotype	Frequency	No. of Informative Families	S	E(S)	Var(S)	Z	P
GSTP1	h1	0.669	211	**2351.772**	**2244.730**	2906.641	1.985	**0.047**
	h2	0.239	197	992.810	1060.992	2675.141	-1.318	0.187
	h3	0.081	95	402.109	436.905	1150.247	-1.026	0.304
	h4	0.012	21	60.813	64.877	155.648	-0.326	0.744
GSTO2	h1	0.652	182	2110.397	2065.569	1999.401	1.003	0.316
	h2	0.23	169	868.217	903.698	1740.319	-0.851	0.395
	h3	0.075	61	323.201	308.839	739.305	0.528	0.597
	h4	0.043	40	153.709	177.419	315.779	-1.334	0.182

4.4.3.2 FBAT Analysis Adjusted for Smoke Exposure

Note: Only summary tables for significant or suggestive results are presented in this results section. The full FBAT result output of the smoke exposure analysis can be found in appendix 10.5.

4.4.3.2.1 Smoke Exposed Individuals

There is no significant association at the 5% level between GST CNV genotypes and the diagnosis of asthma (table 4.23), BHR, percentage predicted FEV_1 or atopy severity score phenotypes in smoke exposed individuals. It is clear from table 4.23, that the number of informative families has been substantially reduced as the majority of the siblings were not exposed to tobacco smoke during pregnancy or early childhood. This will obviously reduce the power of any association tests to detect any significant differences between genotypes in the smoke exposed data subset.

Table 4-23: FBAT Results from Association Analysis Involving *GSTM1*, *GSTT1* CNV Genotypes and the Diagnosis of Asthma on Smoke Exposed Individuals.

Marker	No. of Copies	Allelic Frequency	No. of Informative Families	S	E(S)	Var(S)	Z	P
GSTM1	1	0.262	62	74	73	30.5	0.181	0.856
	0	0.728	62	140	141	30.5	-0.181	0.856
GSTT1	1	0.571	73	127	136	45.3	-1.412	0.158
	0	0.429	73	133	123	45.3	1.412	0.158

Even with reduced statistical power due to a small sample size, a significant association was detected between log age-corrected total IgE levels and the *GSTT1* CNV genotype (P = 0.041). On examination of the test statistic (S) from table 4.24, it is apparent that the presence of a copy of the *GSTT1* gene associates with lower log age-corrected IgE levels in smoke exposed individuals compared to the expected test statistic (E(S)) under the null hypothesis.

Table 4-24: FBAT Results from Association Analysis Involving *GSTT1* CNV Genotypes and the Log Age-Corrected Total IgE Levels on Smoke Exposed Individuals.

Marker	No. of Copies	Allelic Frequency	No. of Informative Families	S	E(S)	Var(S)	Z	P
GSTT1	1	0.571	81	**258.3**	**289.6**	234.1	-2.048	**0.041**
	0	0.429	81	287.8	256.5	234.1	2.048	**0.041**

There is a suggestive trend towards significance at the 5% level between the asthma severity score phenotype variable and the *GSTT1* CNV genotype (table 4.25). Examination of the test statistic values suggests that the presence of a copy of the *GSTT1* gene is protective against the development of more severe asthma than predicted by the null hypothesis test statistic.

Table 4-25: FBAT Results from Association Analysis Involving *GSTT1* CNV Genotypes and Asthma Severity Scores on Smoke Exposed Individuals.

Marker	No. of Copies	Allelic Frequency	No. of Informative Families	S	E(S)	Var(S)	Z	P
GSTT1	1	0.571	76	**535.5**	**584.4**	807.5	-1.719	0.086
	0	0.429	76	596.4	547.5	807.5	1.719	0.086

There is no significant association at the 5% level between *GSTP1* and *GSTO2* haploypes and the diagnosis of asthma, log age corrected IgE levels, BHR, percentage predicted FEV_1 or atopy severity score phenotypes in smoke exposed individuals. In some of HBAT tests there are less than ten informative families for the *GSTP1* h4 haplotype; this is the sample size cut off for calculating a reliable HBAT test statistic and indicates a substantial reduction in power when stratifying by smoke exposure.

4.4.3.2.2 Unexposed Individuals

There is no significant association at the 5% level between GST CNV genotypes and the diagnosis of asthma (table 4.26), log age-corrected IgE levels, BHR, percentage predicted FEV_1 or atopy severity score phenotypes in unexposed individuals.

Table 4-26: FBAT Results from Association Analysis Involving *GSTM1*, *GSTT1* CNV Genotypes and the Diagnosis of Asthma on Individuals Unexposed to Smoke.

Marker	No. of Copies	Allelic Frequency	No. of Informative Families	S	E(S)	Var(S)	Z	P
GSTM1	1	0.262	140	172	182.5	78.3	-1.187	0.235
	0	0.738	152	358	347.5	78.3	1.187	0.235
GSTT1	1	0.571	152	307	315.0	96.5	-0.814	0.415
	0	0.429	152	261	253.0	96.5	0.814	0.415

There is a rare *GSTO2* haplotype (h4; freq = 0.04) that is significantly associated (uncorrected P-value = 0.03) with the diagnosis of asthma phenotype in unexposed siblings (see table 4.27). Comparison of the test statistic (S) against the expected test statistic (E(S)) for the *GSTO2* h4 haplotype shows that there are less h4 haplotypes present than expected under the null hypothesis.

Table 4-27: FBAT Results from Association Analysis Involving *GSTO2* and *GSTP1* Haplotypes and the Diagnosis of Asthma on Individuals Unexposed to Smoke.

Marker	Haplotype	Frequency	No. of Informative Families	S	E(S)	Var(S)	Z	P
GSTP1	h1	0.666	109	255.439	244.932	106.755	1.017	0.309
	h2	0.237	90	113.561	117.734	95.634	-0.427	0.669
	h3	0.085	47	46.439	53.366	41.392	-1.077	0.281
	h4	0.011	11	10.561	9.968	5.102	0.263	0.792
GSTO2	h1	0.655	89	236.521	226.546	83.407	1.092	0.274
	h2	0.241	78	87.521	92.379	70.465	-0.579	0.562
	h3	0.065	29	36.479	32.787	29.890	0.675	0.499
	h4	0.039	15	**6.479**	**15.287**	16.427	-2.173	**0.029**

The *GSTO2* h4 haplotype is significantly associated (uncorrected P value = 0.017; see table 4.28) with log age-corrected total IgE levels in unexposed siblings diagnosed with asthma. Inspection of the FBAT test statistic indicates that individuals with a *GSTO2* h4 haplotype have lower total IgE levels compared to the expected values under the null hypothesis

Table 4-28: FBAT Results from Association Analysis Involving *GSTO2* and *GSTP1* Haplotypes and Log Age-Corrected IgE Levels in Individuals Unexposed to Smoke.

Marker	Haplotype	Frequency	No. of Informative Families	S	E(S)	Var(S)	Z	P
GSTP1	h1	0.669	155	776.733	735.765	665.723	1.588	0.112
	h2	0.239	141	294.872	332.293	563.825	-1.576	0.115
	h3	0.081	73	155.072	159.110	266.567	-0.247	0.804
	h4	0.012	12	17.786	17.296	27.858	0.093	0.926
GSTO2	h1	0.652	129	696.135	672.247	441.729	1.137	0.255
	h2	0.23	123	268.213	287.347	393.406	-0.965	0.334
	h3	0.075	39	111.502	91.438	167.192	1.552	0.120
	h4	0.043	23	**28.222**	**53.041**	107.453	-2.394	**0.017**

Association analysis in the unexposed asthmatic siblings, using the percent predicted FEV_1 phenotype shows a significant association (uncorrected P value = 0.041) towards a *GSTO2* h4 haplotype associating with a lower percentage predicted FEV_1 value (worse lung function) than expected under the null hypothesis (see table 4.29).

Table 4-29: FBAT Results from Association Analysis Involving *GSTO2* and *GSTP1* Haplotypes and Percentage Predicted FEV_1 in Individuals Unexposed to Smoke.

Marker	Haplotype	Frequency	No. of Informative Families	S	E(S)	Var(S)	Z	P
GSTP1	h1	0.669	153	40208.419	39176.305	1093976.398	0.987	0.323
	h2	0.239	139	18713.581	19027.362	1015761.125	-0.311	0.755
	h3	0.081	73	7748.419	8542.338	392043.780	-1.268	0.204
	h4	0.012	12	1067.581	991.995	51541.776	0.333	0.739
GSTO2	h1	0.652	128	37768.353	36189.472	931511.378	1.636	0.101
	h2	0.23	122	14941.353	16388.472	823730.751	-1.594	0.110
	h3	0.075	38	5515.647	4742.528	396095.732	1.228	0.219
	h4	0.043	23	**1922.647**	**2827.528**	196475.953	-2.041	**0.041**

Association analysis in the unexposed asthmatic siblings, using the atopy severity score shows a significant association (uncorrected P value = 0.038) towards a *GSTO2* h4 haplotype associating with a lower positive atopy severity score (less atopic) than expected under the null hypothesis (see table 4.30).

Table 4-30: FBAT Results from Association Analysis Involving *GSTO2* and *GSTP1* Haplotypes and Atopy Severity Score in Individuals Unexposed to Smoke.

Marker	Haplotype	Frequency	No. of Informative Families	S	E(S)	Var(S)	Z	P
GSTP1	h1	0.669	142	525.096	496.779	306.430	1.618	0.105
	h2	0.239	132	183.842	203.121	235.699	-1.256	0.209
	h3	0.081	68	83.262	94.289	144.796	-0.916	0.359
	h4	0.012	12	12.874	10.885	12.166	0.570	0.568
GSTO2	h1	0.652	120	404.091	403.926	185.138	0.012	0.990
	h2	0.23	114	171.281	168.649	159.406	0.208	0.834
	h3	0.075	35	60.025	49.639	56.671	1.380	0.167
	h4	0.043	22	**18.444**	**31.626**	35.573	-2.210	**0.027**

The asthma severity score phenotype association analysis result for unexposed individuals provided a suggestive trend towards significance with *GSTT1* CNV genotype (table 4.31). This suggestive result mirrors the significant association result from the gene only analyses where presence of a *GSTT1* gene is protective against the development of severe asthma.

Table 4-31: FBAT Results from Association Analysis Involving *GSTT1* CNV Genotypes and the Asthma Severity Scores on Individuals Unexposed to Smoke.

Marker	No. of Copies	Allelic Frequency	No. of Informative Families	S	E(S)	Var(S)	Z	P
GSTT1	1	0.571	153	**1169.5**	**1239.1**	1558.1	-1.761	**0.078**
	0	0.429	153	1053.4	965.9	1558.1	1.761	**0.078**

4.4.3.3 Regression Analysis: Smoke Exposure as a Covariate

4.4.3.3.1 Log Age-Corrected Total IgE Levels

No significant association was detected between the GST CNV genotypes and total IgE levels from the model involving the log age-corrected total IgE levels as a function of the sibling's exposure to cigarette smoke (table 4.32).

Table 4-32: FBAT Results from Association Analysis Involving *GSTM1*, *GSTT1* CNV Genotypes and the Log Age-Corrected Total IgE Levels on Asthmatic Siblings Using Smoke Exposure as a Covariate.

Marker	No. of Copies	Allelic Frequency	No. of Informative Families	S	E(S)	Var(S)	Z	P
GSTM1	1	0.262	209	-6.6	3.6	174.5	-0.769	0.442
	0	0.738	209	-14.1	-24.3	174.5	0.769	0.442
GSTT1	1	0.571	230	5.8	4.8	246.0	0.063	0.950
	0	0.429	230	25.1	26.1	246.0	-0.063	0.950

No significant association was detected between the *GSTP1* and *GSTO2* haplotypes and total IgE levels from the model involving the log age-corrected total IgE levels as a function of the sibling's exposure to cigarette smoke (see table 4.33).

Table 4-33: FBAT Results from Association Analysis Involving *GSTO2* and *GSTP1* Haplotypes and the Log Age-Corrected Total IgE Levels on Asthmatic Siblings Using Smoke Exposure as a Covariate.

Marker	Haplotype	Frequency	No. of Informative Families	S	E(S)	Var(S)	Z	P
GSTP1	h1	0.669	215	12.177	-3.571	256.046	0.984	0.325
	h2	0.239	198	-41.516	-26.399	214.664	-1.032	0.302
	h3	0.081	97	-10.825	-11.107	106.278	0.027	0.978
	h4	0.012	20	-6.856	-5.943	15.818	-0.230	0.818
GSTO2	h1	0.652	182	0.758	6.422	188.608	-0.412	0.680
	h2	0.23	168	-14.221	-20.039	163.597	0.455	0.649
	h3	0.075	61	18.734	10.713	64.628	0.998	0.318
	h4	0.043	39	1.746	9.921	36.461	-1.354	0.175

4.4.3.3.2 Bronchial Hyperresponsiveness

Association analysis using the residuals from the model using the bronchial hyperresponsiveness measure as a function of the sibling's exposure to cigarette smoke revealed a suggestive result ($P < 0.1$) that an association may exist with *GSTM1* CNV genotypes (table 4.34).

Table 4-34: FBAT Results from Association Analysis Involving *GSTM1*, *GSTT1* CNV Genotypes and the Measure of BHR on Asthmatic Siblings Using Smoke Exposure as a Covariate.

Marker	No. of Copies	Allelic Frequency	No. of Informative Families	S	E(S)	Var(S)	Z	P
GSTM1	1	0.262	206	**303.9**	**103.1**	13279.5	1.743	**0.081**
	0	0.738	206	-517.3	-316.5	13279.5	-1.743	**0.081**
GSTT1	1	0.571	227	-13.5	-8.9	18189.3	-0.034	0.973
	0	0.429	227	-171.5	-176.1	18189.3	0.034	0.973

As the test statistic is greater than the expected statistic (E(S)) under the null hypothesis, it is possible that larger BHR slope values are associated with the presence of a copy of the *GSTM1* gene in individuals, after taking into account their smoke exposure in early life. As the slope variable negatively correlates with the degree of BHR, the presence of a *GSTM1* gene could be protective against BHR, therefore having more normal airway reactivity.

No significant association was detected between the *GSTP1* and *GSTO2* haplotypes and bronchial hyperresponsiveness variable from the model involving the log age-corrected total IgE levels as a function of the sibling's exposure to cigarette smoke (see table 4.35).

Table 4-35: FBAT Results from Association Analysis Involving *GSTO2* and *GSTP1* Haplotypes and the Measure of BHR on Asthmatic Siblings Using Smoke Exposure as a Covariate.

Marker	Haplotype	Frequency	No. of Informative Families	S	E(S)	Var(S)	Z	P
GSTP1	h1	0.669	212	143.773	42.341	20925.609	0.701	0.483
	h2	0.239	196	184.111	248.991	18031.322	-0.483	0.628
	h3	0.081	95	-85.476	-45.745	6590.921	-0.489	0.624
	h4	0.012	21	41.692	38.512	788.209	0.113	0.909
GSTO2	h1	0.652	181	206.925	163.000	14523.570	0.364	0.715
	h2	0.23	168	195.282	191.258	13965.503	0.034	0.972
	h3	0.075	61	-106.110	-13.334	3686.413	-1.528	0.126
	h4	0.043	40	73.979	29.151	2230.251	0.949	0.342

4.4.3.3.3 Percentage Predicted FEV_1 Values

The results in table 4.36 are suggestive ($P < 0.1$) of an association between the presence of a copy of the *GSTT1* and lower percentage predicted FEV_1 value in sibling's after accommodating for cigarette smoke exposure. It would appear from the highly insignificant P values that there is no association with the *GSTM1* CNV genotype

Table 4-36: FBAT Results from Association Analysis Involving *GSTM1*, *GSTT1* CNV Genotypes and Percent Predicted FEV_1 Values on Asthmatic Siblings Using Smoke Exposure as a Covariate.

Marker	No. of Copies	Allelic Frequency	No. of Informative Families	S	E(S)	Var(S)	Z	P
GSTM1	1	0.262	209	-28.1	19.3	23128.6	-0.311	0.756
	0	0.738	209	205.3	158.0	23128.6	0.311	0.756
GSTT1	1	0.571	230	**-410.5**	**-75.0**	33592.7	-1.830	**0.067**
	0	0.429	230	238.6	-96.9	33592.7	1.830	**0.067**

No significant association was detected between the *GSTP1* and *GSTO2* haplotypes and percent predicted FEV_1 phenotype from the model involving the log age-corrected total IgE levels as a function of the sibling's exposure to cigarette smoke (see table 4.37).

Table 4-37: FBAT Results from Association Analysis Involving *GSTO2* and *GSTP1* Haplotypes and Percent Predicted FEV_1 Values on Asthmatic Siblings Using Smoke Exposure as a Covariate.

Marker	Haplotype	Frequency	No. of Informative Families	S	E(S)	Var(S)	Z	P
GSTP1	h1	0.669	214	-206.042	-90.315	32992.910	-0.637	0.524
	h2	0.239	196	301.720	142.500	30532.270	0.911	0.362
	h3	0.081	97	-216.577	-162.669	9122.720	-0.564	0.572
	h4	0.012	21	25.094	14.680	1416.009	0.277	0.781
GSTO2	h1	0.652	181	-126.132	-228.609	25917.665	0.637	0.524
	h2	0.23	166	-194.613	-46.265	22771.015	-0.983	0.325
	h3	0.075	60	228.880	113.814	7007.174	1.375	0.169
	h4	0.043	40	73.668	142.863	5917.196	-0.900	0.368

4.4.3.3.4 Atopy Severity Scores

No significant association was detected between the GST CNV genotypes and atopy severity scores as a function of the sibling's exposure to cigarette smoke as the P values are non-significant (table 4.38).

Table 4-38: FBAT Results from Association Analysis Involving *GSTM1*, *GSTT1* CNV Genotypes and Atopy Severity Scores on Asthmatic Siblings Using Smoke Exposure as a Covariate.

Marker	No. of Copies	Allelic Frequency	No. of Informative Families	S	E(S)	Var(S)	Z	P
GSTM1	1	0.262	208	-7.1	0.2	123.6	-0.656	0.512
	0	0.738	208	-24.3	-31.6	123.6	0.656	0.512
GSTT1	1	0.571	229	34.5	33.2	182.5	0.098	0.922
	0	0.429	229	16.8	18.2	182.5	-0.098	0.922

Association analysis using the atopy severity score residual shows a significant association (uncorrected P value = 0.045) towards the *GSTO2* h1 haplotype associating with a lower positive atopy severity score (less atopic) than expected under the null hypothesis (see table 4.39).

Table 4-39: FBAT Results from Association Analysis Involving *GSTO2* and *GSTP1* Haplotypes and Atopy Severity Scores on Asthmatic Siblings Using Smoke Exposure as a Covariate.

Marker	Haplotype	Frequency	No. of Informative Families	S	E(S)	Var(S)	Z	P
GSTP1	h1	0.669	214	59.442	50.588	189.996	0.642	0.520
	h2	0.239	194	-14.783	-12.086	155.612	-0.216	0.828
	h3	0.081	95	-17.280	-10.353	71.734	-0.818	0.413
	h4	0.012	21	-4.022	-4.791	8.711	0.261	0.794
GSTO2	h1	0.652	177	**-11.315**	**10.917**	123.433	-2.001	**0.045**
	h2	0.23	164	7.136	-11.153	112.480	1.725	0.084
	h3	0.075	59	10.721	6.050	45.197	0.695	0.487
	h4	0.043	40	3.890	4.619	26.127	-0.143	0.886

4.4.3.3.5 Asthma Severity Scores

All P values from the asthma severity score association analysis using residuals from the regression model as a function of the sibling's exposure to cigarette smoke are outside the limit for suggestive results (P < 0.1; see table 4.40).

Table 4-40: FBAT Results from Association Analysis Involving *GSTM1*, *GSTT1* CNV Genotypes and Asthma Severity Scores on Asthmatic Siblings Using Smoke Exposure as a Covariate.

Marker	No. of Copies	Allelic Frequency	No. of Informative Families	S	E(S)	Var(S)	Z	P
GSTM1	1	0.262	201	-9.8	3.4	64.7	-1.640	0.101
	0	0.738	201	29.1	15.9	64.7	1.640	0.101
GSTT1	1	0.571	217	-0.5	12.7	79.5	-1.479	0.139
	0	0.429	217	10.2	-3.0	79.5	1.479	0.139

Therefore it is unlikely that GST CNV genotypes associate with asthma severity scores. By comparing the test statistic to the expected statistic for the null hypothesis it could be inferred that the presence of either gene is protective against the development of severe asthma. This observation mirrors the findings in the gene only and smoke exposure analyses. No significant association was detected between the *GSTP1* and *GSTO2* haplotypes and the asthma severity score phenotype from the model involving positive asthma severity scores as a function of the sibling's exposure to cigarette smoke (see table 4.41).

Table 4-41: FBAT Results from Association Analysis Involving *GSTO2* and *GSTP1* Haplotypes and Asthma Severity Scores on Asthmatic Siblings Using Smoke Exposure as a Covariate.

Marker	Haplotype	Frequency	No. of Informative Families	S	E(S)	Var(S)	Z	P
GSTP1	h1	0.669	205	3.446	-0.569	78.743	0.452	0.650
	h2	0.239	193	-9.225	-5.411	69.589	-0.457	0.647
	h3	0.081	93	3.458	1.622	29.573	0.338	0.735
	h4	0.012	21	-3.425	-1.389	1.794	-1.520	0.128
GSTO2	h1	0.652	175	1.234	-3.559	60.555	0.616	0.537
	h2	0.23	164	-4.792	1.732	56.233	-0.870	0.384
	h3	0.075	60	0.004	1.024	16.572	-0.251	0.802
	h4	0.043	39	-0.391	-3.142	11.223	0.821	0.411

4.4.4 Post-hoc PBAT power analysis results

Power calculations were performed using odds ratios associated with *GSTM1* and *GSTT1* deletion alleles in a comprehensive meta-analysis of GST polymorphism, smoke-exposure and asthma.[347] The meta-analysis of 2292 asthma patients, 5718 controls in smokers and non-smokers determined that the risk associated with a *GSTM1* and *GSTT1* null deletion allele and asthma was 1.2 (95 % CI: 1.08 – 1.35) and 1.36 (95 % CI: 1.11 – 1.63) respectively. When accounting for smoking status the asthma risk increased to 1.6 (95 % CI: 1.25 – 1.94) for only the *GSTM1* deletion allele. Based on these odds ratios, the *GSTM1* and *GSTT1* gene only and gene-environment analyses in this study are underpowered (50 ≤ 79 % power) to detect an association between asthma and a risk allele with an odds ratio between 1.20 and 1.60. If the *GSTP1* and *GSTO2* risk alleles confer similar odds ratios then it would be likely that the *GSTP1* and *GSTO2* haplotype analyses will also be underpowered in this study.

4.4.5 Summary of Results

Table 4.42 provides a summary of the significant association analyses results previously described in this chapter.

Table 4-42: Summary of the FBAT Association and Regression Analyses

Marker	Individuals tested/ Method	Phenotype	Risk Allele/ Haplotype	Uncorrected P-value
GSTP1	Asthmatic Individuals	↑ Asthma Severity Score	h1 (0.67)	0.047
GSTO2	Asthmatic Individuals	↓Affection	h4 (0.04)	0.046
		↓ Log Age-Corrected IgE Levels	h4 (0.04)	0.01
		↓ % Predicted FEV_1	h4 (0.04)	0.04
	Asthmatic Individuals not exposed to cigarette smoke	↓Affection	h4 (0.04)	0.03
		↓ Log Age-Corrected IgE Levels	h4 (0.04)	0.02
		↓ % Predicted FEV_1	h4 (0.04)	0.04
		↓ Atopy Severity Score	h4 (0.04)	0.03
	Regression	↓ Atopy Severity Score	h1 (0.65)	0.045
GSTM1	-	-	-	-
GSTT1	Asthmatic Individuals	↑ Asthma Severity Score	*GSTT1**0	0.02
	Asthmatic Individuals Exposed to Cigarette Smoke	↑ Log Age-Corrected Total IgE Levels	*GSTT1**0	0.04

4.5 Discussion

In the majority, previous studies of GST polymorphism in asthma have shown that inheritance of a gene variant (or gene deletion) with reduced or no enzyme activity is associated with increased risk of developing asthma. Specifically, whole gene deletions of *GSTM1* and *GSTT1* have been linked with decreased lung function growth and appear to modulate gene-environment interactions including exposure to environmental tobacco smoke, ozone and diesel exhaust in children with asthma. In this chapter, GST gene dosage information was determined by real-time PCR (see chapter 3) in an established UK childhood asthma family cohort and analysed accounting for ETS exposure in early life.

Gene only and smoke exposure adjusted analysis using FBAT revealed no association with physician diagnosed asthma for *GSTM1 GSTP1* or *GSTT1* polymorphisms. A rare *GSTO2* haplotype h4 (freq = 0.04, rs156697 C/rs156699 T: Asp142/intron5) may be protective against the development of asthma (P = 0.046) and was more strongly associated in individuals not exposed to cigarette smoke (P = 0.03). As described earlier, GSTM1 and GSTT1 gene deletions have been previously associated with increased susceptibility to childhood asthma worldwide. Thus, it would appear that *GSTM1* or *GSTT1* gene deletions do not alter susceptibility to asthma in this UK population sample. These findings are in line with a recent HuGE review meta analysis of GST polymorphism and asthma phenotypes which included the GSTM1 and GSTT1 data from this study.[349] The review found little evidence for substantial role of *GST* genes alone in the development of asthma, BHR or wheezing and found considerable heterogeneity in the results, indicating extensive publication bias. Therefore, it is probable that other GST genetic loci, such as *GSTO2*, are more important in the development of asthma in this population.

Glutathione s-transferase Omega 2 detoxifies arsenic compounds from cigarette smoke,[350] recent studies have found *GSTO2* haplotypes, containing the Asp142 residue, are associated with increased Chronic Obstructive Pulmonary Disease (COPD) susceptibility and reductions in adult lung function.[338, 339]

The discovery of a protective *GSTO2* haplotype against asthma development in this study is a novel finding, indicating that *GSTO2* polymorphism could be important in the development of an asthmatic phenotype regardless of exposure to tobacco smoke exposure in early life. This finding warrants further investigation in other genetic studies of childhood asthma with more comprehensive measures of maternal and environmental tobacco smoke exposure.

Several studies have shown lack of association of GST gene deletion with childhood asthma, admittedly some of these studies may lack sufficient power to detect an association because of small sample sizes.[351] Recent work performed at the University of Southampton (in collaboration with University of Bristol and Imperial College London), examining GST polymorphism in the Avon Longitudinal Study of Parents and Children (a large (n > 10,000) population-based birth cohort), has also reported a lack of association of GST gene deletion genotypes with childhood asthma at age 7 - 9.[352] However, GST gene deletion was shown to be associated with deficits in lung function when individuals were exposed to maternal smoking and were also found to modify the risk of childhood asthma associated with prenatal paracetamol exposure.[353, 354] Therefore, it is likely that GST genes interact with environmental factors modulating the asthma risk in susceptible individuals.

A meta-analysis of 14 studies has reported that adults with a combined *GSTM1* and *GSTT1* null genotype were at a significantly higher risk of asthma (OR = 2.15, 95% CI: 1.21-3.71) and that smoking overloads the capacity of either *GSTM1* or *GSTT1* detoxification systems.[320] A combined genotype analysis was not performed in this study, as neither gene was independently associated with the diagnosis of asthma. It is interesting to note that both genes are lowly expressed or absent (based on genotype) in lung tissue, whereas *GSTP1* is highly expressed and reduced activity is associated with lower asthma risk, and airways responsiveness (see section 4.1.4). Saadat et al.[320] commented that glutathione S-transferases have overlapping substrate specificities, and a deficiency in one form may be compensated by another. Therefore, a simultaneous determination of all GST genotypes appears to be necessary for reliable interpretation of the role of GST enzymes in asthma.

Thus, it would be very interesting in the future, for example, to examine a profile of gene polymorphism that alters the expression or function of *GSTP1, GSTO2* and in conjunction with *GSTM1/T1* gene deletion polymorphisms.

A priori power calculations were not performed because further recruitment in this study, was not possible. Post-hoc power calculations, based on adult odds ratios for GST genes, were performed in PBAT[344] and it is possible that low power is responsible for the lack of association of with asthma risk in this study. However, the rare *GSTO2* haplotype h4 was found to be protective against asthma and several asthma intermediate phenotypes were associated with GST polymorphisms. So, the current estimation of the power of this study may be too conservative and these findings may reflect a stronger genetic effect of particular GSTs on asthma in early life. Significant P values described in this chapter were not adjusted for multiple testing by Bonferroni correction as this correction would be too conservative and gene-environment interactions may be missed.[355] However, type 1 error must be considered as a possible explanation of these results.

Log age-corrected total IgE levels were not significantly associated with *GSTM1*, *GSTT1* or *GSTP1* polymorphism in the gene only, linear regression or tobacco smoke unexposed group analyses. A rare *GSTO2* haplotype h4 was significantly associated with decreased log age-corrected IgE in asthmatic children before and after adjusting for individuals unexposed to cigarette smoke (P = 0.01 - 0.02). Interestingly, total IgE levels were found to be significantly higher in asthmatic children possessing a *GSTT1* gene deletion when exposed to familial tobacco smoke (P = 0.041). It is intriguing to find an associaton between *GSTO2* haplotypes and atopy and a novel gene-environment interaction between the presence of *GSTT1* and protection from atopy when exposed to cigarette smoke at a young age.

Atopy severity scores (a specific serum IgE dependant score) lacked any association with *GSTM1*, *GSTP1* and *GSTT1* polymorphism in relation to asthma and/or smoke exposure. Mirroring the total IgE results, decreased atopy severity scores were associated with *GSTO2* h4 haplotype in asthmatic children not exposed to cigarette smoke (P = 0.03). In the regression analysis the common *GSTO2* h1 haplotype (rs156697 T/rs156699 T, freq = 0.65) was found to be weakly associated with decreased atopy severity scores in asthmatic children exposed to cigarette smoke (P = 0.045).

Total IgE levels can be affected by viral infection and cigarette smoke exposure.[356, 357] Therefore, the FBAT analysis controlled for any confounding parental smoke effect on total IgE levels by performing separate exposure analyses. However, confounding from viral infection in study participants at the time of recruitment cannot be ruled out.

The exact mechanism by which GSTs can directly affect IgE levels or atopy severity is unclear as the formation of IgE is complex and involves many processes. Interestingly, other studies have found similar associations to this study, for example, individuals with a low antioxidant capacity have been shown to be at an increased risk of developing allergic asthma[358] and *GSTP1* genotypes have been found to be associated with atopic indices such as, skin test positivity and IgE levels in atopic and non-atopic subjects.[322] Also, a randomised, placebo controlled cross-over study found that deficits in GST anti-oxidant capacity (due to genetic polymorphism) can result in allergic individuals having a more robust increase in IgE levels when exposed to an allergen in combination with diesel exhaust particles, than allergen alone.[318]

The methacholine challenge measure of bronchial hyperresponsiveness (BHR) was not significantly associated with any GST genotype in the gene only, exposure or regression analysis. However, from the regression analysis there is a suggestive trend (P = 0.081) with the presence of a *GSTM1* gene and reduced BHR in smoke exposed children. This trend was not replicated in the separate exposure analyses, most likely because of reduced statistical power from the smaller sample size (208 vs. 147 families). Interestingly, a recent study found that *GSTM1* null variants (homozygous gene deletion) were associated with asthma and greater airflow obstruction in a sub-group analysis of 50 children (10% of the cohort) exposed to intrauterine smoke exposure.[359] Therefore, it is likely that maternal smoking during pregnancy is a greater risk factor for the development of asthma phenotypes in children without a *GSTM1* gene.

In the fetus, normal lung development proceedes through the balanced interplay of cell proliferation and apoptotic events and reactive oxygen species can be transmitted from the mother to the fetus via the placenta.[16, 360] Maternal smoking may overwhelm the fetus' antioxidant capacity, leading to a depletion of fetal GSH, increased cellular damage by ROS.[361]

High levels of ROS have also been shown to dysregulate the ligandin function of certain GST enzymes (*GSTP1 & GSTM1*) leading to loss of the negative regulatory feedback loop involving *JNK1* and *ASK1,* thus activating cellular signalling pathways involved in cell proliferation and apoptosis (through MAPK signalling) in response to oxidative stress.[362-364]

These pathways are critical for the normal course of lung development and growth and therefore deficits in GST enzyme function may not only have direct consequences via protection from cellular damage by ROS (due to maternal smoking), but also indirect consequences resulting from insufficient free cytosolic GSTs to function as ligandins. In early life, exposure to high levels of ROS could alter Th1/Th2 responses, as in mice, glutathione depletion in antigen-presenting cells has been shown to generate more Th2 like response patterns.[365]

GSTP1 genotypes (*GSTP1* Ile105) have previously been associated with increased airways reactivity,[323] although results from other studies provide conflicting evidence.[333] Therefore, the results from this study either contribute to the literature by providing evidence for no association with GST gene polymorphism and BHR. Or, alternatively the data set could lack sufficient power to detect a gene or gene-environment interaction with BHR in smoke exposed asthmatic children, as only 30.7% of individuals were classified as smoke exposed.

As with BHR, previously published studies have provided evidence for reduced lung function in asthmatic children with particular GST genotypes and exposure to ozone in Mexico city (*GSTP1* Ile105).[84] Reductions in MMEF (Maximum Mid Expiratory Flow) correlate with GST gene deletion and are more pronounced in children exposed to ETS *in utero*.[317] No *GSTM1*, *GSTP1* or *GSTT1* genotype was found to be associated with a reduction in % predicted FEV_1 at the 5 % significance level in the current study. There was however, an association between the rare *GSTO2* haplotype h4 and decreased % predicted FEV_1 values in asthmatic children and after adjusting for individuals unexposed to cigarette smoke ($P = 0.04$). This observation replicates the finding, from the Framingham Heart Study genome-wide association of population-based pulmonary function measures, that *GSTO2* genotypes (rs156697 & rs156699) strongly associate with FEV_1 and FVC.[338] In the current study, the *GSTO2* h4 haplotype associations have produced conflicting results, it was found to be protective against asthma and

atopy phenotypes as well as being a risk haplotype for lower lung function in asthmatic children.

It is likely that the h4 haplotype captures information on distant causal polymorphisms that have different roles in asthma pathogenesis and the exact role of *GSTO2* polymorphism may be fully elucidated by extensive haplotype analysis of the *GSTO2* gene region, in conjunction with more comprehensive ETS exposure measures. Nevertheless, the association of *GSTO2* haplotypes with asthma, atopy and lung function phenotypes in asthmatic children remains a novel and interesting preliminary finding because *GSTO2* detoxifies arsenic compounds found in cigarette smoke and deficits in adult lung function correlate with *GSTO2* genotypes.[337, 338]

The gene only analyses involving the asthma severity score phenotype provided the strongest evidence for a role for GST gene deletion polymorphism. Increased asthma severity scores were significantly associated with asthmatic children carrying a *GSTT1* gene deletion (P = 0.015). In this study, presence of a *GSTT1* gene is protective against the development of severe asthma regardless of exposure to tobacco smoke in the home during pregnancy and early life.

Adjusting for parental smoke exposure did not reveal a more significant association (exposed, P = 0.086; unexposed, P = 0.078). Therefore, the protective effect of *GSTT1* in this study does not appear to be modulated or made stronger by tobacco smoke exposure in early life, this is in line with the hypothesis that ETS exposure can overwhelm endogenous anti-oxidant defences or other GSTs compensate for deficiencies in other genes.[320] The common *GSTP1* haplotype h1 (rs1695 A/rs1138272 C: Ile105/Ala114: *GSTP1*A*) was associated with increased asthma severity scores (P = 0.02); this result reflects studies that found the Val105 allele to be associated with milder asthma phenotypes.[202,322,323]

Taken together the strongest associations in this study provide evidence that *GSTT1* and *GSTP1* gene polymorphism may be more important in the development of a severe asthma phenotype in susceptible individuals, than in asthma susceptibility *per se* and a rare *GSTO2* haplotype may mark causal polymorphisms that independently associate with decreased IgE levels and lung function (FEV_1) in asthmatic children.

4.6 Concluding Remarks and Future Work

To summarise, maternal smoking is a known risk factor for childhood asthma, gene polymorphism in detoxifying GST enzymes has been associated with asthma phenotypes and interact with ETS exposures. This work described in this chapter has demonstrated that GSTT1, *GSTP1* and *GSTO2* polymorphism is associated with IgE levels, which may be modified by smoke exposure, and asthma severity phenotypes in a family based study of childhood asthma. The exact biological processes involved remain unclear. However, it is likely that there is a critical time window in prenatal or early post-natal life where maternal smoking or ETS exposure has the strongest detrimental effects on lung health. In humans, ETS exposure (especially maternal smoking) is notoriously difficult to quantify accurately, partly due to recall bias and lack of the ability to monitor levels transmitted to the fetus.

Glutathione (GSH) is the major endogenous antioxidant molecule protecting cellular systems from ROS. Fetal GSH levels decrease during gestation, leaving the fetus particularly sensitive to ROS. During gestation the fetal nutritional requirements must be met entirely by the mother and poor maternal diets have been associated with increased asthma risk. Therefore, it is probable that future lung health in smoke exposed children will depend on the interplay of balanced maternal nutrition, to maintain fetal GSH levels, and interaction between the mothers and fetus' antioxidant capacity to deal with ROS transmitted to the developing lung.

This relationship has the potential to be modelled in animals, where the level and timing of maternal smoking can be tightly controlled and offspring lung tissue can directly examined using morphological and molecular biological techniques at time points throughout development and post-natal life. The remainder of the thesis seeks to establish the suitability of a Developmental Origins of Health and Disease rat model of maternal protein-restriction during pregnancy (Southampton protein-restriction rat model[262]) to model the effects of maternal smoke exposure on offspring lung.

In this model, maternal protein-restriction has been shown to have consequences for offspring organ function and produces many of the phenotypes associated with common adult disease (e.g. hypertension). This model has shown that disease phenotypes may be epigenetically programmed, as they can be transmitted to future generations, this is interesting as there is some evidence to suggest that grand-maternal smoking is associated with increased asthma risk.[366] Before this animal model can be utilised, it is necessary to determine what effects, if any, maternal protein-restriction (PR) has on the lung. Therefore, the research in the following chapters have sought to begin to examine lung physiology, gene and miRNA expression in PR exposed and unexposed control animals.

5 Chapter 5: Protein-Restricted Rat Lung: Gene Expression Analysis

5.1 Introduction

Evidence from epidemiological studies have shown that a sub-optimal intrauterine environment can have long term effects on offspring lung function and may increase susceptibility to future respiratory illnesses such as asthma and COPD (reviewed in Maritz et al, 2005).[367] Low birth weight is associated with respiratory symptoms in children up to 7 years of age and increased asthma risk.[368, 369] COPD mortality in the Hertfordshire 1911 – 1939 birth cohort was associated with low birth weight and weight at one year.[370, 371] Considering that a foetus can also mount immune responses in the womb; that maternal smoking and paracetamol intake is associated with increased asthma risk potentially through a mechanism of increased oxidative stress;[372,354] and that the active EMTU seen in asthmatics is a lung remodelling process similar to embryonic lung development. It is intriguing to think asthma or respiratory disease in general may result, at least in part, from the transmission of maternal environmental risk factors to the fetus that lead to altered immunological and lung development or growth *in utero*. Thus, it would appear that the role of the intrauterine environment may be critical in modulating the risk of developing respiratory illness in later life.

5.1.1 Rational

In humans, maternal diet and environmental exposures are associated with increased risk of respiratory disease. Reduced birth weight (a marker for poor maternal nutrition during pregnancy) is associated with increased risk of asthma in childhood and COPD in later life. Measuring maternal or offspring exposures in humans are notoriously difficult, time consuming, expensive and cannot be altered due to ethical issues. Animal models are routinely used in the study of the developmental origins of health and disease (DOHaD) as they are closed experimental systems, where exposures can be altered and tightly controlled and the effects on offspring phenotype can be rapidly determined due to their shorter life span.

Adaption of an established DOHaD rat model of maternal protein-restriction (Southampton protein-restriction rat model) to include a cigarette smoke exposure would provide a system to examine offspring lung health into old age. However, before this model can be used it is necessary to determine the consequence of maternal protein-restriction alone on offspring lung physiology, metabolism and homeostasis. This work presented in this chapter examines lung gene expression profiles in protein-restricted and control offspring.

5.1.2 Immunological Evidence for Developmental Origins of Asthma

There is a growing body of evidence that asthma originates in early life and even during foetal development. The hygiene hypothesis (see section 1.3.1.1) supports the protective effect of exposure to infection at a critical time window in early life against atopy and asthma. It is apparent that this exposure to bacterial endotoxin is critical to switch the T_H1/T_H2 balance towards a T_H1 phenotype, whereby reducing/ablating IgE mediated responses to allergens.

Recently, it is becoming clear that a new born baby is not immunologically naïve[373] and atopic children have altered immune responses at birth,[374] there is evidence they can even mount immune responses in the womb.[375] Therefore, this critical time window could extend into neonatal life and maternal environment exposures during pregnancy are important in the development of allergy. Immune cells in the foetus can be seen as early as 9 weeks and lymphocytes can be detected in the lung at 14 weeks.[375] The neonate is known to be able to mount antibody responses after exposure of the mother during pregnancy to tetanus toxoid.[376] Swallowing of amniotic fluid is the obvious route for exposure to maternal allergen and this action has probably evolved to provide protection to the foetus against any parasitic infection in the mother.[377] There is a shared genetic susceptibility with in atopic families, but increased circulating IgE levels in mothers' correlate more strongly with raised IgE levels and increased allergic sensitisation in the child than with the fathers IgE level.[378] The foetus is exposed to maternal IgG antibodies during development[379] and this may explain why maternal allergy is a stronger predictor of allergy in the child rather than allergy in the father. This reinforces the role of the environment (i.e. the mother's womb) in the development of allergy in genetically susceptible neonates.

5.1.3 Genetic Evidence for the Developmental Origins of Respiratory Disease Concept

At present, there is not much known about the genetic contribution to the developmental origins of asthma and COPD. Detailed below are examples of possible genes that may directly modify fetal lung development or maternal genes variants that alter fetal exposure to harmful chemicals from tobacco smoke.

5.1.3.1 ADAM33

ADAM33 (Gene ID: 80332; MIM: 607114) was positionally cloned from a genomic wide linkage scan in 460 Caucasian US and UK families (affected sib-pairs) in 2002 and was the first asthma gene identified through this approach.[121] Strongest genetic linkage at 20p13 was seen with a combined asthma and BHR phenotype (LOD, 3.93) and the D20S482 microsatellite with a 35% excess allele sharing. Analysis of polymorphisms in the linkage region revealed that the strongest association was in the region of a novel gene, *ADAM33*.

ADAM33 (a disintergrin and metalloproteinase 33) is part of the ADAM gene family that is a sub-group of a super-family of zinc-dependant metalloproteinases that possess proteolytic activity.[380] Other known functions of *ADAMs* are to promote myogenic fusion and in the release of proliferative growth factors[381, 382] ADAM members have also been shown to have roles in fertilization, muscle development and neurogenesis.[383-386] *ADAM33* is expressed in mesenchymal cells such as sub-epithelial fibroblasts and smooth muscle cells but not in respiratory epithelium or in cells of the immune system.[387] The association of *ADAM33* polymorphism with asthma, BHR and related phenotypes has been studied in several different ethnic populations. The majority have found significant association with *ADAM33* polymorphisms, albeit with different SNPs or haplotypes.[134-141, 388] Non-replication of *ADAM33* association may be explained by differences in population and/or environmental exposures.[389-391]

A meta-analysis of *ADAM33* association data has shown that variation in this gene could account for 50,000 excess asthma cases in the UK alone.[392] Genetic variants of *ADAM33* are associated with asthma, BHR and reduced lung function in early life (age 3 & 5).[393]

ADAM33 protein is known to be expressed in asthmatic airways and in human embryonic lungs[394] and recent study has shown increasing expression of *ADAM33* in mild to severe asthmatics.[395] This gene may also play a role in the decline in lung function in later life and in susceptibility to chronic obstructive pulmonary disease (COPD).[396, 397]

5.1.3.2 Maternal *GSTP1* Genotype Modulates Oxidative Stress Experienced by the Developing Fetus

The chromosomal region (Chr 11q13) containing the *GSTP1* gene has demonstrated strong linkage with asthma phenotypes.[398, 399] During pregnancy the foetus is dependant on the mother to meet its physiological requirements; nutrient delivery and toxin removal is dependant on placental function and maternal factors, such as antioxidant defence mechanisms. *GSTP1* is highly expressed in the lung and maternal genotype (*GSTP1* Val105/ Val105 or Ala114/Val114) has been shown to be predictive of better lung function in asthmatic children.[325] This association is independent of transmission of alleles to the child, maternal ETS exposure and atopy and is stronger than the paternal genotype effect. The authors concluded that this suggested an intrauterine effect of maternal *GSTP1* genotype and hypothesised that the mother's antioxidant defence mechanisms can modulate the transmission of oxidative stress to the foetus.

5.1.4 Nutritional Programming of Respiratory Disease

There is evidence to suggest that measures of restricted growth at birth correlate with reduced lung function in later life.[400] It is thought that restricted growth, altered lung development and a genetic susceptibility to oxidative stress may be a driving force behind the development of respiratory disease in later life. During development, the overriding external environmental cue for the neonate is nutritional intake from the mother via the placenta. Thus, poor maternal nutrition during pregnancy could increase the risk of developing asthma and COPD in late life by nutritional programming (section 1.6.3) during gestation that alters the trajectory of normal lung development and growth.

5.1.4.1 Birth Weight and Lung Function

Prematurely born infants with intrauterine growth restriction (birth weight below 10th percentile for gestational age) are at increased rick of impaired lung function (R_{aw} - airways resistance measure) at follow up, aged 6 -24 months.[401] Observations in infants born small at ≥ 37 weeks gestation have shown that lower rates of fetal growth and higher rates of early infancy weight gain are associated with impaired lung development at 5 – 14 weeks of age.[402] At school age, *in utero* growth restricted children born at term have been shown to have impaired lung function and increased incidence of respiratory symptoms.[403, 404] Independent of birth size and respiratory symptoms, poor airway function (lower quartile for V_{maxFRC}) in early infancy tracks with lower FEV_1/FVC ratios up to age 22.[405] In adulthood, low birth weight and weight gain in the first year of life predict lower adult lung function and worse response to environmental exposures (smoking) that lower lung function.[406] Therefore, IUGR is linked with impaired lung function at birth, infancy childhood and poor lung function in early life tracks into adulthood. Thus, impaired lung development in foetal life could have implications for future development of COPD as these individuals are more likely to develop airways obstruction.[370]

5.1.5 Lung Development in DOHaD Animal Models

Intrauterine growth restriction (IUGR) can induce immediate and long term changes to structure and function of many organs including the lung. Early nutritional, oxygen and endocrine environment may be important for the entire life cycle of the lung. In humans, reduced fetal breathing movements are associated with IUGR, these movements are essential for lung expansion and normal lung growth.[407] Restriction of nutrient supply during fetal life can alter lung architecture, although lung/body weight ratios may not be altered.[367] The timing, severity and type of nutritional restriction determine the nature of alteration in the lungs (see table 5.1).

Table 5-1: Milestones in Human Lung Development.

Changes to lung architecture that have been found in DOHaD animal models

	Week	Lung development stage	Milestones in development	Type of IUGR	Effect of IUGR on lung architecture
Organogenesis	3 - 9	Embryonic	Formation of major airways, bronchial tree, portions of respiratory parenchyma and birth of acinus	60% Nutrient restriction in sheep.[408]	No reduction in lung weight. Up regulation of Glucocorticoid receptor mRNA expression.[408]
	5 - 16	Pseudoglandular			
Differentiation	16 – 26	Canalicular	Air-blood barrier formed. Epithelial differentiation and last generations of lung periphery formed		
	24 - 38	Saccular (start of Surfactant production)	Expansion of air spaces. Surfactant is detectable in amniotic fluid	Starvation in rats.[409] Umbilico-placental embolisation in sheep.[410, 411]	Decrease in number of alveoli[409] Enlarged alveoli with thicker walls[410, 411]
	36 - 39	Alveolar	Secondary septation		
Post-natal	37 - >	Early life growth and lung maturation	Maturation of microvascular and late alveolarisation (ceases by yr 2-3)	Under nutrition in rats by increasing litter size.[412]	Cell division reduced in bronchiolar epithelium[412]

5.1.6 Southampton Protein Restriction Rat Model

The Southampton protein restriction rat model utilises a 50 % reduction in the protein content of feed fed to pregnant rat for the whole of gestation (section 5.3.1). This model has been used to study hypertension, endothelial dysfunction and insulin resistance.[263, 413] A protein-restriction diet during pregnancy has adverse effects on glucose homeostasis in the first generation, second and third generation offspring, without further exposure to the PR diet.[414] Alterations in DNA methylation patterns in hepatic gene promoters in these rats correlate with changes to gene expression and are transmitted up to second generation offspring.[264] It is thought that this is a mechanism for the transmission of a predictive adaptive phenotype between generations.

5.1.7 Maternal Protein-Restriction During Pregnancy

Animal models of under-nutrition recapitulate many of the phenotypes associated with chronic human disease in adulthood, i.e. CVD and diabetes. Milder nutritional manipulations such as a 50% reduction in the protein content of an isocaloric diet fed to mothers during pregnancy results in increased blood pressure in offspring in later life, this is the hallmark phenotype of hypertension.[235, 262] Other alterations in protein-restricted offspring (and in subsequent later generations without further exposure) include, changes to gene promoter DNA methylation levels in the liver, reduced nefron numbers in the kidney and endothelial dysfunction.[263, 264, 415]

A protein-restricted diet will potentially reduce the availability of amino acids and the mother may not be able to meet the developing fetus' requirements during pregnancy. Amino acids are critical for life and have variety of roles in metabolism.(reviewed by Wu, 2009)[416] The maternal-fetal-placental unit is the interface between the mother and fetus.(reviewed by Cetin et al., 2009)[417] A study into the effect of maternal protein restriction on amino acid transfer to the fetus showed that at day 19 the total free fetal amino acids did not change in concentration, apart from a decrease in threonine to 45% of the control values; this reduction was not limiting protein synthesis as fetal growth did not appear to be affected.[418] It is thought that fetal amino acid levels were supplemented by the mobilisation of maternal protein, but availability of threonine may influence growth in late stages of development as it is metabolised to produce glycine.[418] Glycine is important for the synthesis of DNA, DNA methylation and is also important in the synthesis of many compounds required for growth (collagen, haem and glutathione).[419]

In rats, a cross-fostering study of maternal protein-restriction (60 % reduction in casein) during pregnancy and in lactation, demonstrated reduced offspring lung weights at post-natal day 21 and that the lungs were greater in size in proportion to body weight, when compared to controls.[420] In this study, a switch to maternal protein-restriction in lactation alone was sufficient to reduce offspring lung weight. However, these effects were not permanent as at 11 months of age (330 days) the lung weights in offspring were similar to control animals.

5.1.8 Rat and Human Lung Development

Apart from the obvious differences between new-born humans and rats e.g. rats are born to large litters, blind and hairless; there are also differences in the timing of stages of lung development, lung architecture and bronchoconstriction response.[421, 422] However, non-primate models, including rodents, are relevant to human disease and are routinely used in the study of asthma and COPD.[422, 423] Figure 5.1 shows the milestones in human lung development. Full-term rats are born with their lungs in the saccular phase of development (equivalent to human lung at 26–28 wk of gestation), where thinning of interstitium, flattening of epithelium has occurred, alveolar ducts and air sacs are formed and surfactant production is started. In the rat, unlike humans, alveolarisation takes place postnatally in two phases; the first phase corresponds to week 35-40 of gestation in human fetal lung. The second phase includes, maturation of microvasculature and late alveolarisation and can continue up to post natal day 60 (young adulthood in the rat).[424] This difference in the timing of late lung development into postnatal life in the rat provides the opportunity to expose young rats to strict challenges when the lung is still in a phase of developmental plasticity (see section 1.6.4).

5.1.9 Altered Gene Expression in DOHaD Animal Models

As stated earlier in this chapter (section 5.1.1) it is necessary to establish the effect of maternal protein restriction on offspring lung gene expression before we can utilise the Southampton protein-restriction rat model to study the consequence of maternal cigarette smoke exposure on offspring lung health. Previous studies of intrauterine growth restriction (IUGR) in rat and sheep have shown that exposed offspring have altered expression of genes involved in growth and metabolism compared to control offspring. IUGR offspring have shown changes in the expression of growth factors (Insulin-like Growth Factor (*IGF1* and *IGF2*), receptors for steroid hormones (glucocorticoid receptor, *GR* and mineralcorticoid receptor, *MR*), enzymes regulating steroid synthesis (11β-dehydroxysteroid dehydrogenase-1 and 2 (*HSD11B1* & *HSD11B2*) and glucocorticoid responsive ATPase's (*ATP1A1* & *ATP1A2*). Alterations in gene expression have been found in different tissues (including lung, liver, kidney, thymus and hypothalamus) at several time points during gestation, early life and in some cases these changes in gene expression can persist far into adulthood.[425-428]

Furthermore, it has been shown that these changes in gene expression can be programmed at the epigenetic level, through changes in the DNA methylation pattern of gene promoters, and passed on to the next generation of offspring.[429] The genes chosen for gene expression analysis in this chapter are as follows:

5.1.9.1.1 Atp1a1 & 2 and Atp1b1

Sodium/Potassium adenosine triphosphatase subunit gene expression is regulated by corticosteroids and 11β-hydrosteroid dehydrogenase activity. *Atp1a1*, *2* and *b1* mediate Sodium and potassium transport and may have roles in controlling blood pressure, maintaining electrochemical gradient across plasma membrane and oxygen homeostasis.[430] In an IUGR rat model (maternal low protein diet; MLP) examining the *in utero* programming of hypertension increased levels of *Atp1a1* and *b1* mRNA expression was associated with significantly higher levels of glucocorticoid receptor *GR* mRNA and protein in kidney, liver, lung and brain in the MLP offspring at 12 weeks.[428]

5.1.9.1.2 Igf1 & 2

Insulin-like growth factors (IGFs) are cytokines that modulate cell proliferation and differentiation during embryogenesis through interactions with the IGF cell surface receptor (*IGFR1*). Late gestation (last trimester) maternal under-nutrition in rats resulted in significantly lower lung/body weight ratios of the IUGR offspring on postnatal days 7 and 14 and increased expression of *IGF1* and *2* at postnatal days 1 and 28.[425]

5.1.9.1.3 Nr3c1(GR), Nr3c2 (MR), Hsd11b1 & 2

Glucocorticoid and mineralcorticoid hormone levels in the lung are regulated by expression of the glucocorticoid receptor (*GR*) or mineralcorticoid receptor (*MR*) and isoforms of 11β-dehydroxysteroid dehydrogenases (*Hsd11b1* & Hsd11b2) at the levels of gene transcription. Mice with disrupted glucocorticoid receptor genes die shortly after birth with severe lung atelectasis (non-inflated lungs); the action of the glucocorticoid hormone is critical for lung maturation and surfactant production.[431] Mineralcorticoid receptors are found in epithelium in fetal lung and are thought to play an important role in lung liquid re-absorption at birth.[432]

In IUGR fetal ovine lung the mRNA levels of *Hsd11b1* and *GR* were at their most abundant at 140 days gestation, although dietary restriction had no effect on lung weight.[427] In an IUGR rat model (maternal low protein diet; MLP) examining the *in utero* programming of hypertension significantly higher levels of *GR* mRNA and protein were found in kidney, liver, lung and brain in the MLP offspring up to 12 weeks postnatal age; whereas *MR* levels remained the same.[428] Conversely, the levels of *Hsd11b2* were significantly reduced in placenta (14 & 20 days gestation), foetal lung (d20 gestation) and in kidney from new born offspring up to 12 weeks postnatal age.[428] Changes to *GR* promoter DNA methylation in the liver of protein-restricted rat offspring correlated with differential gene expression and were found to be transmitted to the next generation of offspring without further protein restriction.[264]

5.1.9.1.4 Tp53

The transcription factor p53 is a key factor involved in determining cellular response to stress by affecting apoptosis, cell cycle regulation and angiogenesis. It has been hypothesised that changes in p53 expression due to IUGR alters the lung apoptotic /cell proliferation balance during alveolar development resulting in a lung more susceptible to injury and chronic lung disease in later life. An IUGR rat model has demonstrated that IUGR results in decreased localised p53 serine-15P associated with distal lung alveolar septum thickening in neonatal lungs.[433]

5.1.9.1.5 Ucp2

Uncoupling protein 2 is an inner mitochondrial membrane carrier that is highly abundant in the lung, this gene has postulated roles in energy balance, reactive oxygen species production, apoptosis and macrophage mediated immunity.[434] Genetic polymorphism in the *UCP2* gene have been associated with obesity and diabetes risk in children and adolescents.[435, 436] In the ovine lung, increased *UCP2* expression has been shown to peak in offspring, from ewes assigned to the 60% nutrient restricted diet during early to mid gestation compared to the *'feed to appetite'* control diet, at 1 day of age then decreasing at each later sampling age (7, 30 and 180 days) but being still detectable at 6 months (180 days).[427] In offspring from ewes with nutrient restriction in late gestation increased *UCP2* expression persisted until 30 days postnatal age.[427]

5.1.9.2 Oxidative Stress Genes: Glutathione S-Transferases

The expression of glutathione s-transferases will also be examined as; *GSTM1*, *GSTP1* and *GSTT1* polymorphisms have been associated with asthma and COPD. In school children, *GSTM1* and *GSTP1* genotypes are associated with deficits in lung function growth.[319] Children suffering from severe malnutrition have consistently been found to have oxidative stress induced cellular damage and impaired synthesis of the antioxidant glutathione.[437] Under-nutrition in rats during gestation and up to day 21 of life produced a 61% decrease in microsomal glutathione s-transferase activity in lung tissue at day 90 of life. A recovery diet of *ad libitum* feed from day 45 to 90 failed to bring GST activities in line with the activity seen in control animals.[438]

5.1.9.3 Candidate Human Asthma Genes: *Adam33* & *Pcdh1*

Two human candidate asthma genes where included in the panel of genes for expression analysis because of their involvement in asthma in early life:

5.1.9.3.1 Adam33

ADAM33 polymorphism has been shown to be important in determining early-life lung function[393] and in the regulation of fetal lung morphogenesis[394] which could affected by IUGR. Although the exact control mechanism that regulates *ADAM33* expression is not known, promoter methylation does occur in a tissue specific manner.[387] and has shown to be responsible for the suppression of *ADAM33* expression in epithelial cells.[439]

5.1.9.3.2 Pcdh1

Protocadherin 1 (*PCDH1*) polymorphism has also been associated with bronchial hyperresponsiveness in asthma[440] but the exact function of this calcium dependant cell-cell adhesion membrane molecule in the pathogenesis of asthma is unknown. However, it is likely that it may mediate early life development of asthma as there appears to be synergistic interaction between exposure to environmental tobacco smoke *in utero* or first years of life and genetic polymorphism of *PCDH1* in determining asthma susceptibility.

5.1.10 Hypotheses

It is proposed that:

'Maternal protein-restriction during pregnancy leads to altered lung growth and alterations in gene expression that persist into adulthood alter the susceptibility to respiratory disease.'

And that:

'DNA methylation and miRNA expression are the epigenetic mechanisms that control persistent changes to lung gene expression in the offspring of mothers who experienced protein-restriction during pregnancy.'

The work presented in this chapter attempts to establish the validity of the Southampton protein restriction rat model to test the above hypotheses. A panel of genes were selected from a review of the literature and mRNA expression was assayed in archived lung tissue from first (F_1) and second generation (F_2) animals by qPCR. Altered gene expression in adult F_1 offspring lung tissue will indicate that maternal protein restriction during pregnancy in this model produces persistent changes in gene expression that have been previously observed in animal models designed to study the effect of maternal under-nutrition on lung development. Differential gene expression in the F_2 generation may indicate that epigenetic mechanisms such as DNA methylation are controlling the alterations in gene expression in lung tissue from rats exposed to protein-restriction in the F_1 generation.

5.2 Aims

The aim of this work was to:

- Perform morphometric analyses of lung weight and volume to determine whether maternal protein restriction during pregnancy affects organ size.
- Validate the suitability of a TRIzol™ based simultaneous RNA, DNA and Protein isolation protocol for rat lung tissue samples
- Design, optimise and validate real-time PCR assays for mRNA transcripts expressed in rat lung tissue, including a panel of target genes known to be differentially expressed in animal models of intrauterine growth restriction and the Glutathione S-transferases genes *Gstm1*, *Gstp1* and *Gstt1*.
- Test the suitability of three reference genes (*Actb, B2m and Ubc*) for use as endogenous controls for real-time PCR in the rat lung.
- Perform relative quantification analysis on F_1 & F_2 generation protein restricted and control rat lungs and show whether the chosen target genes are differentially expressed in lung tissue and whether this persists in the F2 generation animals, as this would indicate the presence of epigenetically controlled transmission of programmed lung gene expression due to ancestral maternal protein-restriction.

5.3 Methods

5.3.1 Protein Restriction Rat Model

The lung samples used in the following experiments are from an established model of maternal protein-restriction during pregnancy. The lung samples were collected from animals that were part of an ongoing study of endothelial dysfunction at the DoHaD Division (School of Medicine, University of Southampton).[441] All animal handling and harvesting of lungs was performed by Dr. Christopher Torrens and his team. All of the qPCR experiments were performed by myself and the lung morphometry experiments with the assistance of Dr. Christopher Torrens, his team and Miss Shelley Davis (fellow PhD student in the Respiratory Genetics laboratory). The HO project license number for this investigation is 70/6457.

5.3.1.1 Experimental Study Design

Female Wister rats (F_0 generation) were mated and then fed either a Control (C) containing 180g/kg (w/w) casein or an isocaloric Protein Restriction (PR) diet containing 90g/kg (w/w) as described by Langley & Jackson.[262] (cited in Burdge et al)[264] At delivery, the litters were standardised to eight pups (4 male & 4 female). After delivery the dams and pups were fed standard chow (pups were weaned on to the standard chow 28 days after birth). Male offspring (F_1) were killed on postnatal day 120. Lungs were removed immediately, frozen in liquid Nitrogen and stored at 80°C. Figure 5.1 displays the study design for the F_0 & F_1 generation animals. Female F_1 generation offspring (selected by random removal from cage) were mated on postnatal day 125 with males from the control diet litters. At all times the mated F_1 generation females and offspring (F_2 generation) were fed standard chow. Again, the litters were reduced to eight animals. Male offspring (F_2) were killed on day 80 and the lungs harvested and frozen. Lungs from one offspring from each litter were studied.

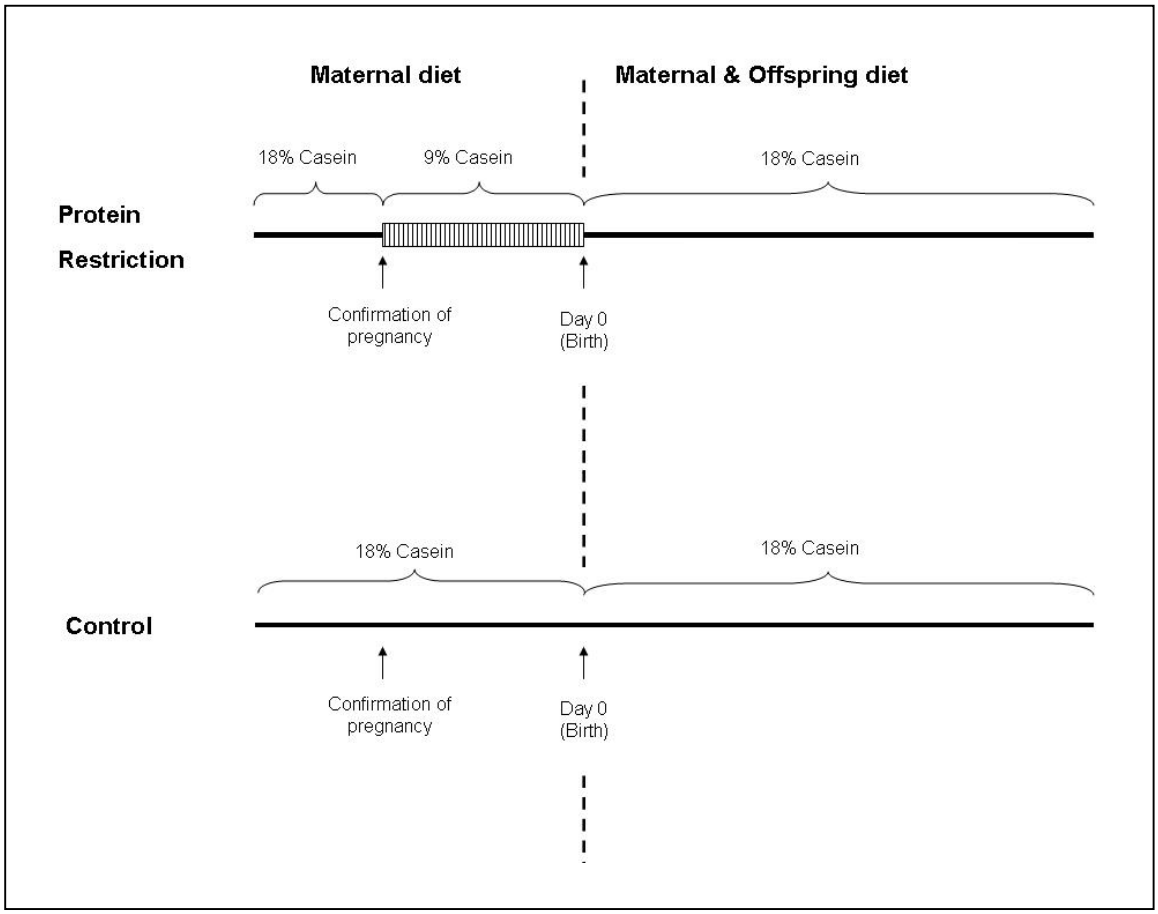


Figure 5-1: Protein Restriction Study Design.

On confirmation of pregnancy by the presence of a vaginal plug expectant mothers are placed on either the protein-restricted or control chow diet. After birth protein-restricted mothers are immediately returned to the original chow diet and litter sizes are standardised. Male offspring are sacrificed at different time points and some of the female offspring are mated at 125 days of age to produce the F_2 generation.

5.3.1.1.1 F_1 Generation Lung Samples

qPCR experiments were performed on RNA isolated from the first generation Protein restricted or control offspring lung samples that were sacrificed on day 120 and 225. Eleven control and seven protein restricted lungs from 120 day old animals were provided by Dr. Christopher Torrens (table 2.7). Six protein-restricted and six control lungs were collected from 225 day old animals that were available at the time of this study.

5.3.1.1.2 F_2 Generation Lung Samples

qPCR experiments were performed on lung RNA isolated from second generation offspring that were sacrificed on day 80. Thirteen frozen lung samples from both groups were provided by Dr. Christopher Torrens (table 2.7).

5.3.2 Lung Morphometry

At the time of this study it was only possible to perform lung morphometry experiments on 225 day old F_1 offspring.

5.3.2.1 Lung Samples

Pregnant Wistar rats were allocated to either control (C, 18% casein) or protein restricted (PR, 9% casein) diet for the duration of the pregnancy. Lung tissue was harvested from F_1 male offspring from separate litters at 225 day of age. Six lungs from each diet were available for lung volume measurements.

5.3.2.2 Lung Volume Measurements

Lungs (left lobe) were fixed by tracheal instillation of formalin under fixed pressure and left lobe volumes were calculated by volume displacement.[442] Left lobe volumes were extrapolated to estimate a total lung volume (Equation 6) and this value was used to calculate lung weight and volume indices for comparison between the PR and control groups (Equation 7 & 8).

Equation 6: Total Lung Volume Extrapolation

$$= \left(\frac{\left(\frac{\text{perfused left lobe weight} - \text{left lobe weight}}{\text{specific gravity of formalin}} \right)}{\text{left lobe weight}} \right) * \text{total lung weight}$$

Equation 7: Lung Weight Index[443]

$$= \frac{\text{total lung weight (kg)}}{\text{body weight (kg)}} * 100$$

Equation 8: Lung Volume Index[443]

$$= \frac{\text{total lung volume (ml)}}{\text{body weight (kg)}}$$

5.3.2.2.1 Lung Fixation Protocol

Whole lungs were dissected free from the animal and excess connective tissue and trachea was removed. The whole lung was weighed, then the left bronchial tube was cut close to the left lung and the weight of the left lung was recorded. A small gauge hypodermic needle on the end of a plastic tube connected to the formalin was inserted into the left bronchial tube and secured with cotton thread (figure 5.2). The value on the formalin reservoir was opened and ran for two minutes to fully inflate and fix the lung tissue. After two minutes the needle was removed from the lung and the cotton thread was used to gently lower the fixed lung into a tared universal tube containing 5 mls of formalin and weight taken.

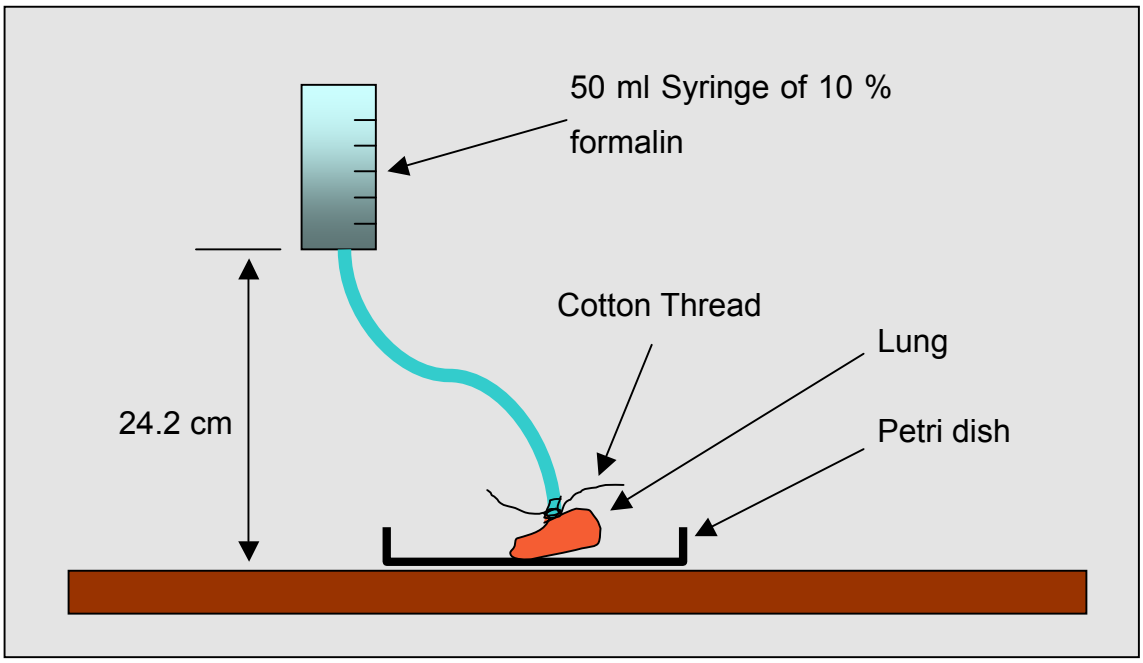


Figure 5-2: Lung Fixation Apparatus

Working on a bench top the base of the syringe of formalin was suspended at the correct height to achieve the correct pressure to inflate and fix the left lung lobe resting in a petri dish.

5.3.2.3 Calculation of Lung Fixing Perfusion Pressure

Instillation of 10 % formalin into the left lung lobe was performed at a fixed pressure of 25 cm of H_2O. This was achieved by calculating the required height of the formalin reservoir as follows:

Perfusion pressure from Chen *et al*[442]: = 25 cm of H_2O (1.36 cm of H_2O is equivalent to 1 mmHg)
= 18.4 mmHg (1 mmHg is equivalent to 133 Pa)
= 2447.2 Pa

Calculation of Hydrostatic pressure [Pa]: = ρgh *Where: g is gravity at 9.81 m/s^2, h is height (m) and ρ is practical gravity of fluid (the practical gravity for 10% formalin is 1030 kg/m^3)*

Therefore:

$2447.2 = 1030 \times 9.81 \times h$
$h = 2447.2 \div 10104.3$
$h = 0.242$ m
$= 24.2$ cm

5.3.2.4 Statistical Analysis

Significant differences in lung weight and volume indices were determined using Mann-Whitney U tests in SPSS v15.

5.3.3 Total RNA, DNA and Protein Isolation Using TRIzol™ Reagent

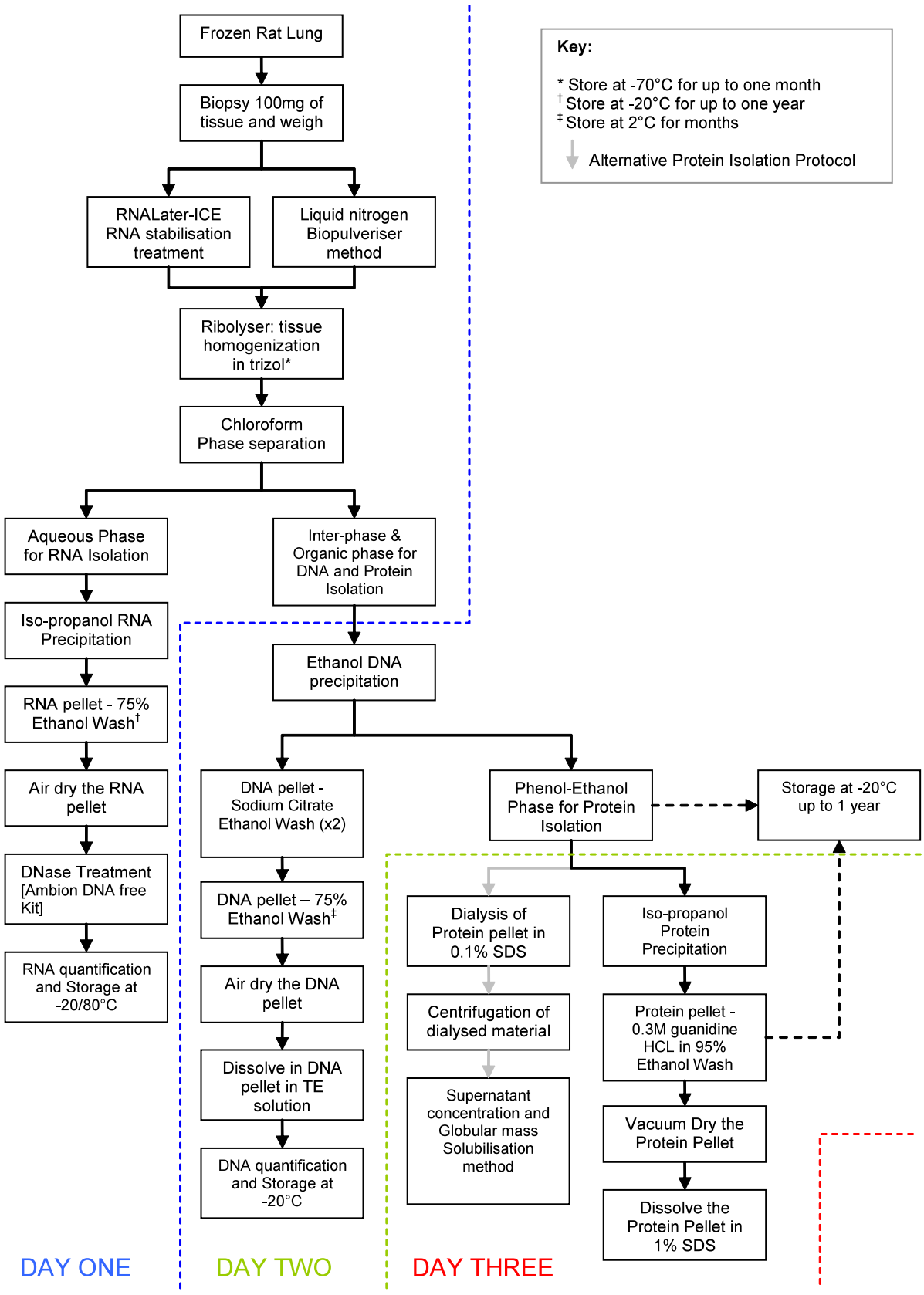
5.3.3.1 Principle of TRIzol™ Reagent Isolation Procedure

As described in section 2.2.3, TRIzol™ reagent allows for the rapid isolation of Nucleic Acids (RNA and DNA) and proteins from the same tissue biopsy. RNA expression and protein level results can be normalised to DNA quantity; allowing for direct comparison of alterations in the genome, transcriptome and proteome.

TRIzol™ is a mono-phasic solution of phenol and guanidine isothiocyanate that has been improved from the single step RNA isolation method developed by Chomczynski and Sacchi.[444] The reagent maintains the integrity of the RNA during sample homogenization and cell lysis. This technique performs well on small (50-100 mg) and large (≥ 1 grm) quantities of tissue and isolates a variety of RNA species of large and small molecular weights suitable for use in northern blots, PCR reactions and other RNA analysis techniques. Figure 5.3 shows an overview of the isolation procedure using TRIzol™. The entire isolation protocol can be completed with in three days with several stages in each isolation procedure where the sample can be conveniently stored. Frequently, issues with low RNA quality are due to active endogenous RNase or contamination that rapidly degrades RNA molecules. It is necessary to control or ablate RNase activity so that any analysis yields relevant and meaningful results. Traditionally, maintaining the tissue at -80°C or lower during storage and homogenisation ensures that any RNase activity is kept to a minimum. TRIzol™ can protect the RNA molecules during the homogenisation step and can be used in conjunction with traditional liquid nitrogen snap-freezing, pulverisation and mechanical cell lysis. Biotechnology companies offer convenient proprietary solutions (such as RNALater-ICE, Ambion) that penetrate the tissue and stabilise RNA molecules by chemically inactivating any RNases while allowing for sample storage at -20°C and handling at room temperature for short period.

Figure 5-3: TRIzol™ Workflow Diagram.

Firstly RNA is isolated from the aqueous layer then DNA can be precipitated with ethanol from the interphase and finally proteins can be pelleted from the phenol-ethanol phase and purified by successive washes or dialysis.


Frozen Rat Lung
Biopsy 100mg of tissue and weigh
RNALater-ICE RNA stabilisation treatment
Liquid nitrogen Biopulveriser method
Ribolyser: tissue homogenization in trizol*
Chloroform Phase separation
Aqueous Phase for RNA Isolation
Inter-phase & Organic phase for DNA and Protein Isolation
Iso-propanol RNA Precipitation
RNA pellet - 75% Ethanol Wash†
Air dry the RNA pellet
DNase Treatment [Ambion DNA free Kit]
RNA quantification and Storage at -20/80°C
Ethanol DNA precipitation
DNA pellet - Sodium Citrate Ethanol Wash (x2)
DNA pellet – 75% Ethanol Wash‡
Air dry the DNA pellet
Dissolve in DNA pellet in TE solution
DNA quantification and Storage at -20°C
Phenol-Ethanol Phase for Protein Isolation
Storage at -20°C up to 1 year
Dialysis of Protein pellet in 0.1% SDS
Centrifugation of dialysed material
Supernatant concentration and Globular mass Solubilisation method
Iso-propanol Protein Precipitation
Protein pellet - 0.3M guanidine HCL in 95% Ethanol Wash
Vacuum Dry the Protein Pellet
Dissolve the Protein Pellet in 1% SDS
DAY ONE
DAY TWO
DAY THREE
Key:
* Store at -70°C for up to one month
† Store at -20°C for up to one year
‡ Store at 2°C for months
Alternative Protein Isolation Protocol

5.3.3.2 Optimisation of the TRIzol Isolation Procedure

In order to maintain RNA integrity prior to isolation two RNA stabilisation methods were initially compared in parallel (see figure 5.4 and section 5.3.3.1). To maximise the nucleic acid and protein yield from each biopsy, a tissue homogenisation step was inserted before the start of the isolation procedure (section 2.2.3.3). No further alterations were made to the manufacturers recommended RNA (section 2.2.3.4), DNA (section 2.2.3.6), protein (section 2.2.3.7) procedures and good laboratory procedures were followed throughout the isolation procedure (section 2.2.3.1).

5.3.3.3 Validation of the RNA Stabilisation Methods

The suitability of the RNA stabilisation methods was examined by qPCR analysis (section 2.2.7.5.2) of *Adam33* expression (normalised by *Actb* expression) in three control rat lung samples (2 biopsies per lung) that were stabilised by either the Liquid Nitrogen-Bio-Pulverizer method or the RNALater-ICE stabilisation treatment (section 2.2.3.2) prior to RNA isolation (2.2.3.4), DNase treatment (section 2.2.3.5) and RT-PCR (section 2.2.4).

5.3.3.4 Isolation of Nucleic Acids and Protein from Lung Samples

Where possible, ~100 mg biopsies were taken from the right lung. In some cases the harvested lungs had been damaged in cold storage and the right lung could not be identified. RNA stabilisation was performed according to section 2.2.3.2.1, the lung tissue was homogenised according to section 2.2.3.3 and RNA (section 2.2.3.4), DNA (section 2.2.3.6) and protein (section 2.2.3.7) isolated from the same lung biopsy. RNA and DNA concentration was measured according to section 2.2.2.

5.3.4 Reverse-Transcription PCR

5.3.4.1 Principle of Reverse-Transcription PCR

Before the RNA can be used as a template in a real-time PCR reaction to measure relative transcript expression levels between the different tissue biopsies, the single stranded molecules must under go a first strand compliment reaction catalysed by a viral reverse-transcriptase.

In this experiment a commercially available reverse-transcriptase kit was used to generate the first strand complimentary DNA (cDNA). cDNA is not susceptible to RNase degradation and is more stable at room temperature than RNA molecules that can be rapidly degraded resulting in altered relative expression levels that will not be representative of any treatment or biological phenotype.

5.3.4.2 Choice of Reverse-Transcription System

The Improm-II™ RT system (Promega, UK) has been proven by the manufacturer to efficiently reverse transcribe both long (up to 8.9 kb) and rare RNA molecules and produces cDNA from up to 1 µg of RNA template per reaction that is suitable for use in real-time PCR (Table 2.8). These qualities where deemed to be advantageous as the experiments will require the amplification of numerous mRNA targets from a single cDNA sample.

5.3.4.3 Improm-II™ Reverse-Transcription System Protocol

All RNA samples were DNase treated (section 2.2.3.5), normalised to 0.5 µg/µl according to section 2.2.2.2, and all RT-PCR reactions were set up according to section 2.2.4.

5.3.5 mRNA qPCR Assay Design

All target gene assays were designed according to section 2.2.5.1. Table 5.2 details the qPCR assay technologies used for the target and reference genes.

Table 5-2: mRNA qPCR Assay Designs

Section 5.3.4.1 explains how mRNA abundance was determined. N.D. = Not Determined.

Gene Name	Target or reference	Assay Type	Intron Spanning?	Adjusted for PCR efficiency	Abundance
Adam33	Target (IUGR marker)	**UPL**	Yes	E-method	Medium
Atp1a1	Target (IUGR marker)	SYBR green	Yes	E-method	High
Atp1a2	Target (IUGR marker)	SYBR green	Yes	LinRegPCR	Medium
GR	Target (IUGR marker)	TaqMan	Yes	E-method	Medium
Gstm1	Target (oxidative stress marker)	SYBR green	Yes	E-method	N.D.
Gstp1	Target (oxidative stress marker)	SYBR green	Yes	E-method	N.D.
Gstt1	Target (oxidative stress marker)	SYBR green	Yes	E-method	N.D.
Hsd11b1	Target (IUGR marker)	SYBR green	Yes	LinRegPCR	High
Hsd11b2	Target (IUGR marker)	SYBR green	Yes	LinRegPCR	Medium
Igf1	Target (IUGR marker)	SYBR green	Yes	LinRegPCR	Medium
Igf2	Target (IUGR marker)	SYBR green	Yes	LinRegPCR	Low
MR	Target (IUGR marker)	SYBR green	Yes	E-method	Medium
Pcdh1	Target (IUGR marker)	**UPL**	Yes	LinRegPCR	Medium
Tp53	Target (IUGR marker)	SYBR green	Yes	LinRegPCR	Medium
Ucp2	Target (IUGR marker)	SYBR green	Yes	LinRegPCR	Low
Actb	Reference (endogenous control)	TaqMan	Yes	E-method	High
B2m	Reference (endogenous control)	TaqMan	Yes	E-method	High
Ubc	Reference (endogenous control)	TaqMan	No	E-method	Medium

5.3.5.1 Target mRNA Abundance

Relative target mRNA abundance (e.g. high, medium and low expression) was determined by running all target gene assay reaction mixes spiked with the same cDNA template and grouping the target amplification plots by Cp values. Figure 5.4 shows that the target gene amplification plots can placed into three groups based on Cp values alone, Cp values < 24 were designated as high abundance, 26-29 Cp values as medium and >30 as low abundance.

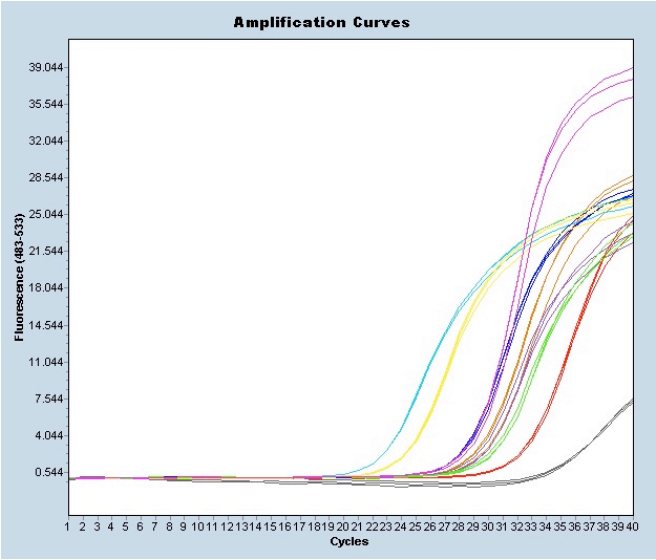


Figure 5-4: Relative Abundance of Target Genes.

Amplification plots of Atp1a1 (yellow lines), Atp1a2 (green), Hsd11b1 (light blue), Hsd11b2 (purple), Igf1 (brown), Igf2 (red), Pcdh1 (violet), Tp53 (dark blue) and Ucp2 (grey)

5.3.6 Optimisation of mRNA qPCR Assays

All mRNA qPCR assay conditions were optimised by performing primer and probe titrations (probe based assays only) and checked for primer specificity by melting curve analysis according to section 2.2.5.2.2.

5.3.7 Validation of mRNA Assays

Standard curve dilutions were performed for all assays (section 2.2.5.3) and the mRNA qPCR reactions plates were setup according to section 2.2.5.4.2. All reaction plates were thermal cycled according to section 2.2.5.4.2.1 and analysed on the Lightcycler 480 real-time PCR machine (appendix 10.3). Where there was an insufficient standard curve range to estimate PCR efficiency a LinRegPCR analysis was performed (section 2.2.5.4.2.2).

5.3.8 Choice of Reference Gene Assays

A minimum of two reference genes are required to determine any variation in reference gene expression between treatment and control group;[445] therefore it was determined to be prudent to include a third gene, just in case one of the chosen reference genes is unsuitable for relative quantification qPCR analysis. *B2m* and *Actb* were in the top five recommended reference genes for numerous rat tissues when examined by TaqMan® low density arrays.[446] *Ubc* was chosen as the third reference gene and all the genes were validated according to section 5.3.7. The *Actb* assay had been previously designed by Dr Fred Anthony (DoHaD Division, School of Medicine, University of Southampton). See table 2.3 for the TaqMan assay designs for *Actb*, *B2m* and *Ubc*.

5.3.9 mRNA qPCR Experiments

5.3.9.1 F_1 Generation Rat Lungs

Three qPCR reaction plates were set up for all the samples described in section 5.3.1.1.1 for all of the target genes in table 5.2 and the three reference genes according to methods section 2.2.5.4.2.

The 225 day lungs were obtained after the first qPCR experiment using the 120 day lung samples. Therefore, in order to reduce reagent costs only the qPCR assays of mRNA transcripts demonstrating different levels between the PR and control lungs were ran on these samples.

5.3.9.2 F_2 Generation Rat Lungs

Three qPCR reaction plates were set up for all the samples described in section 5.3.1.1.2 for the following target genes as these genes were differentially expressed in the F_1 generation lungs:

- *GR*
- *Hsd11b1*
- *Hsd11b2*
- *Igf1*
- *Igf2*
- *Pcdh1*
- *Tp53*
- rUCP

And the oxidative stress genes (*Gstm1, Gstp1* and *Gstt1*) and the three reference genes according to methods section 2.2.5.4.2.

5.3.10 mRNA qPCR Data Analysis

5.3.10.1 Determining the Minimum Detectable Difference in Expression Between Treatment and Control Samples

When using multiple reference genes it is possible to estimate the minimum detectable difference in expression between the protein-restricted and control RNA samples by calculating the ratio of the geometric means of normalised concentration ratios from each reference gene pairing. For example, when using two reference genes you can estimate the minimum detectable difference as follows:

Equation 9: Formula for Calculating the Minimum Detectable Difference in Expression.

$$MDD = \frac{\left(\frac{Conc_{reference\,gene\,1}}{Conc_{reference\,gene\,2}}\right)_{treatment\,group}}{\left(\frac{Conc_{reference\,gene\,1}}{Conc_{reference\,gene\,2}}\right)_{control\ \ group}}$$

Therefore, if the normalised concentration for the treatment group equals the normalised concentration ratio for the control group the minimum detectable difference ratio would be one fold, i.e. you can be relatively certain that a doubling or 50% decrease in expression between the treatment and control group is real as there is no variation between the expression levels of reference genes for the two groups. The minimum detectable difference estimate was calculated in Microsoft® Excel™ for the reference gene combinations used to normalise the F_1 and F_2 generation lung qPCR experiments.

5.3.10.2 Relative Quantification Analysis

Relative quantification analysis was performed according to section 2.2.5.4.2.3 for all target gene reference gene combinations.

5.3.10.3 Statistical Analysis of qPCR Data

Non-parametric Mann-Whitney U tests were performed on the normalised expression data using SPSS 15.0 (SPSS Inc, US). Expression levels were considered to be significantly different at the 5% level (2-tailed P values).

5.4 Results

5.4.1 Lung Morphometry

It is apparent from the comparison of body weights that the PR offspring tend to be smaller than the control animals at 225 day of age (see figure 5.5). However, this reduction in body mass is not statistically significant. Conversely, although the weight of the lungs is similar between the groups, when body weight is accounted for, the lung weight index tends to be larger in PR offspring; a similar observation is also seen with the lung volume index.

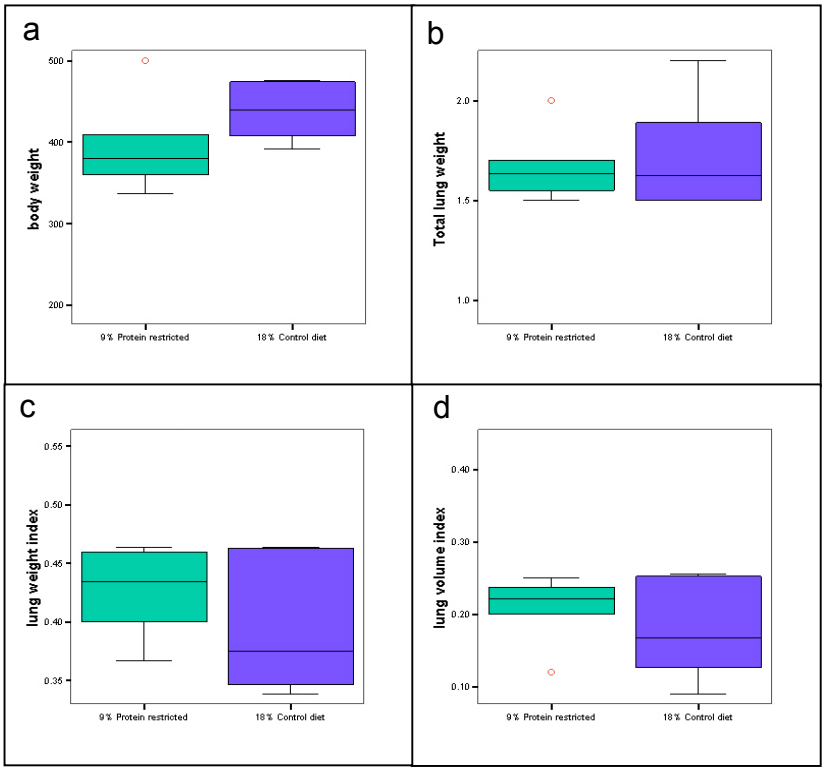


Figure 5-5: Boxplots of F1 225 Day Old Lung Measurements.

(a) body weight in grams (b) total lung weight (c) lung weight index (d) lung volume index of PR (n = 6; green boxes) and control (n = 6; blue boxes) animals. None of the measurements were significantly different at the 5 % level.

The trend of increased lung weight and volume indices in the PR animals was not statistically significant. It is probable that due to the small sample size this experiment lacks sufficient power.

5.4.2 Validation of the RNA Stabilisation Methods: qPCR Results

The results of the qPCR experiment from section 5.3.3.3 show that Liquid Nitrogen-Bio-Pulverizer method of RNA stabilisation out-performed the RNALater-ICE stabilisation treatment for maintaining mRNA integrity during the tissue processing stage of the TRIzol mRNA isolation procedure. *Actb* and *Adam33* amplification of the cDNA template from the Liquid Nitrogen-Bio-Pulverizer method occurred a full 3-4 Cp values before the RNALater-ICE treated cDNA template indicating an approximate 10 fold increase (~3.33 Cp values = 10 fold change in concentration) in reverse-transcribed cDNA template from the same weight of isolated RNA transcripts (Figure 5.6).

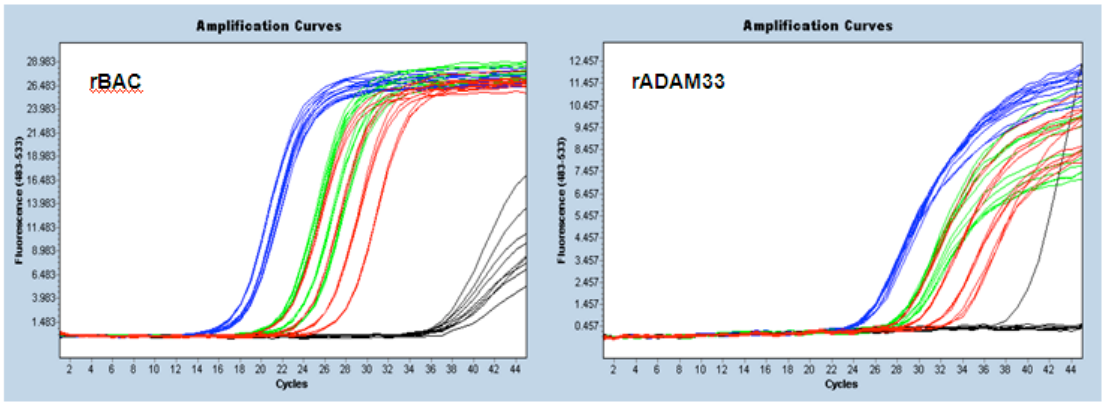


Figure 5-6: Validation of mRNA Stabilisation Methods Amplification Plots.

Blue plots = Liquid Nitrogen-Bio-Pulverizer treatment, Green plots = RNALater-ICE treatment, Red plots = Standard curve analysis, Black plots = negative control reactions. Y-axis is the change in fluorescence and x-axis is the PCR cycle number (Cp value). The Liquid Nitrogen-Bio-Pulverizer treated samples are consistently amplifying earlier than the RNALater-ICE treated samples for both the Actb (left panel) and Adam33 (right panel) genes.

Between mRNA stabilisation treatment groups the *Adam33* relative quantification ratios differed by up to 6 fold in concentration when comparing biopsies from the same lung (Figure 5.7); and the Inter-sample variability in concentration with in groups is less in the Liquid Nitrogen-Bio-Pulverizer treated samples, 1.05 ± 0.158 (mean ± std error) with 26 %CV; than in the RNALater-ICE treated cDNA template, 4.75 ± 1.37 (mean ± std error) with 51 %CV.

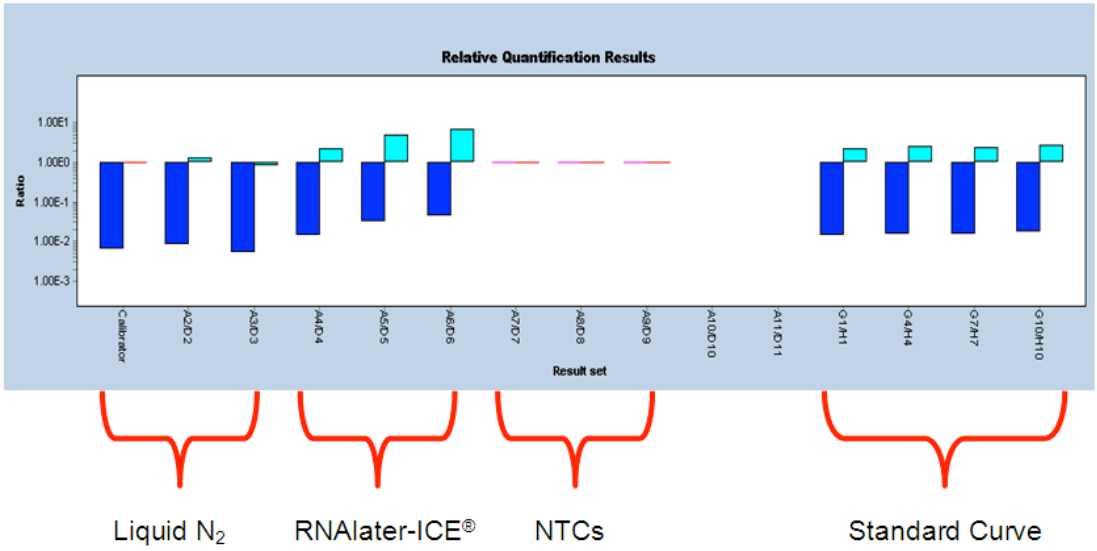


Figure 5-7: Validation of mRNA Stabilisation Methods Relative Expression Chart.

Light blue bars are the calibrator normalised concentration values used to determine the relative quantification of Adam33 to Actb. Y-axis is the log_{10} concentration ratios and each result set is shown on the x-axis. It is clear to see the small change in relative concentration values of ADAM33 in the Liquid Nitrogen-Bio-Pulverizer treated (Liquid N_2) lung biopsy group compared to the RNALater-ICE® treated samples.

From these results it is apparent that the Liquid Nitrogen-Bio-Pulverizer method is superior in maintaining RNA integrity consistently between tissue samples so that any result by qPCR should be a true reflection of transcript expression levels within rat lung.

5.4.3 RNA, DNA and Protein Yields

RNA isolates were obtained from all lung biopsies. RNA yield was from 7 to 100 µg, with 260/280 ratios of 1.97 to 2.11. DNA isolates were obtained from all lung biopsies and stored at -80°C ready for future analysis. DNA yield was from 22 to 160 µg, with 260/280 ratios of 1.7 to 2.0. Large protein pellets (~1 cm diameter) were obtained from all lung biopsy samples. All pellets were frozen in liquid nitrogen according to section 2.2.3.7 ready for future analysis.

5.4.4 Optimisation & Validation of the mRNA qPCR Assays

5.4.4.1 Optimisation of mRNA qPCR Assays

All SYBR green assays were successfully optimised according to section 5.3.6; the optimal primer concentration was determined to be 600 nM for both the forward and reverse primers. The TaqMan® and UPL based probe assay optimal concentrations were 900 nM and 200 nM for the primers and probe, respectively. None of the SYBR green primer sets amplified excessive primer dimer products (figure 5.8) only two assays (*Atp1a1* and *Atp1a2*) were found to generate small amounts of primer dimer product. In order to remove any primer dimer contribution to the SYBR green signal the Lightcycler 480 machine was programmed to acquire the SYBR green data during a short 80°C extension step at each PCR cycle. At this temperature the primer dimer product (max T_m = 78°C) is melted in to single stranded DNA, thus SYBR green molecules cannot bind and fluoresce.

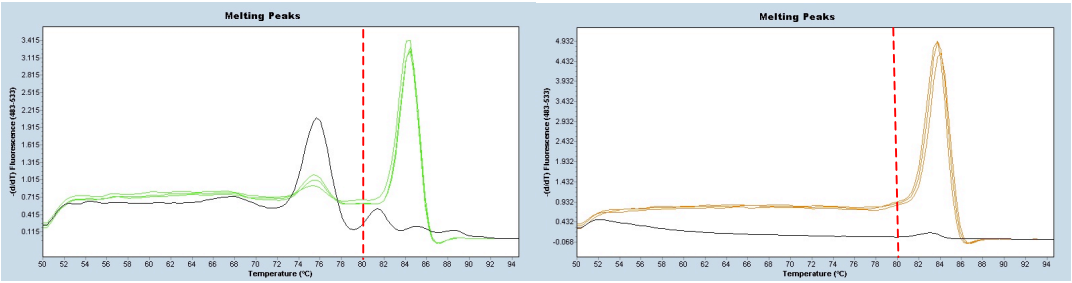


Figure 5-8: SYBRgreen Melting Curve Analysis.

Atp1a2 (left panel, green curve) and IGF1 (right panel, red curve) melting peaks from the Lightcycler 480 machine. It is clear from these melting plots that the Igf1primers forms a single product peak with no primer dimer production in either the cDNA reactions or water control reaction (black curves). While in the Atp1a2 melting plot a primer dimer product peak can be clearly seen in the water control, but it is less pronounced in the cDNA reactions. The red dashed line indicates the temperature at which SYBR green fluorescence is acquired by the Lightcycler 480 machine, to remove any primer dimer contribution to the signal.

5.4.4.2 Validation of mRNA qPCR Assays

Standard curves were performed for all assays as described in section 5.3.7 (see figure 5.9).

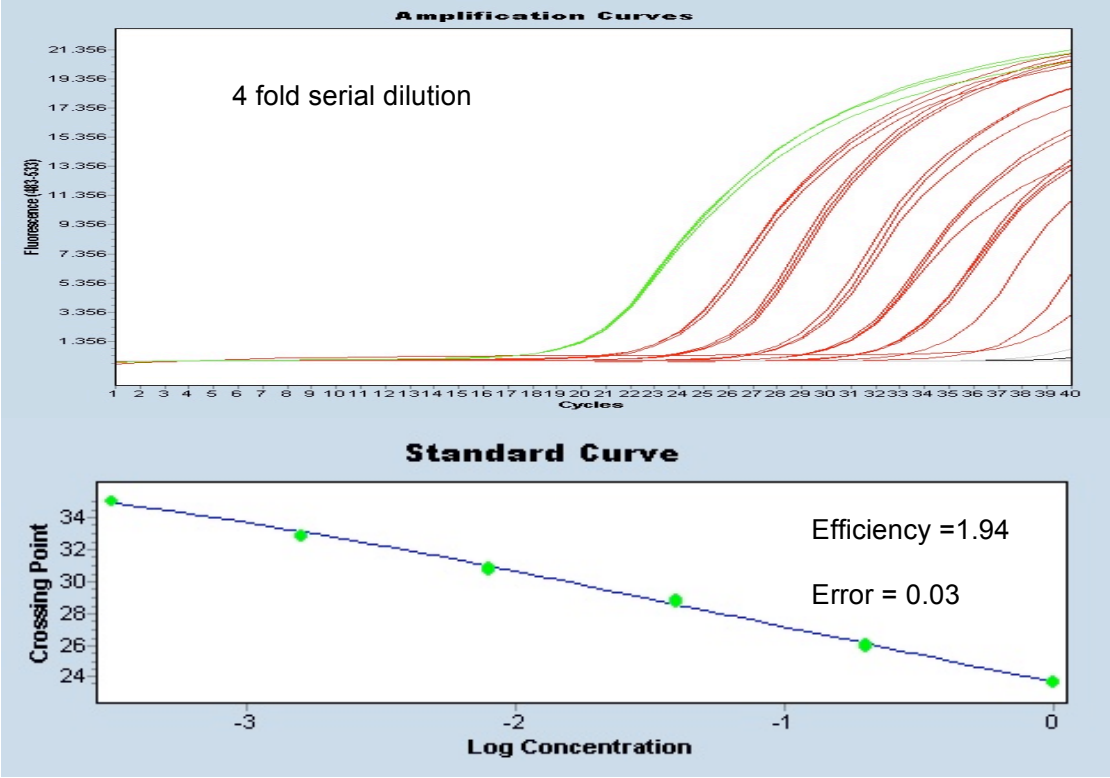


Figure 5-9: *Gstp1* Standard Curve Analysis.

This figure provides an example of a successful standard curve analysis. Each dilution step (4^0 – 4^{-7}) was performed in triplicate 10 µl reactions. The linear range of this assay extends over a 4^0 – 4^{-5} dilution of cDNA template with a high PCR efficiency of 94% before the triplicate wells are no longer reproducible because of the low copy number of template cDNA present in the 4^{-6} & 4^{-7} dilution reactions.

Table 5.3 shows which assays were successfully validated by standard curve analysis, (which generates a PCR efficiency estimate from the slope of the curve) and the other assays that did not generate a standard curve over the required minimum three dilution steps. It would appear that more cDNA template is required to generate an acceptable standard curve for the majority of the SYBR green assays.

Table 5-3: PCR Efficiency Determination

Standard curve or LinRegPCR derived calibrator sample PCR efficiency values.

Gene Name	Assay Type	Adjusted for PCR efficiency	PCR efficiency estimate E (dilution range , std error)
Adam33	UPL	Standard curve	1.98 (4 dil, error 0.013)
Atp1a1	SYBR green	Standard curve	2.03 (4 dil, error 0.009)
Atp1a2	SYBR green	LinRegPCR	1.82
GR	TaqMan	Standard curve	1.83 (5 dil, error 0.06)
Gstm1	SYBR green	Standard curve	1.98 (6 dil, error 0.04)
Gstp1	SYBR green	Standard curve	1.94 (6 dils, error 0.027)
Gstt1	SYBR green	Standard curve	1.92
Hsd11b1	SYBR green	LinRegPCR	1.86
Hsd11b2	SYBR green	LinRegPCR	1.94
Igf1	SYBR green	LinRegPCR	1.85
MR	SYBR green	LinRegPCR	2.0* assumed
Igf2	SYBR green	LinRegPCR	1.81
Pcdh1	UPL	LinRegPCR	1.94
Tp53	SYBR green	LinRegPCR	1.87
rUCP	SYBR green	LinRegPCR	2.0* assumed
Actb	TaqMan	Standard curve	1.96 (4dil, error 0.03)
B2m	TaqMan	Standard curve	1.85 (6 dil, error 0.015)
Ubc	TaqMan	Standard curve	1.89 (5 dil, error 0.01)

Unfortunately, the amount of cDNA template spiked into each reaction is already at the recommended 10% of the final reaction volume. It is well established that >10% volumes of cDNA template has a negative impact on the performance of the PCR reaction. Furthermore, to increase the volume of the cDNA template it would be necessary to use 20 µl reaction volumes and that would double the cost of each reaction setup. Therefore, to maintain performance and minimise cost these assays had to use the LinRegPCR method of PCR efficiency estimation for relative quantification by qPCR.

5.4.4.3 Minimum Detectable Difference in Expression between Treatment and Control Samples

The minimum detectable difference in expression between treatment groups was calculated using the F_1 generation result data set for the three reference genes, according to section 5.3.10.1. Table 5.4 shows the difference as the geometric mean of expression levels of the three possible reference gene combinations between treatment groups. Immediately, it is apparent that the expression of the three reference genes is similar in magnitude but not exactly the same. Therefore, it would not be correct to assume that a difference in the target genes between the groups is due to the treatment alone; because this difference could actually be caused by variation in the reference genes between the treatment and control group samples.

Table 5-4: Fold Expression for all Reference Gene Combinations for the F1 Generation Lung mRNA Samples.

These values are used to estimate the minimum detectable difference in expression between the sample groups.

Group	Geometric mean of expression	SD (std error)
Protein Restriction	5.9	1.5 (0.57)
Control	4.7	1.85 (0.56)

The minimum detectable difference between treatment groups is calculated by simply dividing the expression value of the protein restriction group by the control group:

$$>1.25 \pm 0.81 \text{ (fold expression} \pm \text{SD)}$$

At first glance, it is clear that a 25% or less change in expression in the target gene is most likely to be due variation in the reference genes. The standard deviation for the minimum detectable difference is relatively high; therefore it would be prudent to adjust the difference value by the addition of three standard deviations (covering 99 % of the normal distribution of reference gene variation). With this highly conservative adjustment, the minimum detectable difference between the treatment groups would be a 3.7 fold change in expression.

5.4.5 mRNA qPCR Data Analysis

5.4.5.1 qPCR Performance of 120 Day Old F_1 Generation Lung Samples

IUGR marker results from three out of the eleven 120 day RNA isolates from the control group had to be excluded from analysis because of poor quality qPCR data; all result sets from the seven protein restriction lung isolates were used in data analysis. For the oxidative stress marker panel results from all 120 day lung samples were suitable for data analysis. All of the qPCR data from the 225 day old lungs were suitable for analysis and all control negative controls (RT negatives, water controls, TRIzol™ extraction negatives) from all qPCR runs were free from contamination.

5.4.5.2 IUGR Marker and Candidate Asthma Gene Analysis in 120 Day F_1 Generation Lungs

Figure 5.10 displays the distribution of relative quantification normalised ratios between the protein-restricted and control groups for all of the genes assayed in this experiment. Examination of the chart shows there are only three outlying data points and the inter-sample variability within groups is acceptable for the majority of the genes, with the possible exception of the *Hsd11b1* control group data. Any results for this gene should be interpreted with caution and may need to be repeated in the future to increase confidence in the reported *Hsd11b1* normalised ratio values for the control group.

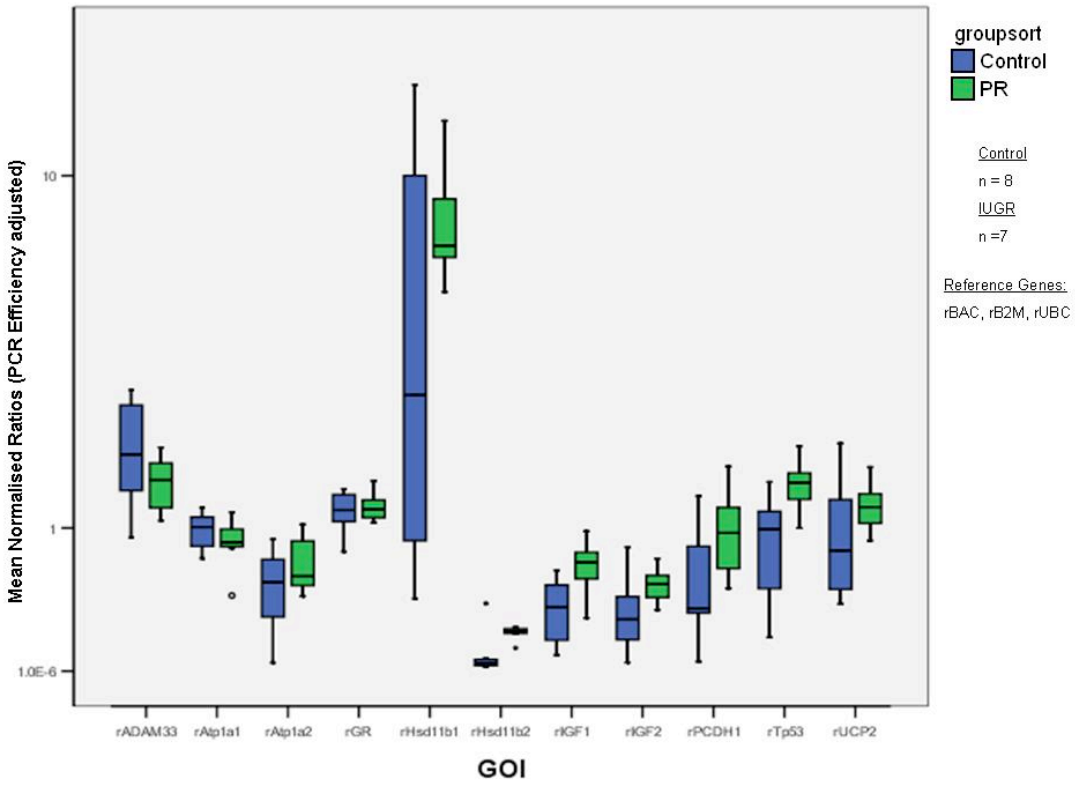


Figure 5-10: 120 Day F_1 Generation Rat Lung qPCR mRNA Expression Results.

Box and whisker plot of the geometric mean of normalised concentration ratio (y-axis) for the IUGR marker genes (x-axis). Boxes display the 25%-75% percentile range, whiskers the 0% & 100% percentiles and the horizontal lines are the median value for the protein-restricted (green boxes) and control lung groups (blue boxes).

Statistical analysis of this data revealed that relative quantification values for 5 genes were significantly higher in the protein-restricted rat lung samples (*Hsd11b2*, *Igf1* & 2, *Pcdh1* and *Tp53*); the *Ucp2* gene also showed a suggestive trend (P = 0.08) towards increased values (see figure 5.11).

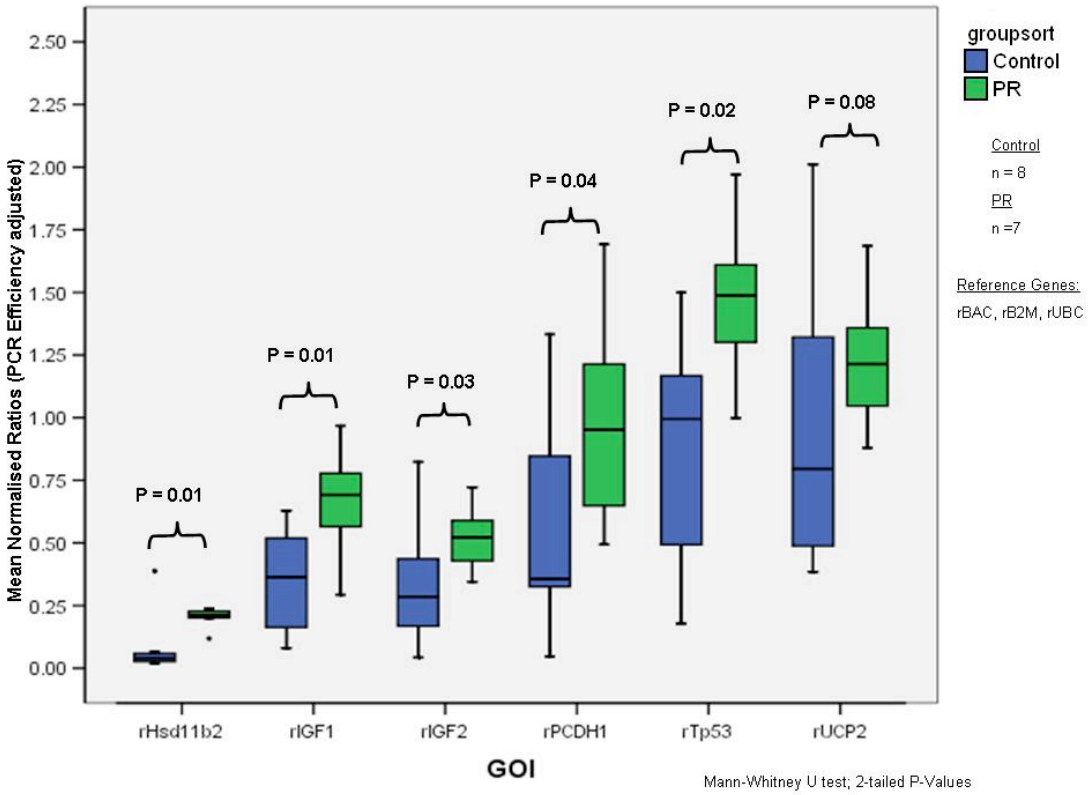


Figure 5-11: 120 Day F_1 Generation Rat Lung qPCR mRNA Differential Expression Results.

Box and whisker plot of the geometric mean of normalised concentration ratio (y-axis) for the IUGR marker genes (x-axis). Boxes display the 25%-75% percentile range, whiskers the 0% & 100% percentiles and the horizontal lines are the median value for the protein-restricted (green boxes) and control lung groups (blue boxes).

The fold difference (i.e. ratio of PR over Control) of median relative quantification values between the two groups is clearly shown in figure 5.12, it is apparent from this chart that *Hsd11b1* could also be increased in the protein-restricted group but this failed to reach statistical significance (P = 0.247).

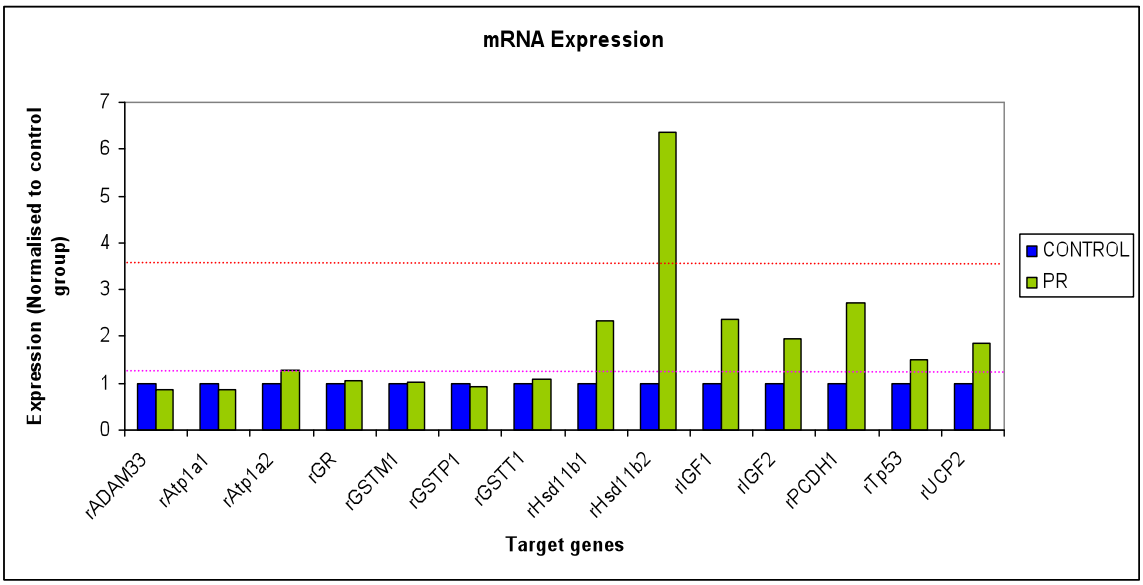


Figure 5-12: Fold Difference in Median Relative Quantification Values.

Minimum detectable difference of 1.25 fold (pink dashed line) and the more conservative adjustment of 3.7 fold (1.25 +3SD; red dashed line). All of the genes with differential mRNA expression levels are above the minimum detectable difference of 1.25 fold (increase greater than 25%).

From figures 5.10 and 5.11 it is clear to see that although *Hsd11b2* appears to be less abundant than other transcripts (i.e. small normalised ratio values compared to other genes) it has the highest level of differential expression between groups and it is possible that this gene is tightly regulated (small distribution of normalised ratio values).

Table 5-5: Fold Expression Levels in F_1 Generation Lung mRNA Isolates

	Hsd11b2	*Igf1*	*Igf2*	*Pcdh1*	*Tp53*
IUGR / CONTROL	**6.3**	**2.4**	**2.0**	**2.7**	**1.5**

Hsd11b2 has the largest fold difference (6.3 fold) between the two groups and is the only gene to remain above the threshold for the most conservative minimum detectable difference adjustment of >3.7 fold (see figure 5.12 and table 5.5).

5.4.5.3 IUGR Marker Gene Analysis in 225 Day F_1 Generation Lungs

Hsd11b2, Igf1, Igf2 and Tp53 mRNA levels in 120 day old rat lung were found to be significantly increased in PR offspring. Therefore, those transcripts (including *Hsd11b1*) were assayed in the 225 day lung RNA samples in order to replicate the 120 day lung findings at a later sampling time point.

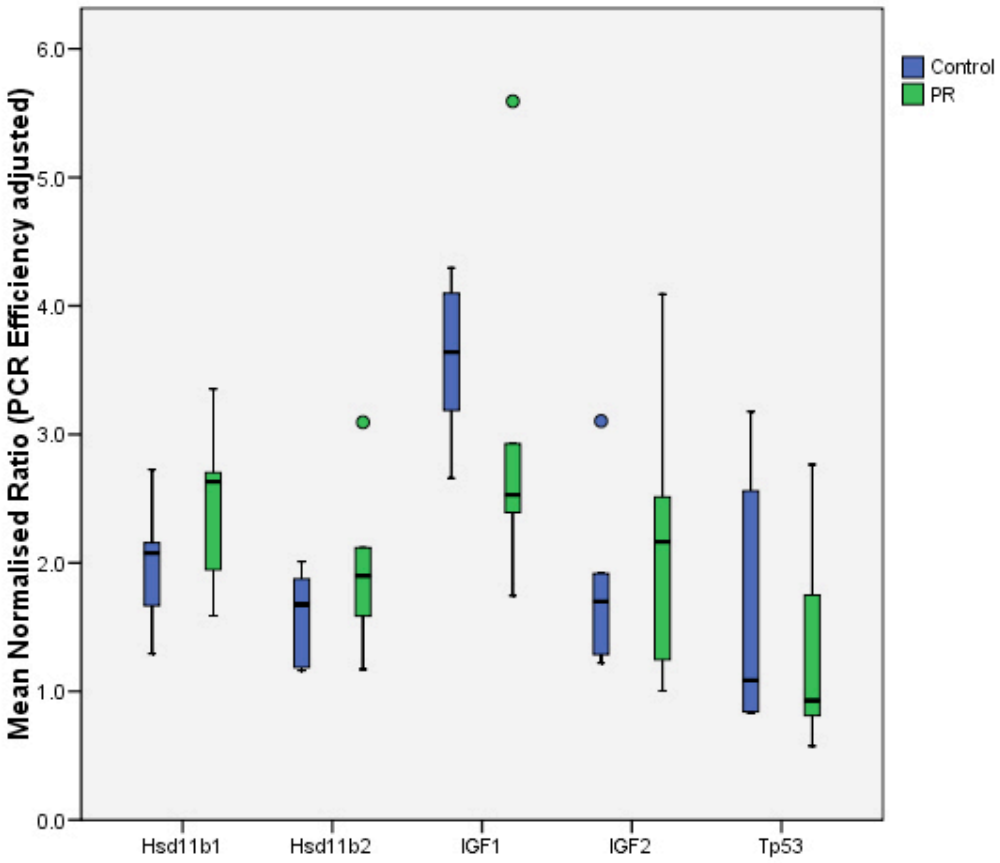


Figure 5-13: IUGR mRNA Expression in 225 Day F_1 Generation Rat Lung

Box and whisker plot of the geometric mean of normalised concentration ratio (y-axis) for the IUGR marker genes (x-axis). Boxes display the 25%-75% percentile range, whiskers the 0% & 100% percentiles and the horizontal lines are the median value for the protein-restricted (green boxes) and control lung groups (blue boxes). PR lung, n =6. C lung , n = 6.

Figure 5.13 shows that at 225 days of age mRNA levels are similar between PR and control lung for *Hsd11b1 & 2, Igf2* and *Tp53*. Furthermore, *Igf1* mRNA levels appear to be lower in control compared to PR lungs, this is the opposite to the 120 day findings where *Igf1* levels were significantly increased 2.4 fold in the PR group.

5.4.5.4 Oxidative Stress Marker Analysis in 120 Day F_1 Generation Lungs

Relative quantification levels of the glutathione s-transferase genes were not significantly different between the F_1 generation protein-restricted and control rat lungs when examined by a Mann-Whitney test (see figure 5.14). The box and whisker plot displays several outlying data points in the *Gstm1* and *Gstp1* data that are not representative of each group and may be caused by qPCR data of borderline quality. The outliers should not impact the statistical test as ranked values are used to test for differences; therefore the outlying data was not excluded from analysis. Although, statistically non-significant (P = 0.15), the geometric mean of *Gstp1* in the protein-restricted group is slightly lower and the inter-sample variability of *Gstp1* is much less when compared to the control lung group; this could indicate that expression of this gene is tightly controlled and slightly down regulated in response to protein restriction.

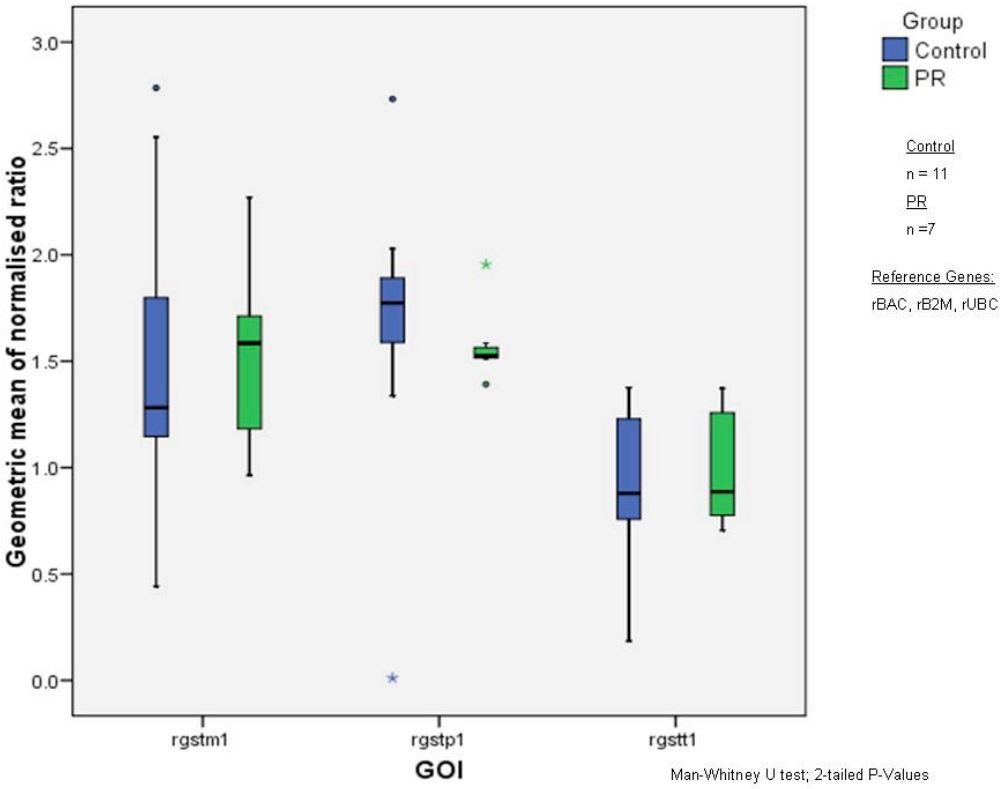


Figure 5-14: GST mRNA Expression in 120 Day F1 Generation Rat Lung.

Box and whisker plot of the geometric mean of normalised ratio (y-axis) for the Gstm1, Gstp1 & Gstt1 genes (x-axis). Boxes display the 25%-75% percentile range, whiskers the 0% & 100% percentiles and the horizontal lines are the median value for the protein-restricted (green boxes) and control lung groups (blue boxes).

5.4.5.5 qPCR Performance of F_2 Generation Lungs

IUGR marker results from two out of the thirteen RNA isolates from protein restricted lung group had to be excluded from analysis because of poor quality qPCR data. Similarly, three control lung data sets had to be excluded from the IUGR marker analysis. For the oxidative stress marker panel results from all samples were suitable for data analysis. All control negative controls (RT negatives, water controls, TRIzol™ extraction negatives) were free from contamination.

5.4.5.6 IUGR Marker & Candidate Asthma Gene Analysis of 80 Day F_2 Generation Lungs

Statistical analysis of the IUGR marker data for the F_2 generation lung samples revealed no significant difference in relative quantification levels between groups in 7 out of the 8 genes that were assayed by qPCR (see figure 5.15). Median *Hsd11b1* relative quantification levels were significantly lower in the protein-restricted group (0.89, range = 0.06 – 3.30) compared to the control group, where the median level was 3.01 (0.27 – 7.59). In the F_1 generation lungs, this gene was significantly higher in the protein-restricted group (6.3 fold increase).

The small numbers of outlying sample data points were not excluded from the analysis, as they are unlikely to generate spurious significant results in a non-parametric statistical test. The disproportionately large inter-quartile ranges of the *Igf2* and *Tp53* geometric mean normalised concentration ratios for both sample groups could indicate a lack of reproducibility of the qPCR data and the results for these genes should be interpreted with caution.

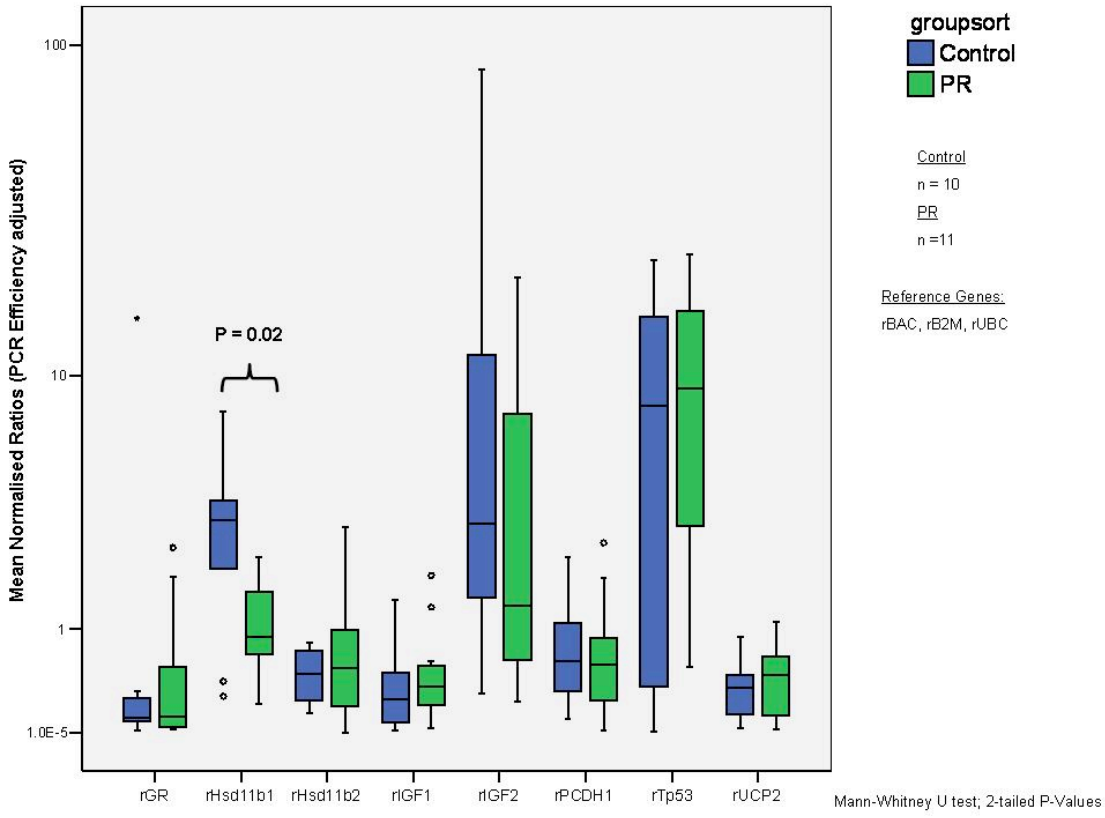


Figure 5-15: F_2 Generation Rat Lung qPCR mRNA Expression Results.

Box and whisker plot of the geometric mean of normalised concentration ratio (y-axis) for the IUGR marker genes(x-axis). Boxes display the 25%-75% percentile range, whiskers the 0% & 100% percentiles and the horizontal lines are the median value for the protein-restricted (green boxes) and control lung groups (blue boxes).

5.4.5.7 Oxidative Stress Marker Analysis of 80 Day F_2 Generation Lungs

Relative quantification levels of the glutathione s-transferase genes were not significantly different between the F_2 generation protein-restricted and control rat lungs when examined by a Mann-Whitney test (data not shown).

The non-significant trend towards lower levels of *Gstp1* mRNA transcripts in the 120 day F_1 generation PR lung samples (see section 5.4.5.4) was not evident in the second generation samples.

5.4.5.8 Replication of 120 F_1 Generation Lung Results

The observation of altered mRNA levels in 120 day F_1 PR lung was not replicated in 225 day F_1 PR lung or in 80 day F_2 PR lung. In order to demonstrate the reproducibility of the findings in 120 day lung, the qPCR experiment was repeated using pooled cDNA templates (i.e. PR or Control) from two further cDNA synthesis reactions. Pooling the cDNA template by treatment helped to reduce reagent costs.

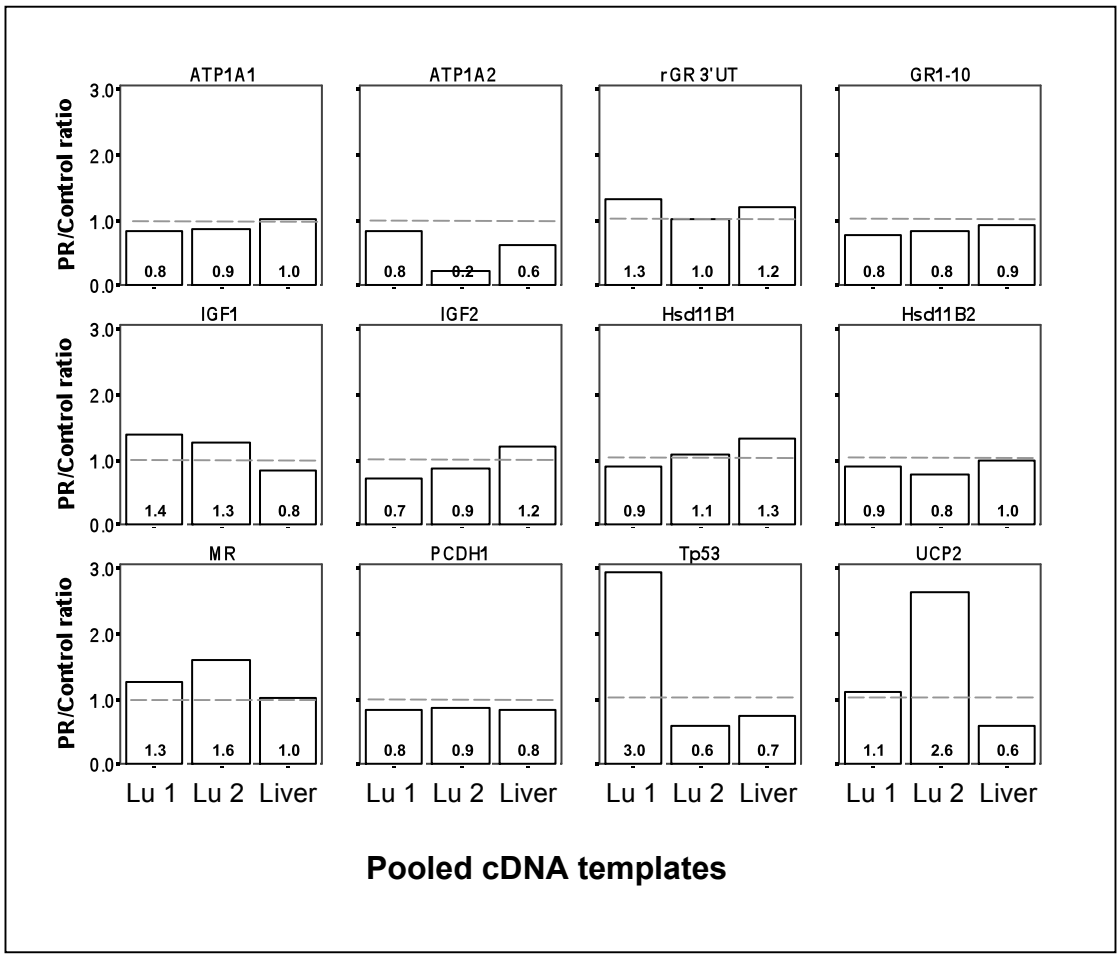


Figure 5-16: Ratio of PR Versus Control mRNA Levels from Pooled 120 Day Lung and Liver cDNA Templates.

The ratios of PR versus control expression do not reflect the original findings in 120d PR lung (see table 5.4) and there are considerable differences in ratios between subsequent cDNA syntheses. Lu1 = 1st lung cDNA pool. Lu1 = 2nd lung cDNA pool. Liver = liver cDNA pool. rGR 3'UT = GR 3'UTR. GR1-10 = GR transcript containing variable exon 1-10.

As liver RNA was available for the same 120 day animals and liver *GR* expression had been previously found to be increased 300% by PR in 6 day old pups;[426] it was reverse-transcribed and included in this experiment.

Figure 5.16 demonstrates the variability in PR/control ratio levels between two separate cDNA syntheses and highlights a substantial lack of replication of the original findings from the same RNA samples. The original 120 day findings that have been marginally reproduced by this experiment, were *Igf1* (ratio of 1.4 & 1.3) and the *Tp53* and *Ucp2* ratios, but only in one out of the two cDNA templates used. Interestingly, *MR* ratios are higher in both pooled lung cDNA templates (ratio of 1.3 & 1.6). The *MR* transcript was not assayed in the original experiment due to delays in assay validation. Previous work using this rat model had demonstrated that the expression of GR was increased 3 fold in 6 day old PR liver,[426] however our results do not replicate those findings.

5.4.6 Summary of Gene Expression Results

Table 5.6 presents a summary of the statistically significant results from the qPCR experiments performed on the F_1 and F_2 generation rat lung mRNA isolates.

Table 5-6: Summary of Gene Expression Results

Panel	Gene	Generation	Expression in PR vs control	Fold change in expression	P value
Asthma candidate	*Pcdh1*	F_1	**UP**	**2.7**	**0.04**
		F_2	-	-	-
IUGR marker	*Hsd11b1*	F_1	-	-	-
		F_2	**DOWN**	**3.4**	**0.02**
	Hsd11b2	F_1	**UP**	**6.3**	**0.01**
		F_2	-	-	-
	Igf1	F_1	**UP**	**2.4**	**0.01**
		F_2	-	-	-
	Igf2	F_1	**UP**	**2.0**	**0.03**
		F_2	-	-	-
	Tp53	F_1	**UP**	**1.5**	**0.02**
		F_2	-	-	-
	rUCP	F_1	**UP**	**1.9**	**0.08**
		F_2	-	-	-

5.5 Discussion

5.5.1 Lung Morphometry

Previous studies of protein restricted offspring have found that at birth, body weight and certain organ sizes are reduced in comparison to the control offspring,[232, 237] although other studies have found no change.[447] As the protein restricted offspring age, there appears to be catch up growth with body weight and organ size similar between PR offspring and controls.[420] No significant differences in body weight or lung weight or lung volume between protein restricted and control offspring were observed in 225 day old offspring in the current study. These findings support previous studies that found no change in the lung weights of older protein restricted offspring.[420] However, some studies have shown reductions in offspring lung size and volume by using more severe deficits in maternal nutrition or altering the time of exposure to protein restriction. A study of maternal under-nutrition in rats (50% reduction in food availability) showed reductions in lung weight/body weight and lung volume/body weight ratios in 42 day old offspring when compared to controls.[442] In 1996 Desai et al. found reduced lung weights in 21 day old rat offspring from dams that were fed a protein-restricted diet during pregnancy and weaning or weaning alone suggesting that post-natal nutrition is important for normal lung growth.[420]

5.5.2 Optimisation & Validation of the TRIzol™ Isolation Protocol

RNA, DNA and protein isolates were successfully obtained from all biopsies. DNA and RNA purity was in the acceptable range of each nucleic acid when determined by absorbance spectrophotometry (260/280 ratio for DNA ≈ 1.8; RNA ≈ 2.0).[448] Therefore, contamination of the nucleic acids with protein or solvent is at a minimum and will not have a detrimental effect on PCR. In the future, an aliquot of all RNA isolations will be processed on an Agilent RNA micro-fluidics chip (Agilent Technologies, CA, USA) to determine the integrity of RNA molecules in the sample. It is well known that RNA integrity is crucial for accurate mRNA analysis by qPCR.[449]

5.5.2.1 mRNA Stabilisation Methods

The liquid Nitrogen (N_2) Bio-pulverizer mRNA stabilisation method was superior compared to the RNAlater-ICE treatment. In general, the liquid N_2 Bio-pulverizer method improved the sensitivity of the qPCR assays ten fold, by virtue of consistent maintenance of RNA integrity. The ultra cold freeze fracture technique enabled rapid permeation of the TRIzol reagent into the tissue, whereby inactivating any endogenous or exogenous RNase activity. Increased inter-sample variability in the RNAlater-ICE treated group and differences in relative quantification ratios for *Adam33* and *Actb* between the two stabilisation methods indicated that RNA integrity was not consistently maintained in the RNAlater-ICE treated biopsies. Differences between the dimensions of each biopsy could account for this variation because the rate of permeation into the tissue is dependant on the size and composition of each biopsy. RNA integrity numbers from future Agilent RNA chip analysis should confirm these conclusions. If RNA integrity remains an issue for the validity of gene expression results qPCR assays can be designed for the 5' and 3' of each transcript of interest and used to ascertain any degradation by deviation from a 1:1 ratio in concentration between the two assays.

5.5.3 qPCR Assay Design, Optimisation and Validation

qPCR assays were successfully designed, optimised and validated by melt curve analysis and either standard curve analysis or PCR efficiency estimation by LinRegPCR. From the standard curve analysis results (section 5.4.4.2) it is apparent that the SYBR green assays did not perform as well as the probe based assays. In the future, performance of these assays can be improved by the inclusion of a fluorescent probe.[450] As the SYBR green assays were originally designed with this in mind, a UPL probe binding site is available in each PCR amplicon, and optimisation should be relatively simple.

5.5.4 Minimum Detectable Difference in Gene Expression

Accurate determination of relative quantification values is limited by variation in reference gene expression between biological replicates in the treatment and control groups. In theory, reference gene expression should be stable and consistent, in practice this may not be the case as experimental treatments may have unexpected consequences on reference gene expression. In this study, variation in the three reference genes was determined by examination of the geometric mean of each gene-pairing in the control and protein-restricted lungs. The minimum detectable difference was initially calculated at 1.25 ± 0.81 fold; when accounting for 99% of the variation in expression this was adjusted to 3.7 fold. This value is relatively high when considering the detected fold differences in gene expression in the F_1 generation lungs (1.5 – 6.3 fold). Therefore, a more robust strategy for testing the stability of reference genes is required; a solution is the GENORM method as it calculates a pair-wise stability index for up to twelve candidate reference genes.[445] In addition to GENORM, inclusion of a panel of genes known to be stably expressed in the rat lung may provide a more reliable normalisation factor for any further analysis.[446]

5.5.5 Gene Expression in F_1 and F_2 Generation Lung Tissue

5.5.5.1 Markers of IUGR

The expression of six markers of IUGR were significantly increased in the 120 day protein restricted lung but not in 225 day F_1 or 80 day F_2 PR lung, when accounting for ~99% of the variation in reference gene expression only *Hsd11b2* (6.3 fold increase) remained above the 3.7 fold threshold. Therefore, it is likely that differential expression of this gene is real and is not an artefact of experimental variation. As discussed earlier (sections 5.4.4.3 and 5.4.5.2), the 3.7 fold threshold may be too conservative. As the difference in expression between groups for the six genes is above the 1.25 fold minimum detectable difference threshold (see figure 5.12), the results of all six genes will be assumed to be real and any implications of the findings will be discussed here.

5.5.5.1.1 Atp1a1 & 2

Atp1a1 & 2 gene expression was unaltered by protein restriction and expression levels remained the same in the next generation. Atp1a1 mRNA levels were previously associated with higher GR mRNA levels in the lungs of MLP (maternal low protein diet) rat offspring at 20 weeks (140 days).[428] As GR levels are also unaffected in the current study at day 120, it is likely that either PR has no effect on lung expression of Atp1a1 & 2 genes or that any affect is transient, and can only be detected at earlier time points. Alternatively, lack of replication with the *Bertram et al*.[428] study mentioned earlier, could indicate a problem with the qPCR experiment for these gene targets or animal model itself. As the lung samples used in both studies came from the Southampton low protein diet rat model housed at the same institution.

5.5.5.1.2 GR and *MR*

Glucocorticoid receptor expression in the lung was also unaffected by protein restriction in this experiment. This result was surprising as GR expression has been shown to be consistently up-regulated in response to protein restriction in several organs and programmed GR expression can be transmitted to the F_2 generation.[264, 428] The GR gene consists of a variable and constant coding region that generates tissue-specific splice forms,[451] the decision was made to design a common primer set to the 3' end of the GR mRNA transcript as the most abundant GR transcript in lung tissue has not been identified. The previous study used a primer set specific for the tenth variable exon ($GR1_{10}$) because it was the most widely expressed splice form and is abundant in their target organ, the liver.[264] The $GR1_{10}$ exon promoter was then found to have altered DNA methylation status in response to PR exposure in the F_1 and F_2 generation offspring.[264] However, profiling GR splice variants expressed in lung tissue in the current work by using a primer set was designed to detect the $GR1_{10}$ exon did not discover any alteration to GR splice form expression between the PR and control lungs and reflects the findings with the 3'UTR assay.

Unfortunately, validation of *MR* assay performance was not completed until after the 120, 225 and 80 day lung experiments. However, the results from the pooled cDNA experiments (testing the reproducibility of real-time PCR data) indicate that Mineralcorticoid receptor expression might be sightly increased in 120 day PR lung (1.3 - 6 fold; see section 5.4.5.8).

A previous study of low maternal protein in rats found that offspring MR expression was unchanged *in ute*ro and early life.[428] Mineralcorticoids are thought to stimulate lung liquid re-absorption by binding the mineralcorticoid receptor on lung epithelium in late-gestation ovine fetal lung and may play a pivotal role in the transition from intrauterine to extrauterine life.[432] However, the programming of rodent hypothalamic-pituitary-adrenal axis is different to sheep and the consequence of increased MR expression in adult PR offspring is unclear.[452] This observation awaits confirmation in late gestation fetal and neo-natal rat lung.

5.5.5.1.3 Hsd11b1 & 2

Hsd11b1 & 2 have refractory effects on glucocorticoid action; *Hsd11b1* catalyses the conversion of cortisone to bioactive cortisol, which can activate glucocorticoid receptors, whereas *Hsd11b2* catalyses the inactivation of cortisol to cortisone and maintains the aldesterone specificity of the minerocorticoid receptor.[453, 454] The programming of hypertension and dysregulation of glucose metabolism are thought to be mediated by inhibition of placental Hsd11b2 activity.[455] However, the consequence of altered glucocorticoid action in the lung by protein restriction is not well understood.

In 2001, Bertram et al.[428] investigated the mechanism of glucocorticoid action the same rat PR model as the current study, reporting that *Hsd11b1* expression was not changed by PR and that *Hsd11b2* mRNA expression was significantly reduced in foetal lung at day 20 of gestation (term = 22days).[428] They also found that *Hsd11b2* mRNA was undetectable in post-natal lung tissue by northern blot analysis.[428] In the current study, *Hsd11b2* expression was significantly increased in the 120 day F_1 generation PR lung (6.3 fold), but not at 225 days of age or in the F_2 generation, where *Hsd11b1* was found to be down regulated 3.4 fold compared to control animals. The decreased *Hsd11b1* mRNA levels in the PR F_2 generation lung tissue in the current study are intriguing. However, they require clarification by further experimentation as no change was observed in the F_1 generation, in fact levels tended to be higher in the F_1 PR group. Utilisation of real-time PCR in the current study has demonstrated that not only is *Hsd11b2* expressed in the post-natal lung, there appears to be a programmed up-regulation of mRNA levels in response to *in utero* protein restriction. *Hsd11b2* has been found to be present in human airways by immunological staining and is localised to vascular endothelial cells in the lamina propria and bronchial epithelium.[456, 457]

In 2003, Susuki et al.[457] concluded that human bronchial epithelial cells possess a putative autoregulatory system for natural and synthetic glucocorticoids and control their own bioactive levels by inducing the expression of Hsd11b2.[457] Therefore, an up-regulation of expression could potentially increase the capacity of Hsd11b2 to down regulate glucocorticoid action in the lung. However, it has been suggested that *Hsd11b2* can act as a oxidoreductase and reactivate synthetic glucocorticoids such as dexamethasone.[458] Furthermore, it has been suggested that Hsd11b2 may modulate the efficacy of systemic glucocorticoids given therapeutically and that steroid resistance in asthmatics could be caused by an adverse balance in metabolism.[456] Therefore, the results from this study suggest that *in utero* protein restriction alters the balance between levels of *Hsd11b1* & *2* and this could affect the control of glucocorticoid action in lung tissue and may be clinically relevant to inhaled corticosteroid treatment in humans. Further examination of *GR*, *MR Hsd11b1* & 2 mRNA levels in earlier time points will clarify the effect of PR on the control of glucocorticoid action in the lung.

5.5.5.1.4 Igf1 & 2

Both Insulin-like growth factor genes were up regulated in response to PR in 120 day first generation animals and increased *Igf1* levels were confirmed by the reproducibility experiment (see section 5.4.5.8). However, expression of *Igf1* appeared to be lower in 225 day F_1 PR lung and similar in the F_2 animals. There is limited understanding of the consequences of IUGR on the IGF system in lung. Chen et al.[425] reported that increased lung *Igf1* mRNA expression can be detected up to postnatal day 28 in rats exposed to maternal under nutrition (50% total nutrient restriction) during the last trimester of pregnancy. Results from this study show that in a rat protein restriction model *Igf1* mRNA expression is still up regulated (2 – 2.4 fold) 120 days after birth but not at 225 days; this is much later than in the previous study. Growth factors like the IGFs are important for the regulation of cell proliferation during lung development. Increased IGF gene expression was significantly up regulated in ovine models of accelerated postnatal lung growth.[459] Deletion of the *Igf1* & 2 genes in mice caused growth retardation and impairments in lung function at birth.[460, 461] and the IGF 1 receptor has been found to be important for lung vascular development in rat embryonic lung explants.[462]

5.5.5.1.5 Tp53

p53 is a transcription factor that plays a pivotal role in determining cellular response to stress and was found to be expressed 1.5 fold more in the 120 day F_1 generation PR lungs compared to controls, and increased *Tp53* levels were confirmed in one out of the two extra cDNA templates in the reproducibility experiment (see section 5.4.5.8) but not at 225 days of age, nor in the F_2 lungs. p53 is involved in alveolar development by regulation of cellular apoptosis. Increased alveolar cell apoptosis is seen in the emphysemic lung and p53 protein levels in lung parenchyma cells are significantly higher in subjects with emphysema.[463]

Uteroplacental insufficiency in rats has revealed an epigenetic programming mechanism through reduced CpG DNA methylation of the p53 gene promoter, resulting in increased p53 expression that correlates with key apoptosis related proteins, renal apoptosis, and reduced glomeruli number in the IUGR kidney.[464] Therefore, the *Tp53* gene is a potential candidate to test whether, the change in gene expression observed in the F_1 PR lung, is programmed by altered gene promoter DNA methylation. It is unlikely that the programmed adaptation is maintained in subsequent generations as the F_2 generation *Tp53* expression was unaltered and its expression levels may alter with age as *Tp53* levels were similar in the 225 day F_1 lungs. Programmed adaptations to p53 expression in response to IUGR may impact COPD susceptibility and modulate disease severity. Smokers expose their lung to excessive levels of oxidative stress and an amplified p53 response in an IUGR individual could leave them predisposed to accelerated alveolar cell apoptosis and by extension, earlier onset or increased severity of emphysema in later life.

5.5.5.1.6 Ucp2

There is a suggestive trend (1.9 fold; P = 0.08) to increased uncoupling protein 2 mRNA levels in the 120 day F_1 generation PR lung but not at 225 days of age or in the F_2 generation. It has been suggested by *Gnanalingham et al, 2005.*[427] who also found increased *Ucp2* expression in response to IUGR in sheep, that increased *Ucp2* abundance in later life could have deleterious consequences in lung function and enhanced susceptibility to infection. *Ucp2* deficient mice have an amplified immune response with greater macrophage phagocytic activity and increased ROS production, one of the phenotypes of these mice is a striking resistance to

infectious microorganisms.[465] These results provide suggestive evidence that programmed up-regulation of *Ucp2* expression by *in utero* protein restriction could result in a dampened immune response in offspring and increased incidence of infection during life. Infection during childhood may impair lung growth and make these individuals more vulnerable to developing COPD on exposure to additional injurious agents (reviewed by Sethi, 2000).[466] Therefore, it would be of interest to examine role of *Ucp2* genetic polymorphism in susceptibility to COPD and asthma.

5.5.5.2 Oxidative Stress Genes

All three GST gene transcripts were present in 120 day old rat lung and the abundance of transcripts was in the following increasing order, *Gstt1* > *Gstm1* > *Gstp1* in the F_1 generation. Interestingly, *Gstp1* expression was slightly less abundant than *Gstm1* expression in the F_2 generation samples and *Gstt1* transcript was detectable in rat lung; whereas some studies have suggested that *GSTT1* is not expressed in human lung,[467] whilst others have found it at low levels.[188, 190]

In both generations, there was no significant difference in expression levels of rGST transcripts between protein-restricted and control groups. Thus, it appears that there is no evidence of alterations in expression of anti-oxidant glutathione s-transferase genes in the lung tissue from animals exposed to protein-restriction during gestation. mRNA expression is not always representative of protein levels, without evidence to suggest that this is the case in this model, the GST expression results could reflect that both the PR and control animals have similar oxidative stress in the lung at 120 days of life. *GSTO1* (Glutathione S-Transferase Omega 1) has recently been found to be expressed in alveolar macrophages, airway and alveolar epithelium and airways secretions.[468] In COPD patients, sputum and lung homogenate levels of *GSTO1* have been found to be lower when compared to non-smokers.[468] Again, as with the IUGR marker genes it would be very interesting to monitor GST gene expression profiles (including *GSTO1*) over time, including earlier and later sample time points (birth (day 0) and old age (day 300)) to determine if response to oxidative stress is affected by protein-restriction in early life. Ideally, incorporation of controlled tobacco smoke exposure into the rat PR model would facilitate the understanding of the effect of an imprudent maternal lifestyle (i.e. maternal ETS and poor nutrition during pregnancy) on risk of respiratory disease in her offspring.

5.5.5.3 Candidate Asthma Genes

As *Adam33* expression was not found to be differentially expressed in the PR F_1 generation lungs at 120 days of life, this gene was not assayed in the F_2 generation samples. *ADAM33* is thought to be important in the developing foetal lung but these results suggest that if PR did alter expression, these effects are transient and are not apparent at 120 days of life. In humans, *ADAM33* is more abundant in the embryonic lung, clustering around developing embryonic bronchi and in smooth muscle bundles, hence involvement of this gene in remodelling and bronchial hyper responsiveness.[394] Expression of mRNA splice variants and immunostaining of normal and asthmatic bronchial biopsies revealed no significant differences.[394] A knockout mouse study found that *ADAM33* is not essential for growth, development or allergic asthma.[469] Recent work has shown that DNA methylation of the *ADAM33* promoter controls its expression in a cell type-specific manner[439] and that the soluble form released from cell membranes of damaged lung epithelium cell results in a disease-related gain of function, that is independent of inflammation.[470] It is now thought that *ADAM33* has a functional role in airway remodelling by the promotion of angiogenesis.[470] Therefore, it would be of interest to examine expression of this gene in earlier sample time points, such as at birth (day 0) or earlier to determine if PR does have any affect on *ADAM33* expression during lung development *in utero,* because of the extensive remodelling of the foetal lung during development.

Pcdh1 expression was significantly increased in the 120 day old PR group (2.7 fold; $P = 0.04$) in the F_1 but not in the F_2 generation. Protocadherin 1 is involved in the maintenance of cell junctions and is expressed in airway epithelial cells and therefore may be important epithelial barrier function.[440] The expression of *Pcdh1* in PR lungs may indicate that there is increased angiogenesis and lung remodelling in these animals into adulthood (120 days). The findings with *Pcdh1* were not replicated in the F_2 generation lungs, suggesting that epigenetic mechanisms are not transmitting enhanced *Pcdh1* expression to the next generation in absence of a protein-restricted environment during pregnancy. In order to test this conclusion, results need to be confirmed by replication in more samples and changes to the expression of key angiogenic genes (e.g. VEGF) would provide extra evidence for alterations to the lung vasculature. Morphological analysis and staining of serial sections of PR rat lung tissue would confirm the presence of increased vascularisation.

5.5.5.4 Reproducibility of 120 Day PR Lung mRNA Expression Results

Attempts to replicate the 120 day PR lung mRNA expression results met with limited success with only the levels of *Igf1, MR, Tp53* and *Ucp2* in either two or all three of the cDNA templates generated from the same RNA isolate. The poor reproducibility of the results reflects the fact that in qPCR reverse-transcription of RNA into cDNA is technically the most variable step in any real-time PCR experiment.[471] As the observed changes in gene expression in this rat model are relatively small in comparison to the inherent experimental variability (e.g. variation in reference gene, mRNA stability and between cDNA templates), it is in some respects, unsurprising that some of the observations are hard to replicate with a small sample size at all sampling time points (i.e. 225 day lung).

Therefore, the gene expression findings presented in this chapter should be interpreted with caution until they are confirmed by further experiments utilising larger sample sizes across a range of sampling time points, with improvements in establishing a reference gene normalisation factor with less background variation (e.g. GENORM normalisation software can use up to twelve reference genes)[445] and using multiple cDNA templates (minimum of two replicates per biological sample) for real-time PCR to account for variation in the efficiency of the RT-PCR step.

5.6 Concluding Remarks and Future Work

In conclusion, the purpose of this work was to determine whether maternal PR alters lung growth and expression of key genes in lung development, establishing a 'baseline' of lung physiology and gene expression, with the ultimate aim of combining this model with an extra maternal exposure, such as tobacco smoke during pregnancy in future studies and monitor offspring respiratory phenotypes. The work presented in this chapter provides supporting evidence for programmed expression of key genes involved in hypertension, dysregulation of glucose metabolism, and mitochondrial energy balance and organ development in the lung, in response to *in utero* protein restriction in the F_1 generation at 120 days of age but not in older animals (225 days). The expression of genes involved in antioxidant defences are unaltered by *in utero* protein restriction and there is no evidence for the trans-generational programming of gene expression in lung. In the rat at 255 days of age there is no evidence to suggest that PR during development has resulted in permanent retardation to organ size or to respiratory function (lung volume).

At the start of this chapter two hypotheses were proposed (section 5.1.10) analysis of mRNA levels by qPCR has revealed that in a rat model of hypertension, a maternal low protein diet during pregnancy has programmed changes in gene expression that could alter the course of normal lung development, vascularisation, growth and later lung function (*Hsd11b2*, *Tp53*, *Pcdh1*, *Igf1* &2 and *Ucp2*) and even dampen immune responses in the offspring (*Ucp2*). Importantly, up-regulation of these genes, programmed *in utero*, have persisted in to adulthood (up to 120 days in the rat) and therefore could have implications for susceptibility to respiratory disease in later life (COPD) as well as in childhood (asthma). The mRNA levels of three key anti-oxidant GST genes are unaltered in 120 day old PR lung, indicating that protein-restriction does not appear to program the expression of cytosolic GST enzymes. In order to fully understand these findings, it is necessary to validate these findings in more sample time points during the life of the PR rats.

Gene expression analysis in other sample time points (time points could include, foetal (20 days gestation), birth (day 0), after weaning (day 21) and old age (day 300) would produce a gene expression profile from development until old age that could be used to examine the consequence of alterations to gene expression at

critical windows of lung development and help understand how gene expression changes as the animals age.

It is thought that epigenetic mechanisms (DNA methylation) could potentially control nutritional programming, but the F_2 gene expression results do not provide any evidence that these candidate genes are under the influence of trans-generational programming at 80 days of age. However, investigation of the epigenetic programming of other genes in PR lung could be performed by gene promoter DNA methylation analysis by real-time PCR melt-curves[472] or by pyrosequencing. The lung DNA, RNA and protein isolations collect in this study would prove to be very useful as any DNA methylation finding could be directly related to mRNA and protein levels from the same sample.

A more detailed approach to the analysis of lung morphology would determine if there are structural changes to the tissue, such as a reductions in alveolar abundance/size and alveolar capillary bed vascularity, which could be detrimental to normal lung function by reducing the gas exchange potential of the organ. Examination of lung tissue morphology by stereological methods would determine if exposure to protein restriction *in utero* alters airways size and number, airways wall thickness, smooth muscle/myofibroblast content or vascularity.

Other areas to address in future work include improvements to the quality control measures for qPCR, such as, accounting for the efficiency of reverse transcription step (by utilising multiple cDNA templates for each biological sample). RNA integrity could be determined by Agilent RNA micro-fluidic chips and confirmed by 5’ & 3’ mRNA qPCR primer sets. Conversion of the SYBRgreen based primer sets into probe based assays will improve specificity and performance of the assays in standard curve analysis. The selection of reference genes could be improved by performing pair-wise assessments using the GENORM method[445] to find the most stably expressed reference genes in the experimental model.

6 Protein-Restricted Rat Lung: micro RNA Expression Analysis

6.1 Introduction

The lung undergoes several stages of development, extending into early life, that are sensitive to alterations in the developmental environment. microRNAs (miRNA) are short nucleotides that can regulate the expression of individual and pathways of genes at the post-transcriptional level. They have an important role in lung development and are differentially expressed during development, to modulate developmental timing, cell fate determination and apoptosis. Currently, little is known about the regulation of miRNA expression. However, dietary components, such as folate have been shown to alter miRNA expression[473] and epigenetic mechanisms, including DNA methylation are known to play an important role in the regulation of miRNA precursor expression.[474]

Therefore, miRNAs responding to fetal nutrition or epigenetically controlled miRNA expression represent novel mechanisms by which predictive adaptive responses can made by the foetus to predict the post-natal nutritional environment and program the lung accordingly during windows of developmental plasticity that continue into early life. The Southampton protein-restriction rat model provides an opportunity to determine whether 'nutritional programming' by maternal protein-restriction during pregnancy alters offspring lung miRNA expression. Furthermore, miRNA target gene predictions, by bio-informatics approaches, should highlight metabolic pathways under the influence of miRNAs.

6.1.1 miRNA Biogenesis and Function

Micro RNA molecules (miRNA) are short molecules (15-25 bp) that are not translated into protein but silence/reduce gene expression by binding to the 3' end of the mRNA gene transcript as part of a RNA-induced Silencing Complex (RISC). (Figure 6.1) This triggers degradation of the mRNA or translationally represses expression of the gene. The degree of miRNA complementary to the mRNA sequence determining whether the target mRNA is cleaved or not. Mature miRNA show near complementarity to the 3' untranslated (3'UTR) region of target mRNA. The complementarity is in the form of a 'seed' sequence of about 7 nucleotides at

the 5’ end of the miRNA molecule to bind the 3’UTR mRNA sequence. The seed sequence is essential to confer regulation of the mRNA target.

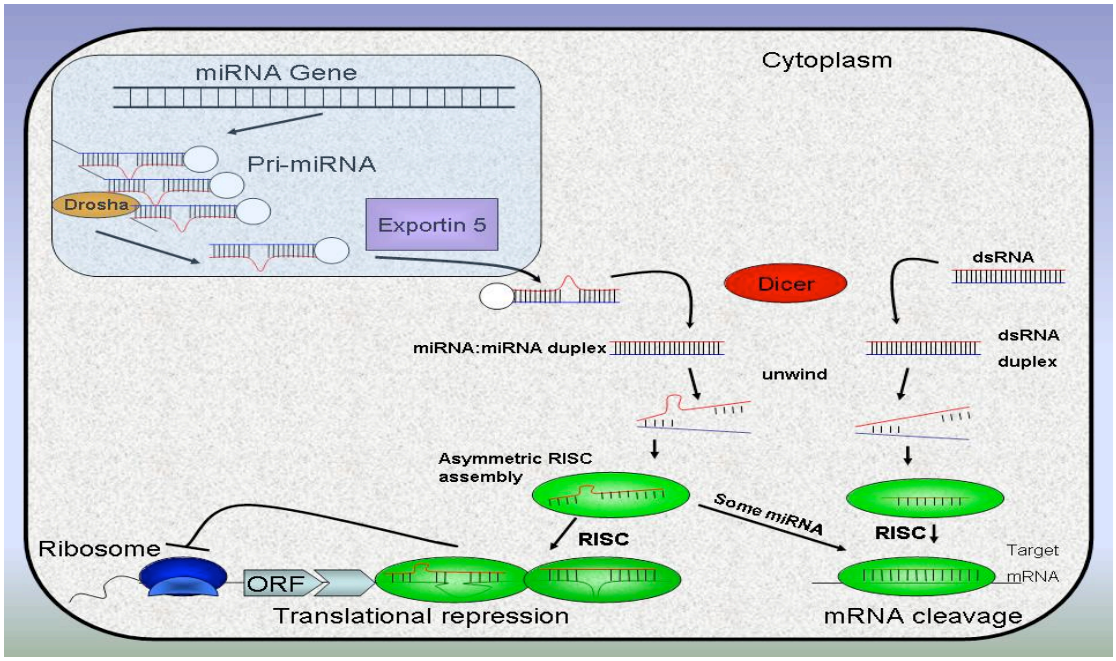


Figure 6-1: microRNA Biogenesis.

The pri-miRNAs are processed into ~70bp pre-miRNAs by Drosha in the nucleus. Exportin 5 transports the pre-miRNA into the cytoplasm, where Dicer processes them into miRNA:miRNA duplexes. Dicer can also process long dsRNA molecules into short interfering RNA (siRNA). One strand of the miRNA:miRNA duplex is assembled into the RNA Induced Silencing Complex(RISC) and the other miRNA strand is targeted for degradation. The miRNA acts on its target by either translational repression or mRNA cleavage; this is dependant on the miRNA complementarily to the target sequence.

The process of miRNA maturation from, (1) polymerase II transcribed Pri-miRNA (2) Cleavage into precursor miRNA (pre-miRNA) by the endonuclease Drosha in the nucleus (3) The importing of pre-miRNA into the cytoplasm by exportin 5 (4) Cleavage of cytoplasmic pre-miRNA by Dicer into miRNA:miRNA duplexes, is tightly regulated at each step.

A recent study of miRNA turnover in *Caenorhabditis elegans* reported degradation of fully processed mature miRNAs by the 5’-3’ exoribonuclease XRN-2; these results suggest the presence of additional layer of regulation of miRNA activity.[475] The authors suggested that regulation of mature miRNA turnover may be important for rapid changes in miRNA expression profiles during developmental transitions and in the homeostasis of miRNA levels. Little is known about how miRNAs are transcribed and what promoter elements control expression. The limited analysis of miRNA promoters has shown they are likely to be under control

of transcription factors, enhancers, silencing elements and chromatin modifications such as DNA methylation, much like any protein-expressing gene (reviewed by Breving et al, 2009).[476] To date, it has been accepted that miRNA species repress gene expression. However, emerging data has shown in some circumstances miRNAs can activate genes by stimulating protein translational activation[477] and the promote expression of genes with complementary promoter sequences.[478] Therefore, as our understanding of this group of short RNA molecules grows, it is likely that more interesting functions of miRNAs will emerge.

6.1.2 Genomics of miRNA

In the human or rat genome about 50% of the miRNA genes are found in clusters on all chromosomes except for the Y chromosome (see Table 6.1). miRNAs in the same cluster are transcribed as a polycistronic transcript. miRNAs in a cluster are often related to each other, possibly as a result of gene duplication of the cluster. Clustered miRNA genes are functionally related and could act on the same gene or gene pathway and many miRNAs target genes are involved in development. The majority of mammalian miRNA genes are located in transcribed units (TUs), consisting of three types; intronic miRNA in protein coding TUs (61%), intronic miRNA in non-coding TUs (18%), and exonic miRNA in non-coding TUs (20%).[479]

Table 6-1: Rattus Norvegicus miRNA Gene Clusters.

Chromosome (strand)	miRNA gene clusters[a] (rno-mir-xxx cluster 1 **|** rno-mir-xxx cluster 2…)
Chr X (+) :	rno-mir-222 rno-mir-221
Chr X (-) :	rno-mir-421 rno-mir-374 **|** rno-mir-363 rno-mir-92-2 rno-mir-19b-2 rno-mir-20b **|** rno-mir-503 rno-mir-322
Chr 1 (+) :	rno-mir-99b rno-let-7e rno-mir-125a **|** rno-mir-194-2 rno-mir-192
Chr 1 (-) :	rno-mir-292 rno-mir-291 rno-mir-290
Chr 2 (+) :	rno-mir-15b rno-mir-16
Chr 3 (-) :	rno-mir-296 rno-mir-298
Chr 4 (-) :	rno-mir-96 rno-mir-183 **|** rno-mir-29a rno-mir-29b-1 **|** rno-mir-141 rno-mir-200c
Chr 6 (+) :	rno-mir-337 rno-mir-540 **|** rno-mir-376b rno-mir-376a rno-mir-300 **|** rno-mir-381 rno-mir-487b **|** rno-mir-382 rno-mir-134 **|** rno-mir-409 rno-mir-412 rno-mir-369
Chr 7 (+) :	rno-let-7c-2 rno-let-7b
Chr 8 (-) :	rno-mir-34c rno-mir-34b
Chr 10 (+) :	rno-mir-497 rno-mir-195 **|** rno-mir-212 rno-mir-132 **|** rno-mir-144 rno-mir-451
Chr 12 (-) :	rno-mir-25 rno-mir-93 rno-mir-106b
Chr 13 (+) :	rno-mir-181a-1 rno-mir-181b-1 **|** rno-mir-194-1 rno-mir-215 **|** rno-mir-29b-2 rno-mir-29c
Chr 15 (+) :	rno-mir-17 rno-mir-18 rno-mir-19a rno-mir-20a rno-mir-19b-1 rno-mir-92-1
Chr 17 (+) :	rno-let-7a-1 rno-let-7f-1
Chr 17 (-) :	rno-mir-24-1 rno-mir-27b rno-mir-23b
Chr 19 (-) :	rno-mir-24-2 rno-mir-27a rno-mir-23a

[a]Inter-miRNA clustering distance = 500bp. red = genic cluster. Data from miRGen database (version 3, 1st Jan 2007)

6.1.3 Methylation and Epigenetically Controlled miRNA Expression

It is thought that 5-10% of mammalian miRNA genes are epigenetically regulated (reviewed by Breving et al, 2009)[476] and with 47% of miRNA genes reported as being associated with CpG islands.[480] The predicted targets of some miRNAs have included transcripts from genes involved in chromatin regulation (miR-520 cluster in human embryonic stem cells)[481] and DNA methylation (miR-29 and DNMT3a &b).[482] Therefore, miRNAs are in a unique position of being subjected to epigenetic control but can also regulate the very same processes, forming a regulatory feedback loop.[483] Miss-expression of miRNAs due to promoter dysregulation by changes in DNA methylation levels is seen in human disease and will be discussed in the next section.

6.1.4 miRNAs in Disease

The over and under-expression of miRNAs has been observed in various diseases, included numerous cancers,[484-486] and many human miRNA genes are located in cancer-associated regions and in gene deletion/amplification and mutation hotspots (reviewed by Rossi et al, 2008).[487] It is thought that miRNA dysregulation is important in cancer because miRNAs are key regulators of cell growth, apoptosis, proliferation and differentiation.

In cancer cell culture and animal model studies, many miRNAs have been shown to have tumour-promoting or suppressive activities, for example miR-21 is over expressed in many cancers and inhibition of miR-21 in HeLa cervical cancer cells reduces tumour cell profileration.[488] The enforced expression of miR-29 in lung cancer cell lines restored normal patterns of DNA methylation (by targeting DNA methyltransferases 3A and 3B), this lead to re-expression of methylation-silenced tumour suppressor genes (*FHIT* & *WWOX*) and inhibited tumorigenicity *in vitro* and *in vivo*.[482] In breast cancer, abberrant DNA hypermethylation occurs early on in tumour development and hypermethylation of CpG islands near miRNA genes is thought to be another mechanism of gene inactivation.[489] In fact, miRNA expression profiles have been demonstrated to be more accurate in predicting poorly differentiated cancer than mRNA expression profiles in the same group of patients.[490] miRNAs have a therapeutic potential in cancer treatment for example, miRNAs with tumour promoting activities could be targeted by miRNA silencing, antisense blocking and miRNA modification (reviewed by Li et al, 2009).[491]

There are now emerging roles for miRNA in other diseases such as, cardiac and vascular malignancies, diabetes, obesity, metabolic syndrome and infection. miRNAs are important regulators of normal vascular development and function, studies in mice with Dicer deletions have demonstrated impairments in blood vessel formation of the developing embryo and altered expression of the angiogenic regulators.[492] miR-126 has been shown to inhibit the expression of vascular cell adhesion molecule1 (*VCAM1*) an important component of leukocyte adhesion to endothelial cells in the early development of atherosclerotic lesions.[493]

Several miRNA species have been indentified as having roles in insulin secretion, islet development, B cell differentiation and can indirectly control glucose and lipid metabolism.(reviewed by Tang et al, 2008)[494] Expression profiling in diabetic animal models have demonstrated altered miRNA levels which have predicted targets in glucose, lipid metabolism and over-expression of miR-29s lead to insulin resistance in adipocytes.[495, 496] Furthermore, the miRNAs expressed in islets, liver and adipose tissue share significant changes in response to genetic obesity in diabetes-susceptible mice.[497]

It is likely that miRNAs can also have direct antiviral effects as one study has demonstrated that IFN-β can modulate the expression of cellular miRNAs that have predicted targets in the hepatitis C virus genome.[498] However, the direct interaction between host miRNA machinery and virus genomes is still not fully understood.

6.1.5 Role of miRNAs in Development

miRNAs were originally indentified in *Caenorhabditis elegans* as genes that were required for the timed regulation of developmental events.[499, 500] Now, hundreds of microRNAs have been characterised and are thought to regulate the expression of 20-30% of human genes, many of which are involved in development and homeostasis.[501] The role of miRNAs in human development are still not fully understood, however forward and reverse genetic experiments in worms, flies, fish and mice have provided insights into the developmental stages where miRNAs are essential for viable life. The most studied miRNAs are *lin-4* and *let-7*, these miRNAs are involved in controlling early and late developmental timing in worms (reviewed by Wienholds et al, 2005).[502] Again, in worms, *lsy-6* and *mir-273* were found to control left/right neuronal asymmetry. In flies, programmed cell death was found to be regulated by *miR-14*.

In zebrafish, *miR-7* was found to regulate Notch signalling which is essential for proper patterning and development of all multicellular organisms and *miR-430* was found to be important in brain morphogenesis. (reviewed by Wienholds et al, 2005)[502]

In 2006, Strauss and colleagues used mouse embryonic stem cells as a model system for the early mammalian embryo and reported the miRNA expression signature against mature heart, lung, liver, kidney and brain tissue miRNA signatures.[503] They observed an overall trend of increasing miRNA complexity from immature embryonic stem cells to mature tissues. The most complex signature came from more functionally complex tissues (brain to liver being least complex). They hypothesed that the interaction between the miRNAs and messenger RNA determines a finely tuned set of mature mRNA for each tissue at the end of development.

6.1.5.1 Lung Development

The role of miRNAs during fetal lung development is not fully understood. In 2007 Williams et al. published miRNA expression profiles of fetal and adult lung in humans and mouse.[504] They found miRNA profiles between humans and mouse to be very similar, both species possessing a maternally imprinted miRNA cluster (miR-154 and miR-335 families) and up-regulation of a subset of miRNAs expressed in adult (miR-26b, 29a, 29b, 142-3p and 187) compared to fetal lung (miR-134, 154, 214, 296, 299, 323, 337 and 370). There were also differences in the spatial expression of miRNAs, with miR-26a and miR-29a localising to airway and alveolar epithelia and miR-154 present in fetal stoma but not in adult lung. In 2008, Ventura et al. demonstrated that mice deficient for the miR-17~92 cluster die after birth with lung hypoplasia and defects in heart and B cells.[505] The authors noted that a previously published work on the over expression of the cluster had demonstrated increased cell proliferation and inhibition of differentiation of the lung epithelium (Lu et al, 2007. Cited in Ventura et al)[505] and concluded that this miRNA cluster played an important role in the control of the proliferation or survival of lung cells.

In rat, microarray analysis of miRNA expression revealed 21 miRNA species that changed during development and that the miRNA could be placed into 4 groups based on the timing of their expression.[506] (see table 6.2)

The authors noted that cluster 1 contained miRNAs that controlled developmental timing (*let-7*), proliferation and differentiation of cells (miR-195). Cluster 2 and 3 expressed miRNA species from the miR-17~92 cluster that, as stated earlier, could promote proliferation and differentiation of lung epithelial cells.

miRNAs in cluster 4 (miR-127 and miR-351), were shown by *In situ* hybridisation to shift from the mesenchymal compartment of the developing lung to epithelial cells. This lead the authors to conclude that these miRNAs may have role in cellular reorganisation. Finally, the over expression of miR-127 at an earlier stage, embryonic day 14, in fetal lung cultures caused defects in lung branching morphogenesis, resulting in larger and fewer terminal buds than untreated cultures.[506] This finding indicates that the timed expression of miRNAs is important for normal lung development to occur.

Table 6-2: Rat Lung Developmental Stages and miRNA Expression.

	Days	Lung development stage	Milestones in development	miRNA Cluster[a]			
				1	2	3	4
Organogenesis	E1 ↓ E13	Embryonic	lobar division	?	?	?	?
	E13 ↓ E18	Pseudoglandular	epithelial tubes of air passage form (no lumen)	L	H	H	L
Differentiation	E18 ↓ E20	Canalicular	Bronchioles are producted and lumen is found in many tubules	L	H	L	L
	E20 ↓ full term	Saccular	Thinning of interstition, flattening of epithelium, alveolar ducts and air sacs are formed. Start of Surfactant production	L	L	L	H
Post-natal	birth (P0) ↓ adult	Alveolar	true alveolar formation	H	L	L	H L

[a] miRNA Cluster 1:let-7b, miR-29a, 23a, 22, 195. Cluster 2: miR-298, 341, 130b, 92. Cluster 3: miR-17, 214, 106b, 93, 290, 20a, 17-5p, 18. Cluster 4: miR-127, 210, 19b, 351. Data adapted from Bhaskaran et al 2009.[506] L = low expression, H= high expression

6.1.6 Hypotheses

It is proposed that:

'Maternal protein-restriction during pregnancy leads to altered lung miRNA gene expression in offspring.'

6.2 Aims

We hypothesised that restricting the amount of protein available to the developing fetus will result in alterations to lung miRNA expression. We aim to test this hypothesis by determining miRNA gene expression levels, by microarray, in lungs from 120 day old protein-restricted and control F_1 generation rat offspring. Furthermore, we will validate any differentially expressed miRNA genes identified by microarray with real-time PCR. We will then perform miRNA gene cluster and miRNA gene target bioinformatics to ascertain whether alterations to miRNA gene expression is independent of shared miRNA gene promoter regions and identify potential gene targets to understand which biological pathways may be under the influence of altered miRNA expression.

6.3 Methods

6.3.1 miRNA Isolation and Purification

Total RNA from lung tissue was extracted using a modified trizol RNA isolation protocol (section 2.2.3). Extractions were carried out utilising ¯20°C isopropanol with 50 mg of glycogen and stored overnight at ¯80°C to precipitate small RNA molecules. RNA quantification was performed according to section 2.2.2.2. For the microarray experiments lung tissue biopsies were sent to Miltenyi Biotec for isolation by standardised in-house protocols.

6.3.2 miRNA Microarray

PIQOR™ miRXplore Microarrays were used to assess global miRNA expression in the lung tissue. Mac Molecular Genomic Services (Miltenyi Biotec GmbH) provided the service and performed the entire tissue processing, total RNA isolation, RNA labeling, microarray hybridization steps and performed quality control analyses for the microarray data. 379 rat miRNA species were targeted on the PIQOR™ miRXplore Microarray (miRNA sequences according to miRBase version 10.0). All oligonucleotides were spotted in quadruplicate.

Controls included on the miRXplore Microarray were:

- 13 Negative mismatch controls
- 36 Positive controls
- 5 Hybridization controls
- 18 Calibration controls

6.3.2.1 Total RNA Samples on the Microarray

7 Protein-restricted and 7 control 120 day F1 lung samples (from section 5.3.1) were analysed by microarray. The total RNA isolates were pooled into two groups (Pool 1 = Protein-restricted; Pool 2 = Control) and labelled with Hy5 dye; then hybridized on to two microarrays, respectively.

6.3.2.2 Microarray Data Normalisation

All Hy5 signals were normalized to a Hy3 labelled miRExplore™ universal reference of synthetic microRNA molecules (supplied by Miltenyi Biotec GmbH) before calculation of the ratio values (RNA Pool 1/ RNA Pool 2) for the miRNA species of interest.

6.3.2.3 Microarray Data Quality Control

As an additional quality filtering step, only hybridisation spots with signals or higher than the 75% percentile of the background signal intensities were used in the calculation of the Hy5/Hy3 ratio. This discriminated questionable results from relevant miRNA results because they were flagged as having very low signal intensities and removed from the final data set. The cut-off value for putative up or down-regulated miRNA expression was determined as a ≤ 1.7 fold change in signal intensity.

6.3.3 Real-Time PCR validation

Real-time PCR miRNA quantification was performed to validate the microarray results. Initially, the miScript system (Qiagen) was chosen to validate the top two over and under-expressed miRNA species. This system utilises a single cDNA synthesis step to generate template for all miRNA species in a single sample. The second real-time PCR miRNA system (miRcury, Exiqon) was used to confirm initial real-time PCR findings from the miScript system and subsequently replaced it as the real-time PCR miRNA system of choice. All miRNA gene quantification levels generated by real-time PCR were normalised to an endogenous control siRNA gene, rRNA 5S.

6.3.3.1 Qiagen miScript System

6.3.3.1.1 Overview

This system combines SYBRgreen dye detection with validated miRNA primer assays to provide rapid determination of hundreds of miRNA molecules from a single cDNA synthesis (Figure 6.2).

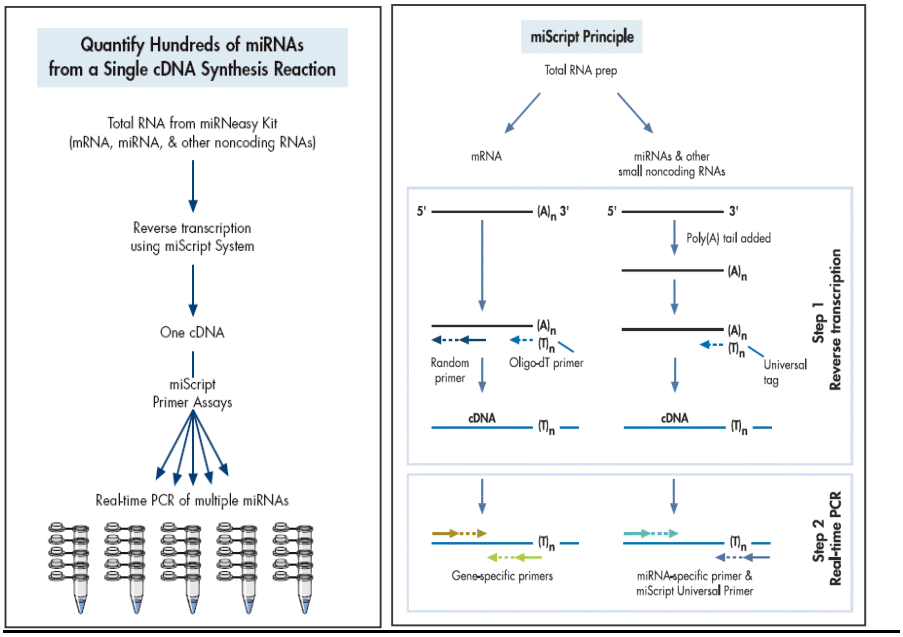


Figure 6-2: Principles of miScript miRNA Real-Time PCR System.

In step 1, the reverse-transcription reaction, miRNAs have a poly(A) tail added to the 3' end of the molecule, prior to hybridisation to a oligo dT primer with a universal tagging sequence to generate cDNA template. In step 2, real-time PCR reaction, the miRNA cDNA template is PCR amplified by a miRNA specific primer and a universal PCR primer. Accumulation of PCR product is detected by hybridisation of the fluorescent SYBRgreen dye to the double-stranded PCR product.

6.3.3.1.2 cDNA Synthesis

cDNA synthesis was performed for each lung sample according to the manufacturers protocol.

6.3.3.1.3 Real-time PCR Setup

PCR reactions were setup in master mixes for miRNAs (RNO-MIR-126, 140, 186 and 674) and the endogenous control gene rRNA 5S as follows:

Table 6-3: miScript MicroRNA PCR Reaction Master Mix.

miRNA/rRNA 5S Control Master mix	x1
miScript SYBRgreen Master mix	10.0 µl
Universal Primer	2.0 µl
microRNA Primer assay	2.0 µl
Water	5.0 µl
cDNA (no dilution)	1.0 µl
Total Volume	20 µl

PCR was performed in triplicate wells and each 384 well reaction plate contained negative control reactions (NTC and RT-ve) and standard curves (section 2.2.5.3).

6.3.3.2 Exiqon miRNA miRcury System

6.3.3.2.1 Overview

This system combines SYBRgreen dye detection with validated mature miRNA primer assays to provide determination of a single miRNA species from a gene-specific cDNA synthesis (Figure 6.3).

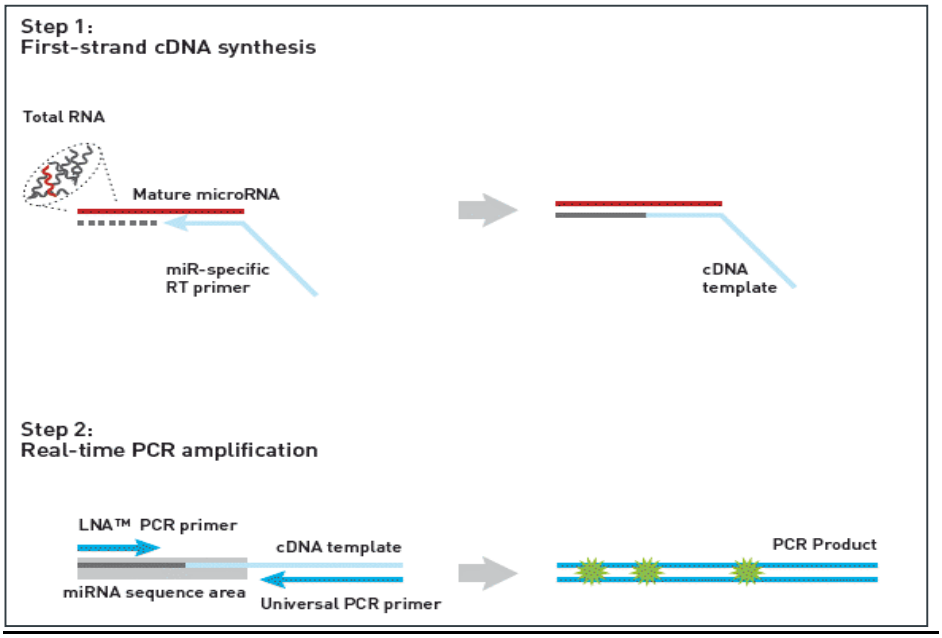


Figure 6-3: Principles of miRcury miRNA Real-Time PCR System.

Mature miRNAs are hybridised to miRNA specific RT primers, which have a universal priming sequence overhang at the 5' end, during cDNA synthesis. In step 2, the cDNA template contains the mature miRNA sequence attached to the universal priming sequence which is targeted for PCR amplification by the Locked Nucleic Acid (LNA) PCR primer and a universal PCR primer. Accumulation of PCR product is detected by hybridisation of the fluorescent SYBRgreen dye to the double-stranded PCR product.

6.3.3.2.2 cDNA synthesis

cDNA synthesis was performed for each of the miRNA species (rno-miR-22*, 126, 153, 181a*, 186, 365 and endogenous control gene rRNA 5S) in duplicate for each of the lung samples according to the manufacturers protocol.

6.3.3.2.3 Real-Time PCR Setup

PCR reactions were setup in master mixes as follows:

Table 6-4: miRcury MicroRNA Real-Time PCR Reaction Master Mixes.

miRNA Master mix	x1
SYBRgreen Master mix	10.0 µl
Universal Primer for PCR	1.0 µl
microRNA LNA PCR primer	1.0 µl
Water	4.0 µl
cDNA (1:10 dilution)	4.0 µl
Total Volume	20 µl

rRNA 5S Control Master mix	x1
SYBRgreen Master mix	10.0 µl
Control PCR primers	2.0 µl
Water	4.0 µl
cDNA (1:10 dilution)	4.0 µl
Total Volume	20 µl

Real-time PCR was performed in duplicate wells and each 384 well reaction plate contained negative control reactions (NTC and RT-ve) and standard curves (section 2.2.5.3).

6.3.4 miRNA Genomics and Gene Cluster Analysis

All miRNAs identified by microarray were searched against the miRbase genome database (www.diana.pcbi.upenn.edu/miRGen/v3/miRGen.html).[507] Changes in putative up or down-regulation (from the microarray data) were correlated to miRNA gene cluster locations to determine if altered miRNA expression could be controlled by cis-elements acting at shared promoter regions or whether there is post-transcriptional regulation of miRNA maturation. Also, similar expression levels with in a miRNA cluster could also act as a way to internally validate the microarray data. In this study, a miRNA gene cluster was defined as two or more miRNA genes less than 500bp apart in non-genic regions, and therefore likely to be a polycistronic transcript, by using Rattus norvegicus chromosomal coordinates (Genome assembly: RGSC3.4) of microRNA genes from miRBase sequence version 13.0 (www.mirbase.org).[508] This search also identified miRNAs located in protein coding gene.

6.3.5 miRNA Target Gene Prediction Analysis

Computational methods are used to predict miRNA targets by scoring base pair matches (Watson-Crick pairing) to gene locations in the genome and there are many websites hosting miRNA target prediction database tools (reviewed by Li et al, 2009).[491] Target predictions have high false positive rates and predicted targets can vary between prediction databases. One method to strengthen a prediction is to compare the results from several databases.

In this study, miRNA target gene predictions were made using the miRNApath database.[509] miRNApath (http://lgmb.fmrp.usp.br/mirnapath/query_mirna.php) is an integration of the miRNA sequence database (miRbase, www.mirbase.org), miRNA target gene database (miRgen, www.diana.pcbi.upenn.edu/miRGen/v3/miRGen.html) and a gene pathway database (KEGG database, www.genome.jp/kegg/pathway.html). miRNApath can interrogate multiple miRNAs simultaneously and produce a list of shared target gene predictions and metabolic pathways.

6.4 Results

6.4.1 miRNA Isolation and Purification

All lung samples yielded high quality total RNA enriched for miRNAs and the total miRNA isolates used on the microarray passed all of the Miltenyi BioTec quality control checks (performed using Agilent 2100 Bioanalyzer RNA microfluidic chips).

6.4.2 miRNA Microarray Results

The lung RNA pools contained 163 miRNA species that returned valid results (See figure 6.4). After normalising the signal intensities for each spot grouping to the synthetic universal reference RNA (ratio values), then dividing the ratio values of RNA pool 1 (protein restricted) by pool 2 (control diet) a total of 43 miRNA species were identified as being different between the two groups. The scatter plot of ratio values revealed 15 miRNA species that were up (1.7 > 16 fold) and 13 miRNA species down (1.7 > 14 fold) in the protein-restricted lung compared to the control lungs. A further 9 miRNA species were found to be only present in the PR lung and 6 miRNAs were only found in the control lung sample pool.

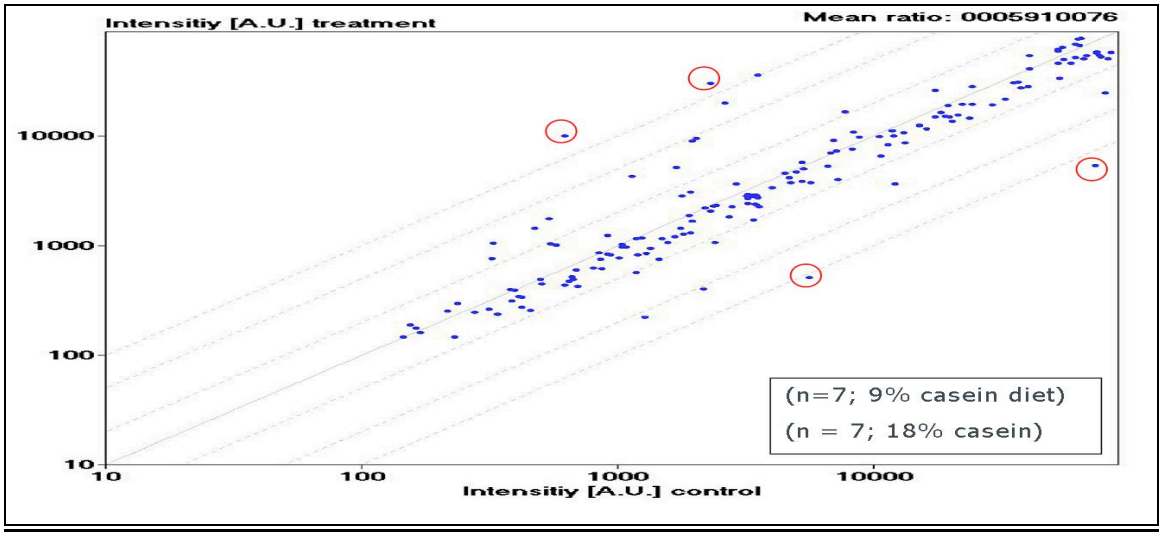


Figure 6-4: Double-Log Scatter Plot of Re-Ratio Values from the Microarray Experiment.

The signal intensities of each spot/miRNA that passed the quality filtering are shown in a double-logarithmic scale, represented by a dot. X-axis: Hy3 signal intensity, y-axis: Hy5 signal intensity. Dashed diagonal lines define the areas of x-fold differential signal intensities. Red circles identify the first four miRNA genes chosen for real-time PCR validation.

Table 6.5 displays all of the miRNA species that are present in lung tissue from both groups but at different levels, and table 6.6 shows miRNA species that were only detected in one group.

Table 6-5: Putative Up or Down Regulated miRNA Species in Protein-Restricted Offspring Lung Tissue.

miRNA name	Increased in PR lung[a]		miRNA name	Decreased in PR lung[a]
RNO-MIR-186	16.14		RNO-MIR-122	0.56
RNO-MIR-126*	12.96		RNO-MIR-101B	0.56
RNO-MIR-218	10.22		RNO-MIR-362-3P	0.53
RNO-MIR-142-3P	7.61		RNO-MIR-181B	0.50
RNO-MIR-22*	4.69		RNO-MIR-151*	0.48
RNO-MIR-347	4.63		RNO-MIR-301A	0.45
RNO-MIR-33;	3.80		RNO-MIR-99A	0.31
RNO-MIR-484	3.26		RNO-MIR-98	0.30
RNO-MIR-130B	3.26		RNO-MIR-23A	0.28
RNO-MIR-124	3.04		RNO-MIR-17-3P	0.19
RNO-MIR-450A-2	3.03		RNO-MIR-181A*	0.18
RNO-MIR-494	2.35		RNO-MIR-674-3P	0.09
RNO-MIR-100	2.15		RNO-MIR-140*	0.07
RNO-MIR-92B	1.90			
RNO-MIR-24-1*	1.78			

[a] *microarray reratio values = fold change in PR lung miRNA levels when compared to control lung data.*

Table 6-6: miRNA Species Found Only in Protein-Restricted or Control Diet Offspring Lung Tissue.

Grouping	miRNA name (rno-miR-xxx)
PR only	LET-7D*, 10B, 20A, 142-5P, 331-5P, 340*, 365, 450A-1, 664
Control only	9*, 129*, 153, 190, 378*, 384-3P

6.4.3 Real-Time PCR Validation

Two real-time PCR systems were used to generate miRNA relative quantification results against the endogenous control gene, rRNA 5S. The real-time PCR relative quantifications results were then compared to the microarray data, corroboration of results would validate the observation.

6.4.3.1 Qiagen miScript miRNA Real-Time PCR System Validation

The top two over and under-expressed miRNA species identified by microarray were assayed using this system. Trizol™ derived total miRNA isolates enriched for miRNAs (isolation carried out in our laboratory) failed to replicate the microarray data (see figure 6.5). All four miRNA species were detected at almost equal levels in the protein-restricted and control diet offspring lung. Assuming that the microarray data is a true reflection of miRNA levels in the tissue of the two groups, this result could indicate a problem with the Trizol™ extraction procedure performed in our laboratory.

Figure 6-5: miScript miRNA Real-Time PCR Validation Using Trizol™ Total RNA Isolates Enriched for miRNAs.

Real-time PCR miRNA levels are similar between the PR (green columns) and control lungs (blue columns).

Therefore, the experiment was repeated on cDNA template generated from the miRNA isolations returned from Miltenyi BioTec (preparations used on the microarray). Figure 6.6 displays real-time PCR data from the microarray RNA preparations, these results are very similar to the data generated by the RNA preparations carried out in our laboratory and do not reflect the original microarray observations.

Figure 6-6: miScript miRNA Real-Time PCR Validation Using Miltenyi Biotec RNA Isolates.

Real-time PCR miRNA levels are similar between the PR (green columns) and control lungs (blue columns).

The lack of replication of microarray-based miRNA levels between the two groups, even when utilising the original microarray RNA preparations, lead to the trial of another real-time PCR system. Lack of replication from a different real-time PCR system would throw the microarray data into question.

6.4.3.2 Exiqon miRNA miRcury Real-Time PCR System Validation

The miRcury real-time PCR system utilises a miRNA species specific cDNA synthesis step, this is unlike the miScript system which generates a cDNA pool of all miRNA species. Unfortunately, at the time of this work Exiqon Inc. did not have a primer assay library containing of all of the miRNA species from the first experiment. Therefore, in this experiment only rno-miR-126 and 186 (both up in PR lung) were quantified by this system. Four extra miRNA species were also quantified by this system, they included rno-miR -22* (up in PR lung; 4.7 fold), 153 (Control lung only), 181a* (down in PR lung; 5.6 fold) and 365 (PR lung only).

Figure 6.7 shows the miRcury real-time quantification results alongside the microarray data; rno-miR -22*, 153 and 181a* did not demonstrate similar levels and in two cases (rno-miR -153 and 181a*) showed a reciprocal fold difference in expression in comparison to the microarray data.

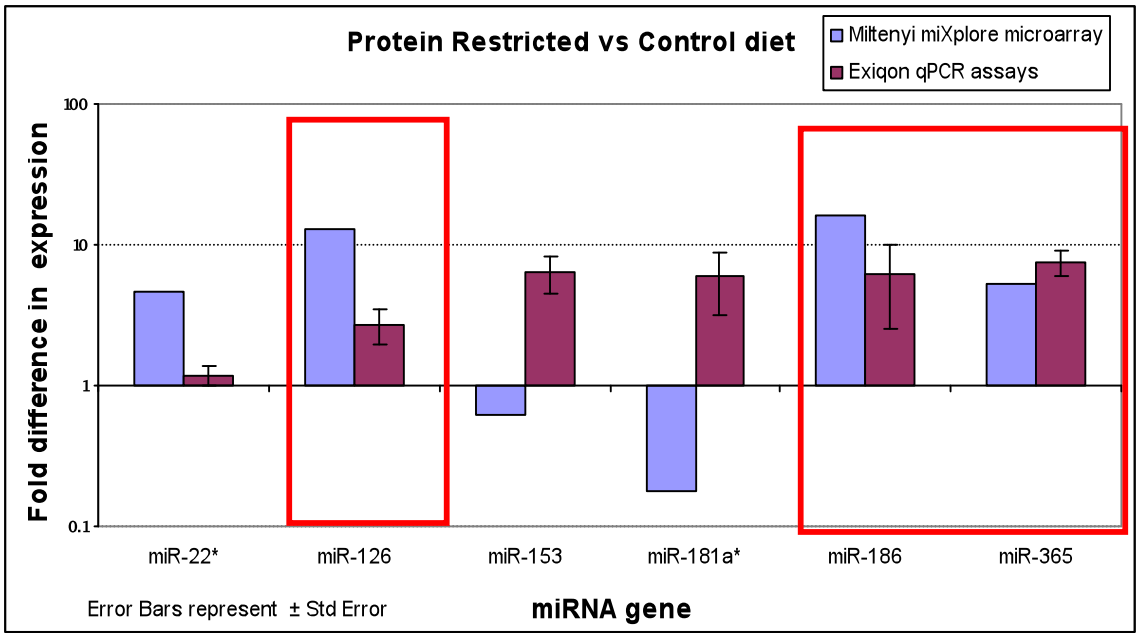


Figure 6-7: miRcury miRNA Real-Time PCR Validation Using Miltenyi Biotec RNA Isolates.

Using the total RNA isolates from the microarray experiment, three miRNA species (red boxes) demonstrated real-time PCR miRNA levels (purple columns) similar to the microarray results (blue columns). Error Bars represent ± Std Error.

However, the remaining three miRNA species results were more encouraging (see table 6.7); with rno-miR-365 being closely replicated by real-time PCR quantification, followed by rno-miR-186 and 126 fold differences that are in the same direction, if not the same magnitude as the microarray results.

Table 6-7: miRNA Species Validated by Real-Time PCR

miRNA species	microarray vs real-time PCR fold difference
rno-miR-365	5.2 vs. 7.5
rno-miR-186	16.1 vs. 6.2
rno-miR-126	13.0 vs. 2.7

6.4.4 miRNA Genomics and Gene cluster Results

Six miRNA genes are located in the introns of protein coding genes and therefore their expression may be controlled by the host gene promoter. (See table 6.8)

Table 6-8: miRNAs in Protein Coding Genes

miRNA name	level in lung	Gene	Role of gene
rno-miR-33	up in PR lung	Srebf2, intron16	*Sterol regulatory element binding transcription factor 2*. This gene encodes a ubiquitously expressed transcription factor that controls cholesterol homeostasis by stimulating transcription of sterol-regulated genes. Cellular cholesterol homeostasis is controlled by Srebf2 and SCAP.[510]
rno-mir-101b	down in PR lung	Rcl1, intron 8	*RNA terminal phosphate cyclase-like 1*. function in pre-rRNA processing and with U3 -snoRNA is thought to control an early step in ribosome biogenesis.
rno-mir-126[a]	up in PR lung	Egfl7, intron 7	*Epidermal growth factor like domain 7*. Important in vascular tubulogenesis. Methylation of host gene promoter controls expression of miR-126 (tumour suppressor) in cancer.[511]
rno-mir-186[a]	up in PR lung	Zfp265; intron 8	*zinc finger protein 265.* is a novel component of spliceosomes
rno-nir-153	present in control lung only	Ptpra2, intron 19	*protein tyrosine phosphatase, receptor type, N polypeptide 2.* may play a role in secretory granule biogenesis and organization or exocytosis
rno-mir-301	down in PR lung	RGD1307084, intron1	Protein coding, function unknown

[a] *validated by real-time PCR*

Another four of the miRNA species identified by microarray are part of a miRNA cluster (<500bp between genes). miR-181a1 and -181b1 form a cluster of miRNAs on chromosome 13 (+), within 270 bp of each other, and both species had lower levels in the PR lungs (5.6 and 2.0 fold) compared to controls.(see figure 6.8)

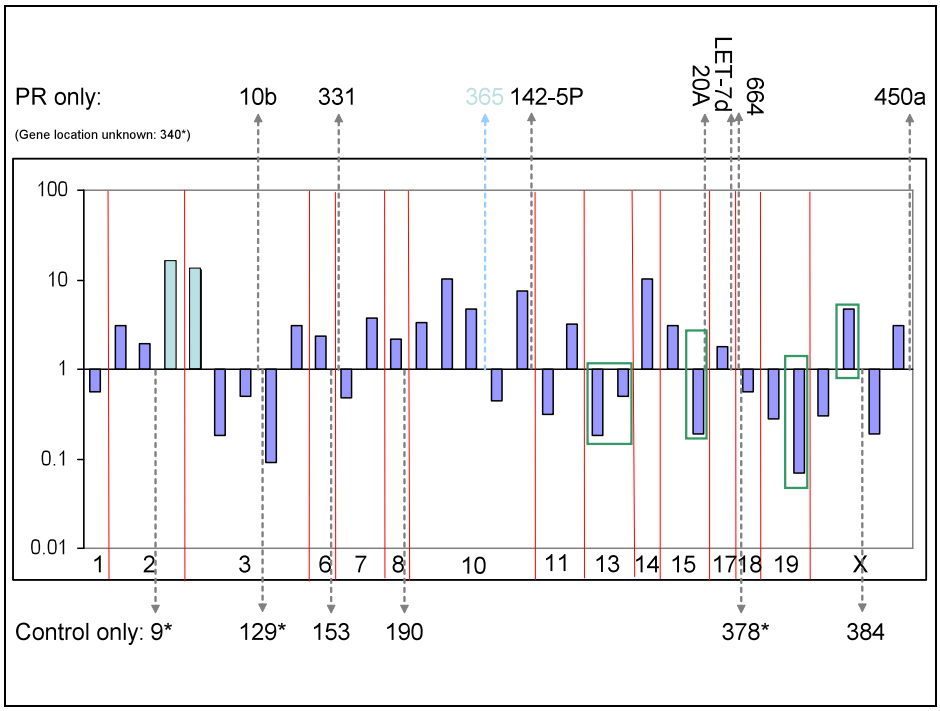


Figure 6-8: miRNA Gene Expression by Chromosome.

This chart shows that the levels (y-axis) of miRNA genes altered in the PR animals are found across the rat genome (Chromsome number on x-axis). Green boxes identify miRNA gene clusters with differential expression. Light blue arrows and columns are the miRNA species validated by real-time PCR.

In the miR-17~92 cluster on Chr 15(+) only miR-17-1 (down in PR lung; 5.3 fold) and 20a (in PR lungs only) were found to be at different levels. This is an interesting observation as the miR-17~92 gene cluster is known to be expressed as a polycistronic transcript. These results indicate that in the PR lungs miR-17-1 may be negatively regulated post pri-miRNA transcription, as it has to be transcribed as part of the polycistronic transcript along with the other members of the gene cluster. This observation is supported by the fact that the other miRNAs (mir-18, 19a, 19b, 92) in the miR-17~92 cluster have similar levels in both lung groups. Furthermore, as miR-20a is only expressed in PR lung this miRNA must be positively regulated or selected after transcription of the miR-17~92 pri-miRNA. In a different miRNA gene cluster on chromosome 19(-), miR-23a was down in PR compared to control lung (3.6 fold), this gene is part of miRNA gene cluster: mir-24-2, 27a and 23a. All of the miRNA genes in this cluster are found in a region of 388 bp and therefore, are likely to be expressed as a polycistronic transcript. Post pri-miRNA transcriptional regulation could also explain the difference in miRNA levels in this gene cluster.

6.4.5 miRNA Gene Target Results

Three miRNApath analyses were performed using: 1) a list of all of the miRNAs species identified by microarray that were increased or present in PR lung. 2) a list of all or decreased/absent (i.e. up/present in control lung) in the PR lung. 3) The miRNA species successfully validated by real-time PCR (miR-126, 186 and 365 all up in PR lung).

The miRNApath analysis results from the increased/ present miRNA list had predicted target genes that are involved in 55 metabolic and 45 pathways were predicted to contain genes targeted in the decreased/absent miRNA list. The top two pathways (i.e. which pathway has the most part of the genes targeted by the miRNAs) for each miRNA list will be presented here.

In the increased/present list in PR lung, 5 miRNA species (*let-7d*, mir-33, 92, 186, 365) targeted 5 genes (*Atp2b2, Camk2d, Itpr2, P2rx3, Gna14*) in the Calcium signalling pathway and four miRNAs (miR-33, 100, 331, 664) targeted 4 genes (*Casp3, Cacnb4, Mapk8ip3, Elk1*) in the MAPK signalling pathway. In the decreased/absent list, 4 miRNAs (miR-98, 99A, 101B, 151) targeted 3 genes (*Cacnb4, Mapk8ip3, Elk1)* in the MAPK signalling pathway and 4 miRNAs (miR-23A, 98, 181A, 153) targeted 2 genes (*Cox7a2 & Ndufs2*) in the oxidative phosphorylation pathway.

6.4.5.1 miRNApath Analyis for the Validated miRNAs

Tables 6.9 shows the predicted gene targets and metabolic pathways in the PR lung, which could be under the influence from increased expression of the three validated miRNA species, miR-126, 186 and 365.

Table 6-9: Predicted Target Genes for rno-mir-126, 186 and 365

miR name	Gene Symbol	Location	Description	Pathways
rno-mir-126	Glul	13q22	glutamate-ammonia ligase (glutamine synthase)	Glutamate metabolism Peptidoglycan biosynthesis Nitrogen metabolism
	Psma2	17q12.1	proteasome (prosome, macropain) subunit, alpha type 2	Proteasome
	Plk1	1q36	polo-like kinase 1 (Drosophila)	Cell cycle
rno-miR-186	Mat1a	16p14	methionine adenosyltransferase I, alpha	Methionine metabolism Selenoamino acid metabolism
	Polm	14q21	polymerase (DNA directed), mu	DNA polymerase
	Rpl24	11q12	ribosomal protein L24	Ribosome
	Gabra4	14p11	gamma-aminobutyric acid (GABA-A) receptor, subunit alpha 4	Neuroactive ligand-receptor interaction
	Atp2b2	4q41.3-	ATPase, Ca++ transporting, plasma membrane 2	Calcium signaling pathway
rno-miR-365	Itpr2	4q44	inositol 1,4,5-triphosphate receptor 2	Calcium signaling pathway Phosphatidylinositol signaling system Gap junction Long-term potentiation Long-term depression GnRH signaling pathway
	Cldn16	11q22	claudin 16	Cell adhesion molecules (CAMs) Tight junction Leukocyte transendothelial migration
	Prkg1	1q43	protein kinase, cGMP-dependent, type 1 (mapped)	Gap junction Long-term depression Olfactory transduction
	Grap2	7q34	GRB2-related adaptor protein 2	T cell receptor signaling pathway
	Psma2	17q12.1	proteasome (prosome, macropain) subunit, alpha type 2	Proteasome

6.5 Discussion

6.5.1 Validation of miRNA Microarray Data

Validation was performed using two commercially available real-time PCR miRNA systems. The miScript system failed to reproduce microarray data for the two most over and under-represented miRNA species, while the miRcury system validated 3 out of 6 miRNA species assayed. The validation success rate stands at only 50 % of the miRNA genes identified by microarray. Data from microarrays is not always reliable; therefore some of the miRNA signals could be false positive results. Furthermore, miRNA are difficult to quantify as the short nucleotide sequences can present problems in maintaining the specificity of hybridisation sequences.[512]

Differences in the real-time PCR system technology can help to explain the failure of one system to replicate any microarray observations. The miScript system uses a global miRNA cDNA synthesis step and does not always target mature miRNA sequence during PCR (personal communication, Qiagen technical support), while the miRcury system amplifies only one miRNA species in duplicate RT-PCR reactions using gene-specific RT-PCR primers and LNA PCR primers that hybridise to the mature miRNA sequence. Reverse transcription is the most variable experimental step in real-time PCR based quantification, gene specific RT-PCRs have been shown to increase sensitivity and duplicate RT-PCR reactions help to estimate reproducibility. (reviewed by Nolan et al, 2006)[449] The locked nucleic acid (LNA) primer technology is also beneficial when using short primer hybridisation sequences as LNA bases have increased binding affinity over standard DNA bases.[513]

6.5.2 Validated miRNA Transcripts

The following three miRNAs were validated by real-time PCR and had increased levels in PR lung tissue.

6.5.2.1 rno-mir-126

miR-126 is the intronic product of the epidermal growth factor like domain 7 (*EGFL7*) gene[514] and this miRNA has been shown to inhibit tumour invasion in lung cancer.[515] miR-126 has also been shown to regulate vascular integrity, developmental angiogenesis[516, 517] and provide control of vascular inflammation by inhibiting *VCAM1* expression (*VCAM1* mediates leukocyte adherence to

endothelial cells) from TNFα activated endothelial cells.[493] In breast cancer, miR-126 was shown to inhibited cell growth by targeting insulin receptor substrate 1 (*IRS1*).[518]

In this rat model, the increased miR-126 levels could be promoting lung blood vessel formation by repressing the expression of *SPRED1*, an intracellular inhibitor of angiogenic signaling, thereby increasing the pro-angiogenic actions of VEGF and FGF.[517] However, this is speculation based on current knowledge of miR-126 action and will have to be assessed by functional studies and morphometric analysis of lung vasculature.

6.5.2.2 rno-mir-186

The function of miR-186 is not fully understood. However, along with mir-150 it has been demonstrated to down regulate the expression of the pro-apoptotic purinergic $P2X_7$ receptor in cancer epithelial cells.[519] In another publication, three miRNAs, including miR-186 were predicted to target over 50% of all lupus genes, indicating that miRNA miR-186 may target immune transcripts.[520] P2X receptors are widely expressed in cells of the immune system with varying functions[521] and can be used as biomarkers for alveolar epithelial type I and type II cells.[522] Therefore, increased levels of miR-186 in PR lung could result in reduced apoptosis of alveolar epithelial cells. In a small study of asthma, attenuated P2X7 pore function has been identified as a risk factor for virus-induced loss of asthma control.[523]

6.5.2.3 rno-mir-365

Even less is known about the function of miR-365, one study has reported changes to miRNA signatures, including miR-365, in different growth arrest states in WI-38 lung fibroblasts.[524] miR-365 appeared to be a key regulator of cell cycle inhibitors, senescence and inverse correlations with p16, p21 and PTEN indicated that it may play an important role in the regulation of cell proliferation and survival axis.[524] Therefore, increased levels of miR-365 may indicate that there is increased cell survival and proliferation in PR lung compared to control lung samples.

6.5.3 Gene Cluster Analysis

Five intronic miRNA genes had altered levels in the rat PR lungs (see table 6.8); this is potentially interesting because as miRNA genomic location is well conserved between species, the relationship between the host gene and miRNA must be functionally important. For example, the host gene of miR-33 (up in PR lung) is *SREBF2* a sterol-sensitive transcription factor that controls cellular cholesterol homeostasis, if miR-33 is co-expressed with the host gene, then activation of *SREBF2* transcription will increase miR-33 levels repressing its predicted target gene *ABCA1*. *ABCA1* is a cholesterol efflux pump in the cellular lipid removal pathway so targeting by miR-33 could increase cellular lipid levels. In the lung, *ABCA1* is involved in basolateral transport of surfactant[525] and knockout studies in mice show major morphological abnormalities (foamy type II pneumocytes, macrophages and cholesterol clefts) that increase in severity with age.[526] It was suggested that this accumulation of foamy cell represented a protective phagocytic response but could also causes chronic stress and impair cellular functions, resulting in the production of cytokines and extracellular degrading enzymes that could injure lung cells.[526] Alternatively, under normal homeostasis miR-33 can act as part of a negative feedback loop to 'fine-tune' the expression of *ABCA1* which can be initiated by its host gene, *SREBF2*.[527] The relationship between miR-126 and *EGFL7* is also interesting and is discussed later as an opportunity for future work (Section 6.7).

miRNA genes within 500 base pairs of each other are likely to be expressed as a single polycistronic transcript. Although based on microarray results from whole lung preparations, lower levels of all mature miRNAs from a single polycistronic transcript or switching on or off an isolated miRNA gene could indicate that the expression of the pri-miRNA is down-regulated by transcription factors or silenced by DNA methylation. Alternatively, the maturation of all miRNAs within a cluster could be regulated to the same extent by post-transcriptional mechanisms. Lower levels of a single or several, but not all miRNAs from a polycistronic transcript could indicate that the maturation of specific miRNA is regulated by post-transcriptional mechanisms. Cluster analysis of the miRNA genes identified a one cluster that could be transcriptionally regulated or reduced by DNA methylation, as both miRNA genes in the miR-181a1~181b1 cluster were found at lower levels in PR lung.

Observations in two other miRNA gene clusters (miR-17~92 and miR-24-2~23a) indicate that post-transcriptional mechanisms are most likely to control the maturation of particular miRNAs from a polycistronic transcript and target others for degradation. Any conclusion drawn from these observations awaits confirmation by detailed functional molecular studies; but it is likely that lung miRNA profiles will vary by tissue/cell type under the control of a complex interaction of epigenetic, transcriptional and post-transcriptional mechanisms.

6.5.4 Predicted miRNA Target Genes and Pathways

The miRNApath analysis predicted that groups of up to five of the differentially expressed miRNA could target multiple genes in the calcium signalling, MAPK signalling and oxidative phosphorylation pathways. Calcium signalling is a major pathway, involved in membrane depolarisation and acts as a secondary messenger in muscle contraction, cellular motility, enzyme activity and ion pumps (reviewed by Petersen, 2005).[528] Mitogen Activated Protein Kinase (MAPK) signalling is involved in action of most of the non-nuclear oncogenes and mediate cellular responses to growth factors.(reviewed by Pearson et al, 2001)[529] For example, the predicted miRNA target gene *MAPK8ip3* interacts with *MAPK8* (a member of the stress-activated c-Jun N-terminal kinases, *JNKs*). *MAPK8ip3* acts as a scaffolding protein to enhance activation of the signalling pathway.[530] The downstream pro-apoptotic molecule, *Caspase3* is also a predicted target of some the miRNAs and another predicted target is *ELK1*, this gene is the nuclear target of ras-raf-MAPK signalling leading to cell proliferation and differentiation.(reviewed by Sharrocks, 2002)[531] Oxidative phosphorylation is the mitochondrial metabolic pathway releasing energy from oxidation of nutrients to produce ATP by the electron transport chain and reactive oxygen species are produced as a by-product (reviewed by Lesnefsky et al, 2006).[532]

Focussing on the some of predicted targets genes of validated miRNAs (miR-126, 186 and 365 up in PR) provides more intriguing insights into the action of miRNAs in the lung. Predicted targets of miR-126 include Glutamine synthetase (*Glul*), this is an enzyme with an essential role in the metabolism of nitrogen and *Glul* mRNA is present in the embryonic day 17 mouse lung bronchioli.[533] Glutamine synthetase catalyses the condensation of glutamate and ammonia to form glutamine, an amino acid essential for growth.[534]

If miR-126 does target glutamine synthetase, then it will provide the first evidence for miRNA involvement in protein metabolism (catabolism) in a DOHaD animal model of maternal protein restriction. In this model, protein-restriction was defined by a 50% reduction in caesin protein as part of an isocaloric diet fed during pregnancy.[262] Caesin accounts for 80% of the protein in cow's milk, and milk is a source of exogenous L-glutamine.[535, 536] Fetal uptake of maternal glutamine across the placenta is the highest of any amino acid and in fetal sheep there is no uptake of maternal glutamate.

Therefore, expectant mothers on the protein-restricted diet could have less glutamine to supply to the fetus via the maternal-fetal-placental unit, although there is some evidence to suggest that the deficit could be made up by the mobilisation of maternal protein.[418] During pregnancy, glutamate requirements must be met by the fetal conversion of glutamine to glutamate by the liver enzyme glutaminase. Fifty percent of the fetal glutamate is sent back to the placenta for conversion into glutamine adding to the placental glutamine pool destined for the fetus. Therefore, is lung glutamine synthetase targeted by miR-126 in order to protect fetal lung glutamate for essential nitrogen signalling and lung growth in development?

Differences in organ weights at birth in a study of maternal protein-restriction indicate that this could be the case as lung growth is not as severely affected (i.e. lung weight is in proportion to body weight) unlike other organs such as pancreas, liver, muscle and spleen demonstrating significantly lower weights in PR animals when compared to control offspring.[420] Affected organs, such liver, do go on to have a period of catch up growth but hepatic function is profoundly affected as a result of maternal protein restriction during pregnancy (cited in Desai et al, 1996).[420] Indicating that although restricted growth may be transient, organ function can be compromised. In humans, premature and intrauterine growth restricted (lower 10th percentile of size for gestational age) infants have poor respiratory function at birth, pulmonary structural abnormalities and go on to experience poorer respiratory function in adulthood. (Reviewed in Maritz et al, 2005)[367] 'Brain-sparing' is an accepted phenomenon in PR models and protects essential neurological functions, is the lung 'spared' as well?

Furthermore, should the timing of exposure to maternal protein-restriction also be considered before interpreting a potential role for miRNA? Rat lung development continues from birth to the first few weeks of life, therefore it is logical that post-

natal maternal diet during lactation could have an effect on offspring lung growth. The observation of significantly lower offspring lung weights from mothers on a low protein diet during the weaning period, indicates that the timing of exposure is important.[420] Therefore, it would be interesting to compare miRNA expression profiles in offspring experiencing exposure to maternal protein-restriction at different times during pregnancy and the weaning period.

The expression of miR-126 in other organs affected by maternal protein restriction awaits determination, however levels of glutamine synthetase enzyme activity have been shown to be increased in the hippocampus, hypothalamus and cerebellum in weanling rats exposed to maternal protein-restriction[262] indicating that gluamine synthetase is sensitive to this exposure and miR-126 could be a way to fine-tune expression of this gene in a tissue-specific manner.

Why are there increased levels of miR-126 in offspring lung 120 days after exposure to maternal protein restriction and are changes in DNA methylation responsible? Understanding the relationship between miR-126 and its host protein-coding gene *Egfl7* could help to answer this question, as in human cancer, the miR-126 transcripts are processed from certain *EGFL7* mRNA splice forms and *EGFL7* expression is known to be controlled by promoter DNA methylation.[511]

Other miR-126 targets include, *Psma2* a sub-unit of the proteosome and *Plk1*. *Plk1*, a potential target for cancer drugs, is known to interact with proteosome sub-units *PSMA3-7* and is a regulator of proteosome activity that is involved in the regulation of cellular division and genomic stability.[537, 538]

miR-186 target genes that are expressed in lung include a neuro-active ligand receptor, *Gabra4*; polymorphism in this gene has been associated with nicotine dependence in humans[539] and the *Atp2b2* gene that is a key membrane protein in the maintenance of intracellular calcium homeostasis.[540] *Mat1a* is also predicted to be a target of miR-186 but this molecule is only expressed in adult liver because of tissue specific DNA hypomethylation.[541] If this miRNA is found to be differentially expressed in PR liver then it could provide a role for miRNA influence on the one-carbon pathway and substrate availability (S-adenosylmethionine) for DNA methyltransferases, potentially leading to alterations to DNA methylation levels.[542]

miR-365 target genes include *Itpr2* and *Prkg1* and *Grap2*. *Itpr2* is a key molecule in mobilising calcium from storage organelles and in the regulation of cell proliferation and Prkg1 is known to interact with *Itpr1*, an inositol trisphosphate

receptor similar to *Itpr2*.[543-545] *Grap2* is a leukocyte specific protein-tyrosine kinase that interacts with CD8 and through the co-stimulatory pathway activates T cells,[546] this molecule contains a glutamine rich linker region with a caspase 3 cleavage site.[547]

6.5.5 Pattern of miRNA Expression

It is tempting to focus on specific predicted gene targets of the differentially expressed miRNA species found in PR lung and qualify gene candidates by *a priori* hypotheses based on potential roles in development, maturation and pathophysioological conditions, without further confirmation by functional studies. The fact that miRNA profiles will differ in time, by cell type and could have many modes of action in numerous pathways that can interact in complex mechanisms, must be taken into account. A more general overview of the significance of altered miRNA expression in lung can be proposed by grouping miRNA expression into patterns.

Irrespective of miRNA target genes and pathways there are three distinct miRNA expression patterns. Group one, consists of miRNAs that were found in both PR and control lungs but at different levels. Group two, consists of miRNAs that were only found in PR lung and the third group are miRNAs found only in control lung. As the exposure (maternal protein-restriction) occurred during fetal development and yet miRNA expression is altered in adulthood (120 days), this poses the question:

'Are miRNA expression profiles altered as a consequence of PR or in response to dysregulation of processes sensitive to PR?'

In short, the pattern of miRNA expression suggests that both have occurred to some extent. Group one must consist of some miRNA genes that are essential for normal organ function (as they are found in PR and control lungs) and alteration of miRNA levels in this group could reflect the miRNA response to dysregulation of gene processes due to PR exposure in development (i.e. compensatory or 'fine tuning' of expressed protein coding gene profiles).

miRNA gene expression from development to adulthood is temporal, and many miRNAs are switched on or off, thereby regulating events such as developmental timing and growth. Group two and three are interesting in this respect as they are found in either PR or control lungs, but not in both.

For example, the presence of let-7d and miR-365 in PR lungs indicates that let-7d could be still regulating events that should have finished in late development and miR-365 could be regulating cell proliferation and survival.

Another factor to take in to consideration when interpreting these results is the role of miRNAs in the aging process.[548] In worms, let-7 expression decreases with age and in mice a maternal protein-restricted diet during pregnancy has been associated with altered offspring longevity.[549] Therefore, miRNA profiles are only likely to provide a snap-shot of miRNA gene regulation at that time, and possibly of events upto that period of life. Also, the role of certain miRNAs may change as the animal ages, further confounding our interpretation of miRNA function.

6.6 Conclusion

In conclusion, this work has demonstrated altered miRNA levels in whole lung preparations from 120 day old offspring of mothers exposed to a protein-restricted diet in pregnancy. The regulation of miRNA gene expression is unknown but is likely to involve, DNA methylation, transcriptional and post-transcriptional regulatory mechanisms. Target gene predictions revealed numerous pathways that could be under the influence of the miRNA, including cellular metabolism (protein metabolism by the glutamate pathway), calcium, nitrogen, MAPK signalling and oxidative phosphorylation. Expression patterns revealed that a set of miRNAs, thought to be important for lung homeostasis, were likely to be altered in response to protein-restriction and two further groups that may have been switched on or off as a consequence of the maternal exposure, these miRNA may regulate developmental timing, cell proliferation and survival. However, these results should be interpreted with caution as miRNA expression at 120 days unlikely to reflect the miRNA profile at birth and in the first weeks of life, where a window of lung developmental plasticity exists.

6.7 Future Work

Only three miRNA species have been validated by real-time PCR. Therefore, further validation work is required for the remaining 37 species identified by microarray. It would be prudent to replicate the miRNA findings in samples from an independent animal experiment and at other sampling time-points. miRNA expression is known to change temporally (development → growth → adulthood) and have tissue specific profiles; none of which was addressed in the current work. As other sampling time-points can be obtained from this rat model it would be interesting to observe miRNA profiles from birth to old age and see if they alter over time. Furthermore, processing of whole lung samples by paraffin-embedding, sectioning (of large and small airways) and Laser Capture Micro-dissection (LCM) would allow for the determination of tissue/cell specific (e.g. airway epithelium or vascular endothelium) expression profiles. Alterations to miRNA profiles in adulthood (120d) may not have persisted from birth (i.e. miRNAs directly altered by or in response to the exposure may change over time) and could actually reflect rescue or maintenance of altered organ homeostasis or function that is programmed by other mechanisms; therefore it is essential to compare miRNA profiles from birth into later life.

In other organs, nutritional programming has shown transgenerational transmission of phenotypes without further exposure, indicating an epigenetic mode of inheritance that may involve changes to DNA methylation. Whether DNA methylation controls miRNA expression in the PR lung and is a trans-generational epigenetic mark could be determined by future studies, assessing miRNA profiles in F_2 and F_3 generation animals (samples under collection), in parallel with genome-wide miRNA gene promoter CpG island DNA methylation analysis by microarray. Alternatively or in parallel with the DNA methylation microarray analysis, *EGFL7* (protein coding gene) and its intra-genic miR-126 could be used as a model system to test the role of DNA methylation controlling miRNA expression. miR-126 (up-regulated in PR lung) is known to be co-expressed with *EGFL7* and has previously been shown to have altered DNA methylation in disease.[511]

The predicted target genes remain that, a prediction. Target gene protein levels need to be determined by western-blotting the protein collected from the Trizol™ preparations. Reduced protein levels would indicate that further functional studies of miRNA action would be warranted; this could be achieved by miRNA gene knockdown experiments *in vivo*.

The ultimate aim of this work is to use this animal model in combination with environmental exposures linked with respiratory disease, such as cigarette smoke. Therefore, it is essential that alterations to miRNA expression through nutritional programming is characterised well and fully understood before interpreting any effects of inhaled risk factors on lung miRNA expression profiles.

7 General Discussion

Chronic non-transmittable human disease, such as asthma, CVD, diabetes, obesity, and COPD are known to have a complex genetic basis that is induced through exposure of genetically susceptible individuals to environmental risk factors (high fat, high sugar diets and smoking etc.).[169, 550]

Common genetic polymorphism associate with altered disease susceptibility and can interact with environmental exposures modifying disease risk. However, genetic polymorphism alone cannot explain all of the heritability of common disease in families or the recent explosion in the common disease epidemic.[551] Changes in lifestyle may be more important, but the exact origin or cause of chronic human disease remains unclear.[552]

Epidemiological studies of low birth weight have shown associations with increased disease risk and mortality in later life.[223-226] This lead to the proposal that impaired fetal growth may leave individuals at higher risk of disease in adulthood.[553] These observations and work using animal models with altered maternal diet have lead to the developmental origins of health and disease hypothesis.[230] Where it is thought that maternal exposures during pregnancy may alter or 'program' the course of fetal development, predisposing offspring to increased risk of adulthood disease if the prenatal exposures do not match exposures in post-natal life.[429]

In respiratory disease such as, asthma and COPD it is known that low birth weight,[370] anthropomorphic measures at birth,[554] and environmental exposures (maternal nutrition and ETS)[58] associate with increased disease risk. Common genetic polymorphism can predict lung function growth,[319] correlate with decline in lung function as the person ages,[406] and deficits in lung function associate with genotype and smoking.[338,556] Smoking is a major risk factor for COPD in later life.[295] However, not all smokers develop COPD and non-smokers can develop COPD.[295] These observations indicate that not everyone is susceptible and certain individuals are predisposed to develop disease without exposure to high levels of tobacco smoke.

Our study began by examining the relationship between important anti-oxidant genes known to detoxify tobacco smoke and early life exposure to tobacco smoke in children with asthma. In the first part of this thesis we investigated the interaction between glutathione s-transferase gene polymorphism, ETS exposure and childhood asthma. We developed copy number variation genotyping methodology to provide full *GSTM1* and *GSTT1* gene dosage information for use in the genetic association tests. We found that common GST polymorphism was associated with increased asthma risk, IgE levels and disease severity scores and indentified the presence of a novel G-E interaction between *GSTT1*0* (gene deletion), ETS exposure and increased IgE levels in asthmatic children.

As respiratory disease is thought to have developmental origins, we then examined whether an established animal model of maternal protein-restriction used in the study of the developmental origins of health and disease hypothesis is relevant for the study of lung health and disease in general. In the second part of the thesis we sought to determine the effect of maternal protein-restriction on offspring lung physiology, gene and miRNA expression. With the aim to use this established DOHaD animal model to investigate the effect of maternal tobacco smoke exposure on offspring lung health. In the DOHaD animal model we found that maternal PR alters gene and miRNA expression profiles in F_1 adult offspring lung tissue. However, these observations were not found in older animals and were not transmitted to the next generation.

7.1 Gene-Environment Interactions in the Early Life Origins of Asthma

Asthma is a chronic respiratory disease involving hypersensitivity to airborne allergens, inflammation, reversible bronchoconstriction and has its origins in early life. Environmental tobacco smoke exposure in early life increases the prevalence and severity of asthma and maternal smoking is associated with reduced lung function in children.[54, 58]

Asthma is a complex genetic disease with incomplete penetrance and common genetic polymorphism has been shown to alter disease susceptibility. For example single nucleotide polymorphisms in *HLA-G, GPRA* and *ADAM33* genes are consistently associated with asthma in several populations.[98]

Glutathione S-transferase genes are important antioxidant enzymes that detoxify harmful combustible tobacco smoke products and deficiencies in enzyme activity due to genetic polymorphism has been associated with increased asthma risk and deficits in lung function.[320, 322]

A gene-environment interaction is the phenotypic effect of the interaction between the environment and a particular gene allele. Gene-environment interactions can explain the inter-individual differences in sensitivity to environmental exposures that are risk factors for disease susceptibility. Several studies have shown presence of gene-environment interactions between Glutathione S-transferase polymorphism and smoke exposure, these interactions can increase the risk of developing asthma and modulate the severity of disease phenotypes in asthmatic children.[317, 331]

This thesis has found no evidence to suggest that *GSTM1*, *GSTP1* or *GSTT1* polymorphism are associated with asthma susceptibility in the UK sample population of families with two asthmatic children. This is in line with recent findings from large birth cohorts and meta-analyses of previous studies.[349, 557]

A possible explanation for the disparity in findings in the literature, include publication bias in favour of small studies with ‘loose’ phenotype definitions of asthma, including wheeze. In our study asthma is clearly defined as doctor diagnosed asthma, requiring asthma medication. Therefore, GST polymorphism may actually be more important for wheeze in childhood, although some studies have found associations with BHR, a hallmark phenotype of asthma.[359, 555, 558, 559]

The findings from this thesis suggest that GST polymorphism is associated with the development of more severe asthma in children regardless of tobacco smoke exposure. This observation fits in with the hypothesis that an inability to efficiently detoxify reactive metabolites such as reactive oxygen species, could enhance the inflammatory response to toxins, leading to a progressive cycle of lung epithelial cell damage, shedding and persistent inflammation.[322] This is relevant, as uncontrolled chronic inflammation in asthmatics is associated with more frequent and severe exacerbations.[560, 561]

The FBAT association tests did reveal a significant negative association with asthma diagnosis, age corrected IgE levels, FEV_1 values and a rare *GSTO2* haplotype. This is an intriguing finding as polymorphism in this gene has previously been independently associated with lung function and COPD susceptibility.[338, 339] Could the *GSTO2* gene be a shared genetic susceptibility locus via the association with lung function? Further genetic association studies of asthma and COPD are required to elucidate this.

The identification of only two gene-environment interactions between *GSTT1*0 and high IgE levels in smoke exposed asthmatics and the *GSTO2* h1 haplotype and reduced atopy severity scores in the regression analysis, highlights the lack of statistical power in the current study. This probably due to the small number of smoke exposed individuals and the fact that the smoke exposure variable lacked any quantifiable measure of the level of smoke exposure.

The significant association with a *GSTO2* haplotype provides the first evidence that this gene may be important in asthma susceptibility. This finding mirrors recent publications that have examined the role of other GST genes in asthma. For example, variation in other GST Mu gene family members (*GSTM2, GSTM3, GSTM4*) have been shown to be determinants of childhood lung function growth and interact with maternal smoking during pregnancy.[562]

The genetic association analyses presented in this thesis were limited by the poorly defined tobacco smoke exposure measure and lack of statistical power due to the small number of exposed individuals. In this study tobacco smoke exposure was simply defined as the presence of a parental smoker in childhood.

Therefore, it was not possible to separate out or quantify the level of exposure that occurred during pregnancy (*in utero*) or in early life (environmental second hand smoke). Also, bias and confounding could also be a potential issue as there was no control over other factors, such as maternal socio-economic status or educational attainment, that are associated with both asthma and smoking.[563, 564]

In the future, the study of gene-environment interactions between ETS exposure and GST polymorphism in asthma would be facilitated by larger family based, birth cohort or case-control sample populations with measures of the level of ETS exposure (for example by the number of pack years of smoking). Potential confounding by other maternal factors associated with both smoking and asthma could be controlled for by performing a 'propensity to smoke' adjustment to the data analysis.[557] A recent study of childhood asthma was able to confirm, by using a propensity score that controlled for 32 confounders, that intrauterine tobacco smoke exposure is associated with lower lung function in childhood and that this was independent of the mother's propensity to smoke.[557]

As the GST enzymes have known to have overlapping substrate affinities and a deficiency in one gene could be compensated by another gene, this could enable an individual to still successfully detoxify tobacco smoke. A conditional logistic regression analysis of a prolife of GST gene polymorphism could be utilised to determine the presence of any epistatic effects between Glutathione S-transferase genes and asthma phenotypes and determine the critical GST genes that modify the risk of smoke exposure.

Finally, we can only fully examine the role of GSTs in the early life origins of asthma by, not only carefully measuring early life environment exposures in relation to polymorphism in genes, but also considering the role of the epigenome all in the same cohort, as prenatal tobacco smoke exposure has been shown to affect global and gene-specific DNA methylation.[562]

7.2 Modelling the Developmental Origins of Respiratory Disease Hypothesis

The interaction between the *in utero* - early life environment, genetic and epigenetic mechanisms in development are thought to predispose individuals to susceptibility to disease in adult life.[228] During development there are critical windows of developmental plasticity that are sensitive to programming mechanisms. This allows the fetus to adapt it's phenotype to the expected external environment, aiding survival in later life, based on nutritional cues from it's mother during pregnancy. Disease is thought to occur when the offspring's phenotype does not match the external environment in later life.[429]

Exposure to under-nutrition via placental ligation, total nutritional restriction or restriction of individual key nutrients such as, protein-restriction during pregnancy, in animal models has recapitulated many of the common chronic human disease phenotypes. Exposure to nutrient restriction early or late in gestation has demonstrated that timing of the exposure is also important, because as the fetus develops there are critical windows of developmental plasticity during organogenesis, where an exposure can have a greater effect.[565]

Transmission of offspring phenotypes to the next and even third generation, without further exposure, have indicated that epigenetic mechanisms, such as DNA methylation, have an important role in the programming of these phenotypes at the germ line level.[264]

Epidemioloigical studies of humans have demonstrated that lower lung function and COPD mortality associates with lower birth weight, a surrogate marker for impaired growth during pregnancy. Maternal diets low in vitamins and grand-maternal and maternal smoking during pregnancy are associated with asthma phenotypes in children.[366, 566, 567] In rats, maternal protein-restriction has been shown to alter the expression of genes involved in glucocorticoid hormone action and reduce lung size in young offspring.[420, 428]

However, little is known about the role of maternal nutrition on offspring lung development, metabolism, homoeostasis and later life function or whether offspring are more susceptible to environmental risk factors for lung disease in later life.

A DOHaD animal model, such as the Southampton protein-restricted rat, represents a convenient experimental system in which to test the developmental origins of respiratory disease hypothesis. The use of an animal model is not without precedence as animals are routinely used in study of respiratory disease; mainly focussing on allergy, immune responses and effects of environmental exposures on lung remodelling processes that underlie the pathogenesis of diseases such as asthma and COPD.

This thesis found that maternal protein restriction did not result in detectable alterations to basic lung measurements of size or volume in 225 day adult rat offspring. This finding is in line with previous studies of the effect of protein restriction or intrauterine growth restriction on organ size in the rat, where differences detected in early life (up to day 42) are likely to be transient and not seen in much older animals (11 months).[420, 442] Our study did find alterations to candidate gene expression and miRNA profiles in 120 day old first generation offspring. This is much later on in the life of the offspring, indicating that this maternal exposure could have consequences for offspring lung metabolism and homeostasis into adulthood. However, there was no evidence to suggest the transmission of altered candidate gene expression to the second generation offspring. This finding is not unexpected as there is no evidence in the literature to suggest a transgenerational transmission of lung phenotypes induced in response to maternal protein restriction in DOHaD animal models.

In humans, grand-maternal smoking is associated with increased asthma risk, therefore it would be interesting to examine whether this environmental respiratory risk factor induces transgenerational transmission of a lung phenotype to the next generation of control diet offspring. As pre-natal cigarette smoke exposure is known to alter global DNA methylation levels,[562] future work should consider the role of DNA methylation in the epigenetic transmission of phenotypes induced by tobacco smoke exposure in early life.

The discovery of altered lung miRNA profiles in PR offspring is a novel finding in this DOHaD animal model and suggests that miRNAs may be important due to their role in developmental timing, stages of lung development and can be induced in response to environmental stimuli.[502]

In cancer, miRNA profiles are more powerful than mRNA profiles in prediction of stage of disease.[490] Therefore, assuming that miRNA expression can be programmed or dysregulated in some way by early life exposures could organ specific miRNA expression profiles predict which individuals are susceptible to common disease in adulthood? A disease-associated SNP guided miRNA association analysis in 15 common diseases, including hypertension, cardio-vascular disease and diabetes, revealed that SNPs in genes encoding miRNAs or in predicted miRNA binding sties in 3'UTR regions can interfere with the function of microRNAs.[568] The findings from this association study reflects our hypothesis that dysregulation of miRNA function could lead to increased susceptibility to disease.

The work presented in this thesis has several limitations. We have sought to define a role for maternal protein restriction by using a candidate gene approach, whereby we chose the genes of interest based on findings from previous studies, to investigate lung gene expression in exposed offspring. With this type of study design there is always the risk of not choosing the relevant genes. It should also be noted that we have only examined gene expression at two adult time points. In order to validate these findings and fully understand the implication of the gene expression results, it will necessary to examine earlier time points in the offspring lung. The implications of altered miRNA expression on predicted target genes is also limited as we have not performed any functional characterisation, such as gene knockdown experiments, to determine which genes the miRNAs actually target.

In the future, it will be necessary to perform gene-expression analysis in a hypothesis independent manner, by using microarrays, to determine global lung gene and miRNA gene expression in samples taken from gestation and early life through to old age. This is particularly relevant to the miRNA observations and should determine whether they are programmed by, or are responding to, effects induced by maternal protein-restriction. Assessment of miRNA expression and function in the lung and in other organs should indicate whether miRNA expression is an important regulatory mechanism that deserves consideration in the study of the developmental origins of health and disease hypothesis.

Previous studies using the Southampton rat model have observed structural changes in organs (e.g. nephron number in kidneys) that could alter normal function and have consequences for disease susceptibility. The lung is a highly structured organ by virtue of its function to provide efficient gas exchange that is necessary for respiration. Analysis of lung morphometry in protein-restricted lungs by stereological techniques should identify the presence of structural alterations that could impede normal function and could lead to increased susceptibility to pathologies induced by respiratory risk factors, such as cigarette smoke, encountered in early or later life.

Although this thesis did not discover any evidence to suggest the transmission of a phenotype to the next generation of offspring, epigenetic mechanisms such as, DNA methylation, are likely to be important for programming of the first generation phenotype. Therefore, it will be necessary to examine DNA methylation in the promoters of programmed genes identified by future microarray experiments and in shared promoter regions of miRNA genes, as miRNA expression can also be regulated by changes in DNA methylation patterns.[569, 489]

7.3 Conclusion

In conclusion, we have shown that GST polymorphism is associated with severity of asthma and atopy in children, including *GSTO2* gene polymorphism which has previously been associated with measures of lung function and COPD may also be important for asthma susceptibility. Using an established animal model used in study of the developmental origins of health and disease hypothesis we have demonstrated that mild maternal protein-restriction during pregnancy results in altered lung gene expression and major changes to miRNA expression profiles in adult offspring. We propose that this model is suitable for the study of environmental respiratory risk factors in relation to the developmental origins of respiratory disease hypothesis.

7.4 Future Directions

The further experimental work required to address the findings from the current project have been discussed at length in the individual chapter and general discussion sections. However, in addition to these possibilities, the results presented in this thesis open up many avenues for further research in other areas including:

- Given the association of *GSTO2* polymorphism with asthma in this thesis. Is *GSTO2* expression different in primary cell lung airway epithelium explants from asthmatic and non-asthmatic patients? Does smoke exposure modify *GSTO2* expression in cell explant models of asthma and COPD?
- Are the changes in gene and miRNA expression in protein-restricted lung due to changes in DNA methylation?
- Are the changes in miRNA expression that were observed in protein-restricted lung, tissue specific or not?
- Does maternal protein restriction in pregnancy in the rat induce a pulmonary hypertension-like phenotype in offspring lung by dysregulation of miRNA expression? In the miRNA gene expression work presented in this thesis, miR-20a was only found to be present in protein-restricted lung. In human cells lines this miRNA has been shown to target bone morphogenetic protein receptor type II (*BMPR2*) mRNA transcripts.[570] Dysregulated expression of this gene is a pathogenetic hallmark of pulmonary hypertension and is thought to lead to remodeling of the pulmonary vascular bed.[571]
- *Pcdh1* expression was found to be increased in protein-restricted rat lung. In mouse embryogenesis, *Pcdh1* expression is especially prominent in blood vessels and in both the developing and adult mouse kidney the glomeruli stained strongly for *Pcdh1* protein.[572] Previous animal studies of the programming of hypertension have used maternal protein restriction in rats and found reductions in nephron number in offspring kidney.[415] Given the role of *Pcdh1* in kidney development, its expression in adult glomeruli and the consequence of maternal protein-restriction on offspring kidney structure. Does maternal protein-restriction program *Pcdh1* expression in highly vascularised organs such as, kidney and placenta?

8 References

1. *Asthma.* World Health Organisation, 2006. Fact sheet N°(307).

2. Asher, M.I., et al., *Worldwide time trends in the prevalence of symptoms of asthma, allergic rhinoconjunctivitis, and eczema in childhood: ISAAC Phases One and Three repeat multicountry cross-sectional surveys.* Lancet, 2006. 368(9537): p. 733-43.

3. Pearce, N., et al., *Worldwide trends in the prevalence of asthma symptoms: phase III of the International Study of Asthma and Allergies in Childhood (ISAAC).* Thorax, 2007. 62(9): p. 758-66.

4. *British guideline on the management of asthma.* Thorax, 2003. 58 Suppl 1: p. i1-94.

5. Quanjer, P.H., et al., *Lung volumes and forced ventilatory flows. Report Working Party Standardization of Lung Function Tests, European Community for Steel and Coal. Official Statement of the European Respiratory Society.* Eur Respir J Suppl, 1993. 16: p. 5-40.

6. Busse, W.W. and R.F. Lemanske, Jr., *Asthma.* N Engl J Med, 2001. 344(5): p. 350-62.

7. Bloemen, K., et al., *The allergic cascade: review of the most important molecules in the asthmatic lung.* Immunol Lett, 2007. 113(1): p. 6-18.

8. Yuksel, H., et al., *Prevalence and comorbidity of allergic eczema, rhinitis, and asthma in a city in western Turkey.* J Investig Allergol Clin Immunol, 2008. 18(1): p. 31-5.

9. Barnes, N., *Most difficult asthma originates primarily in adult life.* Paediatr Respir Rev, 2006. 7(2): p. 141-4.

10. Morwood, K., et al., *Aspirin-sensitive asthma.* Intern Med J, 2005. 35(4): p. 240-6.

11. Pearce, N., J. Pekkanen, and R. Beasley, *How much asthma is really attributable to atopy?* Thorax, 1999. 54(3): p. 268-72.

12. Jeffery, P.K., *Remodeling in asthma and chronic obstructive lung disease.* Am J Respir Crit Care Med, 2001. 164(10 Pt 2): p. S28-38.

13. Fredberg, J.J., *Bronchospasm and its biophysical basis in airway smooth muscle.* Respir Res, 2004. 5: p. 2.

14. Pohunek, P., et al., *Eosinophilic inflammation in the bronchial mucosa in children with bronchial asthma [abstract].* Eur Respir J 2000, 2000. 11(suppl 25): p. 160s.

15. Holgate, S.T., et al., *Epithelial-mesenchymal interactions in the pathogenesis of asthma.* J Allergy Clin Immunol, 2000. 105(2 Pt 1): p. 193-204.

16. Warburton, D., et al., *The molecular basis of lung morphogenesis.* Mech Dev, 2000. 92(1): p. 55-81.

17. Bucchieri, F., et al., *Asthmatic bronchial epithelium is more susceptible to oxidant-induced apoptosis.* Am J Respir Cell Mol Biol, 2002. 27(2): p. 179-85.

18. Bayram, H., et al., *The effect of diesel exhaust particles on cell function and release of inflammatory mediators from human bronchial epithelial cells in vitro.* Am J Respir Cell Mol Biol, 1998. 18(3): p. 441-8.

19. Richter, A., et al., *The contribution of interleukin (IL)-4 and IL-13 to the epithelial-mesenchymal trophic unit in asthma.* Am J Respir Cell Mol Biol, 2001. 25(3): p. 385-91.

20. Holgate, S.T., *Epithelial damage and response.* Clin Exp Allergy, 2000. 30 Suppl 1: p. 37-41.

21. Riedler, J., et al., *Exposure to farming in early life and development of asthma and allergy: a cross-sectional survey.* Lancet, 2001. 358(9288): p. 1129-33.

22. Wickens, K., et al., *Farm residence and exposures and the risk of allergic diseases in New Zealand children.* Allergy, 2002. 57(12): p. 1171-9.

23. Svanes, C., et al., *Childhood environment and adult atopy: results from the European Community Respiratory Health Survey.* J Allergy Clin Immunol, 1999. 103(3 Pt 1): p. 415-20.

24. Hesselmar, B., et al., *Does early exposure to cat or dog protect against later allergy development?* Clin Exp Allergy, 1999. 29(5): p. 611-7.

25. Strachan, D.P., *Hay fever, hygiene, and household size.* Bmj, 1989. 299(6710): p. 1259-60.

26. von Mutius, E., et al., *Skin test reactivity and number of siblings.* Bmj, 1994. 308(6930): p. 692-5.

27. Eduard, W., et al., *Atopic and non-atopic asthma in a farming and a general population.* Am J Ind Med, 2004. 46(4): p. 396-9.

28. Kozyrskyj, A.L., P. Ernst, and A.B. Becker, *Increased risk of childhood asthma from antibiotic use in early life.* Chest, 2007. 131(6): p. 1753-9.

29. Mitchell, E.A., et al., *Risk factors for asthma at 3.5 and 7 years of age.* Clin Exp Allergy, 2007. 37(12): p. 1747-55.

30. Wickens, K., et al., *Antibiotic use in early childhood and the development of asthma.* Clin Exp Allergy, 1999. 29(6): p. 766-71.

31. Cohet, C., et al., *Infections, medication use, and the prevalence of symptoms of asthma, rhinitis, and eczema in childhood.* J Epidemiol Community Health, 2004. 58(10): p. 852-7.

32. Celedon, J.C., et al., *Antibiotic use in the first year of life and asthma in early childhood.* Clin Exp Allergy, 2004. 34(7): p. 1011-6.

33. Celedon, J.C., et al., *Lack of association between antibiotic use in the first year of life and asthma, allergic rhinitis, or eczema at age 5 years.* Am J Respir Crit Care Med, 2002. 166(1): p. 72-5.

34. Foliaki, S., et al., *Antibiotic sales and the prevalence of symptoms of asthma, rhinitis, and eczema: The International Study of Asthma and Allergies in Childhood (ISAAC).* Int J Epidemiol, 2004. 33(3): p. 558-63.

35. Wjst, M., et al., *Early antibiotic treatment and later asthma.* Eur J Med Res, 2001. 6(6): p. 263-71.

36. Macpherson, A.J. and N.L. Harris, *Interactions between commensal intestinal bacteria and the immune system.* Nat Rev Immunol, 2004. 4(6): p. 478-85.

37. Kalliomaki, M., et al., *Probiotics in primary prevention of atopic disease: a randomised placebo-controlled trial.* Lancet, 2001. 357(9262): p. 1076-9.

38. Kalliomaki, M., et al., *Probiotics and prevention of atopic disease: 4-year follow-up of a randomised placebo-controlled trial.* Lancet, 2003. 361(9372): p. 1869-71.

39. Kukkonen, K., et al., *Probiotics and prebiotic galacto-oligosaccharides in the prevention of allergic diseases: a randomized, double-blind, placebo-controlled trial.* J Allergy Clin Immunol, 2007. 119(1): p. 192-8.

40. Noverr, M.C., et al., *Development of allergic airway disease in mice following antibiotic therapy and fungal microbiota increase: role of host genetics, antigen, and interleukin-13.* Infect Immun, 2005. 73(1): p. 30-8.

41. Noverr, M.C., et al., *Role of antibiotics and fungal microbiota in driving pulmonary allergic responses.* Infect Immun, 2004. 72(9): p. 4996-5003.

42. Siegrist, C.A., *Neonatal and early life vaccinology.* Vaccine, 2001. 19(25-26): p. 3331-46.

43. Anderson, H.R., et al., *Immunization and symptoms of atopic disease in children: results from the International Study of Asthma and Allergies in Childhood.* Am J Public Health, 2001. 91(7): p. 1126-9.

44. Maitra, A., et al., *Pertussis vaccination in infancy and asthma or allergy in later childhood: birth cohort study.* Bmj, 2004. 328(7445): p. 925-6.

45. Bernsen, R.M., et al., *Reported pertussis infection and risk of atopy in 8- to 12-yr-old vaccinated and non-vaccinated children.* Pediatr Allergy Immunol, 2008. 19(1): p. 46-52.

46. Cates, C.J., et al., *Vaccines for preventing influenza in people with asthma.* Cochrane Database Syst Rev, 2004(2): p. CD000364.

47. Romieu, I., et al., *Effects of air pollution on the respiratory health of asthmatic children living in Mexico City.* Am J Respir Crit Care Med, 1996. 154(2 Pt 1): p. 300-7.

48. Cody, R.P., et al., *The effect of ozone associated with summertime photochemical smog on the frequency of asthma visits to hospital emergency departments.* Environ Res, 1992. 58(2): p. 184-94.

49. White, M.C., et al., *Exacerbations of childhood asthma and ozone pollution in Atlanta.* Environ Res, 1994. 65(1): p. 56-68.

50. Delfino, R.J., et al., *Daily asthma severity in relation to personal ozone exposure and outdoor fungal spores.* Am J Respir Crit Care Med, 1996. 154(3 Pt 1): p. 633-41.

51. Nel, A.E., et al., *Enhancement of allergic inflammation by the interaction between diesel exhaust particles and the immune system.* J Allergy Clin Immunol, 1998. 102(4 Pt 1): p. 539-54.

52. Baeza-Squiban, A., et al., *Airborne particles evoke an inflammatory response in human airway epithelium. Activation of transcription factors.* Cell Biol Toxicol, 1999. 15(6): p. 375-80.

53. Bonvallot, V., et al., *Organic compounds from diesel exhaust particles elicit a proinflammatory response in human airway epithelial cells and induce cytochrome p450 1A1 expression.* Am J Respir Cell Mol Biol, 2001. 25(4): p. 515-21.

54. Martinez, F.D., M. Cline, and B. Burrows, *Increased incidence of asthma in children of smoking mothers.* Pediatrics, 1992. 89(1): p. 21-6.

55. Eisner, M.D., et al., *Environmental tobacco smoke and adult asthma. The impact of changing exposure status on health outcomes.* Am J Respir Crit Care Med, 1998. 158(1): p. 170-5.

56. Gurkan, F., et al., *Predictors for multiple hospital admissions in children with asthma.* Can Respir J, 2000. 7(2): p. 163-6.

57. Gilliland, F.D., et al., *Maternal smoking during pregnancy, environmental tobacco smoke exposure and childhood lung function.* Thorax, 2000. 55(4): p. 271-6.

58. Gilliland, F.D., Y.F. Li, and J.M. Peters, *Effects of maternal smoking during pregnancy and environmental tobacco smoke on asthma and wheezing in children.* Am J Respir Crit Care Med, 2001. 163(2): p. 429-36.

59. Oryszczyn, M.P., et al., *In utero exposure to parental smoking, cotinine measurements, and cord blood IgE.* J Allergy Clin Immunol, 1991. 87(6): p. 1169-74.

60. Panda, K., et al., *Vitamin C prevents cigarette smoke-induced oxidative damage in vivo.* Free Radic Biol Med, 2000. 29(2): p. 115-24.

61. Aycicek, A., O. Erel, and A. Kocyigit, *Increased oxidative stress in infants exposed to passive smoking.* Eur J Pediatr, 2005. 164(12): p. 775-8.

62. Kosecik, M., et al., *Increased oxidative stress in children exposed to passive smoking.* Int J Cardiol, 2005. 100(1): p. 61-4.

63. Leem, J.H., et al., *Asthma attack associated with oxidative stress by exposure to ETS and PAH.* J Asthma, 2005. 42(6): p. 463-7.

64. Shaheen, S.O., et al., *Prenatal paracetamol exposure and risk of asthma and elevated immunoglobulin E in childhood.* Clin Exp Allergy, 2005. 35(1): p. 18-25.

65. Fryer, A.A., R. Hume, and R.C. Strange, *The development of glutathione S-transferase and glutathione peroxidase activities in human lung.* Biochim Biophys Acta, 1986. 883(3): p. 448-53.

66. Cossar, D., et al., *The alpha and pi isoenzymes of glutathione S-transferase in human fetal lung: in utero ontogeny compared with differentiation in lung organ culture.* Biochim Biophys Acta, 1990. 1037(2): p. 221-6.

67. Chen, T.S., J.P. Richie, Jr., and C.A. Lang, *Life span profiles of glutathione and acetaminophen detoxification.* Drug Metab Dispos, 1990. 18(6): p. 882-7.

68. Placke, M.E., D.S. Wyand, and S.D. Cohen, *Extrahepatic lesions induced by acetaminophen in the mouse.* Toxicol Pathol, 1987. 15(4): p. 381-7.

69. Shaheen, S.O., et al., *Paracetamol use in pregnancy and wheezing in early childhood.* Thorax, 2002. 57(11): p. 958-63.

70. Shaheen, S.O., et al., *Frequent paracetamol use and asthma in adults.* Thorax, 2000. 55(4): p. 266-70.

71. Garcia, E., et al., *Prevalence of and factors associated with current asthma symptoms in school children aged 6-7 and 13-14 yr old in Bogota, Colombia.* Pediatr Allergy Immunol, 2008.

72. Kuschnir, F.C. and A.J. Alves da Cunha, *Environmental and socio-demographic factors associated to asthma in adolescents in Rio de Janeiro, Brazil.* Pediatr Allergy Immunol, 2007. 18(2): p. 142-8.

73. Sharma, S.K. and A. Banga, *Prevalence and risk factors for wheezing in children from rural areas of north India.* Allergy Asthma Proc, 2007. 28(6): p. 647-53.

74. Vlaski, E., et al., *Acetaminophen Intake and Risk of Asthma, Hay Fever and Eczema in Early Adolescence.* Iran J Allergy Asthma Immunol, 2007. 6(3): p. 143-149.

75. Wong, G.W., et al., *Symptoms of asthma and atopic disorders in preschool children: prevalence and risk factors.* Clin Exp Allergy, 2007. 37(2): p. 174-9.

76. Bentur, L., et al., *Salivary oxidative stress in children during acute asthmatic attack and during remission.* Respir Med, 2006. 100(7): p. 1195-201.

77. Katsoulis, K., et al., *Serum total antioxidant status in severe exacerbation of asthma: correlation with the severity of the disease.* J Asthma, 2003. 40(8): p. 847-54.

78. Nadeem, A., et al., *Increased oxidative stress and altered levels of antioxidants in asthma.* J Allergy Clin Immunol, 2003. 111(1): p. 72-8.

79. Robroeks, C.M., et al., *Exhaled nitric oxide and biomarkers in exhaled breath condensate indicate the presence, severity and control of childhood asthma.* Clin Exp Allergy, 2007. 37(9): p. 1303-11.

80. Farchi, S., et al., *Dietary factors associated with wheezing and allergic rhinitis in children.* Eur Respir J, 2003. 22(5): p. 772-80.

81. Shaheen, S.O., et al., *Dietary antioxidants and asthma in adults: population-based case-control study.* Am J Respir Crit Care Med, 2001. 164(10 Pt 1): p. 1823-8.

82. Misso, N.L., et al., *Plasma concentrations of dietary and nondietary antioxidants are low in severe asthma.* Eur Respir J, 2005. 26(2): p. 257-64.

83. Gilliland, F.D., et al., *Children's lung function and antioxidant vitamin, fruit, juice, and vegetable intake.* Am J Epidemiol, 2003. 158(6): p. 576-84.

84. Romieu, I., et al., *Genetic polymorphism of GSTM1 and antioxidant supplementation influence lung function in relation to ozone exposure in asthmatic children in Mexico City.* Thorax, 2004. 59(1): p. 8-10.

85. Weiss, S.T. and A.A. Litonjua, *Maternal diet vs lack of exposure to sunlight as the cause of the epidemic of asthma, allergies and other autoimmune diseases.* Thorax, 2007. 62(9): p. 746-8.

86. Howard, D.J., et al., *Oxidative stress induced by environmental tobacco smoke in the workplace is mitigated by antioxidant supplementation.* Cancer Epidemiol Biomarkers Prev, 1998. 7(11): p. 981-8.

87. Pearson, P.J., et al., *Vitamin E supplements in asthma: a parallel group randomised placebo controlled trial.* Thorax, 2004. 59(8): p. 652-6.

88. Shaheen, S.O., et al., *Randomised, double blind, placebo-controlled trial of selenium supplementation in adult asthma.* Thorax, 2007. 62(6): p. 483-90.

89. Wiener, A.S., I. Zieve, and J.H. Fries, *The inheritance of allergic disease.* Ann Eugen, 1936. 7: p. 141-162.

90. Salter, H.H., *On asthma: its pathology and treatment.* London: John Churchill & Sons, 1868.

91. Spain, W.C. and R.A. Cooke, *Studies in Specific Hypersensitiveness. XI. The Familial Occurrence of Hay Fever and Bronchial Asthma.* Journal of Immunology, 1924. 9: p. 521-569.

92. Clarke, J.R., et al., *Evidence for genetic associations between asthma, atopy, and bronchial hyperresponsiveness: a study of 8- to 18-yr-old twins.* Am J Respir Crit Care Med, 2000. 162(6): p. 2188-93.

93. Koeppen-Schomerus, G., J. Stevenson, and R. Plomin, *Genes and environment in asthma: a study of 4 year old twins.* Arch Dis Child, 2001. 85(5): p. 398-400.

94. Nystad, W., et al., *A comparison of genetic and environmental variance structures for asthma, hay fever and eczema with symptoms of the same diseases: a study of Norwegian twins.* Int J Epidemiol, 2005. 34(6): p. 1302-9.

95. van Beijsterveldt, C.E. and D.I. Boomsma, *Genetics of parentally reported asthma, eczema and rhinitis in 5-yr-old twins.* Eur Respir J, 2007. 29(3): p. 516-21.

96. Duffy, D.L., et al., *Genetics of asthma and hay fever in Australian twins.* Am Rev Respir Dis, 1990. 142(6 Pt 1): p. 1351-8.

97. Laitinen, T., et al., *Importance of genetic factors in adolescent asthma: a population-based twin-family study.* Am J Respir Crit Care Med, 1998. 157(4 Pt 1): p. 1073-8.

98. Ober, C. and S. Hoffjan, *Asthma genetics 2006: the long and winding road to gene discovery.* Genes Immun, 2006.

99. Zhang, J., P.D. Pare, and A.J. Sandford, *Recent advances in asthma genetics.* Respir Res, 2008. 9: p. 4.

100. Moffatt, M.F., et al., *Genetic variants regulating ORMDL3 expression contribute to the risk of childhood asthma.* Nature, 2007. 448(7152): p. 470-3.

101. Howard, T.D., et al., *Identification and association of polymorphisms in the interleukin-13 gene with asthma and atopy in a Dutch population.* Am J Respir Cell Mol Biol, 2001. 25(3): p. 377-84.

102. van der Pouw Kraan, T.C., et al., *An IL-13 promoter polymorphism associated with increased risk of allergic asthma.* Genes Immun, 1999. 1(1): p. 61-5.

103. Tarazona-Santos, E. and S.A. Tishkoff, *Divergent patterns of linkage disequilibrium and haplotype structure across global populations at the interleukin-13 (IL13) locus.* Genes Immun, 2005. 6(1): p. 53-65.

104. Chen, W., et al., *Functional effect of the R110Q IL13 genetic variant alone and in combination with IL4RA genetic variants.* J Allergy Clin Immunol, 2004. 114(3): p. 553-60.

105. Howard, T.D., et al., *Gene-gene interaction in asthma: IL4RA and IL13 in a Dutch population with asthma.* Am J Hum Genet, 2002. 70(1): p. 230-6.

106. Madhankumar, A.B., A. Mintz, and W. Debinski, *Alanine-scanning mutagenesis of alpha-helix D segment of interleukin-13 reveals new functionally important residues of the cytokine.* J Biol Chem, 2002. 277(45): p. 43194-205.

107. Vladich, F.D., et al., *IL-13 R130Q, a common variant associated with allergy and asthma, enhances effector mechanisms essential for human allergic inflammation.* J Clin Invest, 2005. 115(3): p. 747-54.

108. Baldini, M., et al., *A Polymorphism* in the 5' flanking region of the CD14 gene is associated with circulating soluble CD14 levels and with total serum immunoglobulin E.* Am J Respir Cell Mol Biol, 1999. 20(5): p. 976-83.

109. Choudhry, S., et al., *CD14 tobacco gene-environment interaction modifies asthma severity and immunoglobulin E levels in Latinos with asthma.* Am J Respir Crit Care Med, 2005. 172(2): p. 173-82.

110. Eder, W., et al., *Opposite effects of CD 14/-260 on serum IgE levels in children raised in different environments.* J Allergy Clin Immunol, 2005. 116(3): p. 601-7.

111. Leynaert, B., et al., *Association between farm exposure and atopy, according to the CD14 C-159T polymorphism.* J Allergy Clin Immunol, 2006. 118(3): p. 658-65.

112. Nishimura, F., et al., *Failure to find an association between CD14-159C/T polymorphism and asthma: a family-based association test and meta-analysis.* Allergol Int, 2006. 55(1): p. 55-8.

113. Yang, I.A., et al., *Gene-environmental interaction in asthma.* Curr Opin Allergy Clin Immunol, 2007. 7(1): p. 75-82.

114. Simpson, A., et al., *Endotoxin exposure, CD14, and allergic disease: an interaction between genes and the environment.* Am J Respir Crit Care Med, 2006. 174(4): p. 386-92.

115. Vercelli, D., *Learning from discrepancies: CD14 polymorphisms, atopy and the endotoxin switch.* Clin Exp Allergy, 2003. 33(2): p. 153-5.

116. Zambelli-Weiner, A., et al., *Evaluation of the CD14/-260 polymorphism and house dust endotoxin exposure in the Barbados Asthma Genetics Study.* J Allergy Clin Immunol, 2005. 115(6): p. 1203-9.

117. Palmer, C.N., et al., *Common loss-of-function variants of the epidermal barrier protein filaggrin are a major predisposing factor for atopic dermatitis.* Nat Genet, 2006. 38(4): p. 441-6.

118. Marenholz, I., et al., *Filaggrin loss-of-function mutations predispose to phenotypes involved in the atopic march.* J Allergy Clin Immunol, 2006. 118(4): p. 866-71.

119. Palmer, C.N., et al., *Filaggrin null mutations are associated with increased asthma severity in children and young adults.* J Allergy Clin Immunol, 2007. 120(1): p. 64-8.

120. Duetsch, G., et al., *STAT6 as an asthma candidate gene: polymorphism-screening, association and haplotype analysis in a Caucasian sib-pair study.* Hum Mol Genet, 2002. 11(6): p. 613-21.

121. Van Eerdewegh, P., et al., *Association of the ADAM33 gene with asthma and bronchial hyperresponsiveness.* Nature, 2002. 418(6896): p. 426-30.

122. Allen, M., et al., *Positional cloning of a novel gene influencing asthma from chromosome 2q14.* Nat Genet, 2003. 35(3): p. 258-63.

123. Balaci, L., et al., *IRAK-M is involved in the pathogenesis of early-onset persistent asthma.* Am J Hum Genet, 2007. 80(6): p. 1103-14.

124. Nicolae, D., et al., *Fine mapping and positional candidate studies identify HLA-G as an asthma susceptibility gene on chromosome 6p21.* Am J Hum Genet, 2005. 76(2): p. 349-57.

125. Noguchi, E., et al., *Positional identification of an asthma susceptibility gene on human chromosome 5q33.* Am J Respir Crit Care Med, 2005. 172(2): p. 183-8.

126. Vendelin, J., et al., *Characterization of GPRA, a novel G protein-coupled receptor related to asthma.* Am J Respir Cell Mol Biol, 2005. 33(3): p. 262-70.

127. Laitinen, T., et al., *Characterization of a common susceptibility locus for asthma-related traits.* Science, 2004. 304(5668): p. 300-4.

128. Feng, Y., et al., *G protein-coupled receptor 154 gene polymorphism is associated with airway hyperresponsiveness to methacholine in a Chinese population.* J Allergy Clin Immunol, 2006. 117(3): p. 612-7.

129. Kormann, M.S., et al., *G-Protein-coupled receptor polymorphisms are associated with asthma in a large German population.* Am J Respir Crit Care Med, 2005. 171(12): p. 1358-62.

130. Malerba, G., et al., *Chromosome 7p linkage and GPR154 gene association in Italian families with allergic asthma.* Clin Exp Allergy, 2007. 37(1): p. 83-9.

131. Melen, E., et al., *Haplotypes of G protein-coupled receptor 154 are associated with childhood allergy and asthma.* Am J Respir Crit Care Med, 2005. 171(10): p. 1089-95.

132. Zhang, Y., et al., *Positional cloning of a quantitative trait locus on chromosome 13q14 that influences immunoglobulin E levels and asthma.* Nat Genet, 2003. 34(2): p. 181-6.

133. Jang, N., G. Stewart, and G. Jones, *Polymorphisms within the PHF11 gene at chromosome 13q14 are associated with childhood atopic dermatitis.* Genes Immun, 2005. 6(3): p. 262-4.

134. Hirota, T., et al., *Association between ADAM33 polymorphisms and adult asthma in the Japanese population.* Clin Exp Allergy, 2006. 36(7): p. 884-91.

135. Howard, T.D., et al., *Association of a disintegrin and metalloprotease 33 (ADAM33) gene with asthma in ethnically diverse populations.* J Allergy Clin Immunol, 2003. 112(4): p. 717-22.

136. Kedda, M.A., et al., *ADAM33 haplotypes are associated with asthma in a large Australian population.* Eur J Hum Genet, 2006. 14(9): p. 1027-36.

137. Lee, J.H., et al., *ADAM33 polymorphism: association with bronchial hyper-responsiveness in Korean asthmatics.* Clin Exp Allergy, 2004. 34(6): p. 860-5.

138. Noguchi, E., et al., *ADAM33 polymorphisms are associated with asthma susceptibility in a Japanese population.* Clin Exp Allergy, 2006. 36(5): p. 602-8.

139. Sakagami, T., et al., *ADAM33 polymorphisms are associated with aspirin-intolerant asthma in the Japanese population.* J Hum Genet, 2007. 52(1): p. 66-72.

140. Schedel, M., et al., *The role of polymorphisms in ADAM33, a disintegrin and metalloprotease 33, in childhood asthma and lung function in two German populations.* Respir Res, 2006. 7: p. 91.

141. Werner, M., et al., *Asthma is associated with single-nucleotide polymorphisms in ADAM33.* Clin Exp Allergy, 2004. 34(1): p. 26-31.

142. Hardy, G.H., *Mendelian Proportions in a Mixed Population.* Science, 1908. 28(706): p. 49-50.

143. Lewontin, R.C., *The Interaction of Selection and Linkage. I. General Considerations; Heterotic Models.* Genetics, 1964. 49(1): p. 49-67.

144. Devlin, B. and N. Risch, *A comparison of linkage disequilibrium measures for fine-scale mapping.* Genomics, 1995. 29(2): p. 311-22.

145. Wray, N.R., *Allele frequencies and the r2 measure of linkage disequilibrium: impact on design and interpretation of association studies.* Twin Res Hum Genet, 2005. 8(2): p. 87-94.

146. Niu, T., *Algorithms for inferring haplotypes.* Genet Epidemiol, 2004. 27(4): p. 334-47.

147. *The International HapMap Project.* Nature, 2003. 426(6968): p. 789-96.

148. Carlson, C.S., et al., *Selecting a maximally informative set of single-nucleotide polymorphisms for association analyses using linkage disequilibrium.* Am J Hum Genet, 2004. 74(1): p. 106-20.

149. Tease, C., G.M. Hartshorne, and M.A. Hulten, *Patterns of meiotic recombination in human fetal oocytes.* Am J Hum Genet, 2002. 70(6): p. 1469-79.

150. *A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes. The Huntington's Disease Collaborative Research Group.* Cell, 1993. 72(6): p. 971-83.

151. Rommens, J.M., et al., *Identification of the cystic fibrosis gene: chromosome walking and jumping.* Science, 1989. 245(4922): p. 1059-65.

152. Elston, R.C., *The genetic dissection of multifactorial traits.* Clin Exp Allergy, 1995. 25 Suppl 2: p. 103-6.

153. Risch, N. and K. Merikangas, *The future of genetic studies of complex human diseases.* Science, 1996. 273(5281): p. 1516-7.

154. Ellsworth, D.L. and T.A. Manolio, *The emerging importance of genetics in epidemiologic research II. Issues in study design and gene mapping.* Ann Epidemiol, 1999. 9(2): p. 75-90.

155. Hoffjan, S., D. Nicolae, and C. Ober, *Association studies for asthma and atopic diseases: a comprehensive review of the literature.* Respir Res, 2003. 4: p. 14.

156. Holloway, J.W., B. Beghe, and S.T. Holgate, *The genetic basis of atopic asthma.* Clin Exp Allergy, 1999. 29(8): p. 1023-32.

157. Holloway, J.W. and G.H. Koppelman, *Identifying novel genes contributing to asthma pathogenesis.* Curr Opin Allergy Clin Immunol, 2007. 7(1): p. 69-74.

158. Pillai, S.G., et al., *A genome-wide search for linkage to asthma phenotypes in the genetics of asthma international network families: evidence for a major susceptibility locus on chromosome 2p.* Eur J Hum Genet, 2006. 14(3): p. 307-16.

159. Inoue, H. and M. Kubo, *SOCS proteins in T helper cell differentiation: implications for allergic disorders?* Expert Rev Mol Med, 2004. 6(22): p. 1-11.

160. Seki, Y., et al., *Expression of the suppressor of cytokine signaling-5 (SOCS5) negatively regulates IL-4-dependent STAT6 activation and Th2 differentiation.* Proc Natl Acad Sci U S A, 2002. 99(20): p. 13003-8.

161. Nomura, I., et al., *Distinct patterns of gene expression in the skin lesions of atopic dermatitis and psoriasis: a gene microarray analysis.* J Allergy Clin Immunol, 2003. 112(6): p. 1195-202.

162. Schmidt-Weber, C.B., *Gene expression profiling in allergy and asthma.* Chem Immunol Allergy, 2006. 91: p. 188-94.

163. Sugiura, H., et al., *Large-scale DNA microarray analysis of atopic skin lesions shows overexpression of an epidermal differentiation gene cluster in the alternative pathway and lack of protective gene expression in the cornified envelope.* Br J Dermatol, 2005. 152(1): p. 146-9.

164. Wjst, M., et al., *A genome-wide search for linkage to asthma. German Asthma Genetics Group.* Genomics, 1999. 58(1): p. 1-8.

165. Yuyama, N., et al., *Analysis of novel disease-related genes in bronchial asthma.* Cytokine, 2002. 19(6): p. 287-96.

166. Matsuda, A., et al., *Coding SNP in tenascin-C Fn-III-D domain associates with adult asthma.* Hum Mol Genet, 2005. 14(19): p. 2779-86.

167. Morley, M., et al., *Genetic analysis of genome-wide variation in human gene expression.* Nature, 2004. 430(7001): p. 743-7.

168. *A haplotype map of the human genome.* Nature, 2005. 437(7063): p. 1299-320.

169. *Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls.* Nature, 2007. 447(7145): p. 661-78.

170. Frayling, T.M., et al., *A common variant in the FTO gene is associated with body mass index and predisposes to childhood and adult obesity.* Science, 2007. 316(5826): p. 889-94.

171. Gudmundsson, J., et al., *Two variants on chromosome 17 confer prostate cancer risk, and the one in TCF2 protects against type 2 diabetes.* Nat Genet, 2007.

172. Saxena, R., et al., *Genome-wide association analysis identifies loci for type 2 diabetes and triglyceride levels.* Science, 2007. 316(5829): p. 1331-6.

173. Easton, D.F., et al., *Genome-wide association study identifies novel breast cancer susceptibility loci.* Nature, 2007. 447(7148): p. 1087-93.

174. Hunter, D.J., et al., *A genome-wide association study identifies alleles in FGFR2 associated with risk of sporadic postmenopausal breast cancer.* Nat Genet, 2007. 39(7): p. 870-874.

175. Frova, C., *Glutathione transferases in the genomics era: new insights and perspectives.* Biomol Eng, 2006. 23(4): p. 149-69.

176. Ishikawa, T., *The ATP-dependent glutathione S-conjugate export pump.* Trends Biochem Sci, 1992. 17(11): p. 463-8.

177. Rahman, I. and W. MacNee, *Lung glutathione and oxidative stress: implications in cigarette smoke-induced airway disease.* Am J Physiol, 1999. 277(6 Pt 1): p. L1067-88.

178. Lannan, S., et al., *Effect of cigarette smoke and its condensates on alveolar epithelial cell injury in vitro.* Am J Physiol, 1994. 266(1 Pt 1): p. L92-100.

179. Zhao, T., et al., *The role of human glutathione S-transferases hGSTA1-1 and hGSTA2-2 in protection against oxidative stress.* Arch Biochem Biophys, 1999. 367(2): p. 216-24.

180. Yin, Z., et al., *Glutathione S-transferase p elicits protection against H2O2-induced cell death via coordinated regulation of stress kinases.* Cancer Res, 2000. 60(15): p. 4053-7.

181. Rossjohn, J., et al., *Human theta class glutathione transferase: the crystal structure reveals a sulfate-binding pocket within a buried active site.* Structure, 1998. 6(3): p. 309-22.

182. Sheehan, D., et al., *Structure, function and evolution of glutathione transferases: implications for classification of non-mammalian members of an ancient enzyme superfamily.* Biochem J, 2001. 360(Pt 1): p. 1-16.

183. Morel, F., et al., *Gene and protein characterization of the human glutathione S-transferase kappa and evidence for a peroxisomal localization.* J Biol Chem, 2004. 279(16): p. 16246-53.

184. Chelvanayagam, G., M.W. Parker, and P.G. Board, *Fly fishing for GSTs: a unified nomenclature for mammalian and insect glutathione transferases.* Chem Biol Interact, 2001. 133(256).

185. Armstrong, R.N., *Glutathione S-transferases: structure and mechanism of an archetypical detoxication enzyme.* Adv Enzymol Relat Areas Mol Biol, 1994. 69: p. 1-44.

186. Meyer, D.J., *Significance of an unusually low Km for glutathione in glutathione transferases of the alpha, mu and pi classes.* Xenobiotica, 1993. 23(8): p. 823-34.

187. Peters, W.H., et al., *Glutathione and glutathione S-transferases in Barrett's epithelium.* Br J Cancer, 1993. 67(6): p. 1413-7.

188. Raijmakers, M.T., E.A. Steegers, and W.H. Peters, *Glutathione S-transferases and thiol concentrations in embryonic and early fetal tissues.* Hum Reprod, 2001. 16(11): p. 2445-50.

189. Rowe, J.D., E. Nieves, and I. Listowsky, *Subunit diversity and tissue distribution of human glutathione S-transferases: interpretations based on electrospray ionization-MS and peptide sequence-specific antisera.* Biochem J, 1997. 325 (Pt 2): p. 481-6.

190. Beckett, G.J. and J.D. Hayes, *Glutathione S-transferases: biomedical applications.* Adv Clin Chem, 1993. 30: p. 281-380.

191. McIlwain, C.C., D.M. Townsend, and K.D. Tew, *Glutathione S-transferase polymorphisms: cancer incidence and therapy.* Oncogene, 2006. 25(11): p. 1639-48.

192. Cauchi, S., et al., *Haplotype-environment interactions that regulate the human glutathione S-transferase P1 promoter.* Cancer Res, 2006. 66(12): p. 6439-48.

193. Guy, C.A., et al., *Promoter polymorphisms in glutathione-S-transferase genes affect transcription.* Pharmacogenetics, 2004. 14(1): p. 45-51.

194. Sprenger, R., et al., *Characterization of the glutathione S-transferase GSTT1 deletion: discrimination of all genotypes by polymerase chain reaction indicates a trimodular genotype-phenotype correlation.* Pharmacogenetics, 2000. 10(6): p. 557-65.

195. Xu, S., et al., *Characterization of the human class Mu glutathione S-transferase gene cluster and the GSTM1 deletion.* J Biol Chem, 1998. 273(6): p. 3517-27.

196. Widersten, M., E. Holmstrom, and B. Mannervik, *Cysteine residues are not essential for the catalytic activity of human class Mu glutathione transferase M1a-1a.* FEBS Lett, 1991. 293(1-2): p. 156-9.

197. Widersten, M., et al., *Heterologous expression of the allelic variant mu-class glutathione transferases mu and psi.* Biochem J, 1991. 276 (Pt 2): p. 519-24.

198. Seidegard, J., et al., *Hereditary differences in the expression of the human glutathione transferase active on trans-stilbene oxide are due to a gene deletion.* Proc Natl Acad Sci U S A, 1988. 85(19): p. 7293-7.

199. McLellan, R.A., et al., *Characterization of a human glutathione S-transferase mu cluster containing a duplicated GSTM1 gene that causes ultrarapid enzyme activity.* Mol Pharmacol, 1997. 52(6): p. 958-65.

200. Whitbread, A.K., et al., *Characterization of the human Omega class glutathione transferase genes and associated polymorphisms.* Pharmacogenetics, 2003. 13(3): p. 131-44.

201. Strange, R.C. and A.A. Fryer, *The glutathione S-transferases: influence of polymorphism on cancer susceptibility.* IARC Sci Publ, 1999(148): p. 231-49.

202. Strange, R.C., et al., *Glutathione-S-transferase family of enzymes.* Mutat Res, 2001. 482(1-2): p. 21-6.

203. Coggan, M., et al., *Structure and organization of the human theta-class glutathione S-transferase and D-dopachrome tautomerase gene complex.* Biochem J, 1998. 334 (Pt 3): p. 617-23.

204. van Poppel, G., et al., *Increased cytogenetic damage in smokers deficient in glutathione S-transferase isozyme mu.* Carcinogenesis, 1992. 13(2): p. 303-5.

205. Sorensen, M., et al., *Glutathione S-transferase T1 null-genotype is associated with an increased risk of lung cancer.* Int J Cancer, 2004. 110(2): p. 219-24.

206. Wenzlaff, A.S., et al., *GSTM1, GSTT1 and GSTP1 polymorphisms, environmental tobacco smoke exposure and risk of lung cancer among never smokers: a population-based study.* Carcinogenesis, 2005. 26(2): p. 395-401.

207. Ye, Z., et al., *Five glutathione s-transferase gene variants in 23,452 cases of lung cancer and 30,397 controls: meta-analysis of 130 studies.* PLoS Med, 2006. 3(4): p. e91.

208. Tamer, L., et al., *Glutathione S-transferase gene polymorphism as a susceptibility factor in smoking-related coronary artery disease.* Basic Research in Cardiology, 2004. 99(3): p. 223-9.

209. Menegon, A., et al., *Parkinson's disease, pesticides, and glutathione transferase polymorphisms.* Lancet, 1998. 352(9137): p. 1344-6.

210. Bekris, L.M., et al., *Glutathione-s-transferase M1 and T1 polymorphisms and associations with type 1 diabetes age-at-onset.* Autoimmunity, 2005. 38(8): p. 567-75.

211. Masoudi, M., et al., *Genetic polymorphisms of GSTO2, GSTM1, and GSTT1 and risk of gastric cancer.* Mol Biol Rep, 2009. 36(4): p. 781-4.

212. Pongstaporn, W., et al., *Genetic alterations in chromosome 10q24.3 and glutathione S-transferase omega 2 gene polymorphism in ovarian cancer.* J Exp Clin Cancer Res, 2006. 25(1): p. 107-14.

213. Li, Y.J., et al., *Revealing the role of glutathione S-transferase omega in age-at-onset of Alzheimer and Parkinson diseases.* Neurobiol Aging, 2006. 27(8): p. 1087-93.

214. Ozturk, A., et al., *Three SNPs in the GSTO1, GSTO2 and PRSS11 genes on chromosome 10 are not associated with age-at-onset of Alzheimer's disease.* Neurobiol Aging, 2005. 26(8): p. 1161-5.

215. Bolt, H.M. and R. Thier, *Relevance of the deletion polymorphisms of the glutathione S-transferases GSTT1 and GSTM1 in pharmacology and toxicology.* Curr Drug Metab, 2006. 7(6): p. 613-28.

216. Cascorbi, I., *Genetic basis of toxic reactions to drugs and chemicals.* Toxicol Lett, 2006. 162(1): p. 16-28.

217. Hall, A.G., et al., *Expression of mu class glutathione S-transferase correlates with event-free survival in childhood acute lymphoblastic leukemia.* Cancer Res, 1994. 54(20): p. 5251-4.

218. Lo, H.W. and F. Ali-Osman, *Genetic polymorphism and function of glutathione S-transferases in tumor drug resistance.* Curr Opin Pharmacol, 2007. 7(4): p. 367-74.

219. Barker, D.J., et al., *Weight in infancy and death from ischaemic heart disease.* Lancet, 1989. 2(8663): p. 577-80.

220. Stein, C.E., et al., *Fetal growth and coronary heart disease in south India.* Lancet, 1996. 348(9037): p. 1269-73.

221. Brooks, A.A., et al., *Birth weight: nature or nurture?* Early Hum Dev, 1995. 42(1): p. 29-35.

222. Hales, C.N. and D.J. Barker, *The thrifty phenotype hypothesis.* Br Med Bull, 2001. 60: p. 5-20.

223. de Rooij, S.R., et al., *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, 2006. 49(4): p. 637-43.

224. Painter, R.C., et al., *Early onset of coronary artery disease after prenatal exposure to the Dutch famine.* Am J Clin Nutr, 2006. 84(2): p. 322-7; quiz 466-7.

225. Ravelli, A.C., et al., *Obesity at the age of 50 y in men and women exposed to famine prenatally.* Am J Clin Nutr, 1999. 70(5): p. 811-6.

226. Lopuhaa, C.E., et al., *Atopy, lung function, and obstructive airways disease after prenatal exposure to famine.* Thorax, 2000. 55(7): p. 555-61.

227. McMillen, I.C. and J.S. Robinson, *Developmental origins of the metabolic syndrome: prediction, plasticity, and programming.* Physiol Rev, 2005. 85(2): p. 571-633.

228. Gluckman, P.D. and M.A. Hanson, *The developmental origins of the metabolic syndrome.* Trends Endocrinol Metab, 2004. 15(4): p. 183-7.

229. Barker, D.J. and P.M. Clark, *Fetal undernutrition and disease in later life.* Rev Reprod, 1997. 2(2): p. 105-12.

230. Barker, D.J., *Developmental origins of adult health and disease.* J Epidemiol Community Health, 2004. 58(2): p. 114-5.

231. Persson, E. and T. Jansson, *Low birth weight is associated with elevated adult blood pressure in the chronically catheterized guinea-pig.* Acta Physiol Scand, 1992. 145(2): p. 195-6.

232. Kwong, W.Y., et al., *Maternal undernutrition during the preimplantation period of rat development causes blastocyst abnormalities and programming of postnatal hypertension.* Development, 2000. 127(19): p. 4195-202.

233. Langley-Evans, S.C., et al., *Weanling rats exposed to maternal low-protein diets during discrete periods of gestation exhibit differing severity of hypertension.* Clin Sci (Lond), 1996. 91(5): p. 607-15.

234. Ozaki, T., et al., *Dietary restriction in pregnant rats causes gender-related hypertension and vascular dysfunction in offspring.* J Physiol, 2001. 530(Pt 1): p. 141-52.

235. Langley-Evans, S.C., S.J. Welham, and A.A. Jackson, *Fetal exposure to a maternal low protein diet impairs nephrogenesis and promotes hypertension in the rat.* Life Sci, 1999. 64(11): p. 965-74.

236. Payne, J.A., B.T. Alexander, and R.A. Khalil, *Reduced endothelial vascular relaxation in growth-restricted offspring of pregnant rats with reduced uterine perfusion.* Hypertension, 2003. 42(4): p. 768-74.

237. Woods, L.L., et al., *Maternal protein restriction suppresses the newborn renin-angiotensin system and programs adult hypertension in rats.* Pediatr Res, 2001. 49(4): p. 460-7.

238. Steegers-Theunissen, R.P. and E.A. Steegers, *Nutrient-gene interactions in early pregnancy: a vascular hypothesis.* Eur J Obstet Gynecol Reprod Biol, 2003. 106(2): p. 115-7.

239. Bae, S., et al., *Effect of maternal chronic hypoxic exposure during gestation on apoptosis in fetal rat heart.* Am J Physiol Heart Circ Physiol, 2003. 285(3): p. H983-90.

240. Dodic, M., et al., *Altered cardiovascular haemodynamics and baroreceptor-heart rate reflex in adult sheep after prenatal exposure to dexamethasone.* Clin Sci (Lond), 1999. 97(1): p. 103-9.

241. Dodic, M., et al., *Impaired cardiac functional reserve and left ventricular hypertrophy in adult sheep after prenatal dexamethasone exposure.* Circ Res, 2001. 89(7): p. 623-9.

242. Vickers, M.H., et al., *Fetal origins of hyperphagia, obesity, and hypertension and postnatal amplification by hypercaloric nutrition.* Am J Physiol Endocrinol Metab, 2000. 279(1): p. E83-7.

243. Cripps, R.L., et al., *Early life programming of energy balance.* Biochem Soc Trans, 2007. 35(Pt 5): p. 1203-4.

244. Vickers, M.H., et al., *Sedentary behavior during postnatal life is determined by the prenatal environment and exacerbated by postnatal hypercaloric nutrition.* Am J Physiol Regul Integr Comp Physiol, 2003. 285(1): p. R271-3.

245. Davidowa, H., Y. Li, and A. Plagemann, *Altered responses to orexigenic (AGRP, MCH) and anorexigenic (alpha-MSH, CART) neuropeptides of paraventricular hypothalamic neurons in early postnatally overfed rats.* Eur J Neurosci, 2003. 18(3): p. 613-21.

246. Ahima, R.S., D. Prabakaran, and J.S. Flier, *Postnatal leptin surge and regulation of circadian rhythm of leptin by feeding. Implications for energy homeostasis and neuroendocrine function.* J Clin Invest, 1998. 101(5): p. 1020-7.

247. Friedman, J.M. and J.L. Halaas, *Leptin and the regulation of body weight in mammals.* Nature, 1998. 395(6704): p. 763-70.

248. Stocker, C., et al., *Modulation of susceptibility to weight gain and insulin resistance in low birthweight rats by treatment of their mothers with leptin during pregnancy and lactation.* Int J Obes Relat Metab Disord, 2004. 28(1): p. 129-36.

249. Garofano, A., P. Czernichow, and B. Breant, *In utero undernutrition impairs rat beta-cell development.* Diabetologia, 1997. 40(10): p. 1231-4.

250. Holemans, K., et al., *Maternal food restriction in the second half of pregnancy affects vascular function but not blood pressure of rat female offspring.* Br J Nutr, 1999. 81(1): p. 73-9.

251. Wigglesworth, J.S., *Experimental Growth Retardation in the Foetal Rat.* J Pathol Bacteriol, 1964. 88: p. 1-13.

252. Berney, D.M., et al., *The effects of maternal protein deprivation on the fetal rat pancreas: major structural changes and their recuperation.* J Pathol, 1997. 183(1): p. 109-15.

253. Nemeth, A. and G. Langst, *Chromatin higher order structure: opening up chromatin for transcription.* Brief Funct Genomic Proteomic, 2004. 2(4): p. 334-43.

254. Downs, J.A., *Histone H3 K56 acetylation, chromatin assembly, and the DNA damage checkpoint.* DNA Repair (Amst), 2008. 7(12): p. 2020-4.

255. Mosley, A.L. and S. Ozcan, *Glucose regulates insulin gene transcription by hyperacetylation of histone h4.* J Biol Chem, 2003. 278(22): p. 19660-6.

256. Lachner, M. and T. Jenuwein, *The many faces of histone lysine methylation.* Curr Opin Cell Biol, 2002. 14(3): p. 286-98.

257. Wendt, K.D. and A. Shilatifard, *Packing for the germy: the role of histone H4 Ser1 phosphorylation in chromatin compaction and germ cell development.* Genes Dev, 2006. 20(18): p. 2487-91.

258. Lee, J.S., et al., *Histone crosstalk between H2B monoubiquitination and H3 methylation mediated by COMPASS.* Cell, 2007. 131(6): p. 1084-96.

259. Ideraabdullah, F.Y., S. Vigneau, and M.S. Bartolomei, *Genomic imprinting mechanisms in mammals.* Mutat Res, 2008. 647(1-2): p. 77-85.

260. Litonjua, A.A. and D.R. Gold, *Asthma and obesity: common early-life influences in the inception of disease.* J Allergy Clin Immunol, 2008. 121(5): p. 1075-84; quiz 1085-6.

261. Shi, W., S. Bellusci, and D. Warburton, *Lung development and adult lung diseases.* Chest, 2007. 132(2): p. 651-6.

262. Langley-Evans, S.C., D.S. Gardner, and A.A. Jackson, *Maternal protein restriction influences the programming of the rat hypothalamic-pituitary-adrenal axis.* J Nutr, 1996. 126(6): p. 1578-85.

263. Torrens, C., L. Poston, and M.A. Hanson, *Transmission of raised blood pressure and endothelial dysfunction to the F2 generation induced by maternal protein restriction in the F0, in the absence of dietary challenge in the F1 generation.* Br J Nutr, 2008: p. 1-7.

264. Burdge, G.C., et al., *Dietary protein restriction of pregnant rats in the F0 generation induces altered methylation of hepatic gene promoters in the adult male offspring in the F1 and F2 generations.* Br J Nutr, 2007. 97(3): p. 435-9.

265. Ramakers, C., et al., *Assumption-free analysis of quantitative real-time polymerase chain reaction (PCR) data.* Neurosci Lett, 2003. 339(1): p. 62-6.

266. Tellmann, G., *The E-method: a highly accurate technique for gene-expression analysis.* Nature Methods, 2006. (July): p. i-ii.

267. Feuk, L., A.R. Carson, and S.W. Scherer, *Structural variation in the human genome.* Nat Rev Genet, 2006. 7(2): p. 85-97.

268. Levy, S., et al., *The Diploid Genome Sequence of an Individual Human.* PLoS Biol, 2007. 5(10): p. e254.

269. Sharp, A.J., et al., *Segmental duplications and copy-number variation in the human genome.* Am J Hum Genet, 2005. 77(1): p. 78-88.

270. Feuk, L., et al., *Structural variants: changing the landscape of chromosomes and design of disease studies.* Hum Mol Genet, 2006. 15 Spec No 1: p. R57-66.

271. Carson, A.R., et al., *Strategies for the detection of copy number and other structural variants in the human genome.* Hum Genomics, 2006. 2(6): p. 403-14.

272. Bunyan, D.J., et al., *Dosage analysis of cancer predisposition genes by multiplex ligation-dependent probe amplification.* Br J Cancer, 2004. 91(6): p. 1155-9.

273. Markoulatos, P., N. Siafakas, and M. Moncany, *Multiplex polymerase chain reaction: a practical approach.* J Clin Lab Anal, 2002. 16(1): p. 47-51.

274. Brasch-Andersen, C., et al., *Possible gene dosage effect of glutathione-S-transferases on atopic asthma: using real-time PCR for quantification of GSTM1 and GSTT1 gene copy numbers.* Hum Mutat, 2004. 24(3): p. 208-14.

275. Holland, P.M., et al., *Detection of specific polymerase chain reaction product by utilizing the 5'----3' exonuclease activity of Thermus aquaticus DNA polymerase.* Proc Natl Acad Sci U S A, 1991. 88(16): p. 7276-80.

276. Lee, L.G., C.R. Connell, and W. Bloch, *Allelic discrimination by nick-translation PCR with fluorogenic probes.* Nucleic Acids Res, 1993. 21(16): p. 3761-6.

277. Laird, N.M., S. Horvath, and X. Xu, *Implementing a unified approach to family-based tests of association.* Genet Epidemiol, 2000. 19 Suppl 1: p. S36-42.

278. Sasieni, P.D., *From genotypes to genes: doubling the sample size.* Biometrics, 1997. 53(4): p. 1253-61.

279. Tamer, L., et al., *Glutathione-S-transferase gene polymorphisms (GSTT1, GSTM1, GSTP1) as increased risk factors for asthma.* Respirology, 2004. 9(4): p. 493-8.

280. Strange, R.C., et al., *The human glutathione S-transferases: a case-control study of the incidence of the GST1 0 phenotype in patients with adenocarcinoma.* Carcinogenesis, 1991. 12(1): p. 25-8.

281. Zhong, S., et al., *Relationship between the GSTM1 genetic polymorphism and susceptibility to bladder, breast and colon cancer.* Carcinogenesis, 1993. 14(9): p. 1821-4.

282. Lecomte, T., et al., *Glutathione S-transferase P1 polymorphism (Ile105Val) predicts cumulative neuropathy in patients receiving oxaliplatin-based chemotherapy.* Clin Cancer Res, 2006. 12(10): p. 3050-6.

283. Vester, U., et al., *The response to cyclophosphamide in steroid-sensitive nephrotic syndrome is influenced by polymorphic expression of glutathion-S-transferases-M1 and -P1.* Pediatr Nephrol, 2005. 20(4): p. 478-81.

284. Kelly, F.J., *Gluthathione: in defence of the lung.* Food Chem Toxicol, 1999. 37(9-10): p. 963-6.

285. Yao, H., et al., *Redox regulation of lung inflammation: role of NADPH oxidase and NF-kappaB signalling.* Biochem Soc Trans, 2007. 35(Pt 5): p. 1151-5.

286. Kirkham, P. and I. Rahman, *Oxidative stress in asthma and COPD: antioxidants as a therapeutic strategy.* Pharmacol Ther, 2006. 111(2): p. 476-94.

287. Castell, J.V., M.T. Donato, and M.J. Gomez-Lechon, *Metabolism and bioactivation of toxicants in the lung. The in vitro cellular approach.* Exp Toxicol Pathol, 2005. 57 Suppl 1: p. 189-204.

288. Linster, C.L. and E. Van Schaftingen, *Vitamin C. Biosynthesis, recycling and degradation in mammals.* FEBS J, 2007. 274(1): p. 1-22.

289. Burns, J.S., et al., *Low dietary nutrient intakes and respiratory health in adolescents.* Chest, 2007. 132(1): p. 238-45.

290. Wang, X. and P.J. Quinn, *Vitamin E and its function in membranes.* Prog Lipid Res, 1999. 38(4): p. 309-36.

291. Rustow, B., et al., *Type II pneumocytes secrete vitamin E together with surfactant lipids.* Am J Physiol, 1993. 265(2 Pt 1): p. L133-9.

292. Niedernhofer, L.J., et al., *Malondialdehyde, a product of lipid peroxidation, is mutagenic in human cells.* J Biol Chem, 2003. 278(33): p. 31426-33.

293. Del Rio, D., A.J. Stewart, and N. Pellegrini, *A review of recent studies on malondialdehyde as toxic molecule and biological marker of oxidative stress.* Nutr Metab Cardiovasc Dis, 2005. 15(4): p. 316-28.

294. Ohkawa, H., N. Ohishi, and K. Yagi, *Assay for lipid peroxides in animal tissues by thiobarbituric acid reaction.* Anal Biochem, 1979. 95(2): p. 351-8.

295. Yoshida, T. and R.M. Tuder, *Pathobiology of cigarette smoke-induced chronic obstructive pulmonary disease.* Physiol Rev, 2007. 87(3): p. 1047-82.

296. Zheng, M. and G. Storz, *Redox sensing by prokaryotic transcription factors.* Biochem Pharmacol, 2000. 59(1): p. 1-6.

297. Fridovich, I., *Fundamental aspects of reactive oxygen species, or what's the matter with oxygen?* Ann N Y Acad Sci, 1999. 893: p. 13-8.

298. Heeschen, C., et al., *A novel angiogenic pathway mediated by non-neuronal nicotinic acetylcholine receptors.* J Clin Invest, 2002. 110(4): p. 527-36.

299. Carlisle, D.L., et al., *Nicotine signals through muscle-type and neuronal nicotinic acetylcholine receptors in both human bronchial epithelial cells and airway fibroblasts.* Respir Res, 2004. 5: p. 27.

300. Cantin, A.M., et al., *Normal alveolar epithelial lining fluid contains high levels of glutathione.* J Appl Physiol, 1987. 63(1): p. 152-7.

301. Rahman, I., S.R. Yang, and S.K. Biswas, *Current concepts of redox signaling in the lungs.* Antioxid Redox Signal, 2006. 8(3-4): p. 681-9.

302. Antczak, A., et al., *Increased hydrogen peroxide and thiobarbituric acid-reactive products in expired breath condensate of asthmatic patients.* Eur Respir J, 1997. 10(6): p. 1235-41.

303. Krishna, M.T., et al., *Effects of 0.2 ppm ozone on biomarkers of inflammation in bronchoalveolar lavage fluid and bronchial mucosa of healthy subjects.* Eur Respir J, 1998. 11(6): p. 1294-300.

304. Dworski, R., *Oxidant stress in asthma.* Thorax, 2000. 55 Suppl 2: p. S51-3.

305. Sedgwick, J.B., K.M. Geiger, and W.W. Busse, *Superoxide generation by hypodense eosinophils from patients with asthma.* Am Rev Respir Dis, 1990. 142(1): p. 120-5.

306. Hulsmann, A.R., et al., *Oxidative epithelial damage produces hyperresponsiveness of human peripheral airways.* Am J Respir Crit Care Med, 1994. 149(2 Pt 1): p. 519-25.

307. DiFranza, J.R., C.A. Aligne, and M. Weitzman, *Prenatal and postnatal environmental tobacco smoke exposure and children's health.* Pediatrics, 2004. 113(4 Suppl): p. 1007-15.

308. Wang, L. and K.E. Pinkerton, *Detrimental effects of tobacco smoke exposure during development on postnatal lung function and asthma.* Birth Defects Res C Embryo Today, 2008. 84(1): p. 54-60.

309. Gergen, P.J., et al., *The burden of environmental tobacco smoke exposure on the respiratory health of children 2 months through 5 years of age in the United States: Third National Health and Nutrition Examination Survey, 1988 to 1994.* Pediatrics, 1998. 101(2): p. E8.

310. Gilliland, F.D., et al., *Effects of early onset asthma and in utero exposure to maternal smoking on childhood lung function.* Am J Respir Crit Care Med, 2003. 167(6): p. 917-24.

311. Jaakkola, J.J., et al., *Prenatal and postnatal tobacco smoke exposure and respiratory health in Russian children.* Respir Res, 2006. 7: p. 48.

312. Li, Y.F., et al., *Effects of in utero and environmental tobacco smoke exposure on lung function in boys and girls with and without asthma.* Am J Respir Crit Care Med, 2000. 162(6): p. 2097-104.

313. Mannino, D.M., et al., *Health effects related to environmental tobacco smoke exposure in children in the United States: data from the Third National Health and Nutrition Examination Survey.* Arch Pediatr Adolesc Med, 2001. 155(1): p. 36-41.

314. Perera, F.P., et al., *A summary of recent findings on birth outcomes and developmental effects of prenatal ETS, PAH, and pesticide exposures.* Neurotoxicology, 2005. 26(4): p. 573-87.

315. Murray, C.S., et al., *Tobacco smoke exposure, wheeze, and atopy.* Pediatr Pulmonol, 2004. 37(6): p. 492-8.

316. Raherison, C., et al., *In utero and childhood exposure to parental tobacco smoke, and allergies in schoolchildren.* Respir Med, 2007. 101(1): p. 107-17.

317. Kabesch, M., et al., *Glutathione S transferase deficiency and passive smoking increase childhood asthma.* Thorax, 2004. 59(7): p. 569-73.

318. Gilliland, F.D., et al., *Effect of glutathione-S-transferase M1 and P1 genotypes on xenobiotic enhancement of allergic responses: randomised, placebo-controlled crossover study.* Lancet, 2004. 363(9403): p. 119-25.

319. Gilliland, F.D., et al., *Effects of glutathione-S-transferase M1, T1, and P1 on childhood lung function growth.* Am J Respir Crit Care Med, 2002. 166(5): p. 710-6.

320. Saadat, M.A.-L., M., *Genetic Polymorphism of Glutathione S-transferase T1, M1 and Asthma, A Meta-analysis of the Literature.* Pakistan Journal of Biological Sciences, 2007. 10(23): p. 4183-4189.

321. Mak, J.C., et al., *Relationship between glutathione S-transferase gene polymorphisms and enzyme activity in Hong Kong Chinese asthmatics.* Clin Exp Allergy, 2007. 37(8): p. 1150-7.

322. Spiteri, M.A., et al., *Polymorphisms at the glutathione S-transferase, GSTP1 locus: a novel mechanism for susceptibility and development of atopic airway inflammation.* Allergy, 2000. 55 Suppl 61: p. 15-20.

323. Fryer, A.A., et al., *Polymorphism at the glutathione S-transferase GSTP1 locus. A new marker for bronchial hyperresponsiveness and asthma.* Am J Respir Crit Care Med, 2000. 161(5): p. 1437-42.

324. Carroll, W.D., et al., *Effects of glutathione S-transferase M1, T1 and P1 on lung function in asthmatic families.* Clin Exp Allergy, 2005. 35(9): p. 1155-61.

325. Carroll, W.D., et al., *Maternal glutathione S-transferase GSTP1 genotype is a specific predictor of phenotype in children with asthma.* Pediatr Allergy Immunol, 2005. 16(1): p. 32-9.

326. Child, F., et al., *The association of maternal but not paternal genetic variation in GSTP1 with asthma phenotypes in children.* Respir Med, 2003. 97(12): p. 1247-56.

327. Atteslander, P. and B. Schneider, *SAPALDIA Study. Swiss study on Air Pollution and Respiratory Diseases in Adults.* Am J Respir Crit Care Med, 1996. 153(1): p. 462-3.

328. Imboden, M., et al., *Glutathione-S transferase genotype increases risk of progression from bronchial hyperresponsiveness to asthma in adults.* Thorax, 2007.

329. Salam, M.T., et al., *Microsomal epoxide hydrolase, glutathione S-transferase P1, traffic and childhood asthma.* Thorax, 2007. 62(12): p. 1050-7.

330. Mapp, C.E., et al., *Glutathione S-transferase GSTP1 is a susceptibility gene for occupational asthma induced by isocyanates.* J Allergy Clin Immunol, 2002. 109(5): p. 867-72.

331. Palmer, C.N., et al., *Glutathione S-transferase M1 and P1 genotype, passive smoking, and peak expiratory flow in asthma.* Pediatrics, 2006. 118(2): p. 710-6.

332. Romieu, I., et al., *GSTM1 and GSTP1 and respiratory health in asthmatic children exposed to ozone.* Eur Respir J, 2006. 28(5): p. 953-9.

333. Nickel, R., et al., *Association study of Glutathione S-transferase P1 (GSTP1) with asthma and bronchial hyper-responsiveness in two German pediatric populations.* Pediatr Allergy Immunol, 2005. 16(6): p. 539-41.

334. Wenten, M., et al., *In utero smoke exposure, glutathione S-transferase P1 haplotypes, and respiratory illness-related absence among schoolchildren.* Pediatrics, 2009. 123(5): p. 1344-51.

335. Islam, T., et al., *Glutathione-S-transferase (GST) P1, GSTM1, exercise, ozone and asthma incidence in school children.* Thorax, 2009. 64(3): p. 197-202.

336. Li, Y.F., et al., *Glutathione S-transferase P1, maternal smoking, and asthma in children: a haplotype-based analysis.* Environ Health Perspect, 2008. 116(3): p. 409-15.

337. Wang, Y.H., et al., *A significantly joint effect between arsenic and occupational exposures and risk genotypes/diplotypes of CYP2E1, GSTO1 and GSTO2 on risk of urothelial carcinoma.* Toxicol Appl Pharmacol, 2009. 241(1): p. 111-8.

338. Wilk, J.B., et al., *Framingham Heart Study genome-wide association: results for pulmonary function measures.* BMC Med Genet, 2007. 8 Suppl 1: p. S8.

339. Yanbaeva, D.G., et al., *Association of glutathione-S-transferase omega haplotypes with susceptibility to chronic obstructive pulmonary disease.* Free Radic Res, 2009: p. 1-6.

340. Liew, M., et al., *Genotyping of single-nucleotide polymorphisms by high-resolution melting of small amplicons.* Clin Chem, 2004. 50(7): p. 1156-64.

341. Rabinowitz, D. and N. Laird, *A unified approach to adjusting association tests for population admixture with arbitrary pedigree structure and arbitrary missing marker information.* Hum Hered, 2000. 50(4): p. 211-23.

342. Spielman, R.S., R.E. McGinnis, and W.J. Ewens, *Transmission test for linkage disequilibrium: the insulin gene region and insulin-dependent diabetes mellitus (IDDM).* Am J Hum Genet, 1993. 52(3): p. 506-16.

343. Horvath, S., et al., *Family-based association test method: age of onset traits and covariates.* Genet Epidemiol, 2001. 21 Suppl 1: p. S403-8.

344. Horvath, S., et al., *Family-based tests for associating haplotypes with general phenotype data: application to asthma genetics.* Genet Epidemiol, 2004. 26(1): p. 61-9.

345. Lange, C., et al., *PBAT: tools for family-based association studies.* Am J Hum Genet, 2004. 74(2): p. 367-9.

346. Lange, C. and N.M. Laird, *Power calculations for a general class of family-based association tests: dichotomous traits.* Am J Hum Genet, 2002. 71(3): p. 575-84.

347. Saadat, M. and M. Ansari-Lari, *Genetic polymorphism of glutathione S-transferase T1, M1 and asthma, a meta-analysis of the literature.* Pak J Biol Sci, 2007. 10(23): p. 4183-9.

348. Schwartzbaum, J.A., et al., *An international case-control study of glutathione transferase and functionally related polymorphisms and risk of primary adult brain tumors.* Cancer Epidemiol Biomarkers Prev, 2007. 16(3): p. 559-65.

349. Minelli, C., et al., *Glutathione-S-transferase genes and asthma phenotypes: a HuGE systematic review and meta-analysis including unpublished data.* International Journal of Epidemiology, 2009: p. [In press].

350. Whitbread, A.K., et al., *Characterization of the omega class of glutathione transferases.* Methods Enzymol, 2005. 401: p. 78-99.

351. Freidin, M.B., et al., *[Polymorphism of the theta1 and mu1 glutathione s-transferase genes (GSTT1, GSTM1) in patients with atopic bronchial asthma from the West Siberian region].* Mol Biol (Mosk), 2002. 36(4): p. 630-4.

352. Henderson, J., et al., *Effects of glutathione S-transferase polymorphisms on childhood asthma phenotypes and atopy.* European Respiratory Society Annual Congress, 2007. 296s: p. e1850.

353. Henderson, A.J., et al., *Maternal and child glutathione-S-transferase M1 and T1 polymorphisms modify effect of maternal smoking in pregnancy on childhood lung function.* Early Human Development, 2007. 83: p. S97-S97.

354. Shaheen, S.O., et al., *Maternal and child glutathione-S-transferase M1 polymorphism modifies risk of childhood asthma associated with prenatal paracetamol exposure.* Early Human Development, 2007. 83: p. S59-S60.

355. Altman, D., *Practical statistics for medical research.* London: Chapman and Hall, 1992: p. 211-212.

356. Atici, A., et al., *Influence of smoke exposure on serum IgE levels of atopic patients.* Acta Paediatr Jpn, 1994. 36(3): p. 266-7.

357. Frick, O.L., *Effect of respiratory and other virus infections on IgE immunoregulation.* J Allergy Clin Immunol, 1986. 78(5 Pt 2): p. 1013-8.

358. Greene, L.S., *Asthma and oxidant stress: nutritional, environmental, and genetic risk factors.* J Am Coll Nutr, 1995. 14(4): p. 317-24.

359. Rogers, A.J., et al., *The interaction of glutathione S-transferase M1-null variants with tobacco smoke exposure and the development of childhood asthma.* Clin Exp Allergy, 2009. 39(11): p. 1721-9.

360. Myren, M., et al., *The human placenta--an alternative for studying foetal exposure.* Toxicol In Vitro, 2007. 21(7): p. 1332-40.

361. Langley, S.C. and F.J. Kelly, *Depletion of pulmonary glutathione using diethylmaleic acid accelerates the development of oxygen-induced lung injury in term and preterm guinea-pig neonates.* J Pharm Pharmacol, 1994. 46(2): p. 98-102.

362. Dorion, S., H. Lambert, and J. Landry, *Activation of the p38 signaling pathway by heat shock involves the dissociation of glutathione S-transferase Mu from Ask1.* J Biol Chem, 2002. 277(34): p. 30792-7.

363. Ruscoe, J.E., et al., *Pharmacologic or genetic manipulation of glutathione S-transferase P1-1 (GSTpi) influences cell proliferation pathways.* J Pharmacol Exp Ther, 2001. 298(1): p. 339-45.

364. Wang, T., et al., *Glutathione S-transferase P1-1 (GSTP1-1) inhibits c-Jun N-terminal kinase (JNK1) signaling through interaction with the C terminus.* J Biol Chem, 2001. 276(24): p. 20999-1003.

365. Peterson, J.D., et al., *Glutathione levels in antigen-presenting cells modulate Th1 versus Th2 response patterns.* Proc Natl Acad Sci U S A, 1998. 95(6): p. 3071-6.

366. Li, Y.F., et al., *Maternal and grandmaternal smoking patterns are associated with early childhood asthma.* Chest, 2005. 127(4): p. 1232-41.

367. Maritz, G.S., C.J. Morley, and R. Harding, *Early developmental origins of impaired lung structure and function.* Early Hum Dev, 2005. 81(9): p. 763-71.

368. Kiechl-Kohlendorfer, U., et al., *Neonatal characteristics and risk of atopic asthma in schoolchildren: results from a large prospective birth-cohort study.* Acta Paediatr, 2007. 96(11): p. 1606-10.

369. Caudri, D., et al., *Respiratory symptoms in the first 7 years of life and birth weight at term: the PIAMA Birth Cohort.* Am J Respir Crit Care Med, 2007. 175(10): p. 1078-85.

370. Barker, D.J., et al., *Relation of birth weight and childhood respiratory infection to adult lung function and death from chronic obstructive airways disease.* BMJ, 1991. 303(6804): p. 671-5.

371. Syddall, H.E., et al., *Birth weight, infant weight gain, and cause-specific mortality: the Hertfordshire Cohort Study.* Am J Epidemiol, 2005. 161(11): p. 1074-80.

372. Keil, T., et al., *Maternal smoking increases risk of allergic sensitization and wheezing only in children with allergic predisposition: longitudinal analysis from birth to 10 years.* Allergy, 2009. 64(3): p. 445-51.

373. Warner, J., *Developmetnal origins of asthma and related allergic disorders.* Developmental Origins of Health & Disease, 2006. Cambridge University Press: p. p349 - 369.

374. Warner, J.A., et al., *Prenatal origins of allergic disease.* J Allergy Clin Immunol, 2000. 105(2 Pt 2): p. S493-8.

375. Dessaint, J.P. and M. Labalette, *Ontogeny of the immune system with special reference to IgE.* Pediatr Pulmonol Suppl, 1997. 16: p. 3-5.

376. Gill, T.J., 3rd, et al., *Long-term follow-up of children born to women immunized with tetanus toxoid during pregnancy.* Am J Reprod Immunol, 1991. 25(2): p. 69-71.

377. Gitlin, D., et al., *The turnover of amniotic fluid protein in the human conceptus.* Am J Obstet Gynecol, 1972. 113(5): p. 632-45.

378. Ruiz, R.G., D.M. Kemeny, and J.F. Price, *Higher risk of infantile atopic dermatitis from maternal atopy than from paternal atopy.* Clin Exp Allergy, 1992. 22(8): p. 762-6.

379. Platts-Mills, T.A., et al., *The relevance of maternal immune responses to inhalant allergens to maternal symptoms, passive transfer to the infant, and development of antibodies in the first 2 years of life.* J Allergy Clin Immunol, 2003. 111(1): p. 123-30.

380. Yoshinaka, T., et al., *Identification and characterization of novel mouse and human ADAM33s with potential metalloprotease activity.* Gene, 2002. 282(1-2): p. 227-36.

381. Kodama, T., et al., *ADAM12 is selectively overexpressed in human glioblastomas and is associated with glioblastoma cell proliferation and shedding of heparin-binding epidermal growth factor.* Am J Pathol, 2004. 165(5): p. 1743-53.

382. Lafuste, P., et al., *ADAM12 and alpha9beta1 integrin are instrumental in human myogenic cell differentiation.* Mol Biol Cell, 2005. 16(2): p. 861-70.

383. Pan, D. and G.M. Rubin, *Kuzbanian controls proteolytic processing of Notch and mediates lateral inhibition during Drosophila and vertebrate neurogenesis.* Cell, 1997. 90(2): p. 271-80.

384. Rooke, J., et al., *KUZ, a conserved metalloprotease-disintegrin protein with two roles in Drosophila neurogenesis.* Science, 1996. 273(5279): p. 1227-31.

385. Wolfsberg, T.G. and J.M. White, *ADAMs in fertilization and development.* Dev Biol, 1996. 180(2): p. 389-401.

386. Yagami-Hiromasa, T., et al., *A metalloprotease-disintegrin participating in myoblast fusion.* Nature, 1995. 377(6550): p. 652-6.

387. Umland, S.P., et al., *Human ADAM33 messenger RNA expression profile and post-transcriptional regulation.* Am J Respir Cell Mol Biol, 2003. 29(5): p. 571-82.

388. Qiu, Y.M., et al., *[Association between ADAM33 gene polymorphism and bronchial asthma in South China Han population].* Nan Fang Yi Ke Da Xue Xue Bao, 2007. 27(4): p. 485-7.

389. Lind, D.L., et al., *ADAM33 is not associated with asthma in Puerto Rican or Mexican populations.* Am J Respir Crit Care Med, 2003. 168(11): p. 1312-6.

390. Raby, B.A., et al., *ADAM33 polymorphisms and phenotype associations in childhood asthma.* J Allergy Clin Immunol, 2004. 113(6): p. 1071-8.

391. Wang, P., et al., *Lack of association between ADAM33 gene and asthma in a Chinese population.* Int J Immunogenet, 2006. 33(4): p. 303-6.

392. Blakey, J., et al., *Contribution of ADAM33 polymorphisms to the population risk of asthma.* Thorax, 2005. 60(4): p. 274-6.

393. Simpson, A., et al., *Polymorphisms in a disintegrin and metalloprotease 33 (ADAM33) predict impaired early-life lung function.* Am J Respir Crit Care Med, 2005. 172(1): p. 55-60.

394. Haitchi, H.M., et al., *ADAM33 expression in asthmatic airways and human embryonic lungs.* Am J Respir Crit Care Med, 2005. 171(9): p. 958-65.

395. Foley, S.C., et al., *Increased expression of ADAM33 and ADAM8 with disease progression in asthma.* J Allergy Clin Immunol, 2007. 119(4): p. 863-71.

396. Gosman, M.M., et al., *A disintegrin and metalloprotease 33 and chronic obstructive pulmonary disease pathophysiology.* Thorax, 2007. 62(3): p. 242-7.

397. van Diemen, C.C., et al., *A disintegrin and metalloprotease 33 polymorphisms and lung function decline in the general population.* Am J Respir Crit Care Med, 2005. 172(3): p. 329-33.

398. Daniels, S.E., et al., *A genome-wide search for quantitative trait loci underlying asthma.* Nature, 1996. 383(6597): p. 247-50.

399. Palmer, L.J., et al., *Linkage of chromosome 5q and 11q gene markers to asthma-associated quantitative traits in Australian children.* Am J Respir Crit Care Med, 1998. 158(6): p. 1825-30.

400. Edwards, C.A., et al., *Relationship between birth weight and adult lung function: controlling for maternal factors.* Thorax, 2003. 58(12): p. 1061-5.

401. Greenough, A., B. Yuksel, and P. Cheeseman, *Effect of in utero growth retardation on lung function at follow-up of prematurely born infants.* Eur Respir J, 2004. 24(5): p. 731-3.

402. Lucas, J.S., et al., *Small size at birth and greater postnatal weight gain: relationships to diminished infant lung function.* Am J Respir Crit Care Med, 2004. 170(5): p. 534-40.

403. Nikolajev, K., et al., *Effects of intrauterine growth retardation and prematurity on spirometric flow values and lung volumes at school age in twin pairs.* Pediatr Pulmonol, 1998. 25(6): p. 367-70.

404. Rona, R.J., M.C. Gulliford, and S. Chinn, *Effects of prematurity and intrauterine growth on respiratory health and lung function in childhood.* BMJ, 1993. 306(6881): p. 817-20.

405. Stern, D.A., et al., *Poor airway function in early infancy and lung function by age 22 years: a non-selective longitudinal cohort study.* Lancet, 2007. 370(9589): p. 758-64.

406. Canoy, D., et al., *Early growth and adult respiratory function in men and women followed from the fetal period to adulthood.* Thorax, 2007. 62(5): p. 396-402.

407. Harding, R. and S.B. Hooper, *Regulation of lung expansion and lung growth before birth.* J Appl Physiol, 1996. 81(1): p. 209-24.

408. Yakubu, D.P., et al., *Ontogeny and nutritional programming of mitochondrial proteins in the ovine kidney, liver and lung.* Reproduction, 2007. 134(6): p. 823-30.

409. Das, R.M., *The effects of intermittent starvation on lung development in suckling rats.* Am J Pathol, 1984. 117(2): p. 326-32.

410. Maritz, G.S., et al., *Effects of fetal growth restriction on lung development before and after birth: a morphometric analysis.* Pediatr Pulmonol, 2001. 32(3): p. 201-10.

411. Wignarajah, D., et al., *Influence of intrauterine growth restriction on airway development in fetal and postnatal sheep.* Pediatr Res, 2002. 51(6): p. 681-8.

412. Massaro, G.D., L. McCoy, and D. Massaro, *Postnatal undernutrition slows development of bronchiolar epithelium in rats.* Am J Physiol, 1988. 255(4 Pt 2): p. R521-6.

413. Zambrano, E., et al., *Sex differences in transgenerational alterations of growth and metabolism in progeny (F2) of female offspring (F1) of rats fed a low protein diet during pregnancy and lactation.* J Physiol, 2005. 566(Pt 1): p. 225-36.

414. Benyshek, D.C., C.S. Johnston, and J.F. Martin, *Glucose metabolism is altered in the adequately-nourished grand-offspring (F3 generation) of rats malnourished during gestation and perinatal life.* Diabetologia, 2006. 49(5): p. 1117-9.

415. Harrison, M. and S.C. Langley-Evans, *Intergenerational programming of impaired nephrogenesis and hypertension in rats following maternal protein restriction during pregnancy.* Br J Nutr, 2009. 101(7): p. 1020-30.

416. Wu, G., *Amino acids: metabolism, functions, and nutrition.* Amino Acids, 2009. 37(1): p. 1-17.

417. Cetin, I. and G. Alvino, *Intrauterine growth restriction: implications for placental metabolism and transport. A review.* Placenta, 2009. 30 Suppl A: p. S77-82.

418. Rees, W.D., et al., *The effects of maternal protein restriction on the growth of the rat fetus and its amino acid supply.* Br J Nutr, 1999. 81(3): p. 243-50.

419. Jackson, A.A., *The glycine story.* Eur J Clin Nutr, 1991. 45(2): p. 59-65.

420. Desai, M., et al., *Organ-selective growth in the offspring of protein-restricted mothers.* Br J Nutr, 1996. 76(4): p. 591-603.

421. Plopper, C.G. and D.M. Hyde, *The non-human primate as a model for studying COPD and asthma.* Pulm Pharmacol Ther, 2008. 21(5): p. 755-66.

422. Bice, D.E., J. Seagrave, and F.H. Green, *Animal models of asthma: potential usefulness for studying health effects of inhaled particles.* Inhal Toxicol, 2000. 12(9): p. 829-62.

423. Wright, J.L. and A. Churg, *Animal models of cigarette smoke-induced COPD.* Chest, 2002. 122(6 Suppl): p. 301S-306S.

424. Schittny, J.C., S.I. Mund, and M. Stampanoni, *Evidence and structural mechanism for late lung alveolarization.* Am J Physiol Lung Cell Mol Physiol, 2008. 294(2): p. L246-54.

425. Chen, C.M., L.F. Wang, and Y.D. Lang, *Effects of maternal undernutrition on lung growth and insulin-like growth factor system expression in rat offspring.* Acta Paediatr Taiwan, 2007. 48(2): p. 62-7.

426. Lillycrop, K.A., et al., *Dietary protein restriction of pregnant rats induces and folic acid supplementation prevents epigenetic modification of hepatic gene expression in the offspring.* J Nutr, 2005. 135(6): p. 1382-6.

427. Gnanalingham, M.G., et al., *Ontogeny and nutritional programming of uncoupling protein-2 and glucocorticoid receptor mRNA in the ovine lung.* J Physiol, 2005. 565(Pt 1): p. 159-69.

428. Bertram, C., et al., *The maternal diet during pregnancy programs altered expression of the glucocorticoid receptor and type 2 11beta-hydroxysteroid dehydrogenase: potential molecular mechanisms underlying the programming of hypertension in utero.* Endocrinology, 2001. 142(7): p. 2841-53.

429. Godfrey, K.M., et al., *Epigenetic mechanisms and the mismatch concept of the developmental origins of health and disease.* Pediatr Res, 2007. 61(5 Pt 2): p. 5R-10R.

430. Iannello, S., P. Milazzo, and F. Belfiore, *Animal and human tissue Na,K-ATPase in normal and insulin-resistant states: regulation, behaviour and interpretative hypothesis on NEFA effects.* Obes Rev, 2007. 8(3): p. 231-51.

431. Berger, S.A., et al., *Molecular genetic analysis of glucocorticoid and mineralocorticoid signaling in development and physiological processes.* Steroids, 1996. 61(4): p. 236-9.

432. Keller-Wood, M., M. von Reitzenstein, and J. McCartney, *Is the fetal lung a mineralocorticoid receptor target organ? Induction of cortisol-regulated genes in the ovine fetal lung, kidney and small intestine.* Neonatology, 2009. 95(1): p. 47-60.

433. O'Brien, E.A., et al., *Uteroplacental insufficiency decreases p53 serine-15 phosphorylation in term IUGR rat lungs.* Am J Physiol Regul Integr Comp Physiol, 2007. 293(1): p. R314-22.

434. Mattiasson, G. and P.G. Sullivan, *The emerging functions of UCP2 in health, disease, and therapeutics.* Antioxid Redox Signal, 2006. 8(1-2): p. 1-38.

435. D'Adamo, M., et al., *The -866A/A genotype in the promoter of the human uncoupling protein 2 gene is associated with insulin resistance and increased risk of type 2 diabetes.* Diabetes, 2004. 53(7): p. 1905-10.

436. Ochoa, M.C., et al., *Association between obesity and insulin resistance with UCP2-UCP3 gene variants in Spanish children and adolescents.* Mol Genet Metab, 2007. 92(4): p. 351-8.

437. Jahoor, F., et al., *Protein metabolism in severe childhood malnutrition.* Ann Trop Paediatr, 2008. 28(2): p. 87-101.

438. Reen, R.K., G.E. Melo, and T. Moraes-Santos, *Malnutrition sequela on the drug metabolizing enzymes in male Holtzman rats.* J Nutr Biochem, 1999. 10(10): p. 615-8.

439. Yang, Y., et al., *Epigenetic mechanisms silence a disintegrin and metalloprotease 33 expression in bronchial epithelial cells.* J Allergy Clin Immunol, 2008. 121(6): p. 1393-9, 1399 e1-14.

440. Koppelman, G.H., et al., *Identification of PCDH1 as a novel susceptibility gene for bronchial hyperresponsiveness.* Am J Respir Crit Care Med, 2009. 180(10): p. 929-35.

441. Torrens, C., et al., *Maternal protein restriction in the rat impairs resistance artery but not conduit artery function in pregnant offspring.* J Physiol, 2003. 547(Pt 1): p. 77-84.

442. Chen, C.M., L.F. Wang, and B. Su, *Effects of maternal undernutrition during late gestation on the lung surfactant system and morphometry in rats.* Pediatr Res, 2004. 56(3): p. 329-35.

443. Karapolat, S., et al., *Effects of retinoic acid on compensatory lung growth.* J Cardiothorac Surg, 2008. 3: p. 37.

444. Chomczynski, P. and N. Sacchi, *Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction.* Anal Biochem, 1987. 162(1): p. 156-9.

445. Vandesompele, J., et al., *Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes.* Genome Biol, 2002. 3(7): p. RESEARCH0034.

446. Cai, J.H., et al., *Validation of rat reference genes for improved quantitative gene expression analysis using low density arrays.* Biotechniques, 2007. 42(4): p. 503-12.

447. Langley, S.C., R.F. Browne, and A.A. Jackson, *Altered glucose tolerance in rats exposed to maternal low protein diets in utero.* Comp Biochem Physiol Physiol, 1994. 109(2): p. 223-9.

448. Gallagher, S., *Quantitation of nucleic acids with absorption spectroscopy.* Curr Protoc Protein Sci, 2001. Appendix 4: p. Appendix 4K.

449. Nolan, T., R.E. Hands, and S.A. Bustin, *Quantification of mRNA using real-time RT-PCR.* Nat Protoc, 2006. 1(3): p. 1559-82.

450. Bustin, S.A. and T. Nolan, *Pitfalls of quantitative real-time reverse-transcription polymerase chain reaction.* J Biomol Tech, 2004. 15(3): p. 155-66.

451. McCormick, J.A., et al., *5'-heterogeneity of glucocorticoid receptor messenger RNA is tissue specific: differential regulation of variant transcripts by early-life events.* Mol Endocrinol, 2000. 14(4): p. 506-17.

452. Gnanalingham, M.G., et al., *Developmental regulation of the lung in preparation for life after birth: hormonal and nutritional manipulation of local glucocorticoid action and uncoupling protein-2.* J Endocrinol, 2006. 188(3): p. 375-86.

453. Ferrari, P. and Z. Krozowski, *Role of the 11beta-hydroxysteroid dehydrogenase type 2 in blood pressure regulation.* Kidney Int, 2000. 57(4): p. 1374-81.

454. Tomlinson, J.W., et al., *11beta-hydroxysteroid dehydrogenase type 1: a tissue-specific regulator of glucocorticoid response.* Endocr Rev, 2004. 25(5): p. 831-66.

455. Edwards, C.R., et al., *Dysfunction of placental glucocorticoid barrier: link between fetal environment and adult hypertension?* Lancet, 1993. 341(8841): p. 355-7.

456. Orsida, B.E., Z.S. Krozowski, and E.H. Walters, *Clinical relevance of airway 11beta-hydroxysteroid dehydrogenase type II enzyme in asthma.* Am J Respir Crit Care Med, 2002. 165(7): p. 1010-4.

457. Suzuki, S., et al., *Dexamethasone upregulates 11beta-hydroxysteroid dehydrogenase type 2 in BEAS-2B cells.* Am J Respir Crit Care Med, 2003. 167(9): p. 1244-9.

458. Best, R., S.M. Nelson, and B.R. Walker, *Dexamethasone and 11-dehydrodexamethasone as tools to investigate the isozymes of 11 beta-hydroxysteroid dehydrogenase in vitro and in vivo.* J Endocrinol, 1997. 153(1): p. 41-8.

459. Nobuhara, K.K., et al., *Insulin-like growth factor-I gene expression in three models of accelerated lung growth.* J Pediatr Surg, 1998. 33(7): p. 1057-60; discussion 1061.

460. Liu, J.P., et al., *Mice carrying null mutations of the genes encoding insulin-like growth factor I (Igf-1) and type 1 IGF receptor (Igf1r).* Cell, 1993. 75(1): p. 59-72.

461. Powell-Braxton, L., et al., *IGF-I is required for normal embryonic growth in mice.* Genes Dev, 1993. 7(12B): p. 2609-17.

462. Han, R.N., et al., *Insulin-like growth factor-I receptor-mediated vasculogenesis/angiogenesis in human lung development.* Am J Respir Cell Mol Biol, 2003. 28(2): p. 159-69.

463. Morissette, M.C., et al., *Increased p53 Level, Bax/BCL-XL Ratio, and TRAIL Receptors Expression in Human Emphysema.* Am J Respir Crit Care Med, 2008.

464. Pham, T.D., et al., *Uteroplacental insufficiency increases apoptosis and alters p53 gene methylation in the full-term IUGR rat kidney.* Am J Physiol Regul Integr Comp Physiol, 2003. 285(5): p. R962-70.

465. Bai, Y., et al., *Persistent nuclear factor-kappa B activation in Ucp2-/- mice leads to enhanced nitric oxide and inflammatory cytokine production.* J Biol Chem, 2005. 280(19): p. 19062-9.

466. Sethi, S., *Bacterial infection and the pathogenesis of COPD.* Chest, 2000. 117(5 Suppl 1): p. 286S-91S.

467. Hayes, J.D. and R.C. Strange, *Glutathione S-transferase polymorphisms and their biological consequences.* Pharmacology, 2000. 61(3): p. 154-66.

468. Harju, T.H., et al., *Glutathione S-transferase omega in the lung and sputum supernatants of COPD patients.* Respir Res, 2007. 8: p. 48.

469. Chen, C., X. Huang, and D. Sheppard, *ADAM33 is not essential for growth and development and does not modulate allergic asthma in mice.* Mol Cell Biol, 2006. 26(18): p. 6950-6.

470. Puxeddu, I., et al., *The soluble form of a disintegrin and metalloprotease 33 promotes angiogenesis: implications for airway remodeling in asthma.* J Allergy Clin Immunol, 2008. 121(6): p. 1400-6, 1406 e1-4.

471. Stahlberg, A., et al., *Properties of the reverse transcription reaction in mRNA quantification.* Clin Chem, 2004. 50(3): p. 509-15.

472. Smith, E., et al., *Methylation of TIMP3 in esophageal squamous cell carcinoma.* World J Gastroenterol, 2008. 14(2): p. 203-10.

473. Davis, C.D. and S.A. Ross, *Evidence for dietary regulation of microRNA expression in cancer cells.* Nutr Rev, 2008. 66(8): p. 477-82.

474. Chuang, J.C. and P.A. Jones, *Epigenetics and microRNAs.* Pediatr Res, 2007. 61(5 Pt 2): p. 24R-29R.

475. Chatterjee, S. and H. Grosshans, *Active turnover modulates mature microRNA activity in Caenorhabditis elegans.* Nature, 2009. 461(7263): p. 546-9.

476. Breving, K. and A. Esquela-Kerscher, *The complexities of microRNA regulation: mirandering around the rules.* Int J Biochem Cell Biol, 2009.

477. Vasudevan, S., Y. Tong, and J.A. Steitz, *Switching from repression to activation: microRNAs can up-regulate translation.* Science, 2007. 318(5858): p. 1931-4.

478. Place, R.F., et al., *MicroRNA-373 induces expression of genes with complementary promoter sequences.* Proc Natl Acad Sci U S A, 2008. 105(5): p. 1608-13.

479. Kim, V.N. and J.W. Nam, *Genomics of microRNA.* Trends Genet, 2006. 22(3): p. 165-73.

480. Weber, B., et al., *Methylation of human microRNA genes in normal and neoplastic cells.* Cell Cycle, 2007. 6(9): p. 1001-5.

481. Ren, J., et al., *MicroRNA and gene expression patterns in the differentiation of human embryonic stem cells.* J Transl Med, 2009. 7: p. 20.

482. Fabbri, M., et al., *MicroRNA-29 family reverts aberrant methylation in lung cancer by targeting DNA methyltransferases 3A and 3B.* Proc Natl Acad Sci U S A, 2007. 104(40): p. 15805-10.

483. Bowman, R.V., et al., *Epigenomic targets for the treatment of respiratory disease.* Expert Opin Ther Targets, 2009. 13(6): p. 625-40.

484. Coppola, V., R. de Maria, and D. Bonci, *MicroRNAs and prostate cancer.* Endocr Relat Cancer, 2009.

485. Heneghan, H.M., et al., *MicroRNAs as Novel Biomarkers for Breast Cancer.* J Oncol, 2009. 2009: p. 950201.

486. Bandi, N., et al., *miR-15a and miR-16 are implicated in cell cycle regulation in a Rb-dependent manner and are frequently deleted or down-regulated in non-small cell lung cancer.* Cancer Res, 2009. 69(13): p. 5553-9.

487. Rossi, S., et al., *Cancer-associated genomic regions (CAGRs) and noncoding RNAs: bioinformatics and therapeutic implications.* Mamm Genome, 2008. 19(7-8): p. 526-40.

488. Yao, Q., et al., *MicroRNA-21 promotes cell proliferation and down-regulates the expression of programmed cell death 4 (PDCD4) in HeLa cervical carcinoma cells.* Biochem Biophys Res Commun, 2009. 388(3): p. 539-42.

489. Lehmann, U., et al., *Epigenetic inactivation of microRNA gene hsa-mir-9-1 in human breast cancer.* J Pathol, 2008. 214(1): p. 17-24.

490. Lu, J., et al., *MicroRNA expression profiles classify human cancers.* Nature, 2005. 435(7043): p. 834-8.

491. Li, M., et al., *MicroRNAs: control and loss of control in human physiology and disease.* World J Surg, 2009. 33(4): p. 667-84.

492. Yang, W.J., et al., *Dicer is required for embryonic angiogenesis during mouse development.* J Biol Chem, 2005. 280(10): p. 9330-5.

493. Harris, T.A., et al., *MicroRNA-126 regulates endothelial expression of vascular cell adhesion molecule 1.* Proc Natl Acad Sci U S A, 2008. 105(5): p. 1516-21.

494. Tang, X., G. Tang, and S. Ozcan, *Role of microRNAs in diabetes.* Biochim Biophys Acta, 2008. 1779(11): p. 697-701.

495. Herrera, B.M., et al., *MicroRNA-125a is over-expressed in insulin target tissues in a spontaneous rat model of Type 2 Diabetes.* BMC Med Genomics, 2009. 2: p. 54.

496. He, A., et al., *Overexpression of micro ribonucleic acid 29, highly up-regulated in diabetic rats, leads to insulin resistance in 3T3-L1 adipocytes.* Mol Endocrinol, 2007. 21(11): p. 2785-94.

497. Zhao, E., et al., *Obesity and genetics regulate microRNAs in islets, liver, and adipose of diabetic mice.* Mamm Genome, 2009. 20(8): p. 476-85.

498. Pedersen, I.M., et al., *Interferon modulation of cellular microRNAs as an antiviral mechanism.* Nature, 2007. 449(7164): p. 919-22.

499. Wightman, B., I. Ha, and G. Ruvkun, *Posttranscriptional regulation of the heterochronic gene lin-14 by lin-4 mediates temporal pattern formation in C. elegans.* Cell, 1993. 75(5): p. 855-62.

500. Lee, R.C., R.L. Feinbaum, and V. Ambros, *The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14.* Cell, 1993. 75(5): p. 843-54.

501. Bartel, D.P. and C.Z. Chen, *Micromanagers of gene expression: the potentially widespread influence of metazoan microRNAs.* Nat Rev Genet, 2004. 5(5): p. 396-400.

502. Wienholds, E. and R.H. Plasterk, *MicroRNA function in animal development.* FEBS Lett, 2005. 579(26): p. 5911-22.

503. Strauss, W.M., et al., *Nonrestrictive developmental regulation of microRNA gene expression.* Mamm Genome, 2006. 17(8): p. 833-40.

504. Williams, A.E., et al., *Maternally imprinted microRNAs are differentially expressed during mouse and human lung development.* Dev Dyn, 2007. 236(2): p. 572-80.

505. Ventura, A., et al., *Targeted deletion reveals essential and overlapping functions of the miR-17 through 92 family of miRNA clusters.* Cell, 2008. 132(5): p. 875-86.

506. Bhaskaran, M., et al., *MicroRNA-127 modulates fetal lung development.* Physiol Genomics, 2009. 37(3): p. 268-78.

507. Megraw, M., et al., *miRGen: a database for the study of animal microRNA genomic organization and function.* Nucleic Acids Res, 2007. 35(Database issue): p. D149-55.

508. Griffiths-Jones, S., et al., *miRBase: tools for microRNA genomics.* Nucleic Acids Res, 2008. 36(Database issue): p. D154-8.

509. Chiromatzo, A.O., et al., *miRNApath: a database of miRNAs, target genes and metabolic pathways.* Genet Mol Res, 2007. 6(4): p. 859-65.

510. Fan, Y.M., et al., *Expression of sterol regulatory element-binding transcription factor (SREBF) 2 and SREBF cleavage-activating protein (SCAP) in human atheroma and the association of their allelic variants with sudden cardiac death.* Thromb J, 2008. 6: p. 17.

511. Saito, Y., et al., *Epigenetic therapy upregulates the tumor suppressor microRNA-126 and its host gene EGFL7 in human cancer cells.* Biochem Biophys Res Commun, 2009. 379(3): p. 726-31.

512. Sato, F., et al., *Intra-platform repeatability and inter-platform comparability of microRNA microarray technology.* PLoS One, 2009. 4(5): p. e5540.

513. Valoczi, A., et al., *Sensitive and specific detection of microRNAs by northern blot analysis using LNA-modified oligonucleotide probes.* Nucleic Acids Res, 2004. 32(22): p. e175.

514. Musiyenko, A., V. Bitko, and S. Barik, *Ectopic expression of miR-126*, an intronic product of the vascular endothelial EGF-like 7 gene, regulates prostein translation and invasiveness of prostate cancer LNCaP cells.* J Mol Med, 2008. 86(3): p. 313-22.

515. Crawford, M., et al., *MicroRNA-126 inhibits invasion in non-small cell lung carcinoma cell lines.* Biochem Biophys Res Commun, 2008. 373(4): p. 607-12.

516. Fish, J.E., et al., *miR-126 regulates angiogenic signaling and vascular integrity.* Dev Cell, 2008. 15(2): p. 272-84.

517. Wang, S., et al., *The endothelial-specific microRNA miR-126 governs vascular integrity and angiogenesis.* Dev Cell, 2008. 15(2): p. 261-71.

518. Zhang, J., et al., *The cell growth suppressor, mir-126, targets IRS-1.* Biochem Biophys Res Commun, 2008. 377(1): p. 136-40.

519. Zhou, L., et al., *MicroRNAs miR-186 and miR-150 down-regulate expression of the pro-apoptotic purinergic P2X7 receptor by activation of instability sites at the 3'-untranslated region of the gene that decrease steady-state levels of the transcript.* J Biol Chem, 2008. 283(42): p. 28274-86.

520. Vinuesa, C.G., R.J. Rigby, and D. Yu, *Logic and extent of miRNA-mediated control of autoimmune gene expression.* Int Rev Immunol, 2009. 28(3-4): p. 112-38.

521. Wareham, K., et al., *Functional evidence for the expression of P2X1, P2X4 and P2X7 receptors in human lung mast cells.* Br J Pharmacol, 2009. 157(7): p. 1215-24.

522. Chen, Z., et al., *Identification of two novel markers for alveolar epithelial type I and II cells.* Biochem Biophys Res Commun, 2004. 319(3): p. 774-80.

523. Denlinger, L.C., et al., *Attenuated P2X7 pore function as a risk factor for virus-induced loss of asthma control.* Am J Respir Crit Care Med, 2009. 179(4): p. 265-70.

524. Maes, O.C., H. Sarojini, and E. Wang, *Stepwise up-regulation of microRNA expression levels from replicating to reversible and irreversible growth arrest states in WI-38 human fibroblasts.* J Cell Physiol, 2009. 221(1): p. 109-19.

525. Agassandian, M., et al., *Oxysterols trigger ABCA1-mediated basolateral surfactant efflux.* Am J Respir Cell Mol Biol, 2004. 31(2): p. 227-33.

526. McNeish, J., et al., *High density lipoprotein deficiency and foam cell accumulation in mice with targeted disruption of ATP-binding cassette transporter-1.* Proc Natl Acad Sci U S A, 2000. 97(8): p. 4245-50.

527. Wong, J., C.M. Quinn, and A.J. Brown, *SREBP-2 positively regulates transcription of the cholesterol efflux gene, ABCA1, by generating oxysterol ligands for LXR.* Biochem J, 2006. 400(3): p. 485-91.

528. Petersen, O.H., *Ca2+ signalling and Ca2+-activated ion channels in exocrine acinar cells.* Cell Calcium, 2005. 38(3-4): p. 171-200.

529. Pearson, G., et al., *Mitogen-activated protein (MAP) kinase pathways: regulation and physiological functions.* Endocr Rev, 2001. 22(2): p. 153-83.

530. Ito, M., et al., *JSAP1, a novel jun N-terminal protein kinase (JNK)-binding protein that functions as a Scaffold factor in the JNK signaling pathway.* Mol Cell Biol, 1999. 19(11): p. 7539-48.

531. Sharrocks, A.D., *Complexities in ETS-domain transcription factor function and regulation: lessons from the TCF (ternary complex factor) subfamily. The Colworth Medal Lecture.* Biochem Soc Trans, 2002. 30(2): p. 1-9.

532. Lesnefsky, E.J. and C.L. Hoppel, *Oxidative phosphorylation and aging.* Ageing Res Rev, 2006. 5(4): p. 402-33.

533. Lie-Venema, H., et al., *Organ-specific activity of the 5' regulatory region of the glutamine synthetase gene in developing mice.* Eur J Biochem, 1997. 248(3): p. 644-59.

534. Lie-Venema, H., et al., *Regulation of the spatiotemporal pattern of expression of the glutamine synthetase gene.* Prog Nucleic Acid Res Mol Biol, 1998. 61: p. 243-308.

535. Cerbulis, J. and H.M. Farrell, Jr., *Composition of milks of dairy cattle. I. Protein, lactose, and fat contents and distribution of protein fraction.* J Dairy Sci, 1975. 58(6): p. 817-27.

536. Barry, J.M., *The use of glutamine and glutamic acid by the mammary gland for casein synthesis.* Biochem J, 1956. 63(4): p. 669-76.

537. Takaki, T., et al., *Polo-like kinase 1 reaches beyond mitosis--cytokinesis, DNA damage response, and development.* Curr Opin Cell Biol, 2008. 20(6): p. 650-60.

538. Strebhardt, K. and A. Ullrich, *Targeting polo-like kinase 1 for cancer therapy.* Nat Rev Cancer, 2006. 6(4): p. 321-30.

539. Agrawal, A., et al., *Gamma-aminobutyric acid receptor genes and nicotine dependence: evidence for association from a case-control study.* Addiction, 2008. 103(6): p. 1027-38.

540. Brini, M., et al., *A comparative functional analysis of plasma membrane Ca2+ pump isoforms in intact cells.* J Biol Chem, 2003. 278(27): p. 24500-8.

541. Torre, L., et al., *DNA methylation and histone acetylation of rat methionine adenosyltransferase 1A and 2A genes is tissue-specific.* Int J Biochem Cell Biol, 2000. 32(4): p. 397-404.

542. Chiang, P.K., et al., *S-Adenosylmethionine and methylation.* FASEB J, 1996. 10(4): p. 471-80.

543. Schlossmann, J., et al., *Regulation of intracellular calcium by a signalling complex of IRAG, IP3 receptor and cGMP kinase Ibeta.* Nature, 2000. 404(6774): p. 197-201.

544. Mackrill, J.J., et al., *Differential expression and regulation of ryanodine receptor and myo-inositol 1,4,5-trisphosphate receptor Ca2+ release channels in mammalian tissues and cell lines.* Biochem J, 1997. 327 (Pt 1): p. 251-8.

545. Orstavik, S., et al., *Characterization of the human gene encoding the type I alpha and type I beta cGMP-dependent protein kinase (PRKG1).* Genomics, 1997. 42(2): p. 311-8.

546. Ellis, J.H., et al., *GRID: a novel Grb-2-related adapter protein that interacts with the activated T cell costimulatory receptor CD28.* J Immunol, 2000. 164(11): p. 5805-14.

547. Liu, S.K., D.M. Berry, and C.J. McGlade, *The role of Gads in hematopoietic cell signalling.* Oncogene, 2001. 20(44): p. 6284-90.

548. Ibanez-Ventoso, C., et al., *Modulated microRNA expression during adult lifespan in Caenorhabditis elegans.* Aging Cell, 2006. 5(3): p. 235-46.

549. Chen, J.H., et al., *Maternal protein restriction affects postnatal growth and the expression of key proteins involved in lifespan regulation in mice.* PLoS One, 2009. 4(3): p. e4950.

550. Seedorf, U., H. Schulte, and G. Assmann, *Genes, diet and public health.* Genes Nutr, 2007. 2(1): p. 75-80.

551. Manolio, T.A., et al., *Finding the missing heritability of complex diseases.* Nature, 2009. 461(7265): p. 747-53.

552. Gluckman, P.D. and M.A. Hanson, *Living with the past: evolution, development, and patterns of disease.* Science, 2004. 305(5691): p. 1733-6.

553. Barker, D.J., *The developmental origins of chronic adult disease.* Acta Paediatr Suppl, 2004. 93(446): p. 26-33.

554. Gregory, A., et al., *The relationship between anthropometric measurements at birth: asthma and atopy in childhood.* Clin Exp Allergy, 1999. 29(3): p. 330-3.

555. Gilliland, F.D., et al., *Effects of glutathione S-transferase M1, maternal smoking during pregnancy, and environmental tobacco smoke on asthma and wheezing in children.* Am J Respir Crit Care Med, 2002. 166(4): p. 457-63.

556. Pillai, S.G., et al., *A genome-wide association study in chronic obstructive pulmonary disease (COPD): identification of two major susceptibility loci.* PLoS Genet, 2009. 5(3): p. e1000421.

557. Henderson, A.J., et al., *Maternal and child glutathione-S-transferase M1 and T1 polymorphisms modify effect of maternal smoking in pregnancy on childhood lung function.* Early Human Development, 2007. 83: p. S97.

558. Hemmingsen, A., et al., *Simultaneous identification of GSTP1 Ile105-->Val105 and Ala114-->Val114 substitutions using an amplification refractory mutation system polymerase chain reaction assay: studies in patients with asthma.* Respir Res, 2001. 2(4): p. 255-60.

559. Schroer, K.T., et al., *Associations between multiple environmental exposures and Glutathione S-Transferase P1 on persistent wheezing in a birth cohort.* J Pediatr, 2009. 154(3): p. 401-8, 408 e1.

560. Kim, M.S., et al., *Factors for poor prognosis of near-fatal asthma after recovery from a life-threatening asthma attack.* Korean J Intern Med, 2008. 23(4): p. 170-5.

561. Yanagawa, H., et al., *Inhaled steroid therapy and hospitalization for bronchial asthma: trend in Tokushima University Hospital.* J Med Invest, 2003. 50(1-2): p. 72-7.

562. Breton, C.V., et al., *Prenatal tobacco smoke exposure affects global and gene-specific DNA methylation.* Am J Respir Crit Care Med, 2009. 180(5): p. 462-7.

563. de Meer, G., S.A. Reijneveld, and B. Brunekreef, *Wheeze in children: the impact of parental education on atopic and non-atopic symptoms.* Pediatr Allergy Immunol, 2009.

564. Arnedo-Pena, A., et al., *Risk factors and prevalence of asthma in schoolchildren in Castellon (Spain): a cross-sectional study.* Allergol Immunopathol (Madr), 2009. 37(3): p. 135-42.

565. Yiallourides, M., et al., *The differential effects of the timing of maternal nutrient restriction in the ovine placenta on glucocorticoid sensitivity, uncoupling protein 2, peroxisome proliferator-activated receptor-gamma and cell proliferation.* Reproduction, 2009. 138(3): p. 601-8.

566. Devereux, G., *The increase in the prevalence of asthma and allergy: food for thought.* Nat Rev Immunol, 2006. 6(11): p. 869-74.

567. Pattenden, S., et al., *Parental smoking and children's respiratory health: independent effects of prenatal and postnatal exposure.* Tob Control, 2006. 15(4): p. 294-301.

568. Glinsky, G.V., *An SNP-guided microRNA map of fifteen common human disorders identifies a consensus disease phenocode aiming at principal components of the nuclear import pathway.* Cell Cycle, 2008. 7(16): p. 2570-83.

569. Liang, R., D.J. Bates, and E. Wang, *Epigenetic Control of MicroRNA Expression and Aging.* Curr Genomics, 2009. 10(3): p. 184-93.

570. Brock, M., et al., *Interleukin-6 modulates the expression of the bone morphogenic protein receptor type II through a novel STAT3-microRNA cluster 17/92 pathway.* Circ Res, 2009. 104(10): p. 1184-91.

571. Takahashi, H., et al., *Downregulation of type II bone morphogenetic protein receptor in hypoxic pulmonary hypertension.* Am J Physiol Lung Cell Mol Physiol, 2006. 290(3): p. L450-8.

572. Redies, C., et al., *Expression of protocadherin-1 (Pcdh1) during mouse development.* Dev Dyn, 2008. 237(9): p. 2496-505.

573. Rutledge, R.G. and C. Cote, *Mathematics of quantitative kinetic PCR and the application of standard curves.* Nucleic Acids Res, 2003. 31(16): p. e93.

574. Durtschi, J.D., et al., *Evaluation of quantification methods for real-time PCR minor groove binding hybridization probe assays.* Anal Biochem, 2007. 361(1): p. 55-64.

575. Mullis, K., et al., *Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction.* Cold Spring Harb Symp Quant Biol, 1986. 51 Pt 1: p. 263-73.

576. Livak, K.J. and T.D. Schmittgen, *Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method.* Methods, 2001. 25(4): p. 402-8.

577. Pfaffl, M.W., *A new mathematical model for relative quantification in real-time RT-PCR.* Nucleic Acids Res, 2001. 29(9): p. e45.

578. Sagner, G., *Principles, Workflows and Advantages of the New LightCycler Relative Quantification Software.* Biochemica, 2001. 3: p. 15-17.

579. Nordgard, O., et al., *Error propagation in relative real-time reverse transcription polymerase chain reaction quantification models: the balance between accuracy and precision.* Anal Biochem, 2006. 356(2): p. 182-93.

9 Website Bibliography

http://frodo.wi.mit.edu/primer3

http://ihg.gsf.de/cgi-bin/hw/hwa1.pl

http://lgmb.fmrp.usp.br/mirnapath/query_mirna.php

www.appliedbiosystems.com

www.asthmauk.org.uk

www.biostat.harvard.edu/~fbat/default.html

www.brit-thoracic.org.uk

www.coriell.org/

www.diana.pcbi.upenn.edu/miRGen/v3/miRGen.html

www.ecacc.org.uk

www.genome.jp/kegg/pathway.html

www.insightful.com/products/splus/default

www.mirbase.org

www.ncbi.nlm.nih.gov/blast/Blast.cgi).

www.ncbi.nlm.nih.gov/projects/SNP/

www.ncbi.nlm.nih.gov/SNP/snp_blastByOrg.cgi

www.roche-applied-science.com

www.roche-applied-science.com

www.roche-applied-science.com/sis/rtpcr/upl/index.jsp)

www.spss.com

10 Appendices

10.1 Reagent Preparations for Total RNA, DNA and Protein Isolation Using TRIzol™ Reagent:

DEPC treated water preparation:

Draw 18mΩ water into an RNase-free glass nalgene bottle Add diethylpyrocarbonate (DEPC) to 0.01% v/v. Let stand overnight and autoclave.

75% Ethanol preparation (in DEPC treated water for RNA isolation):

Add 75 mls of absolute ethanol to 25 mls of DEPC treated water in a sterile Nalgene glass bottle and mix.

75% Ethanol with glycogen preparation (in DEPC treated water for miRNA enrichment):

Add 75 mls of absolute ethanol to 25 mls of DEPC treated water in a sterile Nalgene glass bottle and mix.

75% Ethanol preparation (For DNA isolation):

Add 75 mls of absolute ethanol to 25 mls of Ultra high quality water (UHQ water) in a sterile Nalgene glass bottle and mix.

10% Ethanol preparation (For DNA isolation):

Add 10 mls of absolute ethanol to 90 mls of UHQ water in a sterile Nalgene glass bottle and mix.

0.1M Sodium Citrate in 10% Ethanol preparation (For DNA isolation):

Weigh out 1.29 grams of Sodium Citrate and add to 50mls of 10% Ethanol in a sterile Nalgene glass bottle and mix.

8mM NaOH preparation (For DNA isolation):

Add 1.0 gram of NaOH to 100 mls of UHQ water in a Nalgene glass bottle and mix to make a 0.25M NaOH solution. Add 3.2 mls of 0.25M NaOH to 96.8 mls of UHQ water in a glass Nalgene bottle and mix.

0.3M Guanidine hydrochloride in 95% ethanol preparation (For Protein isolation):

NB: Perform solution formulation in a fume hood.

Add 14.33 grams of Guanidine hydrochloride to 475 mls of absolute ethanol and 25 mls of UHQ water in a Nalgene glass bottle and mix; this can be stored at 4°C for up to one month.

1% SDS solution preparation (For Protein isolation):

Add 1gram of SDS to 100 mls of UHQ water in a Nalgene glass bottle and mix.

10.2 Robotic Liquid Handling Protocols using the Biomek 3000 Platform

The Biomek 3000 automated liquid handling system maintains DNA sample integrity and facilitates high-throughput genotyping.

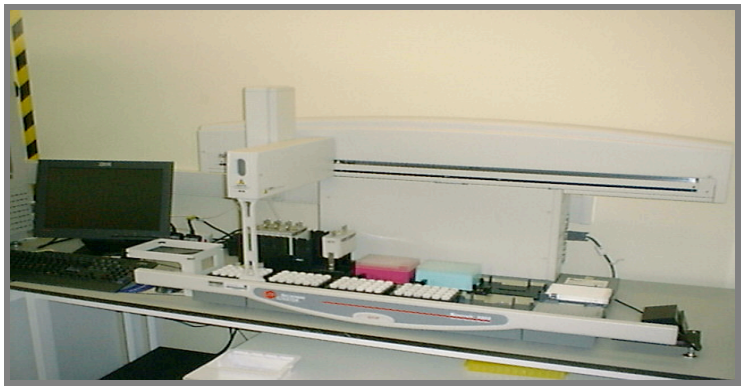


Figure 10-1: Biomek 3000 Robot in the Asthma Genetics Laboratory

Plasticware and Pipette tools

In this project the following plastic-ware was used to store DNA solutions:

Table 10-1: Biomek 3000 Plasticware for DNA Arrays

DNA Array type	Plastic-ware
Working DNA array (10-1 ng/µl)	96 well V-bottomed storage plates (ABgene)
Stock DNA array (100 ng/µl)	96 well deep well plates (ABgene)

To transfer DNA solutions both the 8 channel 200 µl pipette tool (MP200) and the single channel 200 µl liquid sensing pipette tool (P200L) was used and pre-calibrated to ensure accurate volume transfer. Beckman-coulter P250 and P50 robotic tips were used at all times.

DNA Array Dilutions: This robotic program diluted the 100 ng/µl 96 well DNA stock arrays for the GTC cohort, 1 in 10 with water to 10 ng/µl 96 well DNA working arrays using the MP200 pipette tool, p250 and p50 tips. These working arrays could then be used to create the 1 ng/µl 96 well DNA arrays used to supply the DNA template for the GST CNV genotyping assay qPCR reactions.

DNA Working Arrays to 384 Well Reaction Plates: Using this program the Biomek 3000 would transfer 2 µl of the 1 ng/µl 96 well DNA working arrays for the CNV genotyping assays by multi-dispensing into adjacent wells on the 384 well plates. This method used the MP200 pipette tool with p50 tips. By using this robotic method 192 DNA samples could be accurately transferred to the 384 plates without error.

10.3 Use of the Lightcycler480™ Real-Time PCR Machine

All qPCR relative expression experiments were performed on the Lightcycler480™ Real-time PCR machine (Roche Diagnostics, UK). This platform is a third generation real-time PCR machine that has a flexible 96 or the higher-throughput 384 well plate format. The machines unique thermal control and superior optical systems provide highly accurate and reproducible results. The analysis software modules described below allow for the complete analysis of real-time data from copy number variation and mRNA relative expression experiments.

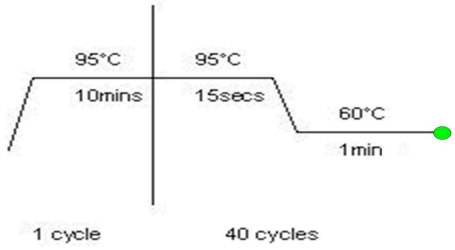
Consumables

In order to generate reproducible data it is essential to use especially designed plastic-ware and plate seals for qPCR applications. In accordance with the manufacturers instructions all experiments were carried out using Roche lightcycler 384 well reaction plates and adhesive plate seals (Roche Diagnostics, UK).

Real-Time PCR Runs

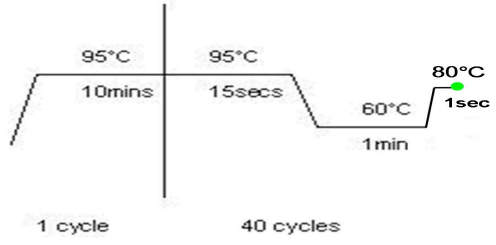
All qPCR relative expression experiments were performed on the Lightcycler480™ Real-time PCR machine (Roche Diagnostics, UK). The TaqMan probe and UPL probe based qPCR assays were thermal-cycled and fluorescence data collected according to the following program:

Figure 10-2: Thermalcycling Conditions for TaqMan and UPL Based qPCR Assays



All SYBRgreen based qPCR assays were ran according to the following program:

Figure 10-3: Thermalcycling Conditions for SYBRgreen Based qPCR Assays



Fluorescent Data was collected at the end of the extension step on a 1 second 80°C hold to ensure that any primer-dimer product was single stranded and therefore did not contribute to any fluorescent signal.

Absolute Quantification Data Analysis

DNA or cDNA templates can be quantified by real-time PCR via a standard curve of a control nucleic acid over a large dynamic range (generally up to 8 orders of magnitude)[573] Crossing point values (Cp values) are determined by a computer controlled algorithm using the second-derivative curve method this produces very reproducible results as Cp values are not user-defined.[574] The analysis module plots the sample Cp values against the standard curve and calculates the concentration via extrapolation against the regression line.

Relative Expression Data Analysis

This analysis module utilises the E-method to generate PCR efficiency adjusted relative expression results for RNA expression experiments.[266] Using standard curve adjusted concentration values for the target and reference genes this module will calculate normalised concentration values relative to the calibrator sample. This allows for comparison of expression results between different target and reference genes and controls for inter-assay variation. The module calculates the crossing point values of the sample target and reference gene assays using the second-derivative curve method[574] and plots them against their respective standard curves, whereby adjusting the concentration values for the PCR efficiency of each assay. Concentration ratios and standard deviations are calculated by the software using equation 10.

Equation 10: Concentration Ratio Formula

$$ConcRatio = \frac{E_{target}^{Cp}}{E_{reference}^{Cp}}$$

Normalised concentration ratios and standard deviations are calculated by dividing the sample concentration ratio by the calibrator sample concentration ratio, see equation 11 below.

Equation 11: Normalised Ratio Formula

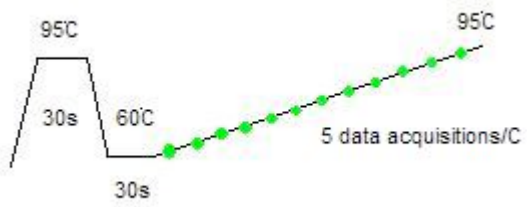
$$NormalisedRatio = \frac{ConcRatio_{sample}}{ConcRatio_{calibrator}}$$

Concentration and Normalised ratio values calculated by the relative quantification software module can be exported for further analysis in Microsoft® excel.

Post PCR Melting Curve Analysis

Melting temperature analysis is performed at the end of the PCR thermalcycling program in SYBRgreen qPCR experiments and monitors the decrease in SYBRgreen fluorescence as the PCR product is heated and denatured (figure 10.4). Melting curve analysis is used to determine that specificity of the SYBRgreen reaction is maintained at all starting DNA template concentrations as the SYBRgreen dye will bind to any double-stranded DNA molecules. A PCR amplicon will melt at a specific temperature; this is the melting temperature (Tm), the point at which 50% of all of the PCR molecules are single-stranded. Therefore a specific qPCR assay will only have one melting temperature. This type of analysis is generally used to determine the specificity of the primers (i.e. one melting curve at the correct Tm) and the presence/absence of primer dimer accumulation at a lower Tm that could cause the results to be inaccurate.

Figure 10-4: SYBRgreen Thermal Melt Program



Colour Compensation Analysis

A colour compensation analysis was performed in accordance with the LightCyler480 operator's manual for qPCR multiplex experiments that used reporter dyes with overlapping emission spectra. Only one colour compensation analysis was required to correct for spectral overlap between the FAM and Yakima Yellow reporter dyes used in the copy number variation experiments in chapters 3 and 4.

10.4 Real-Time PCR Data Analysis

The real-time PCR technique is a well established sensitive fluorescence-based PCR method for quantifying nucleic acids. This technique is primarily used to quantify gene expression (mRNA) by normalising the target gene concentration to a stably expressed reference gene (also known as a 'Housekeeper Gene'). There are several mathematical equations available for calculating normalised gene expression values and all involve large amounts of data analysis after the PCR run. It is important for the researcher to understand any limitations of the mathematical techniques and correctly handle error propagation to produce a valid result data set for analysis.

PCR Efficiency Corrections

PCR is an enzyme driven reaction for amplifying DNA,[575] this reaction is not always 100% efficient, i.e. 1 cycle of PCR does not always result in a doubling of PCR product. The PCR reaction can be described by the following equation:

Equation 12: PCR Reaction Calculation.

Where: X_n = number of DNA molecules at cycle n. n = PCR cycle number. X_o = initial number of molecules. E_x = efficiency value (Probability value, 0-1.0)

$$X_n = X_o(1 + E_x)^n$$

It is impossible to know for certain that there has been a doubling of PCR product at each cycle. Therefore, the efficiency value is a measure of probability that estimates the reaction efficiency over the entire PCR reaction. Knowing the efficiency of each PCR reaction is important for accurate quantification, especially when performing gene expression analysis. If the reactions are amplifying at different rates then they cannot be directly compared without adjusting for the differences in PCR efficiency first.

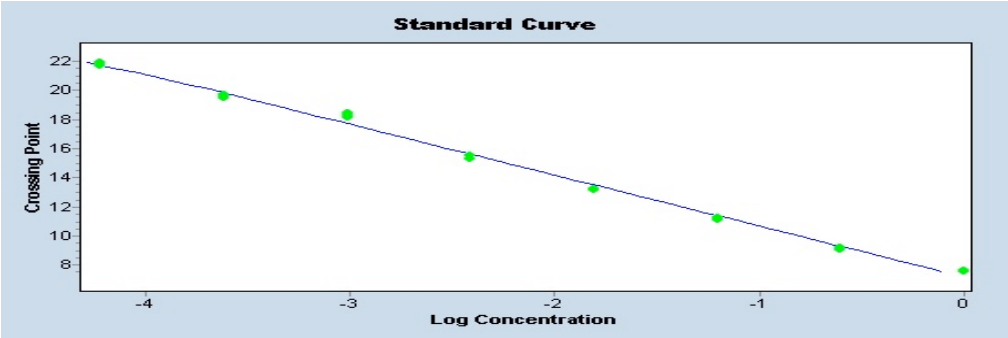
PCR Efficiency Estimation Methods

The efficiency value can be estimated in various ways; the two most common methods are:

1. From the slope of a standard curve (serial dilution of test DNA or cDNA)

Figure 10-5: Standard Curve PCR Efficiency (E) Estimation.

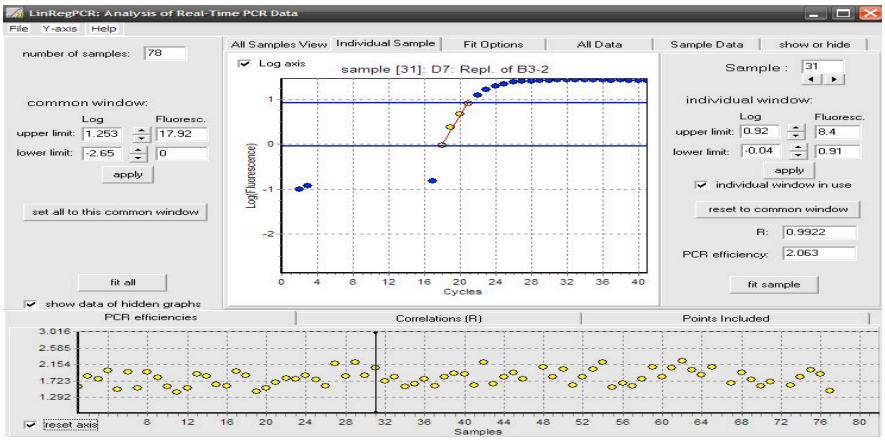
Calculated from an 8 step 4-fold serial dilution of cDNA template; where E = 10 1/s (s is the slope of the standard curve). The PCR efficiency for this assay is 93.3±0.005%.



2. Linear Regression of Fluorescent amplification curve data (LinRegPCR).[265]

Figure 10-6: LinRegPCR Program Screenshot.

A linear regression line is fitted to the exponential part of the amplification plot and the PCR efficiency value is estimated from the slope of the regression line. The PCR efficiency estimate for the PCR assay in this reaction well is 106%, indicating the possible presence of amplified primer-dimer products



Either method produces valid estimates of PCR efficiency and has benefits and disadvantages over one another as listed in table 10.2

Table 10-2: Advantages and Disadvantages of the Different PCR Efficiency Estimation Methods

Efficiency Estimation Method	Advantages	Disadvantages
Standard Curve	Accurate estimate of PCR Efficiency Provides information on dynamic range of assay Sample concentrations can be calculated directly from the curve Included in many Real-Time PCR machine analysis software packages.	Expensive in both reagents and DNA template Uses up space on the reaction plate Standard curve is required for each assay on the plate – time consuming Does not provide efficiency value for every DNA sample Hard to generate a sufficient standard curve for low abundance targets
Linear Regression	Standard curve not required. No added reagent costs Efficiency value for every well on the reaction plate	Intensive external post PCR run data analysis No information on dynamic range of assay Accuracy of PCR efficiency value can dependant on quality of PCR amplification curve data.

Relative Expression Analysis Using Multiple Reference Genes

There are several mathematical equations available to perform relative quantification analyses (target gene versus reference gene).[576-578] The two most widely used methods are as follows:

1. *Comparative Ct ($2^{-\Delta\Delta Ct}$) method:* assumption based relative quantification; the assumption is that the PCR efficiencies are equal for both the target and reference genes in all samples.[576]
2. *Relative standard curve:* Standard curve based relative quantification; adjusts for differences in PCR efficiency of the target and reference gene assays; PCR efficiencies of all samples are not assumed to be equal.[578]

Recent trends in real-time PCR relative quantification practise involves the use of multiple reference genes to control for any variation in reference gene concentrations between sample treatment groups. Thus, improving the accuracy and reproducibility of results.[445] Reference gene combinations (pairs, trios or more) are chosen on the basis of a pair-wise stability index in both the control and treatment sample groups. The geometric mean of the reference gene concentrations can be used as a 'normalisation factor' in relative quantification calculations described earlier (section 1.6.3.3).

There will still be some variation in the concentration of the chosen reference genes between the treatment and control groups, this variation (V = geometric mean of concentration of all possible reference gene combinations with in a group) can be used to determine the 'minimal detectable fold difference' of target gene between the two sample groups ($V_{treatment\ group}$ / $V_{control\ group}$).

Error Propagation in qPCR Relative Expression Analyses

In order to have confidence in real-time PCR quantification results it is essential to determine measurement error, reproducibility (intra/inter-variability) and comparability between different PCR runs by using calibrator samples.[579] Normalisation of all concentration values to a calibrator sample is an efficient way to compare data between runs and measure inter-run variability. Duplicate samples (these can be extraction, RT duplicates and/or simple replicate sample well groupings) are used to determine intra-assay variability accounting for the contribution of all experimental, RT-onwards or pipetting and instrument steps respectively. The way the standard deviation of the result for each well grouping is calculated (Equations 13 & 14) is dependant on the relative expression calculation method used:

Equation 13: Comparative Ct Method Standard Deviation.

Where S_1 = standard deviation of target values, S_2 = standard deviation of reference values.

$$s = \sqrt{{s_1}^2 + {s_2}^2}$$

Equation 14: Relative Standard Curve Standard Deviation.

Where CV_1 = standard deviation of target value/mean target value and CV_2 = standard deviation of reference value/mean reference value.

$$s = \sqrt{{cv_1}^2 + {cv_2}^2}$$

10.5 Full FBAT Result Outputs

Which genes are responsible for asthma?

Marker	No. of Copies	Allelic Frequency	No. of Informative Families	S	E(S)	Var(S)	Z	P
GSTM1	1	0.262	199	252	260.8	110.8	-0.839	0.401
GSTM1	0	0.738	199	508	499.2	110.8	0.839	0.401
GSTT1	1	0.573	224	450	468.7	146.2	-1.544	0.123
GSTT1	0	0.427	224	414	395.3	146.2	1.544	0.123

Results from Association Analysis Involving the Markers and the Diagnosis of Asthma

Marker	No. of Copies	Allelic Frequency	No. of Informative Families	S	E(S)	Var(S)	Z	P
GSTM1	1	0.262	208	496.8	527.8	545.4	-1.327	0.185
GSTM1	0	0.738	208	981.1	950.1	545.4	1.327	0.185
GSTT1	1	0.571	229	870.1	901.7	770.7	-1.136	0.256
GSTT1	0	0.429	229	817.5	785.9	770.7	1.136	0.256

Results from Association Analysis Involving the Markers and the log Age-Corrected Total IgE Variable, Involving the Asthmatic Siblings

Marker	No. of Copies	Allelic Frequency	No. of Informative Families	S	E(S)	Var(S)	Z	P
GSTM1	1	0.262	206	3726.1	3643.7	31267.2	0.467	0.641
GSTM1	0	0.738	206	6099.2	6181.7	31267.2	-0.467	0.641
GSTT1	1	0.571	227	5693.0	6021.7	41501.4	-1.614	0.107
GSTT1	0	0.429	227	5195.2	4866.4	41501.4	1.614	0.107

Results from Association Analysis Involving the Markers and the Measure of Bronchial Responsiveness, Involving the Asthmatic Siblings

Marker	No. of Copies	Allelic Frequency	No. of Informative Families	S	E(S)	Var(S)	Z	P
GSTM1	1	0.262	209	25933	26876.5	1003691.8	-0.942	0.346
GSTM1	0	0.738	209	50851	49907.5	1003691.8	0.942	0.346
GSTT1	1	0.571	230	44261	46361.0	1325268.0	-1.824	0.068
GSTT1	0	0.429	230	41147	39047.0	1325268.0	1.824	0.068

Results from Association Analysis Involving the Markers and Percentage Predicted FEV1 Values, Involving the Asthmatic Siblings

Marker	No. of Copies	Allelic Frequency	No. of Informative Families	S	E(S)	Var(S)	Z	P
GSTM1	1	0.262	187	317.0	336.4	272.1	-1.180	0.238
GSTM1	0	0.738	187	617.6	598.1	272.1	1.180	0.238
GSTT1	1	0.571	207	600.6	620.1	420.8	-0.951	0.341
GSTT1	0	0.429	207	531.9	512.4	420.8	0.951	0.341

Results from Association Analysis Involving the Markers and Atopy Severity Scores, Involving the Asthmatic Siblings

Marker	No. of Copies	Allelic Frequency	No. of Informative Families	S	E(S)	Var(S)	Z	P
GSTM1	1	0.262	201	1027.4	1076.8	1873.3	-1.142	0.254
GSTM1	0	0.738	201	2059.1	2009.7	1873.3	1.142	0.254
GSTT1	1	0.571	217	1705.1	1823.5	2351.4	-2.441	0.015
GSTT1	0	0.429	217	1631.7	1513.4	2351.4	2.441	0.015

Results from Association Analysis Involving the Markers and Asthma Severity Scores, Involving the Asthmatic Siblings

Does exposure to smoke affect this? The Smoke Exposed Children

Marker	No. of Copies	Allelic Frequency	No. of Informative Families	S	E(S)	Var(S)	Z	P
GSTM1	1	0.262	69	165.7	171.8	145.3	-0.506	0.613
GSTM1	0	0.738	69	271.0	264.9	145.3	0.506	0.613
GSTT1	1	0.571	81	258.3	289.6	234.1	-2.048	0.041
GSTT1	0	0.429	81	287.8	256.5	234.1	2.048	0.041

Results from Association Analysis Involving the Markers and the log Age-Corrected Total IgE Variable, In Exposed Asthmatic Siblings

Marker	No. of Copies	Allelic Frequency	No. of Informative Families	S	E(S)	Var(S)	Z	P
GSTM1	1	0.262	68	1302.1	1236.2	10220.7	0.651	0.515
GSTM1	0	0.738	68	1947.5	2013.4	10220.7	-0.651	0.515
GSTT1	1	0.571	80	1573.6	1685.8	13562.7	-0.963	0.336
GSTT1	0	0.429	80	1901.2	1789.0	13562.7	0.963	0.336

Results from Association Analysis Involving the Markers and the Measure of Bronchial Responsiveness, In Exposed Asthmatic Siblings

Marker	No. of Copies	Allelic Frequency	No. of Informative Families	S	E(S)	Var(S)	Z	P
GSTM1	1	0.262	69	8551	8444.5	291184.3	0.197	0.844
GSTM1	0	0.738	69	14537	14643.5	291184.3	-0.197	0.844
GSTT1	1	0.571	81	13320	14231.0	432673.5	-1.385	0.166
GSTT1	0	0.429	81	14560	13649.0	432673.5	1.385	0.166

Results from Association Analysis Involving the Markers and Percentage Predicted FEV1 Values, In Exposed Asthmatic Siblings

Marker	No. of Copies	Allelic Frequency	No. of Informative Families	S	E(S)	Var(S)	Z	P
GSTM1	1	0.262	57	101.5	106.8	78.6	-0.604	0.546
GSTM1	0	0.738	57	180.7	175.3	78.6	0.604	0.546
GSTT1	1	0.571	68	186.6	204.5	136.7	-1.533	0.125
GSTT1	0	0.429	68	188.0	170.0	136.7	1.533	0.125

Results from Association Analysis Involving the Markers and Atopy Severity Scores, In Exposed Asthmatic Siblings

Marker	No. of Copies	Allelic Frequency	No. of Informative Families	S	E(S)	Var(S)	Z	P
GSTM1	1	0.262	64	316.4	321.1	490.3	-0.211	0.833
GSTM1	0	0.738	64	548.5	543.8	490.3	0.211	0.833
GSTT1	1	0.571	76	535.5	584.4	807.5	-1.719	0.086
GSTT1	0	0.429	76	596.4	547.5	807.5	1.719	0.086

Results from Association Analysis Involving the Markers and Asthma Severity Scores, In Exposed Asthmatic Siblings

Unexposed

Marker	No. of Copies	Allelic Frequency	No. of Informative Families	S	E(S)	Var(S)	Z	P
GSTM1	1	0.262	150	330.9	355.8	404.6	-1.238	0.216
GSTM1	0	0.738	150	709.5	684.6	404.6	1.238	0.216
GSTT1	1	0.571	163	623.8	624.0	544.6	-0.009	0.993
GSTT1	0	0.429	163	533.7	533.5	544.6	0.009	0.993

Results from Association Analysis Involving the Markers and the log Age-Corrected Total IgE Variable, In Unexposed Asthmatic Siblings

Marker	No. of Copies	Allelic Frequency	No. of Informative Families	S	E(S)	Var(S)	Z	P
GSTM1	1	0.262	147	2024.1	2407.5	21080.8	0.115	0.909
GSTM1	0	0.738	147	4151.7	4168.3	21080.8	-0.115	0.909
GSTT1	1	0.571	159	4119.3	4335.9	27944.1	-1.296	0.195
GSTT1	0	0.429	159	3294.0	3077.3	27944.1	1.296	0.195

Results from Association Analysis Involving the Markers and the Measure of Bronchial Responsiveness, In Unexposed Asthmatic Siblings

Marker	No. of Copies	Allelic Frequency	No. of Informative Families	S	E(S)	Var(S)	Z	P
GSTM1	1	0.262	150	17382	18432	723058.5	-1.235	0.217
GSTM1	0	0.738	150	36314	35264	723058.5	1.235	0.217
GSTT1	1	0.571	163	30941	32130	899113.5	-1.254	0.210
GSTT1	0	0.429	163	26587	25398	899113.5	1.254	0.210

Results from Association Analysis Involving the Markers and Percentage Predicted FEV1 Values, In Unexposed Asthmatic Siblings

Marker	No. of Copies	Allelic Frequency	No. of Informative Families	S	E(S)	Var(S)	Z	P
GSTM1	1	0.262	137	215.5	229.6	195.1	-1.010	0.312
GSTM1	0	0.738	137	436.9	422.8	195.1	1.010	0.312
GSTT1	1	0.571	148	414.0	415.5	288.0	-0.093	0.926
GSTT1	0	0.429	148	344.0	342.3	288.0	0.093	0.926

Results from Association Analysis Involving the Markers and Atopy Severity Scores, In Unexposed Asthmatic Siblings

Marker	No. of Copies	Allelic Frequency	No. of Informative Families	S	E(S)	Var(S)	Z	P
GSTM1	1	0.262	145	711.0	755.7	1393.3	-1.199	0.231
GSTM1	0	0.738	145	1510.7	1465.9	1393.3	1.199	0.231
GSTT1	1	0.571	153	1169.5	1239.1	1558.1	-1.761	0.078
GSTT1	0	0.429	153	1053.4	965.9	1558.1	1.761	0.078

Results from Association Analysis Involving the Markers and Asthma Severity Scores, In Unexposed Asthmatic Siblings

Using Residuals

Marker	No. of Copies	Allelic Frequency	No. of Informative Families	S	E(S)	Var(S)	Z	P
GSTM1	1	0.262	209	-6.6	3.6	174.5	-0.769	0.442
GSTM1	0	0.738	209	-14.1	-24.3	174.5	0.769	0.442
GSTT1	1	0.571	230	5.8	4.8	246.0	0.063	0.950
GSTT1	0	0.429	230	25.1	26.1	246.0	-0.063	0.950

Results from Association Analysis Involving the Markers and the log Age-Corrected Total IgE Variable Involving the Asthmatic Siblings, using Smoking As a Covariate

Marker	No. of Copies	Allelic Frequency	No. of Informative Families	S	E(S)	Var(S)	Z	P
GSTM1	1	0.262	206	303.9	103.1	13279.5	1.743	0.081
GSTM1	0	0.738	206	-517.3	-316.5	13279.5	-1.743	0.081
GSTT1	1	0.571	227	-13.5	-8.9	18189.3	-0.034	0.973
GSTT1	0	0.429	227	-171.5	-176.1	18189.3	0.034	0.973

Results from Association Analysis Involving the Markers and the Measure of Bronchial Responsiveness Variable Involving the Asthmatic Siblings, using Smoking As a Covariate

Marker	No. of Copies	Allelic Frequency	No. of Informative Families	S	E(S)	Var(S)	Z	P
GSTM1	1	0.262	209	-28.1	19.3	23128.6	-0.311	0.756
GSTM1	0	0.738	209	205.3	158.0	23128.6	0.311	0.756
GSTT1	1	0.571	230	-410.5	-75.0	33592.7	-1.830	0.067
GSTT1	0	0.429	230	238.6	-96.9	33592.7	1.830	0.067

Results from Association Analysis Involving the Markers and Percentage Predicted FEV1 Values Variable Involving the Asthmatic Siblings, using Smoking As a Covariate

Marker	No. of Copies	Allelic Frequency	No. of Informative Families	S	E(S)	Var(S)	Z	P
GSTM1	1	0.262	208	-7.1	0.2	123.6	-0.656	0.512
GSTM1	0	0.738	208	-24.3	-31.6	123.6	0.656	0.512
GSTT1	1	0.571	229	34.5	33.2	182.5	0.098	0.922
GSTT1	0	0.429	229	16.8	18.2	182.5	-0.098	0.922

Results from Association Analysis Involving the Markers and Atopy Severity Scores Variable the Involving Asthmatic Siblings, using Smoking As a Covariate

Marker	No. of Copies	Allelic Frequency	No. of Informative Families	S	E(S)	Var(S)	Z	P
GSTM1	1	0.262	201	-9.8	3.4	64.7	-1.640	0.101
GSTM1	0	0.738	201	29.1	15.9	64.7	1.640	0.101
GSTT1	1	0.571	217	-0.5	12.7	79.5	-1.479	0.139
GSTT1	0	0.429	217	10.2	-3.0	79.5	1.479	0.139

Results from Association Analysis Involving the Markers and Asthma Severity Scores Variable the Involving Asthmatic Siblings, using Smoking As a Covariate

10.6 Publications Arising From This Thesis

Clinical Chemistry 55:9
1680–1685 (2009)

Copy-Number Variation Genotyping of *GSTT1* and *GSTM1* Gene Deletions by Real-Time PCR

Matthew J. Rose-Zerilli,[1,2*] Sheila J. Barton,[3] A. John Henderson,[4] Seif O. Shaheen,[5] and John W. Holloway[1,2]

BACKGROUND: Structural variation in the human genome is increasingly recognized as being highly prevalent and having relevance to common human diseases. Array-based comparative genome-hybridization technology can be used to determine copy-number variation (CNV) across entire genomes, and quantitative PCR (qPCR) can be used to validate de novo variation or assays of common CNV in disease-association studies. Analysis of large qPCR data sets can be complicated and time-consuming, however.

METHODS: We describe qPCR assays for *GSTM1* (glutathione *S*-transferase mu 1) and *GSTT1* (glutathione *S*-transferase theta 1) gene deletions that can genotype up to 192 samples in duplicate 5-μL reaction volumes in <2 h on the ABI Prism 7900HT Sequence Detection System. To streamline data handling and analysis of these CNVs by qPCR, we developed a novel interactive, macro-driven Microsoft Excel® spreadsheet. As proof of principle, we used our software to analyze CNV data for 1478 DNA samples from a family-based cohort.

RESULTS: With only 8 ng of DNA template, we assigned CNV genotypes (i.e., 2, 1, or 0 copies) to either 96% (*GSTM1*) or 91% (*GSTT1*) of all DNA samples in a single round of PCR amplification. Genotyping accuracy, as ascertained by familial inheritance, was >99.5%, and independent genotype assignments with replicate real-time PCR runs were 100% concordant.

CONCLUSIONS: The genotyping assay for *GSTM1* and *GSTT1* gene deletion is suitable for large genetic epidemiologic studies and is a highly effective analysis system that is readily adaptable to analysis of other CNVs.

The advent of microarray-based comparative genomic-hybridization technologies and whole-genome sequencing has revealed unexpectedly heterogeneous structural variation (deletions, duplications, inversions, and translocations) in the human genome *(1–3)*. These variations may have a greater effect on disease susceptibility than previously thought; therefore, it is important for any genetic study to consider the presence of structural variation in the region of interest and to have suitable technologies for detecting them *(4)*. Current technologies such as microarray-based comparative genomic hybridization and multiplex ligation-dependent probe amplification facilitate the detection of structural variation across the genome and at smaller chromosomal regions, respectively *(5, 6)*. Glutathione *S*-transferases (GSTs)[6] (EC 2.5.1.18) are phase II detoxification enzymes found free in the cytosol. These enzymes catalyze the conjugation of glutathione to many electrophilic substrates (including free radicals, xenobiotics, and physiological metabolites), producing stable and more soluble compounds that can then be excreted or compartmentalized by phase III enzymes *(7)*. The families of genes that encode GSTs are highly polymorphic in the human population *(8)*. Complete deletion of the *GSTM1*[7] (glutathione *S*-transferase mu 1) and *GSTT1* (glutathione *S*-transferase theta 1) genes occurs in approximately 50% and 20% of Caucasians, respectively. Copy-number variation (CNV) has a dosage effect on the concentrations of the GSTM1 (MIM 138350) and GSTT1 (MIM 600436) proteins that affects an individual's ability to detoxify compounds efficiently and provide protection from oxidative stress *(9, 10)*.

Polymorphisms for GST gene deletion have consequences for the detoxification of xenobiotics *(11)*. Single or combined deletions in GST-encoding genes

[1] Respiratory Genetics Group, Human Genetics, [2] Infection, Inflammation and Immunity, and [3] Public Health Sciences and Medical Statistics, Community Clinical Sciences Divisions, School of Medicine, University of Southampton, Southampton, UK; [4] Department of Community-Based Medicine, University of Bristol, Bristol, UK; [5] The National Heart and Lung Institute, Imperial College London, London, UK.

* Address correspondence to this author at: Respiratory Genetics Group, Human Genetics Division, University of Southampton, Duthie Bldg. (MP 808), Southampton General Hospital, Tremona Rd., Southampton SO16 6YD, UK. Fax +02380-794264; e-mail mjrz@soton.ac.uk.

Received February 10, 2009; accepted June 23, 2009.
Previously published online at DOI: 10.1373/clinchem.2008.120105

[6] Nonstandard abbreviations: GST, glutathione *S*-transferase; CNV, copy-number variation; ECACC, European Collection of Cell Cultures; FBAT, family-based association test; Cq, quantification cycle; qPCR, quantitative PCR.

[7] Human genes: *GSTM1*, glutathione *S*-transferase mu 1; *GSTT1*, glutathione *S*-transferase theta 1; *ALB*, albumin.

increase the susceptibility to lung cancer in smokers *(12)* and in nonsmokers exposed to household environmental tobacco smoke *(13)*. There is also an increased risk of cardiovascular disease in smokers who have a GST deficiency *(14)*. Polymorphisms for GST-encoding gene deletions and environmental tobacco smoke contribute to the development of childhood asthma *(15)* and may make asthmatic children more susceptible to the deleterious effects of ozone *(16)*.

Gross deletions in *GSTT1* and *GSTM1* genes are created by separate equal or unequal recombination events involving crossing-over between 2 highly homologous repeat regions that flank each gene *(9, 10)*. These separate recombination events produce deletion-junction regions of up to several kilobases that have very high homology (>98%) to the flanking repeat regions. This phenomenon severely restricts the design of unique primer sequences for use in standard PCR methods to identify the presence of a deletion. Sprenger et al. described a multiplex PCR method that uses unique priming sites in the *GSTT1* deletion junction *(9)*. This method appears to have limited suitability for use in large genetics studies because 2 steps, PCR amplification followed by agarose gel electrophoresis, are required to visualize the 1.46-kb deletion junction and the 466-bp gene-specific PCR products. Furthermore, unbiased coamplification of products of different sizes can be problematic, especially when the quality of the DNA is variable, because the product with the higher PCR efficiency may be preferentially amplified *(17)*.

Brasch-Andersen et al. showed that it is possible to use the increased sensitivity of real-time PCR assays to provide dosage data for the *GSTT1* and *GSTM1* genes *(18)*. This method can distinguish between individuals with 2, 1, or 0 copies of a gene. Because *GSTT1* or *GSTM1* CNV correlates with altered enzyme activity *(9, 19)*, analysis in a dose-dependent manner would best describe any association with disease outcome.

We developed new real-time PCR assays that use minor groove–binding hydrolysis probe technology on the 384-well ABI Prism 7900HT Sequence Detection System (Applied Biosystems) for higher sample throughput, and we used a novel interactive macro-driven Microsoft Excel® spreadsheet for analyzing the CNV data.

Materials and Methods

QUANTITATIVE PCR DESIGN AND VALIDATION

We used Primer Express v2.0 software (Applied Biosystems) to design novel minor groove–binding hydrolysis probe assays for target genes *GSTT1* and *GSTM1* and for reference gene *ALB* (albumin). We subsequently conducted a BLAST search (http://www.blast.ncbi.nlm.nih.gov) to evaluate the specificity of these assays. We performed PCR reactions with asymmetric concentrations of primers (0.3–0.9 μmol/L) and probe titrations (200–50 nmol/L) to lower the detection limit and to increase the cost-effectiveness of the assays, respectively. Triplicate wells of simplex (*GSTT1*, *GSTM1*, or *ALB*) and multiplex (i.e., *GSTT1* plus *ALB* or *GSTM1* plus *ALB*) reactions containing a series of 2-fold dilutions of pooled DNA samples (20–0.156 ng/reaction) from the European Collection of Cell Cultures (ECACC) were used to determine the PCR efficiencies of each assay. We performed a melting-curve analysis of the PCR to assess the specificity of the primers. Each 10-μL reaction with SYBR Green I contained 5 mmol/L $MgCl_2$, 0.2 mmol/L of each deoxynucleoside triphosphate, 30 g/L SYBR Green I (dissolved in DMSO), 0.25 U of GoldStar *Taq* DNA polymerase (Eurogentec), 10 ng of pooled control DNA template, and primers to the final concentrations detailed in Table 1 of the Data Supplement that accompanies the online version of this article at http://www.clinchem.org/content/vol55/issue9. To evaluate genotyping accuracy, we used replicate sample wells and conducted transmission disequilibrium tests with family-based association test (FBAT) v1.4 *(20)* as an inheritance check for the family-based cohort.

DNA TEMPLATES

ECACC Human Random Control DNA samples (HRC-1) (http://www.hpacultures.org.uk/collections/ecacc.jsp) were used to optimize reaction conditions, to determine the stability of reaction plate storage, and to validate subsequent experiments as positive controls. Details of the ECACC HRC-1 DNA samples used as controls are described in the user manual for CNV analysis in the online Data Supplement. We conducted *GSTT1* and *GSTM1* CNV experiments with 1478 DNA samples previously collected for an asthma study of sibling pairs from 341 affected Caucasian families in the UK *(21)*. Samples were obtained with informed consent after ethics approval from the Southampton and South West Hampshire Local Research Ethics Committee and the Portsmouth and South East Hampshire Local Research Ethics Committee.

PROTOCOL

We transferred 2 μL of 1-ng/μL genomic DNA in duplicate to a 384-well plate. Each plate also contained 3 positive controls and a minimum of 4 negative controls (water). Before adding the reaction mixture, we dried the DNA for 10 min at 80 °C on a PCR thermal cycler.

We used 5 μL for each multiplex assay, and each well contained the following: 2.5 μL of 2× Precision-R MasterMix (PrimerDesign), which contained 5 mmol/L $MgCl_2$, 200 μmol/L of each deoxynucleoside triphosphate, 0.025 U/μL *Taq* polymerase, and 6-carboxy-X-rhodamine reference dye; the necessary volume of

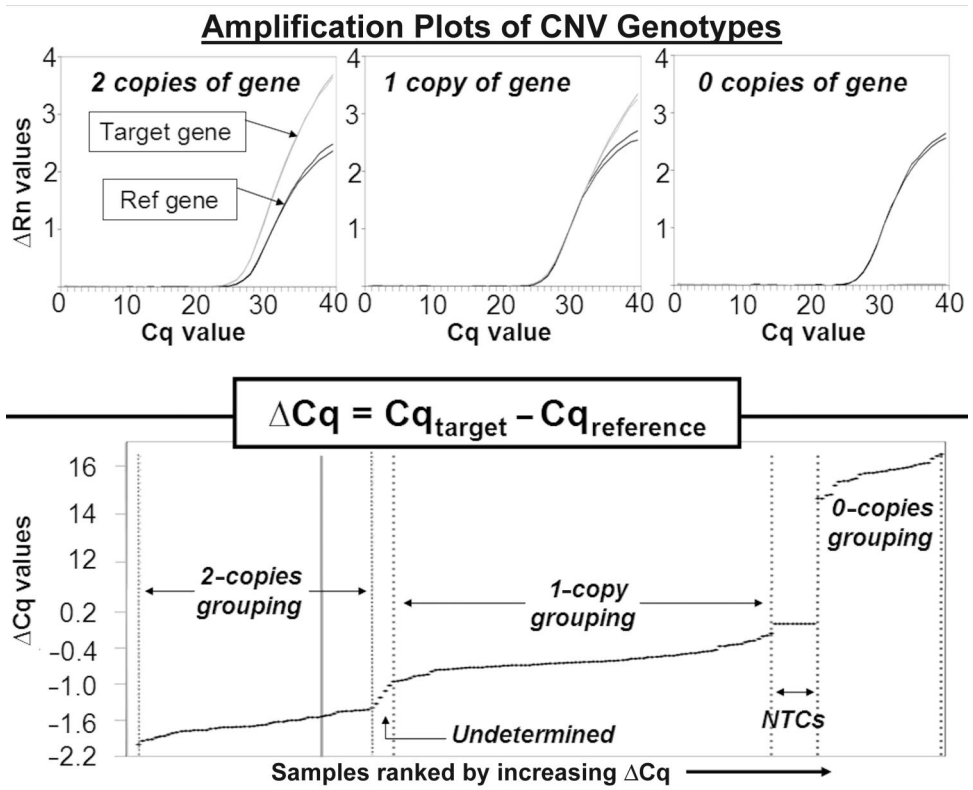


Fig. 1. CNV analysis spreadsheet.

Real-time PCR data are imported into the analysis spreadsheet, and the macro calculates the difference between the target and reference Cq values with the indicated formula. ΔCq values are displayed in a ranked scatter graph (bottom chart), allowing assignment of CNV-genotype groupings by moving the vertical dotted lines. The solid vertical line highlights any desired sample and displays the PCR curves; this feature allows determination of the validity of any genotype assignment within the grouping. The top 3 graphs display plots of PCR curves for each of the 3 possible CNV genotypes. ΔRn, reporter signal normalized to ROX; Ref, reference; NTCs, no-template controls.

primer/probe mix; and PCR-grade water to volume. Table 1 in the online Data Supplement lists primer and probe sequences and reagent concentrations in the reactions. All PCR reaction mixes were prepared in bulk with 10% extra volume for pipetting and were transferred to the 384-well plate with an electronic multichannel pipettor. Primer/probe mixes were also prepared in advance to reduce the interassay imprecision attributable to setting up each PCR reaction mix.

The plates were heat-sealed (Thermo-Sealer; ABgene) to stop any evaporation from occurring during the DNA-rehydration step (30 min at 4 °C protected from light, with periodic mixing) and during thermal cycling. PCR amplification and collection of real-time fluorescence data were performed on the ABI Prism 7900HT Sequence Detection System. The thermal profile consisted of incubation at 95 °C for 10 min followed by 40 cycles of 95 °C for 15 s and 60 °C for 60 s without 9600-emulated temperature ramping. Fluorescence data were collected during the 60 °C annealing/extension cycles, and quantification cycle (Cq) values were calculated with the automatic Cq-analysis settings on an absolute quantification run (SDS Plate Utility v2.1; Applied Biosystems).

CNV ANALYSIS SPREADSHEET

We designed a macro-based Microsoft Excel spreadsheet to analyze real-time PCR data for gene dosage (Fig. 1). The results (Cq values) and files of clipped data (normalized fluorescence values from the reporter dyes at the end of each extension phase) generated by the SDS v2.1 software for each plate run can be imported into the spreadsheet. Each duplicate data point is visualized in a ranked *xy* scatter graph (ΔCq value vs samples ranked by increasing ΔCq). The ΔCq value ($\Delta Cq = Cq_{target} - Cq_{reference}$) for each duplicate can be

compared directly with the plots of real-time amplification curves generated by each grouping of sample wells. The interactive sliding windows allow semiautomated assignment of genotype groups and outliers on the basis of the grouping of ΔCq values on the ranked *xy* scatter graph. Importantly, interrogation of the plots of real-time amplification curves permits rapid determination of the validity of any genotype call and allows manual correction. The spreadsheet automatically calculates CNV and frequencies of a gene present or absent, deviation from the Hardy–Weinberg equilibrium, and several QC parameters (i.e., mean ΔCq value and SD of each CNV group). The interactive and color-coded 384- and 96-well arrays of the plate layout generated by the macro also allows postanalysis visualization of the well position for any genotype call, water control, or reaction failure. This utility allows monitoring for any problematic samples or experimental controls. It also identifies any possible effects of the plate edges on data quality. Finally, the macro exports all analyzed data to a new sheet in the spreadsheet in a database-compatible format (Microsoft Access®). The CNV-analysis spreadsheet is available in the online Data Supplement.

CALCULATIONS OF THE HARDY–WEINBERG EQUILIBRIUM

These calculations were performed on the parental genotypes with the online version of the Definetti program (http://ihg.gsf.de/cgi-bin/hw/hwa1.pl). A Pearson χ^2 P value <0.05 was considered to indicate significant deviation of the genotype frequency from the Hardy–Weinberg equilibrium.

Results

ASSAY OPTIMIZATION

The detection limits of the quantitative PCR (qPCR) assays were substantially improved with the use of asymmetric concentrations of primers (see Table 1 in the online Data Supplement). These conditions reduced Cq values by up to 1.38 Cq. Titrations of the probe concentration revealed that a final probe concentration of 150 nmol/L for each assay could be used without decreasing the signal strength (lower signal-to-noise ratio) or the detection limit (increase in Cq value). An analysis of the SYBR Green I melting curve revealed single melting peaks for all assays (see Fig. 1 in the online Data Supplement).

Under multiplex conditions and the optimized asymmetric primer concentrations and probe concentrations described above, the calibration curves (see Fig. 2 in the online Data Supplement) demonstrated that the multiplex amplification reactions were linear over the entire range of DNA template concentrations (20–0.156 ng); all values for PCR reaction efficiency ($E = 10^{1/s}$, where s is the slope of the calibration curve) were between 1.95 and 1.99, values close to the theoretical maximum value (i.e., $E = 2.0$, in which 1 Cq is equivalent to a doubling of the PCR product with each PCR cycle). The interassay variation in PCR efficiency was <1% in both multiplex reactions (E_{GSTM1} and E_{ALB}, 0.65%; E_{GSTT1} and E_{ALB}, 0.01%), confirming that both PCR products in the multiplex were amplifying at nearly the same rate.

Table 1. ΔCq values for GST CNV genotypes.[a]

Gene	CNV genotype	Samples, n	Mean ΔCq	ΔCq SD	Mean difference: $\Delta Cq_{2\ copies} - \Delta Cq_{1\ copy}$
GSTM1	2	5	−1.25	0.09	1.03
	1	64	−0.22	0.16	
	0	112	12.90	1.14	
GSTT1	2	54	−1.67	0.15	1.06
	1	90	0.61	0.18	
	0	30	15.66	0.49	

[a] Mean ΔCq values for each genotype group for 176 DNA samples.

The SD for the mean ΔCq ($\Delta Cq = Cq_{target} - Cq_{reference}$) determined at each DNA dilution along the calibration curve was used to estimate the experimental variation in the ΔCq value used to determine the copy number for the gene of interest. Over the range of DNA dilutions (20–0.156 ng), the mean (SD) ΔCq was −0.72 (0.1) for the *GSTM1* assay and −0.54 (0.08) for the *GSTT1* assay.

The initial genotyping results for the 2 genes obtained for 176 DNA samples (Table 1) demonstrated that the sensitivity of the qPCR assay was sufficient so that ΔCq values could be used to distinguish a change in gene copy number. The mean difference in ΔCq between samples assigned a 2-copies or 1-copy genotype was 1.03 for *GSTM1* and 1.06 for *GSTT1*. These values were very close to the theoretical 1 Cq value for a doubling in concentration (i.e., 1 extra copy of the gene). In practice, this value varied slightly between plate runs (0.85–1.20), although this variation did not affect genotyping accuracy because the data from each plate were analyzed independently in the analysis spreadsheet.

GENOTYPING RESULTS

Genotyping accuracies for the *GSTM1* (99.9%) and *GSTT1* (99.5%) assays were determined by dividing the number for non-Mendelian inheritances detected with the FBAT program by the number of child genotypes (n = 770 for *GSTM1*; n = 735 for *GSTT1*). All of the results indicating non-Mendelian inheritance (n = 1

Table 2. Genotype frequencies.[a]

Group	Gene	CNV frequency, n (%) 2 Copies	1 Copy	0 Copies	HWE *P*
Entire cohort	*GSTT1*	415 (31)	690 (51)	247 (18)	NA
	GSTM1	98 (7)	534 (38)	786 (55)	NA
Parents only	*GSTT1*	193 (31)	321 (52)	103 (17)	0.117
	GSTM1	42 (6)	250 (39)	356 (55)	0.831

[a] Calculations of the Hardy–Weinberg equilibrium (HWE) were not performed on the full family cohort because the individuals are related; the parents' frequencies were analyzed separately as a surrogate of that of the general population and found to be in HWE. NA, not applicable.

for *GSTM1*; n = 4 for *GSTT1*) were due to child "2 copies" genotype calls when the parental genotype combination permitted only 0- or 1-copy offspring. To test the interassay reproducibility of genotype assignments, we re-genotyped 92 DNA samples for both CNVs in a separate qPCR run; comparison of the 2 independent genotype calls demonstrated 100% concordance. Table 2 shows the full cohort and parent genotype frequencies for both gene CNVs and a lack of statistical significance in Pearson tests for deviation from the Hardy–Weinberg equilibrium. The rates for undetermined CNV calls (i.e., 2, 1, and 0 copies) were 4% for *GSTM1* and 9% for *GSTT1*, and 2% and 3%, respectively, for the gene absent/present call.

ASSAY PLATE STABILITY

Testing performed at the 24-h, 48-h, and 72-h time points with the ECACC HRC-1 DNA panel showed that sealed reaction plates could be stored protected from light at room temperature for 24 h before the real-time PCR run without any degradation in data quality (data not shown).

Discussion

This new methodology enhances the study of gene CNV in the large sample sizes required in genetic-association studies. This study is not the first application of real-time PCR to assay *GSTM1* and *GSTT1* CNV; several other protocols have been described in the literature *(18, 22)*. These methods are not effective, however, for processing the large sample sizes required in genetic-association studies, because the low sample density on reaction plates caused by the use of triplicate wells and larger reaction volumes (10 μL) result in a high genotyping cost per sample. With our method, 96 samples can be genotyped for 2 CNVs in duplicate wells on a single reaction plate (or 192 samples for 1 CNV) in <2 hours with a proportional 66% savings in the costs for 2× Precision-R MasterMix (5 μL vs 15 μL), primers, and probe. The small amount of DNA template used in the assay conserves DNA template (4 ng DNA per CNV). Diluting the template also dilutes any potential PCR inhibitors present in the sample, and drying the template reduces sample cross-contamination. We determined 2 ng DNA to be optimum in our genotyping experiments because this quantity falls in the central part of the range of calibrator dilutions where Cq values have a direct relationship to the log of the DNA template concentration, thus providing accurate quantification. Minor groove–binding hydrolysis probes were used because they permitted smaller reaction volumes than required with standard hydrolysis probes without compromising the assay's detection limit or reproducibility *(23)*. The calibration curves demonstrate that the multiplex PCR reactions are highly efficient ($E \geq 1.95$) and amplify at nearly the same rate with a large dynamic range. The calibration curves cover a DNA template concentration range of 7 "orders of magnitude" (i.e., 2^0–2^{-7}: 20–0.156 ng DNA per PCR reaction).

The combination of assay optimization and use of a novel analysis spreadsheet enabled these assays to report 97%–98% of the individuals in our family cohorts with gene present/absent genotype calls and 91%–96% of the individuals with full CNV genotypes (i.e., 0, 1, or 2 copies) with high accuracy (≥99.5%, as determined by evaluating inheritance in the family cohort). Success rates for calling full CNV genotypes approached those routinely achieved by standard technologies for calling genotypes for single-nucleotide polymorphisms (≥95%) *(24)*. A subsequent round of genotyping for the unassigned samples (between 6% and 9% for *GSTM1* and *GSTT1*) increases the genotyping success rate to the rates achieved with the genotyping technologies used for single-nucleotide polymorphisms while still obtaining a greater proportional savings in costs compared with previously described GST CNV protocols *(18, 22)*. The detection of a small number of non-Mendelian inheritances in our cohort of families may reflect a possible limitation of the manual sliding-window method to determine the end of the 2-copies ΔCq grouping and the start of the single-copy grouping (i.e., 1-copy individuals could be erroneously assigned to the 2-copies group). The start and end of the 2 groupings of genotypes on the ranked *xy* scatter graph are usually distinct, however, and barring any outlying data points (outliers can easily be distinguished by examining the quality of the data for real-time PCR plots for the well grouping), the FBAT-determined inheritance results could simply reflect the true genotyping error of the technique (i.e., sporadic contamination not detected by the negative controls and chance results). In support of this hypothesis, the genotypes of re-genotyped samples were 100% concordant,

and the lack of any discordant genotypes indicated that the manual sliding-window method for assigning genotype groups was sound.

The protocol described in this report has advantages over previously published methodologies *(18, 22)*, including a higher sample density per reaction plate, a smaller reaction volume, and compatibility with liquid-handling robotic platforms for DNA aliquoting, reaction plate setup, and loading of plates onto the real-time PCR machine. The analysis spreadsheet provides a complete analysis, QC, troubleshooting capabilities, and a data-handling solution to CNV genotyping on the ABI 7900HT Sequence Detection System. The macro-based Excel spreadsheet eliminates the considerable bottleneck involved in processing real-time PCR data and may be applicable to CNV genotyping of other genes from any genome.

Author Contributions: *All authors confirmed they have contributed to the intellectual content of this paper and have met the following 3 requirements: (a) significant contributions to the conception and design, acquisition of data, or analysis and interpretation of data; (b) drafting or revising the article for intellectual content; and (c) final approval of the published article.*

Authors' Disclosures of Potential Conflicts of Interest: *Upon manuscript submission, all authors completed the Disclosures of Potential Conflict of Interest form. Potential conflicts of interest:*

Employment or Leadership: None declared.
Consultant or Advisory Role: None declared.
Stock Ownership: None declared.
Honoraria: None declared.
Research Funding: M.J. Rose-Zerilli, The British Lung Foundation; Asthma, Allergy & Inflammation Research Charity; and the Infection, Inflammation and Immunity Division, School of Medicine, University of Southampton. J.W. Holloway, The Medical Research Council UK; The British Lung Foundation; Asthma UK; and the Asthma, Allergy & Inflammation Research Charity. S.O. Shaheen, Asthma UK.
Expert Testimony: None declared.

Role of Sponsor: The funding organizations played no role in the design of study, choice of enrolled patients, review and interpretation of data, or preparation or approval of manuscript.

References

1. Feuk L, Carson AR, Scherer SW. Structural variation in the human genome. Nat Rev Genet 2006; 7:85–97.
2. Levy S, Sutton G, Ng PC, Feuk L, Halpern AL, Walenz BP, et al. The diploid genome sequence of an individual human. PLoS Biol 2007;5:e254.
3. Sharp AJ, Locke DP, McGrath SD, Cheng Z, Bailey JA, Vallente RU, et al. Segmental duplications and copy-number variation in the human genome. Am J Hum Genet 2005;77:78–88.
4. Feuk L, Marshall CR, Wintle RF, Scherer SW. Structural variants: changing the landscape of chromosomes and design of disease studies. Hum Mol Genet 2006;15:R57–66.
5. Carson AR, Feuk L, Mohammed M, Scherer SW. Strategies for the detection of copy number and other structural variants in the human genome. Hum Genomics 2006;2:403–14.
6. Bunyan DJ, Eccles DM, Sillibourne J, Wilkins E, Thomas NS, Shea-Simonds J, et al. Dosage analysis of cancer predisposition genes by multiplex ligation-dependent probe amplification. Br J Cancer 2004;91:1155–9.
7. Frova C. Glutathione transferases in the genomics era: new insights and perspectives. Biomol Eng 2006;23:149–69.
8. McIlwain CC, Townsend DM, Tew KD. Glutathione *S*-transferase polymorphisms: cancer incidence and therapy. Oncogene 2006;25:1639–48.
9. Sprenger R, Schlagenhaufer R, Kerb R, Bruhn C, Brockmoller J, Roots I, Brinkmann U. Characterization of the glutathione *S*-transferase *GSTT1* deletion: discrimination of all genotypes by polymerase chain reaction indicates a trimodular genotype-phenotype correlation. Pharmacogenet 2000;10:557–65.
10. Xu S, Wang Y, Roe B, Pearson WR. Characterization of the human class Mu glutathione *S*-transferase gene cluster and the *GSTM1* deletion. J Biol Chem 1998;273:3517–27.
11. Bolt HM, Thier R. Relevance of the deletion polymorphisms of the glutathione *S*-transferases *GSTT1* and *GSTM1* in pharmacology and toxicology. Curr Drug Metab 2006;7:613–28.
12. Sorensen M, Autrup H, Tjonneland A, Overvad K, Raaschou-Nielsen O. Glutathione *S*-transferase T1 null-genotype is associated with an increased risk of lung cancer. Int J Cancer 2004;110: 219–24.
13. Wenzlaff AS, Cote ML, Bock CH, Land SJ, Schwartz AG. GSTM1, GSTT1 and GSTP1 polymorphisms, environmental tobacco smoke exposure and risk of lung cancer among never smokers: a population-based study. Carcinogenesis 2005;26:395–401.
14. Tamer L, Ercan B, Camsari A, Yildirim H, Cicek D, Sucu N, et al. Glutathione *S*-transferase gene polymorphism as a susceptibility factor in smoking-related coronary artery disease. Basic Res Cardiol 2004;99:223–9.
15. Kabesch M, Hoefler C, Carr D, Leupold W, Weiland SK, von Mutius E. Glutathione *S* transferase deficiency and passive smoking increase childhood asthma. Thorax 2004;59:569–73.
16. Romieu I, Sienra-Monge JJ, Ramirez-Aguilar M, Moreno-Macias H, Reyes-Ruiz NI, Estela del Rio-Navarro B, et al. Genetic polymorphism of *GSTM1* and antioxidant supplementation influence lung function in relation to ozone exposure in asthmatic children in Mexico City. Thorax 2004;59: 8–10.
17. Markoulatos P, Siafakas N, Moncany M. Multiplex polymerase chain reaction: a practical approach. J Clin Lab Anal 2002;16:47–51.
18. Brasch-Andersen C, Christiansen L, Tan Q, Haagerup A, Vestbo J, Kruse TA. Possible gene dosage effect of glutathione-*S*-transferases on atopic asthma: using real-time PCR for quantification of *GSTM1* and *GSTT1* gene copy numbers. Hum Mutat 2004;24:208–14.
19. McLellan RA, Oscarson M, Alexandrie AK, Seidegard J, Evans DA, Rannug A, Ingelman-Sundberg M. Characterization of a human glutathione *S*-transferase mu cluster containing a duplicated *GSTM1* gene that causes ultrarapid enzyme activity. Mol Pharmacol 1997;52:958–65.
20. Laird NM, Horvath S, Xu X. Implementing a unified approach to family-based tests of association. Genet Epidemiol 2000;19(Suppl 1):S36–42.
21. Van Eerdewegh P, Little RD, Dupuis J, Del Mastro RG, Falls K, Simon J, et al. Association of the ADAM33 gene with asthma and bronchial hyperresponsiveness. Nature 2002;418:426–30.
22. Covault J, Abreu C, Kranzler H, Oncken C. Quantitative real-time PCR for gene dosage determinations in microdeletion genotypes. Biotechniques 2003;35:594–6, 8.
23. Kutyavin IV, Afonina IA, Mills A, Gorn VV, Lukhtanov EA, Belousov ES, et al. 3′-minor groove binder-DNA probes increase sequence specificity at PCR extension temperatures. Nucleic Acids Res 2000;28:655–61.
24. Tindall EA, Speight G, Petersen DC, Padilla EJ, Hayes VM. Novel Plexor SNP genotyping technology: comparisons with TaqMan and homogenous MassEXTEND MALDI-TOF mass spectrometry. Hum Mutat 2007;28:922–7.

Int. J. Epidemiol. Advance Access published January 4, 2010

Published by Oxford University Press on behalf of the International Epidemiological Association *International Journal of Epidemiology* 2009;1–24
© The Author 2009 doi:10.1093/ije/dyp337

Glutathione-*S*-transferase genes and asthma phenotypes: a Human Genome Epidemiology (HuGE) systematic review and meta-analysis including unpublished data

Cosetta Minelli,[1]*† Raquel Granell,[2†] Roger Newson,[3] Matthew J Rose-Zerilli,[4] Maties Torrent,[5] Sue M Ring,[2] John W Holloway,[4] Seif O Shaheen[3] and John A Henderson[6]

Accepted 6 October 2009

Background Oxidative stress is thought to be involved in the pathogenesis of asthma. Glutathione-*S*-transferase (GST) enzymes, which play an important role in antioxidant defences, may therefore influence asthma risk. Two common deletion polymorphisms of *GSTM1* and *GSTT1* genes and the *GSTP1* Ile105Val polymorphism have been associated with asthma in children and adults, but results are inconsistent across studies.

Methods Systematic review and meta-analysis of the effects of *GST* genes on asthma, wheezing and bronchial hyper-responsiveness (BHR), with inclusion of unpublished data from three studies, including the large Avon Longitudinal Study of Parents and Children (ALSPAC). Random effect or fixed effect models were used as appropriate, and sensitivity analyses were performed to assess the impact of study characteristics and quality on pooled results.

Results The meta-analyses of *GSTM1* ($n = 22$ studies) and *GSTT1* ($n = 19$) showed increased asthma risk associated with the null genotype, but there was extreme between-study heterogeneity and publication bias and the association disappeared when meta-analysis was restricted to the largest studies. Meta-analysis of *GSTP1* Ile105Val ($n = 17$) and asthma suggested a possible protective effect of the Val allele, but heterogeneity was extreme. Few studies evaluated wheezing and BHR and most reported no associations, although weak evidence was found for positive associations of *GSTM1* null

[1] Institute of Genetic Medicine, EURAC Research, Bolzano, Italy.
[2] Department of Social Medicine, University of Bristol, Bristol, UK.
[3] Respiratory Epidemiology and Public Health Group, National Heart and Lung Institute, Imperial College London, London, UK.
[4] Divisions of Human Genetics and Inflammation, Infection & Immunity, School of Medicine, University of Southampton, Southampton, UK.
[5] Ib-salut Menorca Health Area and Fundació Caubet-CIMERA, Balearic Islands, Spain.
[6] Department of Community-based Medicine, University of Bristol, Bristol, UK.
* Corresponding author. Institute of Genetic Medicine, EURAC research Viale Druso 1, 39100 Bolzano, Italy. Tel: +39 0471 055 526; Fax: +39 0471 055 599;
E-mail: cosetta.minelli@eurac.edu
† These authors contributed equally to this work.

and *GSTP1* Val allele with wheezing and a negative association of *GSTP1* Val allele with BHR.

Conclusions Our findings do not support a substantial role of *GST* genes alone in the development of asthma. Future studies of large size should focus on interactions of *GST* genes with environmental oxidative exposures and with other genes involved in antioxidant pathways. Quality of study conduct and reporting needs to be improved to increase credibility of the evidence accumulating over time.

Keywords Meta-analysis, systematic review, glutathione-*S*-transferase genes, *GSTM1* gene, *GSTT1* gene, *GSTP1* gene, asthma, wheezing, bronchial responsiveness, The Avon Longitudinal Study of Parents and Children (ALSPAC)

Introduction

Asthma is characterized by chronic airway inflammation and oxidative stress in the lungs has been implicated in its pathogenesis.[1] Sources of oxidant injury are reactive oxygen and nitrogen species generated by activated inflammatory cells and bronchial epithelial cells and inhalation of atmospheric pollutants, notably tobacco smoke and oxidant gases, including ozone, sulphur dioxide and nitrogen oxides.[2] These are countered by enzymatic and non-enzymatic antioxidants, including dietary antioxidants, such as vitamins C and E, and glutathione, a major protective antioxidant in the lungs that also has a role in regulation of inflammatory responses.[3] A family of enzymes, the glutathione-*S*-transferases (GSTs), has the general function of conjugating glutathione with electrophilic substances that are capable of generating free radicals, thus leading to detoxification of their effects. Genetic polymorphisms associated with reduced activity of GSTs are therefore of interest in the study of disease susceptibility. Two common deletion polymorphisms of *GSTM1* and *GSTT1* genes have been associated with asthma in children and adults.[4–7] The Val allele of the *GSTP1* Ile105Val polymorphism, although associated with reduced glutathione activity,[8] has surprisingly been reported to be highly protective for asthma and airway responsiveness in adults.[5,9,10] However, the Val105 allele has also been reported to increase the risk of asthma[11] and to increase susceptibility to the effects of ozone on breathing difficulties in children with asthma.[12] Others have reported no associations between this *GSTP1* polymorphism and asthma in children and adults.[4,8,13] Many of the studies were relatively small or confined to selected populations. A systematic review of the literature can help understand the reasons for the heterogeneity in study results, and the pooling of results across similar studies in a meta-analysis can overcome the problem of limited statistical power. With this aim, we have conducted a systematic review and meta-analysis on the associations between *GSTM1*, *GSTT1* and *GSTP1* polymorphisms and asthma phenotypes in children and adults, with inclusion of unpublished data from a large UK cohort study in children, a UK family-based study in children and young adults and a Spanish cohort study in children.

Methods

Primary studies

Methods of the three primary studies included in the meta-analyses are briefly summarized below. Further details are reported in the Supplementary Appendix 1 available as supplementary data at *IJE* online.

ALSPAC

The Avon Longitudinal Study of Parents and Children (ALSPAC) is a longitudinal, population-based birth cohort study that recruited 14 541 pregnant women in 1991–92, with 14 062 liveborn children.[14,15] Ethical approval for the study was obtained from the ALSPAC Law and Ethics Committee and the Local Research Ethics Committees.

The study collected data on asthma and wheezing in children at age 7.5 years, based on a questionnaire sent to mothers. Children from multiple births were excluded from the analyses. Asthma was defined as a report of doctor diagnosis of asthma ever together with reported asthma, wheeze or asthma treatment in the previous 12 months. Wheezing was defined as a positive response to the question 'Has your child had any wheezing with whistling on his/her chest when he/she breathed in the past 12 months?'. Spirometry was performed at 8.5 years, and bronchial hyper-responsiveness (BHR) to methacholine was defined as a 20% reduction from baseline (post-saline) Forced Expired Volume in 1 s (FEV_1) at a cumulative dose of methacholine $\leqslant 1.2$ mg. In the mothers' population, asthma was defined as a positive response to the question 'Have you ever had asthma?',

from a self-completion questionnaire administered during pregnancy.

The majority of the children's DNA samples were extracted from cord blood or venous blood collected at age 7 years with a small number extracted from venous blood collected at 43–61 months. The majority of maternal DNA samples were extracted from blood taken during pregnancy and a small number were buccal DNA extracted from mouthwash samples.[16] *GSTP1* polymorphism Ile105Val was determined using competitive allele specific polymerase chain reaction (PCR) system (KASPar). Copy number variation (CNV) of *GSTT1* and *GSTM1* deletion polymorphisms was analysed using a real-time PCR method.[17] Genotyping was performed blind to the outcome status of individuals.

The genetic effects of the three genes were evaluated using between-genotype comparisons (variant homozygotes vs wild homozygotes and heterozygotes vs wild homozygotes). Since the samples of mothers and children were not independent, the odds ratios (ORs) used in the meta-analysis for the two groups were estimated by analysing them within the same logistic regression model, and calculating confidence intervals (CIs) using Huber variances clustered by pregnancy.

AMICS-INMA study

The Menorca Asthma Multicentre Infant Cohort Study (AMICS) is a population-based birth cohort included in the Spanish environment and childhood research network (INMA study). In 1997–98, the study recruited 482 children at birth from 492 pregnant women resident in the island of Menorca, Spain. The study has been described previously.[18,19] Ethical approval was obtained from the Balearic Islands Research Ethics Committee.

Information on asthma and wheezing in children was collected through telephone questionnaires administered yearly to mothers, for children up to the age of 6 years. Presented here are data based on the questionnaire at age 6 years. Asthma was defined as doctor-diagnosis of asthma in the previous 12 months, and wheezing as a positive response to the question 'Has your child had any wheezing with whistling on his/her chest when he/she breathed in the past 12 months?'.

Most of the DNA samples (87%) were extracted from blood obtained at age 4 years, and the rest from saliva collected at age 6 years. *GSTM1* and *GSTT1* gene deletions were detected by multiplex PCR, with a method modified from Arand and coll[20] (Supplementary Appendix 1 available as supplementary data at *IJE* online). *GSTP1* polymorphism Ile105Val was determined using pyrosequencing (Biotage, Uppsala, Sweden). All assays were performed blind to the outcome status of the children.

GSTM1 and *GSTT1* genetic effects were evaluated assuming a recessive model (OR for the null genotype), whereas the genetic effect of *GSTP1* was assessed using between-genotype comparisons.

Southampton study

The Southampton Asthma Cohort is a family study that started in 1997 and recruited 341 families (1508 individuals) from Southampton, Portsmouth, Bournemouth and the Isle of Wight, UK.[21] Inclusion criteria required two siblings (5–21 years) diagnosed with asthma and currently using asthma medication. Reported here are the associations of *GSTM1* and *GSTT1* deletion polymorphisms with asthma, defined as physician's diagnosis and current use of medication, and bronchial responsiveness to methacholine, only measured in participants with a baseline FEV_1 $\geqslant 70\%$ predicted.

DNA from blood samples was analysed for CNV of *GSTM1* and *GSTT1* genes, using quantitative real-time PCR (Supplementary Appendix 1 available as supplementary data at *IJE* online). Genotyping was performed blind to the outcome status.

Data were analysed using Family Based Association Tests (FBAT version 1.4 software; http://www.biostat.harvard.edu/_fbat.htm), assuming an additive genetic model with a null hypothesis of 'no association and no linkage' between presence/absence of the gene and asthma phenotypes.[22]

Systematic review

This systematic review was conducted following a protocol in accordance with the Human Genome Epidemiology Network (HuGENet) guidelines.[23] Studies were identified through electronic search of MEDLINE, EMBASE, ISI Web of Science and HuGENet up to February 2008, using a comprehensive search strategy (Supplementary Figure S1 available as supplementary data at *IJE* online). Included were population-based and family-based studies examining the effect of *GST* polymorphisms on asthma phenotypes. Although the original search strategy included atopy, this article focuses on the association of *GST* genes with asthma, wheezing and BHR. No restrictions were placed on language and type of report, with inclusion of conference abstracts. When multiple reports were available for a single study, only the most recent article or that with the largest sample size was included. Additional studies were identified through cross-checking of reference lists of all relevant studies, including previous reviews and editorials. Two reviewers (C.M. and R.G.) independently extracted data using a pre-piloted form, with a third reviewer available for arbitration (J.W.H.). Authors of eligible primary studies were contacted for further information whenever the data required by the meta-analysis were not fully reported in the article.

Study quality of primary studies was evaluated based on the HuGENet guidelines[23] and STrengthening the REporting of Genetic Association studies (STREGA) recommendations for the reporting

of genetic association studies.[24] We also assessed the epidemiological credibility of the results of our meta-analyses by use of an index recently proposed for assessing cumulative evidence in genetic associations (Venice criteria), which classifies epidemiological credibility of the results of a meta-analysis as 'strong', 'moderate' or 'weak', based on three elements: (i) amount of evidence, (ii) extent of replication and (iii) protection from bias.[25]

Meta-analysis

The meta-analyses were performed on the population-based studies, and included unpublished results from the ALSPAC and AMICS-INMA studies, together with those of studies identified through the systematic review.

Most studies evaluated *GSTM1* and *GSTT1* as presence/absence of gene deletion, so that meta-analyses of these polymorphisms were performed using a single OR (null vs present). Results for *GSTP1* Ile105Val are reported as two ORs, Ile/Val vs Ile/Ile and Val/Val vs Ile/Ile. Given the absence of *a priori* evidence on the genetic model for this polymorphism, a 'genetic model-free' approach was also used to pool results across genotype groups.[26] This approach does not assume a specific genetic model but estimates it from the data, with the only assumption that the unknown genetic model is constant across studies. The genetic model-free approach was implemented within a Bayesian framework and assuming non-informative prior distributions for all parameters.[27] Details of the model are reported in the Appendix 1.

In the presence of at least five studies, meta-analyses were performed using random effects models, which account for between-study heterogeneity. If less than five studies were included, meta-analysis was performed only in the presence of small between-study heterogeneity ($I^2 < 25\%$), using fixed-effect models. The presence of between-study heterogeneity was investigated using the Q test and its magnitude estimated using the I^2 statistic, which measures the proportion of variation across studies due to genuine differences rather than random error.[28] Possible causes of heterogeneity were investigated by subgroup analyses based on geographic location, population age (adults vs children), size of the study and definition of asthma. Additional sensitivity analyses were performed by excluding studies with poor methodological quality. Studies were considered of poor quality if no definition of the disease outcomes was provided and either the description of the study sample (e.g. selection of cases/controls for case–control studies) was incomplete or there was no mention of Hardy–Weinberg equilibrium (HWE). Although many more items were considered in the quality assessment, we chose to limit exclusion to those studies where poor quality seemed beyond any reasonable doubt.

For each study, deviation from HWE was tested using the exact test, and the magnitude of the departure measured using the disequilibrium coefficient. Studies with large deviations from HWE were further investigated for possible methodological problems, including population stratification and genotyping errors.[29] The presence of small-study bias, which is a proxy for publication bias, was investigated both graphically using funnel plots and formally using Begg's and Egger's tests.

All analyses were performed using Stata for Windows v10 (Stata Corp, College Station, TX, USA), with the exception of the model-free approach, which was implemented in WinBUGS 1.4.[30] For the Bayesian analyses, 95% credible intervals (i.e. range in which the probability that the parameter lies within this interval is 95%) were calculated in place of 95% CIs.

Results

Primary studies

ALSPAC

For analyses on asthma in children, data were available from 5327, 5300 and 5330 children for *GSTM1*, *GSTT1* and *GSTP1*, respectively; corresponding figures for wheezing were 5991, 5957 and 5992, and for bronchial responsiveness were 3430, 3402 and 3417. For analyses on asthma in mothers, data were available from 7049, 7017 and 7262, respectively. The allele frequencies of *GSTM1*, *GSTT1* and *GSTP1* were similar to those reported in other UK populations[31,32] and in a White, non-Hispanic US population.[33] All gene polymorphisms were in HWE. Overall, there was no evidence of substantial association of the three genes with asthma phenotypes. The results for *GSTM1* and *GSTT1*, analysed as CNV with three-level genotype data, showed no association for *GSTM1*, whereas a possible small protective effect on asthma of the *GSTT1* null allele could not be excluded in mothers (OR: 0.71; 95% CI: 0.57–0.90 and 0.84; 0.63–1.12, for heterozygotes and null homozygotes, respectively, compared with wild-type homozygotes). However, evidence of a possible protective effect was weak in children (0.91; 0.76–1.11 and 0.89; 0.70–1.13). In children, the OR of *GSTP1* Ile/Val vs Ile/Ile for asthma was 1.16 (95% CI 0.98–1.37) and for wheezing 1.25 (1.05–1.49), but such association was not shown for *GSTP1* Val/Val vs Ile/Ile, and the finding was not replicated in mothers. Most subjects in both children's and mothers' groups were White (96 and 98%, respectively), and results did not change when excluding non-White subjects from the analyses.

AMICS-INMA study

Data were available from 411 and 404 children for analyses on asthma and wheezing, respectively. *GSTP1* genotype was in HWE. No association was found between *GSTM1* or *GSTT1* and asthma or

wheezing (Figures 2, 3 and 5). *GSTP1* showed no association with asthma, but was associated with wheezing (Figures 4 and 5). The ORs for *GSTP1* and wheezing were 2.32 (1.03–5.23) and 5.01 (1.72–14.58) for Ile/Val and Val/Val vs Ile/Ile, respectively, which suggests that the increased risk of wheezing associated with the Val allele might follow an additive genetic model.

Southampton study

Among the 341 families, the analyses on asthma included 199 and 224 informative families for *GSTM1* and *GSTT1*, respectively; these figures were 206 and 227 for bronchial responsiveness. No association was found for asthma or bronchial responsiveness with either *GSTM1* or *GSTT1*, although the presence of *GSTT1* null allele was associated with an increased severity score in patients with asthma ($P = 0.015$).

Systematic review and meta-analysis

The process of inclusion and exclusion of studies is presented in Figure 1. Our initial search strategy identified 804 articles, among which the full text of 87 articles was retrieved for more detailed evaluation. Of the 30 eligible articles, 2 reported on family-based studies and 28 articles reported on 26 population-based studies. We contacted authors of 12 articles for further information on genotype counts and outcome definition, and received replies from 5.

Of the 26 eligible population-based studies, we could include in the meta-analyses data from 25 studies

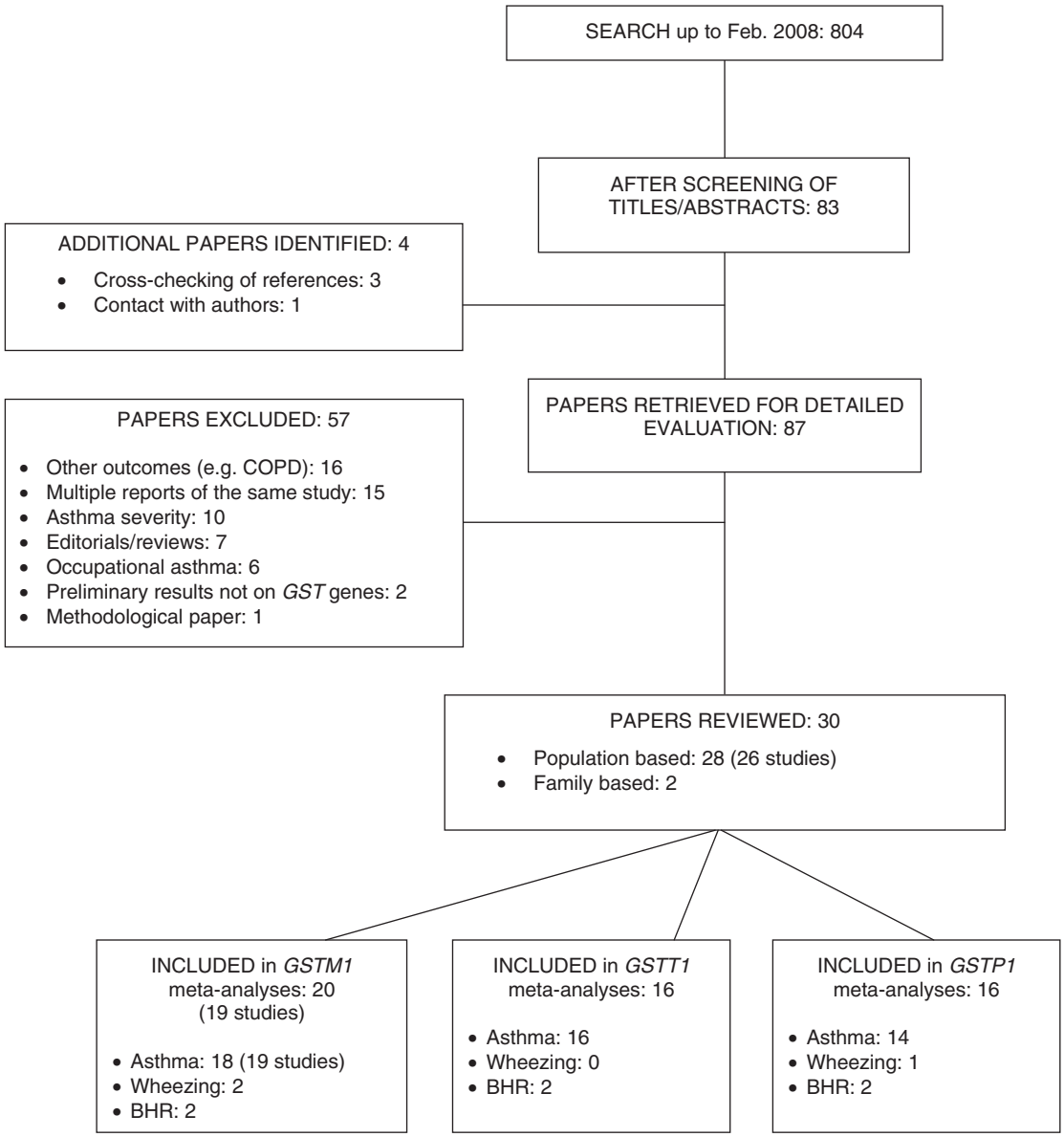


Figure 1 Flow chart of the inclusion and exclusion of published articles in the review

(reported in 27 articles), which included 2 cohorts, 2 cross-sectional and 21 case–control studies. After inclusion of the three primary studies (ALSPAC mothers, ALSPAC children, AMICS-INMA) association with asthma was evaluated in 22 studies on *GSTM1*, 19 on *GSTT1* and 17 on *GSTP1*. Asthma was defined as doctor-diagnosed asthma in 19, 16 and 12 studies on *GSTM1*, *GSTT1* and *GSTP1*, respectively. Wheezing was evaluated in four studies on *GSTM1*, two on *GSTT1* and three on *GSTP1*. BHR, defined in all studies as a FEV_1 reduction of $\geqslant 20\%$ after methacholine challenge, was considered in three studies on *GSTM1*, three on *GSTT1* and three on *GSTP1*. Overall, 10 studies were performed in children, 7 in adults, 6 in both, whereas for 2 studies the information was not available. In terms of geographical location, 13 studies were performed in Asia, 8 in Europe, 2 in North Africa and 2 in North America. Case–control studies varied in sample size, from less than 100 to approximately 1000. On the other hand, both cross-sectional studies had around 3000 subjects and three out of five cohort studies had large sample size of 4400–7200. Other characteristics of the population-based studies included in the meta-analyses are summarized in Table 1, and genotype counts by disease outcome are reported in Table 2.

Two family studies were included in the systematic review (Table 1), but only one provided complete

Table 1 Characteristics of the studies evaluating the effects of *GST* genes on asthma risk, wheezing and BHR

Study, year	Study population	Location, ethnicity	Adults/ children	Sample size[a]	Definition of disease outcome
Case–control studies ($n=21$)					
Fryer, 2000[9]	Cases: from patient database Controls: healthy volunteers	UK, European	Adults	171	**Atopic asthma** diagnosed by a physician [Criteria: a history of wheezing, cough, dyspnoea and/or chest tightness; spirometric demonstration of airflow obstruction reversible with ab-agonist bronchodilator (>15% change in FEV_1); and positive atopic status) **BHR** FEV_1 reduction >20% after methacholine]
Chung, 2002[59]	Cases and controls: from cohort of civil servants and their families	Korea, Asian	N/A	99	**Asthma** diagnosed by a physician (Criteria: symptoms of wheezing/ chest tightness and FEV_1 reduction >20% after methacholine challenge)
Freidin, 2002[60]	Cases: hospitalized patients Controls: healthy volunteers	Russia, Asian	N/A	126	**Atopic asthma** diagnosed by a physician (Criteria: N/A)
Sideleva, 2002[61]	Cases: hospitalized patients Controls: healthy volunteers	Russia, European	Both	199	**Atopic asthma** diagnosed by a physician (Criteria: the commonly accepted clinical laboratory examination and analysis of the external respiration function[62])
Vavilin, 2002;[39] Safronova, 2003[38]	Cases: hospitalized patients Controls: healthy volunteers	Russia, European	Children	237	**Asthma** diagnosed by a physician (Criteria: N/A)
Aynacioglu, 2004[41]	Cases: consecutive asthma outpatients Controls: healthy volunteers	Turkey, Asian	Adults	475	**Asthma** diagnosed by a physician (Criteria: ECRHS protocol; medical history, physical examination, lung function tests, chest X-rays, SPT, TOT IgE)
Saadat, 2004[6]	Cases: N/A Controls: healthy volunteers	Iran, Asian	Both	170	**Asthma** diagnosed by a physician [Yes to two or more criteria: (i) history of wheezing/breathlessness; (ii) airflow obstruction of >15% FEV_1; and (iii) PEF >20%]
Tamer, 2004[11]	Cases: hospitalized patients Controls: healthy volunteers	Turkey, Asian	Adults	204	**Asthma** diagnosed by a physician (Criteria: American Thoracic Society statement[63])
Zhang, 2004[36]	Cases: hospitalized patients Controls: N/A	China, Asian	Adults	120	**Asthma** diagnosed by a physician (Criteria: N/A)

(Continued)

Table 1 Continued

Study, year	Study population	Location, ethnicity	Adults/ children	Sample size[a]	Definition of disease outcome
Lee, 2005, 2008[10,64]	Cases: school children from three communities Controls: healthy random sample from same three communities	Taiwan, Asian	Children	397	**Asthma** parent-reported (positive response to the question, 'Has a physician ever diagnosed your child as having asthma?') **Wheezing** parent-reported (lifetime history)
Nickel, 2005[13]	Cases: from Multicenter Allergy Study Controls: healthy volunteers	Germany, European	Children	205	**Asthma** diagnosed by a physician (Criteria: one or more episodes of wheezing during previous 12 months)
Oh, 2005[37]	Cases: recruited at an allergy clinic Controls: healthy subjects	Korea, Asian	Both	273	**Asthma** diagnosed by a physician (Criteria: N/A)
Arbag, 2006[65]	Cases: cases of nasal polyposis Controls: healthy individuals from same geographic location and same ethnicity as cases	Turkey, Asian	Both	133	**Asthma** diagnosed by a physician (Criteria: N/A; cases were receiving regular inhaled corticosteroids and beta2-adrenoreceptor stimulants for use as required)
Ercan, 2006[1]	Cases: random sample from cohort of asthmatic children Controls: healthy school children	Turkey, Asian	Children	563	**Asthma** diagnosed by a physician [Criteria: history of intermittent wheezing and the presence of reversible airway obstruction as defined by at least a 12% improvement in FEV_1 following bronchodilator administration, therapeutic response to anti-asthma treatment, or an abnormal result in methacholine bronchoprovocation test ($PC_{20} < 8$ mg/dl)]
Holla, 2006[66]	Cases and controls are unrelated subjects selected from questionnaires	Czech Republic, European	Both	637	**Asthma** diagnosed by a physician (Criteria: asthma symptoms and use of antiasthma medication according to the GINA guidelines[67])
Plutecka, 2006[35]	Cases: N/A Controls: healthy volunteers	Poland, European	Adults	496	**Asthma** not defined
Abdel-Alim, 2007[40]	Cases: children attending outpatient clinic of allergy and immunology Controls: matched healthy children enrolled in this work	Egypt, North African	Children	90	**Asthma** diagnosed by a physician (Criteria: GINA guidelines[68])
Hanene, 2007[5]	Cases and controls from a representative region	Tunisia, North African	Children	347	**Asthma** not defined
Kamada, 2007a[69]	Cases: recruited from a medical centre Controls: healthy volunteers	Japan, Asian	Both	980	**Asthma** diagnosed by a physician (Criteria: American Thoracic Society[70])
Kamada, 2007b[69]	Cases: recruited from a medical centre Controls: healthy volunteers	Japan, Asian	Children	289	**Asthma** diagnosed by a physician (Criteria: American Thoracic Society[70])
Mak, 2007[8]	Cases: recruited from asthma clinic Controls: random healthy subjects from Hong-Kong population	China, Asian	Adults	626	**Asthma** diagnosed by a physician (Criteria: cough/wheeze/chest tightness and airflow obstruction of $\geqslant 15\%$ FEV_1)
Cohort studies ($n = 5$)					
Imboden, 2007[71]	SAPALDIA—a prospective multicentre study on adult Swiss general population, investigating environmental and genetic effects on lung	Switzerland, European	Adults	4422	**Asthma** self-reported (positive response to 'Have you ever had asthma?' and 'Was this confirmed by a doctor?') **BHR** FEV_1 reduction >20% after methacholine

(Continued)

Table 1 Continued

Study, year	Study population	Location, ethnicity	Adults/ children	Sample size[a]	Definition of disease outcome
Schroer 2009[42]	CCAAPS study on allergy and air pollution in infants with at least one atopic parent	US, mixed ethnicity	Children	498	**Wheezing** parent-reported (parental report of the child wheezing at least once at 2 years in last 12 months)
ALSPAC mothers	ALSPAC—a UK population-based birth cohort, with pregnant women recruited in 1990	UK, European	Adults	7262	**Asthma** self-reported (history of asthma reported in a questionnaire administered during pregnancy)
ALSPAC children	ALSPAC—children recruited at birth and followed up to investigate their health, behaviour and development	UK, European	Children	5330	**Asthma** parent-reported (positive response to 'Has a doctor ever said that your child has asthma?' at 7 years) **Wheezing** parent-reported (current history) **BHR** FEV_1 reduction >20% after methacholine
AMICS-INMA	AMICS—a population-based birth cohort included in the Spanish environment and childhood research network (INMA study)	Spain, European	Children	428	**Asthma** parent-reported (positive response to 'Has your child had any wheezing with whistling on his/her chest when he/she breathed in the past 12 months?' at age 6 years) **Wheezing** parent-reported (current wheeze at 6 years)
Cross-sectional studies ($n=2$)					
Salam 2007;[7] Gilliland, 2002[33]	Cross-sectional study as part of the Children Health Study, a Californian ongoing cohort study of school children in 12 communities	US, Mixed ethnicity	Children	3081	**Asthma** parent-reported DDA (Criteria: N/A) **Wheezing** parent-reported (lifetime history)
Kabesch 2004[53]	Cross-sectional study part of ISAAC project, assessing prevalence of asthma and allergy in school children	Germany, European	Children	3005	**Asthma** parent-reported DDA (Criteria: N/A)
Family studies ($n=3$)					
David 2003[34]	Family study based on case–parent triad design and focused on gene–environment interactions	Mexico, Latin American	Children	218 families	**Asthma** diagnosed by a physician (Criteria: included skin-prick and pulmonary function testing)
Brasch-Andersen 2004[4]	Family study that recruited atopic families with asthmatic children in two separate samples. Sample A: asthma; Sample B: atopic asthma	Denmark, European	Children	246 families	(Sample A) **Asthma** diagnosed by a physician (Criteria: clinical symptoms plus positive methacholine challenge) (Sample B) **Atopic asthma** diagnosed by a physician (Criteria: clinical symptoms plus positive specific IgE measurement)
Southampton	UK multicentre family study that recruited families with two siblings diagnosed with asthma	UK, European	Children and young adults	341 families	**Asthma** self and parent reported (Criteria: three positive responses to: 'Have you ever had asthma?' and, 'Was this confirmed by a doctor?' and, 'Have you used any medicines to treat asthma, or any breathing problems, at any time in the last 12 months') **Bronchial responsiveness** methacholine dose response measure

N/A: not available; DDA: doctor-diagnosed asthma; GINA: global initiative on asthma.
[a]When results are reported for more genotypes or outcomes, the largest sample size is reported.
ECRHS: european community respiratory health survey; SPT: skin prick tests; TOT: total; SAPALDIA: swiss cohort study on air pollution and lung and heart diseases in adults; CCAAPS: cincinnati childhood allergy and air pollution study; ISAAC: international study of asthma and allergies in childhood

Table 2 Summary of genotype frequencies by disease status in the population-based studies

	GSTM1 (null/present) [CNV: 0/1/2 copies]		*GSTT1* (null/present) [CNV: 0/1/2 copies]		*GSTP1* (Ile105Ile/Ile105Val/Val105Val)		
Study, year	Affected	Non-affected	Affected	Non-affected	Affected	Non-affected	Disequilibrium coefficient (*P*-value)[a]
Asthma							
Fryer, 2000[9]	72/53	24/20	24/103	7/37	54/66/7	10/26/8	−0.046 (*0.36*)
Chung, 2002[59]	16/17	31/35	21/12	34/32	25/8/0	51/15/0	−0.013 (*0.59*)
Freidin, 2002[60]	52/17	42/15	24/45	15/42			
Sideleva, 2002[61]	83/26	43/47	73/36	21/69			
Vavilin, 2002[39] Safronova, 2003[38]	52/48	44/60	26/74	12/92	54/77/10	32/61/3	−0.091 (*<0.01*)
Aynacioglu, 2004[41]					109/93/8	134/99/32	0.026 (*0.04*)
Kabesch, 2004[53]	148/120	1394/1343	42/226	477/2258			
Saadat, 2004[6]	42/43	20/65	55/30	35/50			
Tamer, 2004[11]	64/37	42/61	27/74	25/78	33/45/23	49/46/8	−0.013 (*0.64*)
Zhang, 2004[36]	49/11	18/42	43/17	7/53			
Lee, 2005[10]	49/33	97/87			62/18/2	112/64/8	−0.004 (*1.00*)
Nickel, 2005[13]					38/39/6	54/60/8	−0.031 (*0.14*)
Oh, 2005[37]					102/45/7	75/39/5	−0.000 (*1.00*)
Arbag, 2006[65]	12/19	47/55	11/20	24/78			
Ercan, 2006[1]	124/179	100/151	64/246	48/202	174/116/22	128/98/25	0.013 (*0.36*)
Holla, 2006[66]	166/140	166/165	59/247	73/258			
Plutecka, 2006[35]					137/123/26	94/95/21	−0.006 (*0.75*)
Abdel-Alim, 2007[40]					32/4/14	9/2/29	0.163 (*<0.01*)
Hanene, 2007[5]	84/37	111/115	41/80	67/159	63/45/13	78/107/41	0.007 (*0.68*)
Imboden, 2007[71]	73/71	2249/2029	25/119	797/3481			
Kamada, 2007a[69]	178/189	326/287					
Kamada, 2007b[69]	57/57	95/80					
Mak, 2007[8]	161/150	185/130	144/167	168/147	207/94/11	214/91/9	−0.001 (*1.00*)
Salam, 2007[7]	240/215	1226/1300	91/363	537/1992	171/230/65	1043/1229/343	−0.004 (*0.37*)
ALSPAC children	419/333	2419/2156	124/621	786/3769	307/360/92	1985/2006/585	0.005 (*0.10*)
	[419/249/49]	[2419/1689/306]	[124/313/227]	[786/1917/1276]			
ALSPAC mothers	239/223	3523/3064	74/386	1069/5488	337/377/105	2684/2935/755	−0.003 (*0.26*)
	[239/179/28]	[3523/2388/435]	[74/164/148]	[1069/2787/1782]			

(Continued)

Table 2 Continued

Study, year	*GSTM1* (null/present) [CNV: 0/1/2 copies]		*GSTT1* (null/present) [CNV: 0/1/2 copies]		*GSTP1* (Ile105Ile/Ile105Val/Val105Val)		
	Affected	Non-affected	Affected	Non-affected	Affected	Non-Affected	Disequilibrium Coefficient (*P*-value)[a]
AMICS-INMA	11/7	224/169	4/14	74/319	8/8/2	177/183/33	−0.016 (*0.17*)
Wheezing							
Gilliland, 2002[33]	455/529	816/1010					
Lee, 2008[64]	128/85	99/85					
Schroer, 2009[42]					40/41/18	186/155/58	0.03 (*0.008*)
ALSPAC children	358/289	2831/2513	114/530	898/4415	256/330/68	2297/2364/677	0.002 (*0.41*)
AMICS-INMA study	19/17	211/157	7/29	71/297	9/20/7	174/167/27	−0.015 (*0.17*)
BHR							
Fryer 2000[9]	30/22	24/20	7/46	7/37	21/29/3	10/26/8	−0.046 (*0.36*)
Imboden 2007[71]	327/298	1995/1802	115/510	707/3090	253/266/48	1792/1650/358	−0.004 (*0.17*)
ALSPAC children	278/246	1515/1391	81/437	477/2407	210/251/64	1206/1321/365	−0.001 (*0.85*)

[a]HWE tested (exact test) among controls in case–control studies, and in the whole population otherwise.

results[4] and no meta-analysis could be performed. The study by Brasch-Andersen and colleagues,[4] consisting of 246 Danish atopic families (452 asthmatic children plus their parents) recruited in two separate samples, evaluated the risk of asthma and atopic asthma associated with the three genes. *GSTM1* and *GSTT1* were analysed as CNV with full gene-dosage information available. The study found strong evidence of an association of the *GSTM1* homozygous deletion with asthma ($P < 0.0005$), which became even stronger when limiting the analyses to atopic asthma ($P < 0.00005$). An association of *GSTT1* with asthma was also found, but under an additive model ($P = 0.019$), with similar results obtained for atopic asthma ($P = 0.021$), whereas no association was found with *GSTP1*. The study by David and colleagues,[34] consisting of 218 Mexican case–parent triads, focused on the interaction between *NQO1* (Pro187Ser polymorphism) and *GSTM1* genes in increasing the risk of asthma in children exposed to high level of ozone (Mexico City). The study, which evaluated *GSTM1* in terms of null genotype (homozygous deletion), did not assess the effect of *GSTM1* alone, but showed a protective effect of the *NQO1* Ser allele in children with null *GSTM1* genotype.

Quality of the studies included

We assessed the quality of the 27 articles, reporting on the 25 studies included in the meta-analyses, and the 2 family-based articles. Three articles[6,35,36] did not describe the study population (selection of cases and/or controls for case–control studies), and six articles[5,35–39] did not provide a definition of the outcome considered (asthma in all cases). HWE was reported to be assessed in less than half of the articles (46%). All studies described the genotyping methods used. Genotyping error was evaluated in a quarter of them, mainly by repeating the genotyping in 5–15% of the samples, and no errors were found in most cases. However, only 14% of the studies reported blinding of genotyping to the clinical status of the subjects. Among the population-based studies, 35% of the articles discussed the issue of population stratification, although only a few discussed the potential problem in some detail.

Four studies were identified as being of poor quality (no outcome definition plus incomplete description of study population/no assessment of HWE). All of them were case–control studies with total sample size less than 500, and they were published in Russian,[39] Chinese,[36] Korean[37] and international[35] journals.

Table 2 shows deviation from HWE for *GSTP1* gene in four studies of relatively small size (less than 500), including three case–control studies of rather poor quality,[38,40,41] and a cohort study[42]. Only two of them[41,42] performed quality control of genotyping and found no genotyping error, and none of them reported blinded genotyping. Only one of them[42]

mentioned the possibility that population stratification might have occurred.

Meta-analyses on GSTM1 deletion polymorphism

A total of 19 published and 3 primary studies evaluated the association of *GSTM1* and asthma, including a total of 4416 affected and 23 902 non-affected individuals. The meta-analysis in Figure 2a shows an increased risk of asthma associated with the *GSTM1* null genotype (pooled OR 1.28; 95% CI 1.09–1.52). However, large between-study heterogeneity was observed ($I^2 = 76\%$), which could not be explained by age (adults/children) or ethnicity, as approximated by study geographical location (continent). These subgroup analyses, however, showed that most heterogeneity was present in studies on adults and in those performed in Asia (Supplementary Figure S2 available as supplementary data at *IJE* online). Similarly, exclusion of two studies with poor quality[36,39] reduced the pooled OR to 1.19 (1.03–1.38) with little reduction of the I^2 (69%). The subgroup analysis by asthma definition (Figure 2b) could not explain heterogeneity either, although it showed that heterogeneity was limited to studies where asthma was diagnosed by a physician and where asthma was not defined in the article, whereas it was absent among studies with self-reported asthma. However, studies using self-reported asthma included those with the largest sample size, therefore it is difficult to disentangle the impact of asthma definition from that of study size on the observed heterogeneity. In fact, the funnel plot in Figure 2c strongly supports the presence of small-study bias. When the meta-analysis was repeated by limiting inclusion to the nine studies at the top of the funnel, the heterogeneity dropped from 76 to 38% and the OR became very close to 1 (1.02; 0.92–1.13).

As for the association of *GSTM1* null genotype with wheezing and BHR, only four and three studies were available, respectively. As shown in Figure 5a, for both outcomes results were very homogeneous across studies. The meta-analysis showed no effect on BHR (1.01; 0.90–1.15), and only a possible small effect on wheezing (1.08; 0.99–1.19).

Meta-analyses on GSTT1 deletion polymorphism

A total of 16 published and 3 primary studies evaluated the association of *GSTT1* and asthma, including 3852 affected and 22 880 non-affected individuals. Similarly to *GSTM1*, the meta-analysis in Figure 3a shows an OR of 1.39 (1.09–1.77), but with extreme between-study heterogeneity ($I^2 = 81\%$). As for *GSTM1*, heterogeneity was absent within the subgroup of studies with self-reported asthma (Figure 3b), and small-study bias was the only factor substantially explaining heterogeneity. When restricting the meta-analysis to the nine largest studies at the top of the funnel plot in Figure 3c, the OR became 0.93 (0.84–1.03) and the heterogeneity disappeared ($I^2 = 0\%$). On the contrary, exclusion of the two studies with poor quality[36,39] reduced the pooled OR to 1.19 (0.97–1.44), but did not reduce substantially the degree of heterogeneity ($I^2 = 70\%$). Subgroup analyses by age and ethnicity did not provide additional insight on the causes of heterogeneity (Supplementary Figure S3 available as supplementary data at *IJE* online).

Similarly to *GSTM1*, the association of *GSTT1* with wheezing and BHR (Figure 5b) showed homogeneous results across studies (two and three studies included, respectively). The meta-analyses showed no genetic effect for either outcome (wheezing 1.05; 0.86–1.30; BHR 0.96; 0.81–1.13).

Meta-analyses on GSTP1 Ile105Val polymorphism

A total of 14 published and 3 primary studies with 3363 affected and 14 442 non-affected individuals were available for the meta-analysis of *GSTP1* Ile/Val vs Ile/Ile and asthma, whereas 13 published and 3 primary studies with 2160 affected and 9034 non-affected individuals were available for the contrast Val/Val vs Ile/Ile (Figure 4; Supplementary Figure S4 and S5 available as supplementary data at *IJE* online). The meta-analyses in Figure 4 show no clear association of *GSTP1* with asthma, although the two pooled estimates point to a possible protective effect of the Val allele, with an OR of 0.93 (0.82–1.06) for Ile/Val vs Ile/Ile and 0.79 (0.57–1.08) for Val/Val vs Ile/Ile. Although between-study heterogeneity was moderate for Ile/Val vs Ile/Ile ($I^2 = 39\%$), it was large for Val/Val vs Ile/Ile (72%), so that the result for the latter is difficult to interpret. Heterogeneity could not be explained by small-study bias (Supplementary Figure S5c available as supplementary data at *IJE* online), and was not reduced by removing the two studies with poor quality[35,37] either. Ethnicity did explain some heterogeneity (Supplementary Figure S5b available as supplementary data at *IJE* online), and exclusion of the two North African studies (Egypt and Tunisia) brought the ORs close to 1 (Ile/Val vs Ile/Ile 0.98; 0.87–1.10, and Val/Val vs Ile/Ile 0.94; 0.71–1.24) and reduced I^2 to 59%. Pooling of the two African studies showed a marked protective effect (Ile/Val vs Ile/Ile 0.52; 0.33–0.83, and Val/Val vs Ile/Ile 0.24; 0.09–0.69). The subgroup analysis by asthma definition showed similar results as for *GSTM1* and *GSTT1*, with no heterogeneity within studies with self-reported asthma (Figure 4c and d).

Only three studies were available for the association of *GSTP1* with wheezing and BHR, and meta-analysis could only be performed for the comparison of Ile/Val vs Ile/Ile where the heterogeneity was small (Figure 5c and d). For the Ile/Val vs Ile/Ile comparison there is a suggestion of an increased risk associated with the Val allele for both outcomes. Effects in opposite directions are shown for the Val/Val vs Ile/Ile comparison, with the Val allele having a protective effect on BHR as with asthma, but being associated

with increased risk of wheezing. These findings need to be interpreted with caution since they are driven by two relatively small studies, the study by Fryer[9] and the AMICS-INMA study, respectively.

The results of the analyses based on the model-free approach were similar to those based on between-genotype comparisons, giving an OR of 0.93 (0.77–1.03) for Ile/Val vs Ile/Ile and 0.77 (0.49–1.14) for Val/Val vs Ile/Ile. The model-free approach provided an estimate of λ, the parameter indicating the underlying genetic model, of 0.27 (0.01–0.61), where 0 corresponds to the recessive, 0.5 to the co-dominant and 1 to the dominant model.[26] Despite the wide CI, the estimate of λ seems to exclude a dominant genetic model and suggests either a co-dominant or a recessive model.

Strength of the evidence from the meta-analyses

When applying the Venice criteria to assess credibility,[25] the cumulative evidence reviewed does not clearly support a genetic effect for any of the three *GST* genes on asthma. In fact, although the amount of evidence is good for all three meta-analyses, particularly those on *GSTM1* and *GSTT1*, the degree of heterogeneity across studies and the likely presence of publication bias make the credibility of the evidence weak. Moreover, when limiting the meta-analyses to the largest studies, positive findings shown by other studies could not be replicated, suggesting that a genetic effect is likely to be either absent or very small.

Discussion

Overall, the evidence synthesized in this review does not support a substantial role of *GST* genes on asthma phenotypes in either children or adults, although small effects cannot be excluded and it is possible that these genes act on airway disease through interaction with environmental exposures or other genes. The weakness of the available evidence in supporting an effect of *GST* genes on asthma is mainly due to publication bias and large heterogeneity of study results. Interestingly, the results from the three

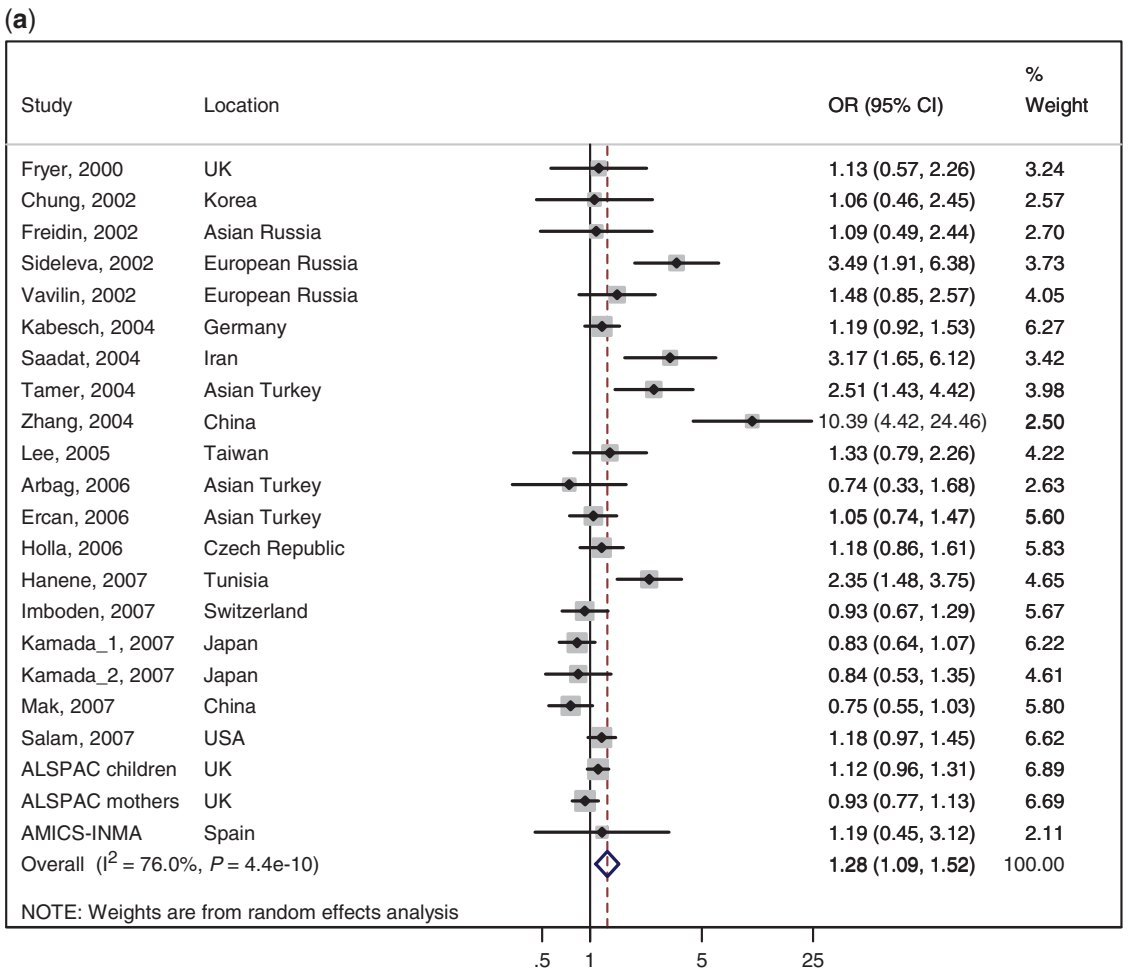
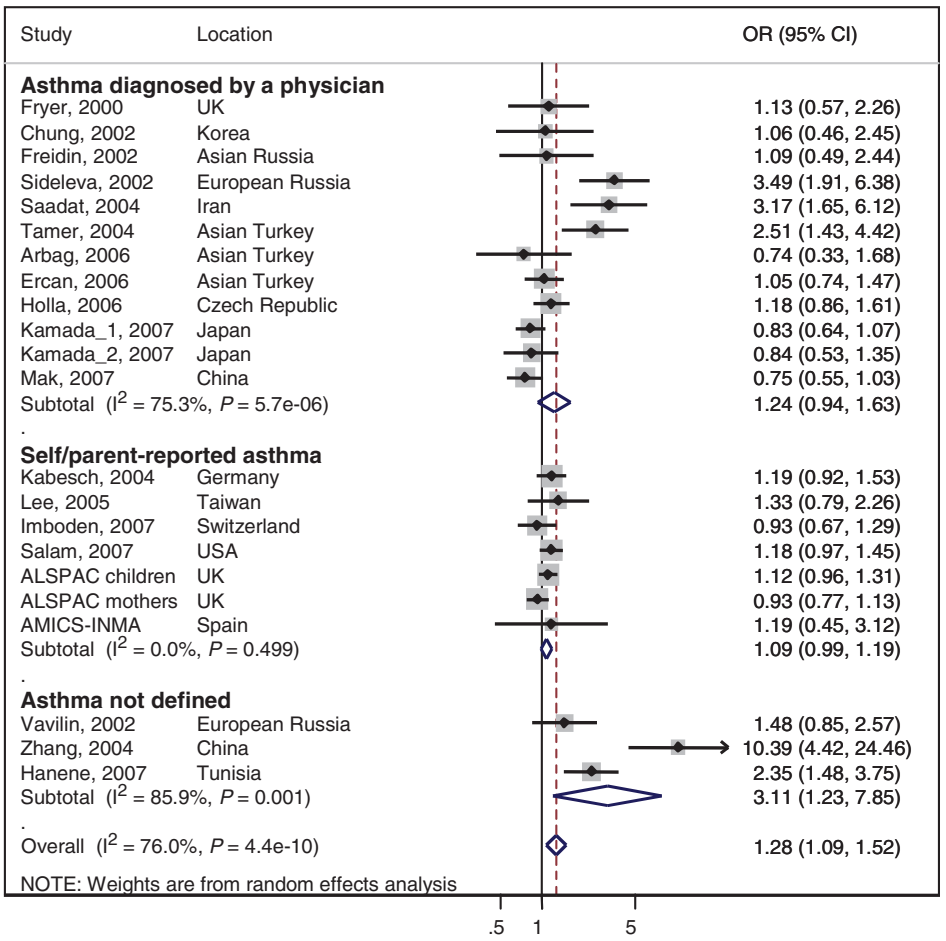


Figure 2 Random effects meta-analysis of *GSTM1* (null vs present) effect on asthma: (a) forest plot for the main analysis; (b) forest plot for the subgroup analysis by asthma diagnosis; (c) funnel plot, where studies classified as 'small' or 'large' in the sensitivity analyses are represented by dots and triangles, respectively (Egger's test: $P = 0.035$; Begg's test: $P = 0.135$)

(**b**)



(**c**)

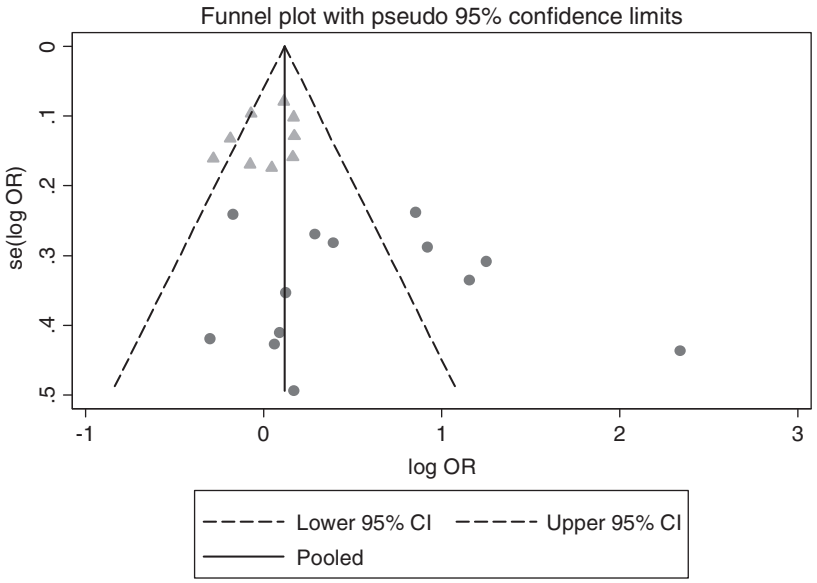


Figure 2 Continued

unpublished studies on *GST* genes and asthma phenotypes were negative. The only exception was the association between *GSTP1* and wheezing (but not asthma) in the AMICS-INMA study, which suggested an increased risk for the Val allele. It is important to note that the ALSPAC, which provided negative results consistently for all *GST* genes and asthma phenotypes, was the largest study included in all the meta-analyses. In fact, the plan for the systematic review and meta-analysis presented here originated with the aim of putting the results of the ALSPAC cohort into context, and understand the reasons for the apparent disagreement with previous published studies. The lack of evidence of an important role of *GST* genes in the development of asthma, wheezing and BHR is in agreement with negative findings on lung function in children for all three *GST* genes from the ALSPAC (data not shown), although an association between *GSTM1* and *GSTP1* genes and lung function in childhood has been previously suggested.[43]

Although the meta-analyses on *GSTM1* and *GSTT1* with inclusion of all studies showed an increased risk of asthma associated with the two null genotypes, the presence of small-study bias and extreme heterogeneity in study results make the credibility of these findings very low. Despite our efforts to be as inclusive as possible (comprehensive search strategy, avoidance of language restrictions, inclusion of conference abstracts), the meta-analyses showed clear absence of small studies with negative results, suggesting the presence of publication bias. This is an important issue in meta-analysis of genetic association studies, which needs to be highlighted and addressed. We did so by repeating the analyses with inclusion of only the largest studies, and we found loss of the observed associations, with ORs very close to 1. Positive results were not only confined to small studies, but also to studies with poor quality, the exclusion of which caused similar loss of the observed associations. In support of an absence of an association of *GSTM1* and *GSTT1* with airway disease are the negative findings of the meta-analyses on wheezing and BHR. An exception might be the association between *GSTM1* and wheezing, where the

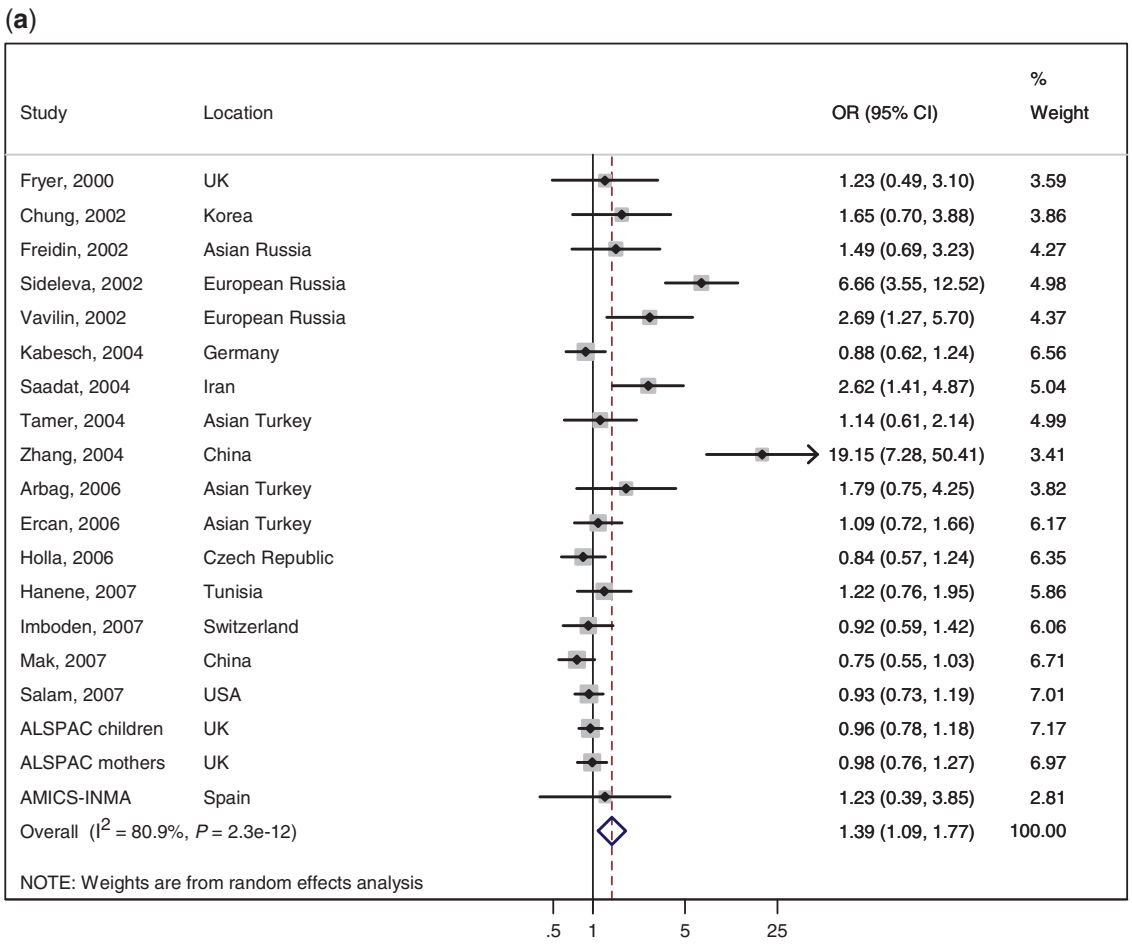


Figure 3 Random effects meta-analysis of *GSTT1* (null vs present) effect on asthma: (a) forest plot for the main analysis; (b) forest plot for the subgroup analysis by asthma diagnosis; (c) funnel plot, where studies classified as 'small' or 'large' in the sensitivity analyses are represented by dots and triangles, respectively (Egger's test: $P = 0.003$; Begg's test: $P = 0.001$)

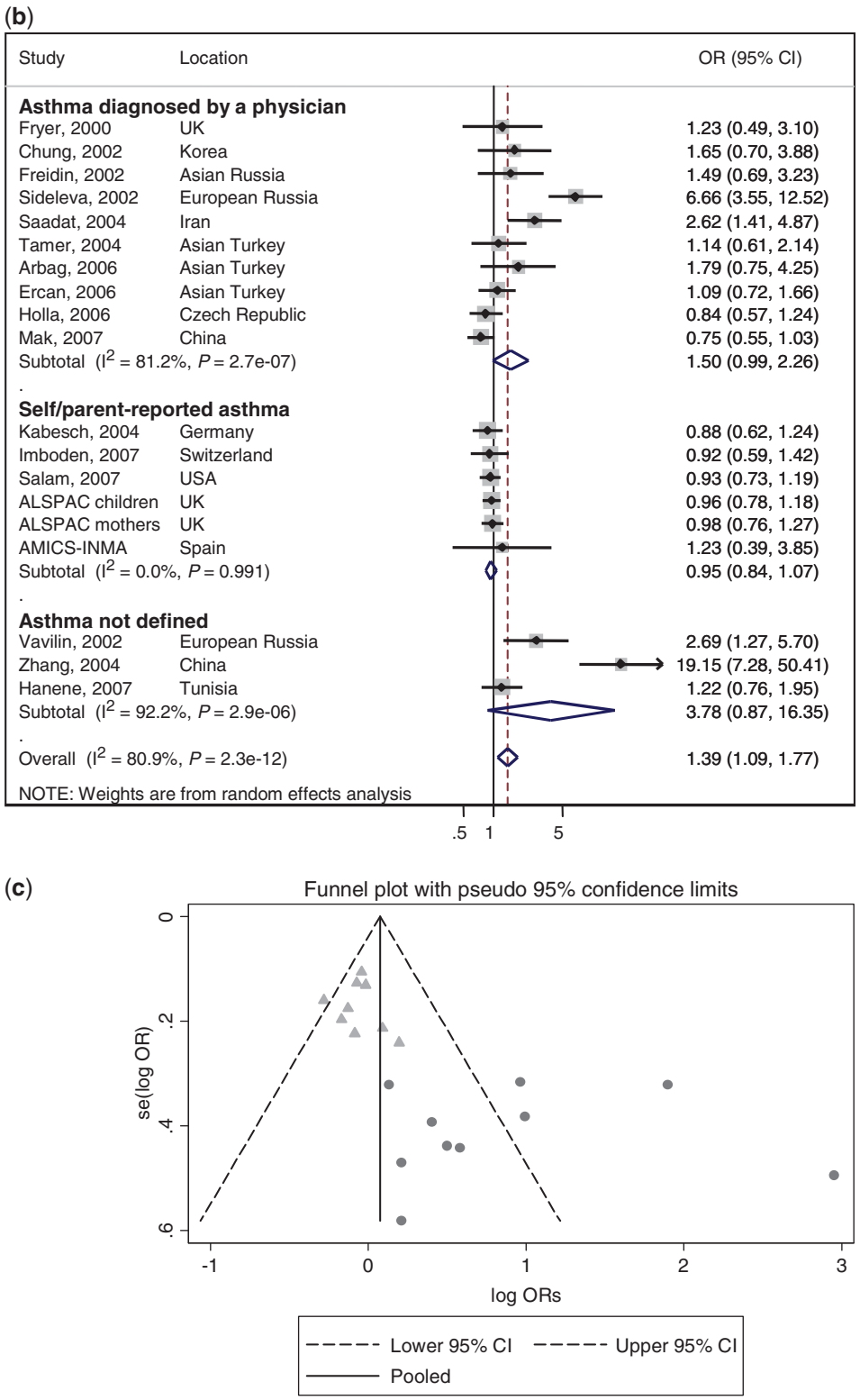


Figure 3 Continued

meta-analysis suggests a possible small effect of the null genotype in increasing the risk of ~10%. Although the studies included in the meta-analyses on wheezing and BHR were only few and the results supported only weak evidence of associations, the consistency in results and narrow CIs of the ORs give some credibility to these findings. An important methodological point regarding studies on *GSTT1* or

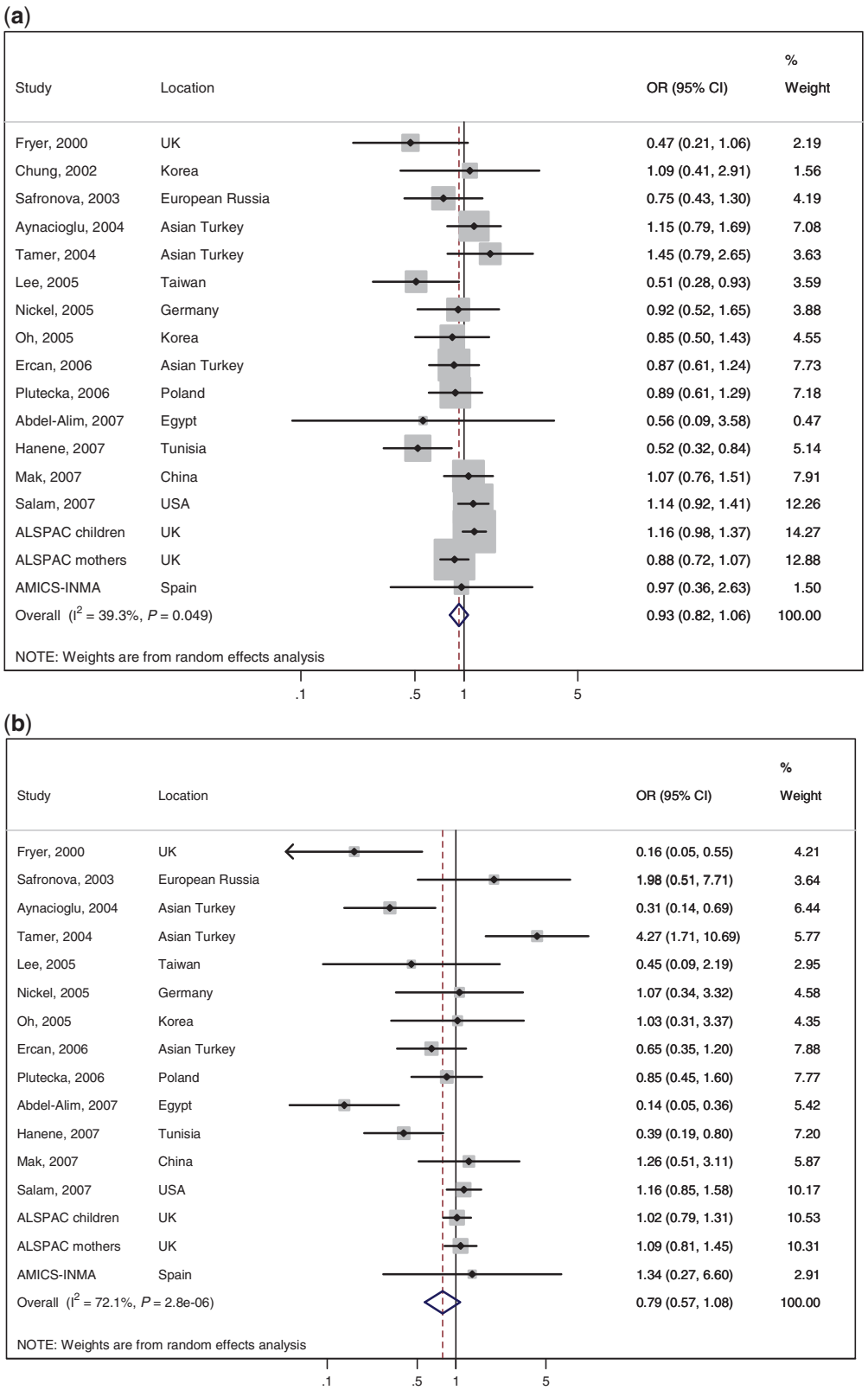


Figure 4 Random effects meta-analysis of *GSTP1* effect on asthma: (a) Ile/Val vs Ile/Ile; (b) Val/Val vs Ile/Ile; (c) subgroup analysis by asthma diagnosis for Ile/Val vs Ile/Ile; (d) subgroup analysis by asthma diagnosis for Val/Val vs Ile/Ile

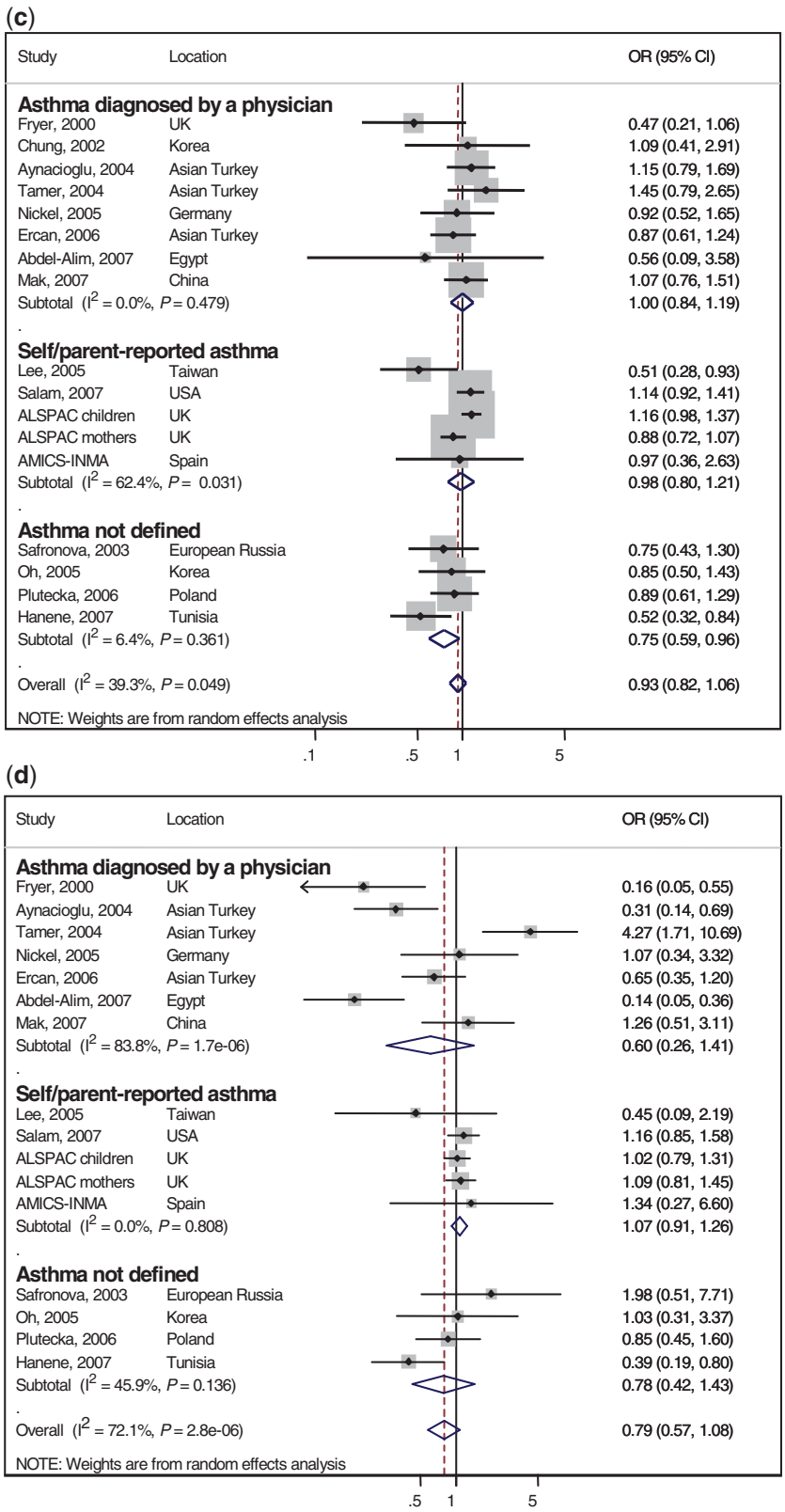


Figure 4 Continued

GSTM1 is that methods currently used to evaluate these genes cannot distinguish between genotypes with one or two copies of the gene. By classifying the genotype as 'present' or 'null' they imply a recessive model (one or two copies vs absence of the risk allele), which may not reflect the true underlying genetic model and thus may not provide a valid and accurate estimate of the genetic risk. *GSTT1* or *GSTM1* CNVs (also known as 'gene dosage') are correlated with altered enzyme activity, and analysis in a dose-dependent manner would best describe any disease outcome association.[44,45] Brasch–Andersen and colleagues have shown that it is possible to utilize the increased sensitivity of real-time PCR assays to provide dosage of *GSTT1* and *GSTM1*.[4] A similar approach

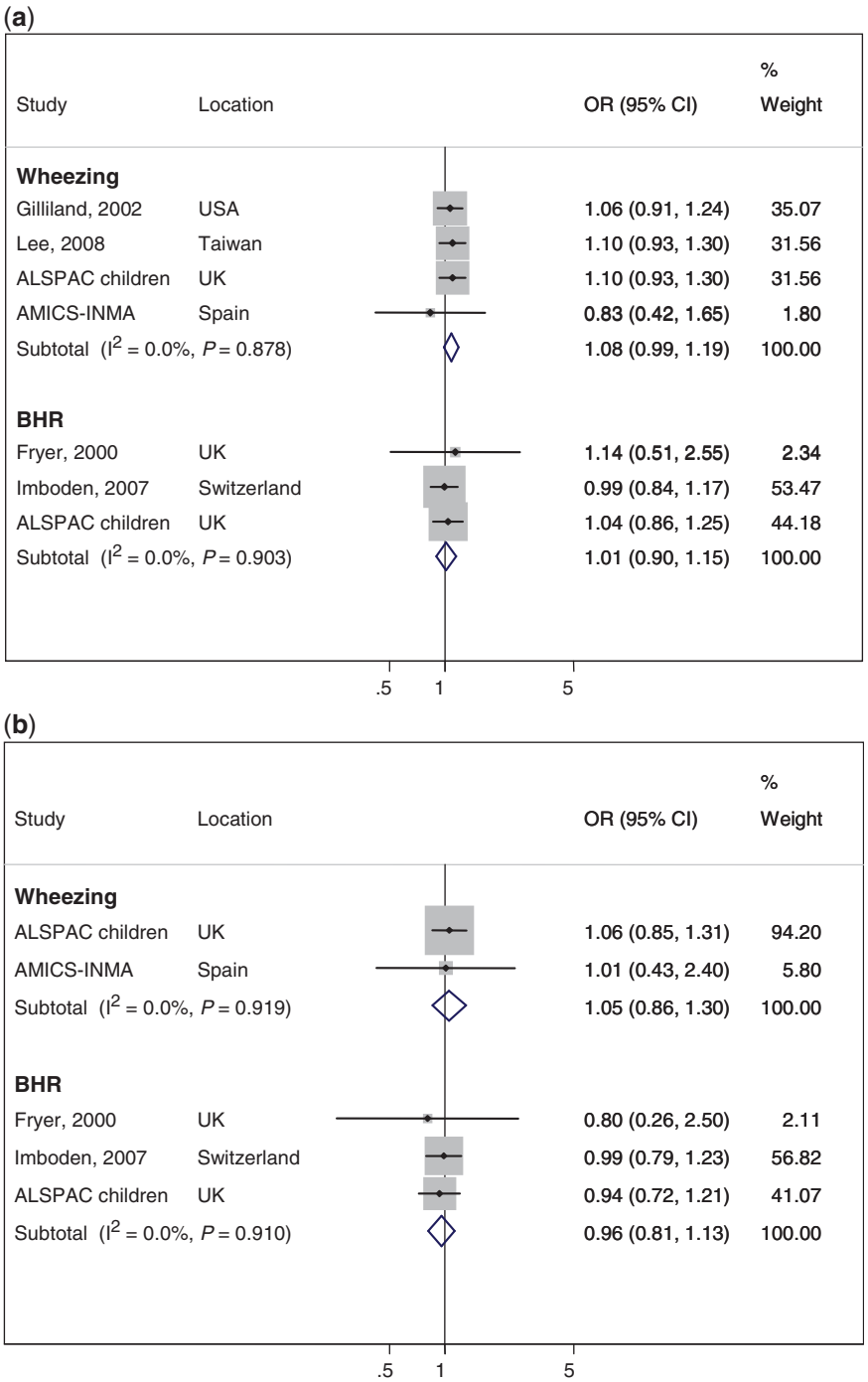


Figure 5 Results on wheezing and BHR for: (a) *GSTM1*; (b) *GSTT1*; (c) *GSTP1* Ile/Val vs Ile/Ile; (d) *GSTP1* Val/Val vs Ile/Ile. Fixed effects meta-analysis was performed for *GSTM1*, *GSTT1* and *GSTP1* Ile/Val vs Ile/Ile

was utilized to obtain gene-dosage information in the ALSPAC cohort.[17]

The meta-analysis on *GSTP1* Ile105Val polymorphism and asthma suggested a possible, weak protective effect of the Val allele. The Val substitution in the *GSTP1* gene is associated with altered substrate affinities compared with the Ile105 wild-type allele[46] and heterozygosity has also been associated with reduced risk of chronic obstructive pulmonary disease.[47] Although in the meta-analysis on *GSTP1* and asthma there was no clear evidence of small-study bias, heterogeneity in study results was large and limited the interpretability of the pooled estimates. Subgroup analyses showed a potential role of ethnicity in explaining such heterogeneity, with a protective effect of the Val allele limited to the two North African studies, where the association was very strong. This finding might be explained by the presence of gene–environment or gene–gene interactions in these study populations. *GST* genes effects on

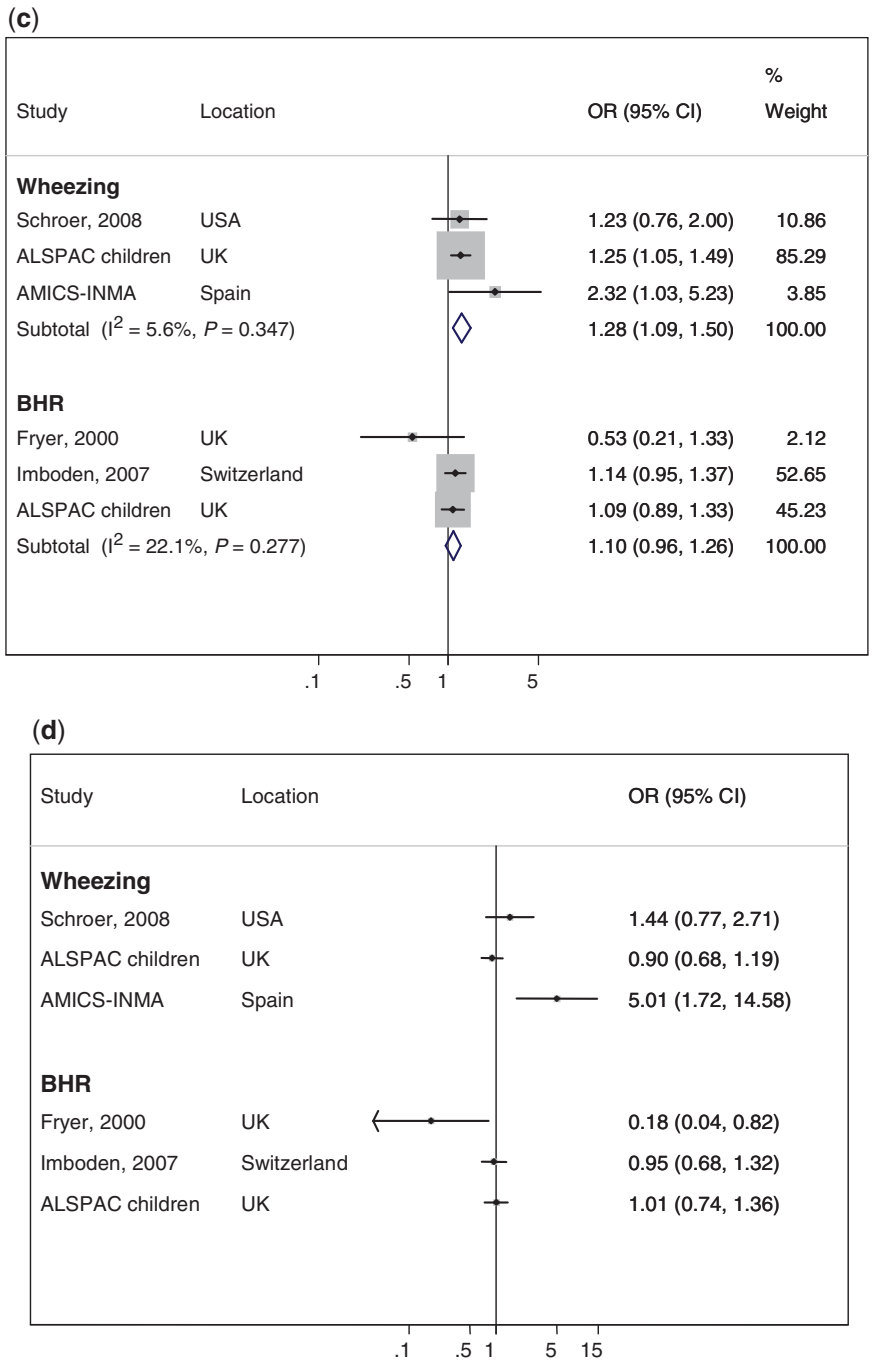


Figure 5 Continued

asthma risk may be modified by exposures that produce lung oxidative injury, such as smoking, air pollution and infections. For example, one might speculate that smoking in North African countries is a major problem, and the common waterpipe smoking seems to produce even more free radicals than cigarette smoking.[48] The value of such speculations is limited in the absence of supporting evidence, and bias associated with selective reporting or poor quality might be an alternative explanation for the positive findings of the two North African studies. Future research in these populations should collect information on factors potentially interacting with *GST* genes in modifying asthma risk, including oxidative exposures. Only few studies evaluated the effects of *GSTP1* on wheezing and BHR, and the findings were in opposite directions for the two outcomes. The Val allele was protective for BHR, in agreement with the results for asthma, but associated with increased risk of wheezing. These effects were suggested by two relatively small studies, whereas the large ALSPAC cohort did not provide convincing evidence of any association with either outcome.

Failure to account for environmental exposures in our meta-analyses might partly explain not only the heterogeneity of results across studies, but also the overall negative findings. Strong environmental effects on asthma phenotypes could mask modest genetic effects and, more importantly, gene–environment interactions could make the effects of *GST* genes become substantial only in the presence of oxidative exposures and not detectable at a population level.[49] Passive smoking, ambient air pollution and endotoxin or other pathogen-associated molecules are good candidates for gene–environment interactions in asthma, particularly with genes involved in antioxidant defence, such as the *GST* genes.[49–51] In our review, variation in exposure to these environmental factors across studies is likely to have happened given the diverse geographical setting of the studies included, and gene–environment interactions might partly explain the large heterogeneity observed. Moreover, there is evidence that antioxidant supplementation can modify these gene–environment interactive effects,[52] so that the nutritional status of the study population could represent an additional source of heterogeneity. The evaluation of gene–environment interactions is problematic due to the lack of power of statistical tests for interactions and the high measurement error present in the assessment of most environmental exposures. In fact, despite the strong biological rationale, results from the literature on gene–environment interactions in asthma remain inconclusive.[51] Only a small proportion of the studies reviewed in this article reported on gene–environment interactions, including environmental tobacco smoke (ETS),[33,39,53] *in utero* ETS[33,53] and air pollution,[7,54] and the same environmental exposure was assessed in different ways across studies. Researchers interested in the effects of *GST* genes should be encouraged to collect information on relevant environmental exposures and carefully choose the methods to measure them. Standardization of methods for environmental exposure assessment and full reporting of the interactions tested will allow the pooling of data across studies and to reach adequate power to detect interactions. The study of possible interactions between *GST* genes and environmental exposures on asthma is important not only for our understanding of the etiopathogenesis of this complex disease, but also for its possible implications for public health. Asthma prevalence is high and increasing over time worldwide, and *GST* genes polymorphisms are frequent in the population, particularly the *GSTM1* null genotype, with reported frequencies in different ethnic groups varying from 18 to 66%.[55] The possibility of implementing simple measures such as antioxidant diet supplementation in subjects at risk makes this topic worthwhile for further research.

GSTs Mu-1, Theta-1 and Pi-1 have overlapping substrate specificities, so that a deficiency in one isoform may be compensated by another. A meta-analysis of 14 studies showed that adults with a combined *GSTM1* and *GSTT1* null genotype were at a higher risk of asthma (OR 2.15, 95% CI 1.21–3.71) and suggested that smoking overloads the capacity of either *GSTM1* or *GSTT1* detoxification systems.[56] We investigated the effect of the combined null genotype using published data from 16 studies and unpublished data from the ALSPAC and AMICS-INMA study (data not shown). Although we found an association with increased asthma risk (OR 1.39, 95% CI 1.16–1.67), the extreme between-study heterogeneity ($I^2 = 88\%$) and presence of publication bias greatly limit the credibility of these findings. Moreover, the results from the ALSPAC, which represented the largest sample in the meta-analysis, were negative. On the other hand, *GST* genes could interact with genes coding for other detoxifying enzymes induced in response to oxidative stress. Supporting this hypothesis is some evidence of interaction between *GSTM1* null genotype and *NQO1* Pro187Ser polymorphism on asthma.[34] Future large studies evaluating *GST* genes in addition to other antioxidant genes are needed to provide evidence on gene–gene interactive effects on asthma.

This systematic review only focussed on asthma risk, and did not consider the possible association of *GST* genes with asthma severity in patients affected by the disease. A finding from the Southampton study was the association of the *GSTT1* null allele with an increased severity score in patients with asthma (data not shown). There is evidence suggesting that *GST* genes, in particular *GSTM1* and *GSTP1*, might also interact with air pollution and tobacco smoke exposures in exacerbating respiratory symptoms and decreasing lung function in asthmatic individuals.[12,32,57]

Despite our attempt to investigate causes of heterogeneity in study results in the meta-analyses presented, we were limited by the availability of relevant information in the published reports. For example, although we did not find evidence of effect modification by ethnicity and age at onset, we cannot exclude the contribution of these factors to the observed heterogeneity, given that we used crude proxies for them (study geographical location for ethnicity, and age of study population instead of age at onset for childhood vs adult asthma). Differences in asthma definition may also have played a role in generating the observed heterogeneity. Asthma diagnosed by a physician, self-reported doctor-diagnosed asthma and self-reported history of asthma differ in sensitivity and specificity.[58] Moreover, asthmatic individuals identified through questionnaire in a population-based study may have lower severity than patients recruited at a clinic. Although asthma definition could not explain heterogeneity in our meta-analyses, heterogeneity was not present within the subgroup of studies that used self-reported asthma, where effects were closer to the null hypothesis compared with studies that used physician-diagnosed asthma and those that did not report how asthma was defined. However, interpretation of this finding is difficult, since we could not disentangle the possible effect of asthma definition from that of study design and study size. Physician-diagnosed asthma was used by relatively small case–control studies, whereas self-reported asthma was used by large cohort and cross-sectional studies. The presence of publication bias, for example, could in theory explain the fact that heterogeneity was limited to the smallest studies, which happened to be case–control studies with physician-diagnosed asthma.

In conclusion, the evidence reviewed in this article does not support a substantial role of *GST* genes alone in the development of asthma. However, given the potential for interactions between *GST* genes and environmental toxins known to cause oxidative damage to the lungs, future research should be planned to explore gene–environment interactions. Large studies with accurate measurement of the environmental exposure are needed in order to reach adequate power to detect such interactions. Similarly, further research should evaluate possible interactions between *GST* genes and other genes involved in the antioxidant pathway. Future studies will also need to improve their methodological quality and the reporting of their findings, in order to increase the credibility of the evidence accumulating over time.

Supplementary data

Supplementary data are available at *IJE* online.

Funding

S.O.S. is an Asthma UK Senior Research Fellow. R.B.N. is supported by the UK Department of Health Policy Research Programme. The UK Medical Research Council, the Wellcome Trust and the University of Bristol provide core support for ALSPAC. The ALSPAC genetics study was funded by the British Lung Foundation. The Southampton family analysis was supported by the Asthma, Allergy and Inflammation Research Trust. The AMICS-INMA Cohort was supported by the Spanish Ministry of Health (grants FIS 97/0588, 00/0021-02, G03/176 and PI06-1756) and the genetics study performed at the PRBB (Barcelona) under the direction of Rafael de Cid and Mariona Bustamante (FIS PI041436). Funding to pay the Open Access publication charges for this article was provided by The Wellcome Trust.

Acknowledgements

The authors are very grateful to Dr LK Siew, Dr E Leszkowicz and Dr S Volkov for the translation of articles that were only available in foreign languages, and to the authors of the articles used for the meta-analysis for providing additional data. The authors are extremely grateful to all individuals and families who took part in the ALSPAC, Southampton and AMICS-INMA studies, and they thank the members of the three study teams for all their help.

Conflict of interest: None declared.

KEY MESSAGES

- Oxidative stress is involved in the pathogenesis of asthma, and glutathione-S-transferase (GST) enzymes may therefore influence asthma risk.
- Common polymorphisms of the *GSTM1, GSTT1* and *GSTP1* genes have been associated with asthma in children and adults, but published results are inconsistent across studies.
- Our meta-analyses do not support a substantial role of the *GST* genes alone in the development of asthma, and they suggest presence of publication bias.
- Future studies should focus on possible interactions of the *GST* genes with environmental oxidative exposures and with other genes involved in the antioxidant pathway.

References

[1] Ercan H, Birben E, Dizdar EA *et al*. Oxidative stress and genetic and epidemiologic determinants of oxidant injury in childhood asthma. *J Allergy Clin Immun* 2006;**118:** 1097–104.

[2] Rahman I, Biswas SK, Kode A. Oxidant and antioxidant balance in the airways and airway diseases. *Eur J Pharmacol* 2006;**533:**222–39.

[3] Rahman I, MacNee W. Oxidative stress and regulation of glutathione in lung inflammation. *Eur Respir J* 2000;**16:** 534–54.

[4] Brasch-Andersen C, Christiansen L, Tan Q, Haagerup A, Vestbo J, Kruse TA. Possible gene dosage effect of glutathione-S-transferases on atopic asthma: using real-time PCR for quantification of *GSTM1* and *GSTT1* gene copy numbers. *Hum Mutat* 2004;**24:**208–14.

[5] Hanene C, Jihene L, Jamel A, Kamel H, Agnes H. Association of *GST* genes polymorphisms with asthma in Tunisian children. *Mediat Inflamm* 2007:19564.

[6] Saadat M, Saadat I, Saboori Z, Emad A. Combination of *CC16*, *GSTM1*, and *GSTT1* genetic polymorphisms is associated with asthma. *J Allergy Clin Immun* 2004;**113:**996–98.

[7] Salam MT, Lin PC, Avol EL, Gauderman WJ, Gilliland FD. Microsomal epoxide hydrolase, glutathione S-transferase P1, traffic and childhood asthma. *Thorax* 2007;**62:**1050–57.

[8] Mak JC, Ho SP, Leung HC *et al*. Relationship between glutathione S-transferase gene polymorphisms and enzyme activity in Hong Kong Chinese asthmatics. *Clin Exp Allergy* 2007;**37:**1150–57.

[9] Fryer AA, Bianco A, Hepple M, Jones PW, Strange RC, Spiteri MA. Polymorphism at the glutathione S-transferase *GSTP1* locus. A new marker for bronchial hyperresponsiveness and asthma. *Am J Resp Crit Care Med* 2000;**161:**1437–42.

[10] Lee YL, Hsiue TR, Lee YC, Lin YC, Guo YL. The association between glutathione S-transferase P1, M1 polymorphisms and asthma in Taiwanese schoolchildren. *Chest* 2005;**128:** 1156–62.

[11] Tamer L, Calikoglu M, Ates NA *et al*. Glutathione-S-transferase gene polymorphisms (*GSTT1*, *GSTM1*, *GSTP1*) as increased risk factors for asthma. *Respirology* 2004;**9:**493–98.

[12] Romieu I, Ramirez-Aguilar M, Sienra-Monge JJ *et al*. *GSTM1* and *GSTP1* and respiratory health in asthmatic children exposed to ozone. *Eur Respir J* 2006; **28:**953–59.

[13] Nickel R, Haider A, Sengler C *et al*. Association study of Glutathione S-transferase P1 (*GSTP1*) with asthma and bronchial hyper-responsiveness in two German pediatric populations. *Pediatr Allergy Immu* 2005;**16:**539–41.

[14] Golding J, Pembrey M, Jones R. ALSPAC—the Avon Longitudinal Study of Parents and Children. I. Study methodology. *Paediatr Perinat Epidemiol* 2001;**15:** 74–87.

[15] Pembrey M. The Avon Longitudinal Study of Parents and Children (ALSPAC): a resource for genetic epidemiology. *Eur J Endocrinol* 2004;**151:**U125–29.

[16] Jones RW, Ring S, Tyfield L *et al*. A new human genetic resource: a DNA bank established as part of the Avon longitudinal study of pregnancy and childhood (ALSPAC). *Eur J Hum Genet* 2000;**8:**653–60.

[17] Rose–Zerilli MJ, Barton S, Henderson AJ, Shaheen SO, Holloway JW. Copy-Number Variation Genotyping of *GSTT1* and *GSTM1* Gene Deletions by Real-Time PCR. *Clin Chem* 2009;**55:**1680–85.

[18] Polk S, Sunyer J, Munoz-Ortiz L *et al*. A prospective study of Fel d1 and Der p1 exposure in infancy and childhood wheezing. *Am J Respir Crit Care Med* 2004;**170:** 273–78.

[19] Torrent M, Sunyer J, Garcia R *et al*. Early-life allergen exposure and atopy, asthma, and wheeze up to 6 years of age. *Am J Respir Crit Care Med* 2007;**176:**446–53.

[20] Arand M, Muhlbauer R, Hengstler J *et al*. A multiplex polymerase chain reaction protocol for the simultaneous analysis of the glutathione S-transferase *GSTM1* and *GSTT1* polymorphisms. *Anal Biochem* 1996;**236:**184–86.

[21] Van Eerdewegh P, Little RD, Dupuis J *et al*. Association of the *ADAM33* gene with asthma and bronchial hyperresponsiveness. *Nature* 2002;**418:**426–30.

[22] Horvath S, Xu X, Laird NM. The family based association test method: strategies for studying general genotype—phenotype associations. *Eur J Hum Genet* 2001;**9:**301–6.

[23] Little J HJ. *HuGE Review Handbook, Version 1.0*. Ottawa, Ontario, Canada: HuGENet Canada Coordinating Centre (http://www.hugenet.ca/). *The HuGENetTM* 2006 (February 2008, date last accessed).

[24] Little J, Higgins JP, Ioannidis JP *et al*. STrengthening the REporting of Genetic Association studies (STREGA): an extension of the STROBE Statement. *Ann Intern Med* 2009; **150:**206–15.

[25] Ioannidis JP, Boffetta P, Little J *et al*. Assessment of cumulative evidence on genetic associations: interim guidelines. *Int J Epidemiol* 2008;**37:**120–32.

[26] Minelli C, Thompson JR, Abrams KR, Thakkinstian A, Attia J. The choice of a genetic model in the meta-analysis of molecular association studies. *Int J Epidemiol* 2005;**34:**1319–28.

[27] Minelli C, Thompson JR, Abrams KR, Lambert PC. Bayesian implementation of a genetic model-free approach to the meta-analysis of genetic association studies. *Stat Med* 2005;**24:**3845–61.

[28] Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. *Stat Med* 2002;**21:**1539–58.

[29] Minelli C, Thompson JR, Abrams KR, Thakkinstian A, Attia J. How should we use information about HWE in the meta-analyses of genetic association studies? *Int J Epidemiol* 2008;**37:**136–46.

[30] Spiegelhalter DJ, Thomas A, Best NG *et al*. *Winbugs Version 1.4 User's Manual*. Cambridge, United Kingdom: Medical Research Council Biostatistics Unit, Institute of Public Health, University of Cambridge, 2003.

[31] Carroll WD, Lenney W, Child F, Strange RC, Jones PW, Fryer AA. Maternal glutathione S-transferase GSTP1 genotype is a specific predictor of phenotype in children with asthma. *Pediatr Allergy Immu* 2005;**16:**32–39.

[32] Palmer CN, Doney AS, Lee SP *et al*. Glutathione S-transferase M1 and P1 genotype, passive smoking, and peak expiratory flow in asthma. *Pediatrics* 2006;**118:** 710–16.

[33] Gilliland FD, Li YF, Dubeau L *et al*. Effects of glutathione S-transferase M1, maternal smoking during pregnancy, and environmental tobacco smoke on asthma and wheezing in children. *Am J Respir Crit Care Med* 2002;**166:** 457–63.

[34] David GL, Romieu I, Sienra-Monge JJ *et al*. Nicotinamide adenine dinucleotide (phosphate) reduced:quinone oxidoreductase and glutathione S-transferase M1 polymorphisms and childhood asthma. *Am J Respir Crit Care Med* 2003;**168:**1199–204.

[35] Plutecka H, Sanak M, Dziedzina S, Mastalerz L, Szczeklik A. Reactive oxygen species metabolism and allergy—The molecular link. *Allergy Clin Immunol Int* 2006;**18:**158–64.

[36] Zhang YQ, Sun BY, Dai JJ *et al*. Studies on the genetic diathesis of asthma bronchial [Article in Chinese]. *Yi Chuan* 2004;**26:**147–50.

[37] Oh JM, Kim SH, Suh CH *et al*. Lack of association of glutathione S-transferase P1 Ile105Val polymorphism with aspirin-intolerant asthma. *Kor J Intern Med* 2005; **20:**232–36.

[38] Safronova OG, Vavilin VA, Lyapunova AA *et al*. Relationship between glutathione S-transferase P1 polymorphism and bronchial asthma and atopic dermatitis. *Bull Exp Biol Med* 2003;**136:**73–75.

[39] Vavilin VA, Makarova SI, Liakhovich VV, Gavalov SM. Polymorphic genes of xenobiotic-metabolizing enzymes associated with bronchial asthma in genetically predisposed children [Article in Russian]. *Genetika* 2002;**38:** 539–45.

[40] Abdel-Alim SM, El-Masry MM, Aziz M, El-Bassiony SO, Aly AA. Association of glutathione-S-transferase P1 genotypes with susceptibility to bronchial asthma in children. *Arch Med Sci* 2007;**3:**200–7.

[41] Aynacioglu AS, Nacak M, Filiz A, Ekinci E, Roots I. Protective role of glutathione S-transferase P1 (*GSTP1*) Val105Val genotype in patients with bronchial asthma. *Br J Clin Pharmacol* 2004;**57:**213–17.

[42] Schroer KT, Biagini Myers JM, Ryan PH *et al*. Associations between multiple environmental exposures and glutathione S-transferase P1 on persistent wheezing in a birth cohort. *J Pediatr* 2009;**154:**401–8.

[43] Carroll WD, Lenney W, Jones PW *et al*. Effects of glutathione S-transferase M1, T1 and P1 on lung function in asthmatic families. *Clin Exp Allergy* 2005;**35:**1155–61.

[44] McLellan RA, Oscarson M, Alexandrie AK *et al*. Characterization of a human glutathione S-transferase mu cluster containing a duplicated *GSTM1* gene that causes ultrarapid enzyme activity. *Mol Pharmacol* 1997; **52:**958–65.

[45] Sprenger R, Schlagenhaufer R, Kerb R *et al*. Characterization of the glutathione S-transferase *GSTT1* deletion: discrimination of all genotypes by polymerase chain reaction indicates a trimodular genotype-phenotype correlation. *Pharmacogenetics* 2000;**10:**557–65.

[46] Zimniak P, Nanduri B, Pikula S *et al*. Naturally occurring human glutathione S-transferase GSTP1-1 isoforms with isoleucine and valine in position 104 differ in enzymic properties. *Eur J Biochem* 1994;**224:**893–99.

[47] Bentley AR, Emrani P, Cassano PA. Genetic variation and gene expression in antioxidant related enzymes and risk of COPD: a systematic review. *Thorax* 2008;**63:** 956–61.

[48] Maziak W, Ward KD, Afifi Soweid RA, Eissenberg T. Tobacco smoking using a waterpipe: a re-emerging strain in a global epidemic. *Tob Control* 2004;**13:**327–33.

[49] von Mutius E. Gene-environment interactions in asthma. *J Allergy Clin Immunol* 2009;**123:**3–11.

[50] Islam T, Berhane K, McConnell R *et al*. Glutathione-S-transferase (*GST*) *P1*, *GSTM1*, exercise, ozone and asthma incidence in school children. *Thorax* 2009;**64:** 197–202.

[51] London SJ, Romieu I. Gene by Environment Interaction in Asthma. *Annu Rev Public Health* 2009;**30:**55–80.

[52] Romieu I, Sienra-Monge JJ, Ramirez-Aguilar M *et al*. Genetic polymorphism of *GSTM1* and antioxidant supplementation influence lung function in relation to ozone exposure in asthmatic children in Mexico City. *Thorax* 2004;**59:**8–10.

[53] Kabesch M, Hoefler C, Carr D, Leupold W, Weiland SK, von Mutius E. Glutathione S transferase deficiency and passive smoking increase childhood asthma. *Thorax* 2004; **59:**569–73.

[54] Lee YL, Lee YC, Guo YL. Associations of glutathione S-transferase P1, M1, and environmental tobacco smoke with wheezing illness in school children. *Allergy* 2007; **62:**641–47.

[55] Carlsten C, Sagoo GS, Frodsham AJ, Burke W, Higgins JP. Glutathione S-transferase M1 (*GSTM1*) polymorphisms and lung cancer: a literature-based systematic HuGE review and meta-analysis. *Am J Epidemiol* 2008; **167:**759–74.

[56] Saadat M, Ansari-Lari M. Genetic polymorphism of glutathione S-transferase T1, M1 and asthma, a meta-analysis of the literature. *Pak J Biol Sci* 2007;**10:** 4183–89.

[57] Gilliland FD, Li YF, Gong H Jr, Diaz-Sanchez D. Glutathione S-transferases M1 and P1 prevent aggravation of allergic responses by secondhand smoke. *Am J Respir Crit Care Med* 2006;**174:**1335–41.

[58] Toren K, Brisman J, Jarvholm B. Asthma and asthma-like symptoms in adults assessed by questionnaires. A literature review. *Chest* 1993;**104:**600–8.

[59] Chung HW, Ahn TH, Kim SY, Kim TY, Paek DY. Polymorphisms of *GSTM1*, *GSTT1*, *GSTP1 NATII* and *CYP1A1* and the susceptibility to asthma. *Kor J Genet* 2002;**24:**259–66.

[60] Freidin MB, Bragina E, Ogorodova LM, Puzyrev VP. Polymorphism of the theta1 and mu1 glutathione s-transferase genes (*GSTT1*, *GSTM1*) in patients with atopic bronchial asthma from the West Siberian region [Article in Russian]. *Mol Biol (Mosk)* 2002;**36:**630–34.

[61] Sideleva OG, Ivashchenko TI, Orlov AV *et al*. Comparative analysis of polymorphous genes of glutathione-S-transferases in north-western populations of Russia—Healthy and patients with bronchial asthma. *J Evol Biochem Physiol* 2002;**38:**285–90.

[62] Ovcharenko SI, Filippov VV. The complex assessment of disordered bronchial patency in chronic obstructive lung diseases [Article in Russian]. *Terap Arkh* 1990;**62:** 63–66.

[63] Standards for the diagnosis and care of patients with chronic obstructive pulmonary disease (COPD) and asthma. This official statement of the American Thoracic Society was adopted by the ATS Board of Directors, November 1986. *Am Rev Respir Dis* 1987;**136:**225–44.

[64] Lee YL, Gilliland FD, Wang JY, Lee YC, Guo YL. Associations of FcepsilonRIbeta E237G polymorphism with wheezing in Taiwanese schoolchildren. *Clin Exp Allergy* 2008;**38:**413–20.

[65] Arbag H, Cora T, Acar H, Ozturk K, Sari F, Ulusoy B. Lack of association between the glutathione-s-transferase genes (*GSTT1* and *GSTM1*) and nasal polyposis. *Rhinology* 2006;**44:**14–18.

[66] Holla LI, Stejskalova A, Vasku A. Polymorphisms of the *GSTM1* and *GSTT1* genes in patients with allergic diseases in the Czech population. *Allergy* 2006;**61:**265–67.

[67] Global Strategy for Asthma Management and Prevention. Global initiative for asthma. *NHLBI/WHO:NIH Publication no. 01-3659* 2002.

[68] Global Strategy for Asthma Management and Prevention. *NHLBI/WHO Workshop Report. National Institute of Health* 2003, Vol. 2, pp. 4–10.

[69] Kamada F, Mashimo Y, Inoue H *et al*. The *GSTP1* gene is a susceptibility gene for childhood asthma and the *GSTM1* gene is a modifier of the GSTP1 gene. *Int Arch Allergy Immunol* 2007;**144:**275–86.

[70] Bethesda. *National Institutes of Health* 1997;Publ 97–4051.

[71] Imboden M, Rochat T, Brutsche MH *et al*. Glutathione-S transferase genotype increases risk of progression from bronchial hyperresponsiveness to asthma in adults. *Thorax* 2008;**63:**322–28.

Appendix 1

WinBUGS code for the genetic model-free approach

Model

```
{

        for(i in 1:17) {
case[i,1:3] ~ dmulti(p.case[i,], N.case[i])
control[i,1:3] ~ dmulti(p.ctrl[i,], N.ctrl[i])

p.ctrl[i,1] <- 1/(1+beta[i,1]+beta[i,2])
p.ctrl[i,2] <- beta[i,1]/(1+beta[i,1]+beta[i,2])
p.ctrl[i,3] <- beta[i,2]/(1+beta[i,1]+beta[i,2])

beta[i,1] ~ dgamma(0.001,0.001)
beta[i,2] ~ dgamma(0.001,0.001)

p.case[i,1] <- 1/(1+theta[i,1]+ theta [i,2])
p.case[i,2] <- theta [i,1]/(1+ theta[i,1]+ theta [i,2])
p.case[i,3] <- theta [i,2]/(1+ theta[i,1]+ theta [i,2])

theta[i,1] <- beta[i,1]*exp(lambda*delta[i])
theta[i,2] <- beta[i,2]*exp(delta[i])
delta[i] ~ dnorm(d, prec)

                        }

        d ~ dnorm(0.0,1.0E-6)
        lambda ~ dbeta(0.5,0.5)
    prec <- 1/var
    var <- pow(sd,2)
    sd ~ dnorm(0,1)I(0,)
        OR_Gg <- exp(lambda*d)
        OR_GG <- exp(d)

}
```