THE ROLE OF GENES AND ENVIRONMENT ON FETAL GROWTH

Thesis presented for the degree of Doctor of Philosophy in the Faculty of Population Health Sciences, University College London

Dr Sara Louise Hillman

SIGNED DECLARATION

I, Sara Louise Hillman confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

ACKNOWLEDGEMENTS

I want to thank my supervisor, Dr David Williams, for the opportunity given to me to pursue this PhD. I am enormously grateful for his unwavering support and guidance. Under his tutelage I have developed so much as a clinical academic and I hope this work is just the beginning of an extremely productive partnership.
I am also grateful to my secondary supervisor, Professor Donald Peebles; his sense and reasoning have kept the work focussed and his support has been much appreciated.

I was fortunate enough to be awarded funding for the project from Wellbeing of Women through a Research Training Fellowship. I am also grateful to the UCL Biomedical Centre who provided seed money to get the project underway and the BMA Joan Dawking award, who continue to fund this work.

There are several people I need to mention that made the project possible. I am indebted to their expertise and kindness. At, UCL, special mention must be made of Dr Mike Okorie who taught me flow mediated dilation and Dr Chiara Bacchelli and Dr Louise Ocaka who made exome sequencing a possibility. I also want to thank the team at the Blizard, especially Dr Melissa Smart and Chris Matthews, who got me up to speed on DNA methylation and helped tremendously with the practicalities of running and analysing my samples.

I must thank the patients and families that agreed to take part in the study. I also want to thank the staff at UCH; the midwives in MFAU and FMU who helped me recruit; the staff on Labour Ward who helped me collect samples and the doctors who assisted me; Dr Melissa Whitten, Dr Tamara Kubba and Dr Anna Jeffery Smith.

Finally I want to thank my family; my husband Guy who is my rock and a pillar of support; our girls, Molly and Isla, who have broadened our world more than we thought possible and ‘the mums’ who have made having twins and completing a PhD a possibility.

I could not have done it without you all. X

ABSTRACT

Fetal growth is influenced by the *in utero* environment and genetic factors inherited from both parents. Poor fetal growth leading to low birth weight is associated with insulin resistance and type-2 diabetes in later life.

The fetal programming of adult disease hypothesis suggests that growth-restricted fetuses make enduring physiological adaptations that predispose to diabetes in later life. The fetal insulin hypothesis suggests that poor fetal growth and diabetes are two phenotypes of genetically determined insulin resistance. Under these circumstances, an insulin resistant fetus cannot optimise insulin-mediated growth and is predisposed to diabetes in later life. Environmental and genetic influences come together through epigenetic modifications, for example DNA methylation, that alter gene expression without altering the nucleotide sequence.

The first aim of this thesis was to investigate whether men who fathered pregnancies complicated by fetal growth restriction had an insulin resistant phenotype at the time of the index pregnancy. A case-control study showed that men who fathered growth-restricted offspring have pre-clinical insulin resistance and are more likely to smoke than fathers of normal grown offspring. This observation supports the concept that an insulin resistant genotype inherited from a father could manifest as poor fetal growth in offspring.

I then investigated the mechanisms through which paternal insulin resistance might be inherited by a growth-restricted fetus. I studied DNA extracted from the cord blood of growth-restricted offspring using whole exome sequencing to identify novel gene variants and those known to be associated with type-2 diabetes. I validated findings with Sanger sequencing and Taqman genotyping in all family members.

Using the Illumina Human 450 BeadChip, I found marked differences in genome wide DNA methylation of fetal cord blood and placental samples from growth restricted compared with normal grown offspring.

Future work is aimed at investigating the functional consequences of genetic and epigenetic differences to identify targets for treatment and prophylaxis against fetal growth restriction and diabetes.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

Ala	alanine
BP	blood pressure
CRP	c-reactive protein
dbSNP	single nucleotide polymorphism database
DMP	differentially methylated position
DNP	dinitrophenyl
ESP	exome sequencing project
EVS	exome variant server
FGR	fetal growth restriction
FMD	flow mediated dilation
GP	general practitioner
HbA1c	glycated haemoglobin
HDL	high density lipoprotein cholesterol
HM450	Infinium Human Methylation 450 array
HOMA	homeostasis model assessment
HOMA-IR	homeostasis model assessment insulin resistance
HOMA-β	homeostasis model assessment beta cell function
IGT	impaired glucose tolerance
IR	insulin resistance
IUGR	intrauterine growth restriction
LDL	low density lipoprotein cholesterol
MAF	mean allele frequency
MODY	maturity onset diabetes of the young
MVP	most valuable position

NTC	negative controls
OGTT	oral glucose tolerance test
OMIM	online Mendelian inheritance in man
PET	pre eclampsia
PolyPhen	Polymorphism Phenotyping
Ponderal index	newborn weight/length^3 measure of leanness
PWV	pulse wave velocity
SD	standard deviation
SGA	small for gestational age
SH	Sara Hillman
SIFT	sorting intolerant from tolerant variants
SNP	single nucleotide polymorphism
T2DM	type 2 diabetes
TG	triglycerides
WES	whole exome sequencing
WGA	whole genome amplification

Chapter 1

Introduction

Chapter 1: INTRODUCTION

1.1 Fetal growth

In utero growth from a single cell to a fully formed neonate is the most fundamental yet complex process we undergo. Fetal growth and development are driven by genes inherited from both mother and father and the interplay of those genes with the *in utero* environment. Of particular importance is the interaction between maternal and fetal genes at the utero-placental interface, as these dictate placental development. Failure of the placenta to invade the uterine placental bed is a common cause of reduced fetal growth in the developed world. In the developing world, environmental factors such as maternal malnutrition and disease more commonly influence fetal growth (Bodsky D et al., 2004; Maulik 2006b & 2006c).

When the gene-environment interaction is sub-optimal, a fetus will adapt to optimise its growth potential. Fetal adaptations are however inevitably a compromise and result in reduced birth weight. Furthermore, as *in utero* growth is so rapid and profound, fetal adaptations may have a commensurately profound and enduring effect on its' life-long health. This thesis explores how genes and environment interact to result in severely compromised fetal growth.

1.2 Consequences of poor fetal growth

Given optimal conditions, a fetus follows a growth trajectory *in-utero* that allows it to fulfil its own unique growth potential and reach its expected birth weight. In contrast, fetal growth restriction (FGR) occurs when a fetus is unable to reach its growth potential (Pallotto E et al., 2006). FGR should be considered a syndrome, with multi-factorial origins which can broadly be divided into maternal, fetal or placental causes (Figure 1.1).

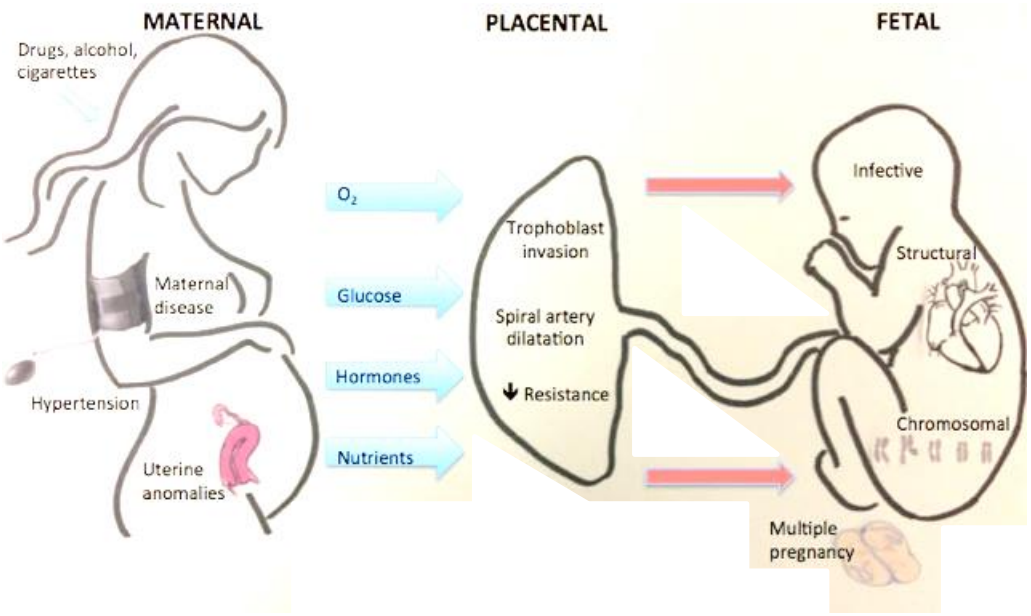


Fig 1.1 Known causes of fetal growth restriction and normal role of placenta

1.3 Fetal Growth Restriction

FGR is the second leading cause of perinatal morbidity and mortality affecting around 5% of all babies (which equates to over 8 million worldwide each year) (Neerhof MG, 1995 & Bernstein IM et al., 2000). It poses a significant health and socio-economic problem in both developed and developing countries. Problems relate to immediate *in-utero* and birth complications, increased neonatal problems, childhood and adolescent development problems and disease in later life.

1.3.1 Immediate consequences of fetal growth restriction

Stillbirth accounts for up to 1 in 200 pregnancies in the United Kingdom (CEMACH, 2009) of which a significant proportion will be affected by growth restriction. A fetus with growth restriction has an increased risk of intrauterine death (OR 7; 95% CI 3.3-15.1) (Froen JF et al., 2004), with the severity of FGR related to stillbirth risk independently of gestational age (Piper JM et al., 1996 and Divon MY, 1998).

Perinatal mortality is also increased in FGR offspring up to 10 fold compared to similar gestational aged babies (Lackman F et al., 2001 and McIntire DD et al., 1999). Within growth restricted fetuses those with more severe pathology (as determined by fetal ultrasound patterns of reversed or absent Doppler blood flow patterns) have been shown to have the highest perinatal mortality (Hackett GA et al., 1997 and Simchen MJ et al., 2000).

1.3.2 Long term consequences of poor fetal growth

Poor fetal growth and subsequent low birth weight are risk factors for the future development of impaired glucose tolerance (IGT), metabolic syndrome, type 2 diabetes (T2DM) and cardiovascular disease (Hales CN et al., 1991 and Barker DJP et al., 1993). Through the reporting of birth weight records, of men and women in middle and later life living in Hertfordshire, it was identified that the incidence of impaired glucose tolerance or T2DM in these men (mean age 64 years) fell from 40% if born at less than 2.5kg to less than 14% if born at more than 4.3kg (Figure 1.2). In these same men, blood pressure also inversely correlated with reported birth weight (Barker DJ et al., 1993). This trend was maintained when reported weight at age 1 was assessed in these men in relation to IGT and T2DM.

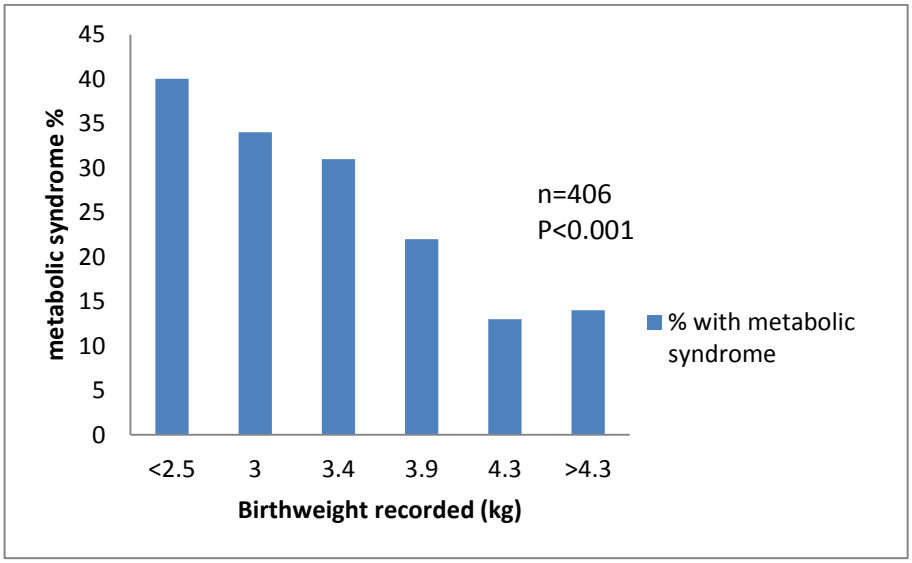


Figure 1.2 Reproduced data from Hertfordshire cohort of men aged 64 with metabolic syndrome or T2DM in relation to birth weight (Barker DJ et al., 1993)

Many subsequent studies have confirmed the link between birth size and future cardio-metabolic disease. An inverse correlation was seen between a measure of leanness the 'ponderal index' (calculated from birth weight/length^3) at birth and subsequent insulin concentrations in 1333 Swedish men aged 50 and presence of diabetes in these men at aged 60 (Lithell HO et al., 1996).

A prospective study of 1258 men from South Wales revealed those born in the lowest birth weight quartile had the highest incidence of coronary heart disease during a 10 year follow up (Frankel S et al., 1997).

1.4 Role of fetal insulin in growth

Fetal insulin, a growth factor, acting mainly in the third trimester, is one of the key determinants of fetal growth (Constância M et al., 2012). An example of the importance of insulin in fetal growth can be seen in gestational diabetes where fetuses are often macrosomic. High maternal glucose levels pass easily across the placenta to cause elevated fetal glucose levels, which results in over-production of insulin by the fetus and increased fetal growth (Pedersen, 1952).

In diabetic pregnancies with good maternal blood glucose control, macrosomia is less commonly seen compared with poorly controlled gestational diabetes (Naylor CD et al., 1993). In normal pregnancy a direct correlation can be seen between maternal blood glucose and offspring birth weight (Tallarigo L et al., 1986 and Sermer N et al., 1995), providing evidence that glucose homeostasis mechanisms are vital in fetal growth.

One of the ways fetal insulin is thought to be regulated is by insulin growth factor-1 (Constância M et al., 2012). In humans, fetal serum *IGF-1* levels have been shown to correlate with birth weight with lower levels observed in FGR babies (Maulik D, 2006). Low levels of *IGF-1* and elevated levels of *IGF* binding protein-1 (*IGFBP-1*) have consistently been observed in FGR infants (Wang HS et al, 1991).

Transgenic mice overexpressing *IGFBP-1* demonstrate poor growth (Crossey PA et al., 2002) and abnormal trophoblast differentiation and provide further evidence that insulin and its regulators has a role in placental development and fetal growth.

1.5 The fetal insulin hypothesis

The importance of insulin and insulin pathways as mediators of fetal growth is explored further in the fetal insulin hypothesis (Hattersley AT and Tooke JE, 1999). In the reverse of the situation previously described, where fetuses are macrosomic as a result of maternal hyperglycaemia, the fetal insulin hypothesis provides a mechanism whereby genetically inherited fetal insulin resistance leads to poor growth and low birth weight. If fetal insulin mediated growth can reflect maternal glycaemia then fetal genetic factors which regulate the ability of the fetus to secrete insulin and the ability of fetal tissues to be sensitive to the effects of insulin (i.e. the degree of insulin resistance) must also influence fetal growth.

A genetic defect in either the sensing of the maternal glucose by the fetal pancreas, or insulin secretion by the fetal pancreas, or the action of the insulin on the insulin-dependent tissues, would all result in reduced fetal growth (Figure 1.3). It is this mechanism of fetal insulin resistance and or beta cell dysfunction that underlies the hypothesis and provides insight into how a growth restricted fetus may also be predisposed to later T2DM development.

Figure 1.3 The fetal insulin hypothesis. In response to maternal glucose, fetal genetics influence glucose and insulin sensing by fetal pancreas and alter insulin response which is critical in fetal growth. Reproduced from: Frayling T M, and Hattersley A T Br Med Bull 2001; 60:89-101

1.6 Evidence of a genetic influence on fetal growth

Despite growth restriction being the extreme phenotype, evidence that further supports a genetic influence in fetal growth is based on studies performed across the spectrum of birth weight. There is a strong correlation in family birth weights between siblings (Tanner JM et al., 1972). Estimates of heritability suggest fetal genes explain around 30% of the normal variation in birth weight and birth length (Lunde A et al., 2007).

1.6.1 Imprinted genes

Genomic imprinting is a process through which the expression of a gene is dependent on the sex of the parent from which it was inherited (Morrison IM et al., 1998). Imprinted genes have been implicated in growth disorders (Haycock PC et al., 2009). An example of an imprinted gene disorder resulting in a growth restricted phenotype is transient neonatal diabetes mellitus (TNDM). Paternal inheritance of an imprinted gene (*PLAGL1*) on chromosome 6 has been found in offspring affected by TNDM (Temple IK et al., 1996*).*

In addition to the diabetic phenotype, infants suffering from TNDM have markedly reduced birth weights (<1st centile) (Shield JPH, 1996). This rare genotype is a good example of a gene mutation that can result in two apparently disparate phenotypes – diabetes and fetal growth restriction. Over 60 imprinted genes have been identified (Morrison IM et al., 1998) and in many cases a phenotype has been identified which is associated with altered fetal growth (e.g. overgrowth in Prader-Willi syndrome and poor growth in TNDM and Silver-Russell syndrome). However, these are rare examples and it seems unlikely that imprinted genes have a significant role in the common variation seen in birth weight or the majority of cases of fetal growth restriction.

1.7 Monogenic diabetes, birth weight and T2DM risk

Further evidence for a genetic influence in fetal growth arises through the study of monogenic conditions which occur as a result of a defective single gene (Lenderman HM, 1995). Maturity-onset diabetes of the young (MODY) is a form monogenic diabetes characterised by single autosomal dominant mutations in a range of individual genes accounting for 1-2% of non-insulin dependent diabetes in Europe (Lenderman HM, 1995).

Whilst MODY is a genetically heterogeneous collection of primarily autosomal dominant mutations on a number of genes, up to 80% of cases can be attributed to heterozygous mutations on one of four genes; *GCK, HNF1α, HNF1β* and *HNF4α* (Ellard S et al., 2008). About 14% of diagnosed cases are mutations within the gene for the enzyme gluckokinase (GCK) (Stride A et al., 2002).

Glucokinase is a hexokinase enzyme that phosphorylates glucose to glucose-6-phosphate which within beta cells is closely linked to insulin secretion (Meglasson MD et al., 1986). Mutations in this gene first provided evidence for the fetal insulin hypothesis and showed that genetic inheritance of a single gene mutation in *GCK* from father to offspring could result in low birth weight.

Over 100 mutations within the gene have been described (Stride A et al., 2002). Heterozygous presentation usually results in mild hyperglycaemia often only detected during screening for gestational diabetes. This gene mutation was one of the first studied and able to provide a plausible mechanism for a single genotype causing pleiotropic effects that leads to both low birth weight and diabetes.

Maternal *GCK* heterozygotes provide a hyperglycaemic environment to the fetus which in turn increases its insulin secretion with a resultant higher birth weight. Fetal *GCK* heterozygotes fail to secrete insulin efficiently for fetal growth with a resultant decrease in birth weight. In the presence of maternal-fetal heterozygotes the normal increase in insulin secretion (to the maternal hyperglycaemic environment) fails in the fetus, resulting in no difference in birth weight from wild-type controls.

As shown in Figure 1.4, mother-offspring pairs carrying the glucokinase mutation showed that on average, birth weight is decreased by 533g if the *GCK* mutation is present in the fetus but not mother. Conversely, when the *GCK* mutation is present in the mother and not the fetus, birth weight is increased by 601g. When a mutation is present in both mother and offspring birth weight is unaffected (Hattersley et al., 1998). *GCK* mutations cause impaired glucose sensing, which reduces an affected individual's ability to secrete insulin and results in stable fasting hyperglycaemia throughout life (Froguel P et al., 1993).

Figure 1.4 The effect of mutations in the glucokinase gene on centile birth weight from 59 offspring pairs (M+ mother with mutation: F+ fetus with mutation) (Frayling T M, Hattersley a T Br Med Bull 2001; 60:89-101).

1.7.1 *GCK* mutations and evidence provided through animal models

In mice, fetuses heterozygous for a *GCK* mutation resulted in significantly lower birth weight replicating what has been identified in humans (Terauchi Y et al., 2000). Metabolites of the insulin pathway could also be investigated more easily than in humans. Low birth weight in the presence of gluckokinase mutation was also associated with insulin receptor substrate 1 (*IRS-1*) deficiency. *IRS-1* is the major substrate of the insulin receptor tyrosine kinase and peripheral insulin resistance has been shown in *IRS-1* deficient mice (Tamemoto H et al., 1994). The study confirmed the human subject data but also went on to provide direct evidence of a link to insulin and lower birth weight in the offspring.

1.7.2 Other MODY genes that support a negative influence on fetal growth

A similar birth weight profile has also been seen in conjunction with mutations in other genes associated with MODY including *HNF1β*. Neonates studied with *HNF1β* mutations were on average 900g lighter than negative controls (Edghill EL et al., 2006).

Aside from the gluckokinase mutation almost all known MODY mutations lie in transcription factor genes responsible for gene expression and tissue development. In these mutations the exact mechanisms leading to diabetes is less clear but it is apparent that there is a defect in insulin sensing and subsequent beta cell function. The different phenotypes or time of disease onset in MODY may be explained by the different mechanisms of beta cell dysfunction that are occurring. Not all mutations may be able to have an impact *in-utero* and only those that do are likely to influence birth weight.

However, monogenic forms of diabetes form only a small proportion (2-5%) of all diabetes cases, and are unlikely to explain the majority of growth restriction or susceptibility to T2DM. T2DM is characterized by both insulin secretory dysfunction and peripheral insulin resistance (Taylor S et al., 1994).

T2DM development is more complex than monogenic forms of diabetes with individual susceptibility influenced by genetic variation and environmental exposures encountered during life (Stumvoll M et al., 2005). However, there is evidence for a genetic influence in T2DM. Siblings of parents who have diabetes have a 3- 4 fold risk of developing T2DM over the general population (Kobberling J et al., 1982) and T2DM is more concordant in monozygotic as opposed to dizygotic twins (Newman B et al., 1987). Whilst environment is also shared in most of these instances, further evidence of a genetic basis is drawn from studies where despite environmental similarities risk of diabetes development is altered by genetic admixture. An example of this is the Pima Indian community where Caucasian genetic admixture altered the risk of developing diabetes (Williams RC et al., 2000).

1.8 Genetic influence of T2DM risk alleles

Genome wide association studies (GWAS) which provide strong, reproducible associations between common genetic variants and disease state have yielded some useful information about genetic risk loci in T2DM. Currently there are around 40 published loci reproducibly associated single nucleotide polymorphisms (SNPs) (McCarthy M et al., 2010).

However, the influence of each of these genes is modest. The largest effect in relation to T2DM is seen with *TCFL7* (transcription factor 7-like-2), with a per allele odds ratio of 1.4. The 15% of population who carry two risk alleles are twice more likely to develop T2DM than the 40% population who do not carry the alleles (Zeggini E et al, 2007). It has been shown that individual smaller effect sizes can combine to generate higher individual risk of developing T2DM (Weedon MN et al., 2006).

Genetic mutations in monogenic diabetes have been shown to have a large effect size on fetal growth (Table 1.1) with common variants found in polygenic T2DM having smaller effects (Table 1.2). It is unclear if these influences may be additive. Given the potential genetic basis for both T2DM and the observed and potentially interlinked association between MODY; genetic loci that predispose to T2DM (particularly those implicated in the rarer inherited monogenic diseases associated with low birth weights, such as MODY) are of interest in FGR.

Gene	Genetic defect	Adult clinical trait	BW decrease (g)	BW increase (g)	Reference
GCK (rs1799884)		Fasting glucose		64[c]	Weedon et al. 2005
	Het mutation		531[d]		Frayling et al., 2001
HNF1α		No[b]			Hegele et al. 1999
HNF1β (rs757219)		T2DM	No		Winckler et al.2007
	Het mutation		900		
HNF4α (rs1884614)		T2DM	No		Hanson et al. 2005
	Het mutation			790[e]	Pearson et al. 2007
KCNJ11 (rs5215/19)		T2DM	No[a]		
	Het mutation		TNDM[f]		Slingerland et al. 2006

[a] Once maternal genotype accounted for
[b] Specific mutation in Oji-Cree population predisposes to early onset T2DM
[c] Driven by maternal genotype [d] Neutral if present in fetus and mother
[e] According to recalled birth weight [f] With resultant low birth weight

Table 1.1 Genes known to cause monogenic diabetes and subsequent birth weight effect (when assessed in small sample cohorts with specific heterozygous mutations) and the relationship of these genes with common T2DM risk loci and birth weight. The example of *GCK* shows that despite a heterozygous genotype resulting in decreased birth weight, when the recognised T2DM risk allele is present, birth weight is increased and thought to be driven by maternal genotype.

Gene	Common T2DM SNP	Adult Clinical trait	Per risk allele effect size on Birth Weight (g)*		References	
			Decrease	Increase		
CDKAL1	rs10946398	T2DM	21		Andersson et al. 2010	
CDKN2A-2B	rs10811661	T2DM		11	Freathy et al. 2009	Andersson et al. 2010
HHEX-IDE	rs1111875	T2DM	14		Andersson et al. 2010	
IGF2BP2	rs440290	T2DM		4	Freathy et al. 2009	Andersson et al. 2010
SLC30A8	rs13266634	T2DM		12	Freathy et al. 2009	Andersson et al. 2010
ADCY5	rs9883204	T2DM	30		Freathy et al. 2010	
CCNL1	rs900400	No known risk	42		Freathy et al. 2010	
TCFL7	rs7903146	T2DM		18	Freathy et al. 2007	Andersson et al. 2010
WFS1	rs10010131	T2DM	NA[a]		Freathy et al. 2010	
FTO	rs8050136	T2DM	NA		Freathy et al. 2010	
PPARG	rs1801282	T2DM	NA		Freathy et al. 2010	
NOTCH2	rs10923931	T2DM	NA		Freathy et al. 2010	
THADA	rs7578597	T2DM	NA		Freathy et al. 2010	
JAZF1	rs864745	T2DM	NA		Freathy et al. 2010	

***BW (birth weight) Adjusted for sex and gestational age**

[a] Once maternal genotype accounted for

Table 1.2 Common T2DM risk SNP loci where an effect on birth weight has been assessed in a large sample cohorts

1.8.1 Overlap between rare glucose metabolism disorders and T2DM risk alleles

There is overlap between rare gene mutations resulting in disorders of glucose metabolism and common variants of these genes that are also implicated in T2DM risk.

Common variants on the most widely studied MODY gene *GCK* have not been robustly identified in T2DM populations (Lotfi K et al, 1997) but a mutation upstream from the promoter region has been associated with increased fasting glucose (Weedon MN et al, 2005).

Of the four common MODY genes, mutations in the *HNF1α* gene account for about 50% of recognised MODY cases (Stride A et al., 2002). This gene codes for a key member of a transcription factor network that is essential for the development and function of the pancreatic cell making it a reasonable target to implicate in T2DM risk (Odom DT et al., 2004). A variant of *HNF1α*, resulting in a p.G319S change, is known to predispose to T2DM in the Oji-Cree population of Canada (Hegele RA et al., 1999) but common variants in Caucasian populations have not replicated this finding (Urhammer SA et al., 1997). Despite the p. I27L polymorphism (rs1169288) being shown to predict future T2DM in elderly obese men (HR 1.2 [95% CI 1.1–1.3], P = 0.0002) (Holmkvist J, 2006), large scale studies of common variants of the *HNF1α* gene have failed to find an association with T2DM risk (Weedon MN et al., 2005 & Winckler W et al., 2005).

A gene variant in *HNF1β* is a rarer cause of MODY. It is a gene important for the development of the pancreas and is associated with characteristic phenotypes like cystic kidney disease, liver dysfunction, and abnormal urogenital tract development (Winkler Wet al., 2007). Polymorphisms in *HNF1β* (rs757210) have been reported in association with type 2 diabetes in Caucasians but would seem overall to contribute very modestly to T2DM risk (Winckler W, 2007).

The hepatocyte nuclear factor (*HNF*)-4α regulates expression of genes involved in glucose metabolism and insulin secretion, including activation of expression of the insulin gene [Bartoov-Shifman R et al., 2002]. In a Danish population a common variant (rs1884614) near the P2 promoter in *HNF4* has been shown to be associated with an increased risk of T2DM (OR 1.14) as well as a subtle increase in post-OGTT plasma glucose levels (Hansen SK et al., 2005). These findings have also been replicated in other populations albeit with lower odds ratios (Love-Gregory LD et al., 2004 & Silander K et al., 2004).

As better understanding of genes involved in MODY pathogenesis emerges, new genes and avenues of research have opened up with regards to T2DM risk. Other rarer genes thought to now fall within the MODY spectrum include: *PDX; INS; CEL; NEUROD1; WFS1* and *KCNJ11*, and have been studied in relation to T2DM risk. A common polymorphism of *KCNJ11* (p.E23K) is associated with T2DM in Caucasians (Hani EH et al., 1998). Studies of the p.E23K variant show associations with impaired glucose induced insulin release and increased BMI (Nielsen EM et al., 2003).

Mutations in the gene *WFS1,*cause Wolfram syndrome, which is characterized by diabetes insipidus and juvenile-onset non-autoimmune diabetes mellitus (Inoue H et al., 1998). In a pooled analysis of UK populations comprising 9,533 cases and 11,389 controls, SNPs in *WFS1* were strongly associated with T2DM diabetes risk (Sandhu MS et al., 2007).

As discussed previously, lower birth weight is a recognised phenomenon in those with MODY gene mutations. Several of these genes have polymorphic variations that increase adult risk of T2DM. This understanding makes it possible that polygenic loci within MODY genes may be responsible for both a growth restricted phenotype and T2DM presentation.

1.9 Birth Weight Genome Wide Association Studies (GWAS)

The GWAS approach taken to identify T2DM risk alleles has also been taken in relation to birth weight. To date there has been limited success in identifying robust and reproducible associations.

A meta-analysis performed by The Early Growth Genetics (EGG) Consortium identified two robust genetic associations with birth weight (Freathy RM et al., 2010). One was contained within a cluster of correlated SNPs near the *CCNL1* and *LEKR1* genes (rs900400). Each risk allele present in a fetus for SNP rs900400 was associated with a 42g [95% confidence interval 35–48 g] lower birth weight. However, it is unclear how these SNPs function to influence fetal growth.

SNP rs900400 was also associated with lower birth weight in a study of preterm neonates, arguing against a gestational age effect (Ryckman KK et al., 2012).

The second cluster of SNPs identified with birth weight centred on the *ADCY5* gene. Each additional fetal risk allele for SNP rs9883204 is associated with 30 g (95% CI: 23–38 g) lower birth weight (Freathy RM et al., 2010). SNP rs9883204 is in linkage disequilibrium with SNP rs11708067 which is also known to predispose to T2DM (Dupuis J et al., 2010).

In both instances, despite the evidence of their involvement in fetal growth, the causal mechanisms that influence fetal growth are not yet known. The *ADCY5* locus has pleiotropic effects on glucose regulation and T2DM in adulthood as well as fetal growth and provides some evidence of a genetic association between lower birth weight and subsequent T2DM.

1.10 Candidate approach linking T2DM risk alleles and birth weight

In conjunction with GWAS, candidate gene approaches have also been undertaken attempting to identify T2DM risk alleles that may be influencing birth weight. As *TCFL7* has been robustly associated with T2DM, it was hypothesised that it may play an important role in fetal growth. However, GWAS and candidate approaches have failed to confirm any birth weight association with *TCF7L2*. In a meta-analysis of 24,053 individuals fetal inheritance of the T2DM risk allele SNP rs7903146 resulted in an 18g (95% CI 7-29g) increase in birth weight (Freathy RM et al., 2007). However, following adjustment for maternal genotype, no birth weight differences could be identified.

The candidate approach has been more successful for other T2DM risk alleles. 19,200 individuals were assessed for the presence of risk alleles within the genes *CDKAL1* (SNP rs10946398) and *HHEX-IDE* (SNP rs1111875). Birth weight was decreased by 21g and 14g per risk allele present respectively (Freathy RM et al., 2009). It is likely that maternal risk alleles provide an opposing effect that increases birth weight. These findings have been replicated in other cohorts from around the world (Zhao J et al., 2009 & Pulizzi N et al., 2009). However, these and other large studies have on the whole failed to identify any association between birth weight and other T2DM common variants in *PPARG, KCNJ11, SLC30A8, IGF2BP2* or *CDKN2A/2B.*

Interestingly, although effect size alone is small it was shown that carrying four risk alleles for both *CDKAL1* and *HHEX-IDE* resulted in a birth weight decrease of 80g in the 4% of population carrying all versus the 8% carrying none (Freathy RM et al., 2009). This suggests that multiple gene variants have a synergistic or additive effect that can significantly influence fetal growth and birth weight.

1.11 The paternal influence on fetal growth

1.11.1 Epidemiological studies

Epidemiological work provides evidence for a paternal influence on fetal growth leading to low birth weight. Studies of Pima Indians support a genetic influence on low birth weight and T2DM transmitted from the father (Lindsay RS et al., 2000). Pima Indians are unusual in that their diabetes rate is one of the highest in the world. In this study, 50% of adult men and women had diabetes. The data revealed that men who had diabetes, when the mother did not, fathered children whose birth weight was lower than those who were not diabetic or when the mother was diabetic (Lindsey RS et al., 2000) (Figure 1.5).

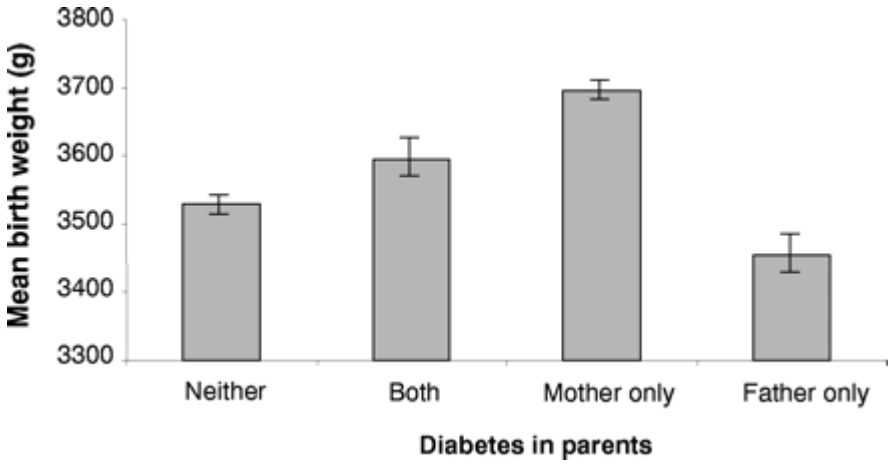


Figure 1.5 Offspring birthweight in parents with diabetes. Reproduced from Lindsay RS et al., 2000

Further epidemiological studies have validated these findings in different populations confirming that men who develop diabetes in later life were more likely to have fathered low birth weight offspring (Davey Smith G et al., 2004; Wannamethee SG et al., 2004, and Hypponen E et al., 2003).These fathers are also at increased risk of cardiovascular disease (Davey Smith G et al., 2007).

1.11.2 Phenotypic studies

Phenotypic and functional measures have also shown a correlation between father and offspring. A study of infants born in Southampton showed a correlation between a father's birth weight and height and the birth weight, length and placental weight of his offspring (Godfrey KM et al., 1997).

It has also been shown that men who fathered small for gestational age infants were more likely to be obese and have larger waist circumferences, but in this study no measure of insulin resistance was made (McCowan LM et al., 2011).

The Exeter Family Study of Childhood Health (EFSOCH) studied almost 1000 normal grown offspring and their fathers, and identified that paternal insulin resistance was inversely correlated with fetal insulin concentrations (Shields BM et al., 2006) but they did not find an association between offspring birth weight and paternal insulin resistance (Knight B et al., 2006).

1.12 The environmental influences on fetal growth

Genotypic differences do not account for all the influence on fetal growth. The maternal environment is also important. Furthermore, the placenta acts as the interface between mother and fetus providing essential nutrients and oxygen. Placental dysfunction is thought to underlie a significant proportion of FGR. A normally functioning placenta relies on a complex series of changes at the vascular interface that produces a low resistance, high capacitance utero-placental circulation allowing for effective oxygen and nutrient transfer between mother and fetus. Maintenance of placental function is energy intensive. Under physiological conditions, the placenta consumes as much as 40% of O_2 and 70% of glucose supplied to the uterus (Meschia G, 1987 & Carter AM, 2000). In the presence of a sub-optimal placental environment one of the consequences will be poor fetal growth.

1.12.1 Placental dysfunction and fetal growth

The placenta is of both maternal and paternal origin and shares the same genotype as the fetus. Placental function is directly linked to fetal development and health, and as such placental dysfunction is implicated in many pregnancy-associated diseases including; pre-eclampsia and recurrent miscarriage as well as underlying a significant proportion of fetal growth restriction. The growth of the placenta and fetus are intrinsically linked, suggesting a common *in-utero* environment benefits both. However, fetal growth, especially towards the later stages of pregnancy outstrips that of the placenta with placental growth plateauing in mid-gestation while fetal growth continues exponentially at a rate of 1.5%/day to term (Molteni RA et al., 1978).

Immune interactions have also been shown to have a role in placental function. In animals, maternal uterine Natural Killer cells (uNK) remodel maternal decidual blood vessels for a healthy pregnancy (Leonard S et al., 2006). Adverse immune interactions between placental antigens, which are specific haplotypes of HLA-C and inherited from both mother and father, and receptors on maternal uNK (KIR) have been associated with the placental syndromes; pre-eclampsia (Hiby S et al., 2004) and recurrent miscarriage (Hiby S et al., 2008).

The development of a low resistance, high capacitance circulation is dependent on efficient spiral artery remodelling at the materno-fetal placental interface. Deficient spiral artery remodelling within the first trimester during embryonic organogenesis seems to be highly associated with the subsequent development of placental insufficiency. Spiral arteries have been shown to be 8 times less likely to be remodelled at high altitude, presumably secondary to low oxygen levels, and birth weight has been shown to decrease with increasing altitude (Zamudio S, 2003). The failure of spiral arteries to develop and cause a reduction in vascular resistance may lead to reduced oxygen tension in the placenta and is hypothesised to limit the differentiation of the invading trophoblast (Tissot van Patot et al., 2010).

Placentas of FGR pregnancies are characterized by reduced syncytiotrophoblast surface area, increased thickness of the exchange barrier formed by the trophoblast and fetal capillary endothelium and an increase in placental apoptosis (Ishihara N et al., 2002 and Mayhew TM et al., 2000). This complex series of maladaptations affect the overall ability of the placenta to provide sufficient oxygen and nutrients to the fetus which in turn affect the ability of the fetus to reach its growth potential *in-utero*. Depending on the degree of placental dysfunction, some fetuses may be able to adapt to the environment and not exhibit reduced growth whilst others will fail to thrive. However, direct cause and effect is not clear and despite being intrinsically linked to eventual poor growth, the underlying mechanisms that cause placental dysfunction and compromised fetal growth may differ.

1.12.2 Maternal nutrition - The 'thrifty phenotype' hypothesis

The 'thrifty phenotype' suggests that the maternal *in-utero* environment is the driving force between poor fetal growth and later development of T2DM (Barker DJP et al., 1993). In the 'thrifty phenotype' exposure to poor nutrition in-utero produces permanent changes in glucose-insulin metabolism and sensing.

Those *in-utero* adaptations predispose an individual to T2DM and metabolic syndrome through increased insulin resistance and decreased insulin sensitivity. Insulin resistance as seen in T2DM is seen as a persistent effect from fetal life in response to adverse conditions and having to maintain glucose levels to the brain at the expense of glucose transport to the muscles for growth (Philips DIW, 1996).

One of the mechanisms through which the environment might influence the placenta and fetus is through epigenetic modifications.

1.13 Environmental influence on gene expression through epigenetic modifications

Epigenetics refers to molecular mechanisms which establish and maintain mitotically stable patterns of gene expression that do not alter the genomic DNA sequence. Epigenetic mechanisms enable developing organisms to produce disparate cellular phenotypes from the same genotype (Jirtle RL et al., 2007). These modifications take the form of DNA methylation, chromatin remodelling and covalent modifications to histones which package DNA into chromatin and include; methylation, acetylation, phosphorylation and ubiquitination of histones (Feinberg AP, 1983).

The best understood example of epigenetic modification is DNA methylation, a covalent addition of a methyl (CH3) group to the nucleotide cytosine. It is maintained during cell division in mammals only at the dinucleotide C-G by the enzyme DNA methyltransferase- 1 (Feinberg AP, 1983).

Epigenetic events may explain the relationship between an individual's genetic background, the environment, aging and disease. Whilst a DNA sequence always remains the same, cells in a specific tissue have the ability to vary their epigenetic state and hence gene expression, through their life course.

The subsequent development of T2DM following a pregnancy that resulted in FGR may follow modifications to genes that were made *in-utero* by a fetus attempting to optimise its growth.

Individual genetic loci methylation levels can be investigated on candidate genes but recent technological advances have meant that the whole genome can be investigated using high throughput methylation arrays. These arrays can simultaneously investigate the methylation status of 27,000 or 450,000 individual CpG sites across the human genome.

1.14 Epigenetics and maternal diet

Maternal diet is one of the ways in which the supply of nutrients vital to DNA methylation may be affected. Famine provides an extreme example of disruption to the supply of nutrients essential in the methylation pathway. The Dutch Hunger Winter famine of 1944-45 is unusual in several ways; first, that the famine was imposed on a previously well-nourished population; second, that there was a sudden onset and relief from the famine; and third, that despite the adversities of the war, midwives and doctors continued to offer professional obstetric care and kept detailed records of the course of pregnancy, the delivery and the size and health of the baby at birth. (Roseboom T et al., 2006). These unique set of circumstances allowed the study of under-nutrition restricted to pregnancy alone, unencumbered by the effects of chronic under nutrition with all its accompanying influences, such as vulnerability to infection. Individuals who were exposed to the famine in-utero have been shown to have lower birth weights (Lumey LH et al., 1992.) and, furthermore exposure to famine during any stage of gestation was associated with glucose intolerance later in life. More coronary heart disease, a more atherogenic lipid profile, disturbed blood coagulation and more obesity among those exposed to famine in early gestation was also noted (Roseboom T et al., 2006).

The environmental factor that influences CpG methylation is methionine, the source of methyl groups, via a pathway dependent on folic acid (Giovannucci E et al., 2004). Individuals born during the Dutch famine born 60 years ago were later found to have less DNA methylation of the imprinted *IGF2* gene compared with their unexposed same sex siblings (Heijmans BT et al., 2008). This observation supports the hypothesis that exposure to adverse environmental factors *in-utero* could permanently alter epigenetic marks which might have a bearing in adult disease risk.

1.15 Placental methylation

The human placenta has a strikingly different DNA methylation profile relative to other somatic tissues (Christensen BC et al., 2009). This includes low global DNA methylation compared to healthy somatic tissue (Gama-Sosa MA et al., 1983). The functional role of lower methylation is not fully understood but may include the regulation of genome 'plasticity' and gene regulation. Whilst the overall pattern of methylation seen in the placenta is of hypomethylation there is also evidence of a programmed and reproducible change in methylation that occurs with increasing gestational age. Global methylation increases as a placenta matures (Gama-Sosa MA et al., 1983 & Macaulay EC et al., 2011) and genes involved in immune regulation are highly prevalent amongst those whose methylation changes with gestation (Novakovic B et al, 2011). These dynamic changes may help explain the time-of-exposure-dependent effects of certain environments or stresses on fetal health at different gestational ages (Laplante DP et al., 2004) but could also reflect a change in cell composition of placental tissue (Dancause KN et al., 2011).

1.15.1 Methylation of imprinted genes in the placenta

Imprinted genes are highly conserved and expressed in the placenta. These genes are effectively driven by epigenetic regulation theorized to be controlled by differential methylation of the relevant genes (Haycock PC et al., 2009). It has been hypothesized that sporadic loss-of-imprinting induced through methylation change could occur in human placentas and contribute to abnormal placental development and consequentially decreased fetal growth (Maccani MA et al., 2009).

Within the placenta regulation of imprinted gene expression appears to be less stable than in the fetus itself. This allows the placenta to better adapt to changing physiological environments but with potential adverse consequences such as poor growth (Novakovic B et al., 2010 & Nelissen EC et al., 2011).

Paternally expressed insulin-like growth factor 2 (*IGF2*) gene, is imprinted and is a major modulator of both placental and fetal growth (Constancia M et al., 2002). *IGF-2* is co-ordinately regulated by a differentially methylated CTCF binding region

known as imprinting control region 1 (ICR1) or differentially methylated region 2 (DMR2) along with H19 promoter activation (Murrell A et al., 2004 & Engel N et al., 2004). In growth restricted placentas the ICR1 region has been found to be significantly hypomethylated compared with normal grown placentas, leading to reduced expression of *IGF2* Decreased placental methylation at the *H19/IGF2* imprinting control region is associated with normotensive intrauterine growth restriction but not pre-eclampsia. (Bourque DK et al., 2010). Furthermore, previous studies have reported a reduction in *IGF2* expression in placentas from pregnancies associated with FGR (McMinn J et al., 2006) and complete loss of placental *IGF2* expression is associated with fetal growth restriction in mice (Fowden AL et al., 2006).

Reduced placental *IGF2* expression leads to a reduction in size of all placental layers and alters the diffusional exchange characteristics of the placenta (Sibley CP et al., 2004). In human pregnancies, reduced exchange surface area, and likely reduced transfer capacity of the placenta, has been noted in growth restricted placentas (Mayhew TM et al., 2003).

1.15.2 Methylation of *IGF2* control regions in placenta and cord blood

Studies have investigated DNA methylation of the regulators of *IGF2* in relation to other tissues and postnatal cord blood. The ICR1 region was found to be similarly methylated in placenta and cord blood (around 50% which would be expected for monoallelic expression) but reduced in the DMR2 region in placenta and normally methylated in cord blood (Guo L et al., 2008).

However, these results have not been reproduced. A further investigation found no methylation difference in these regions between FGR and normal grown placentas or between cord blood samples (Tabano S et al., 2012).

1.16 DNA Methylation in placental diseases

DNA hypomethylation at gene enhancer regions has been identified in a study that investigated the methylation status of placentae from pregnancies affected by pre-eclampsia (PET). Using the Infinium Human Methylation 450 BeadChip®, 20 placentas from preterm PET pregnancies across a range of gestational ages were compared to 20 pre-term placentae unaffected by pre-eclampsia. However, sub-group analysis revealed no identifiable DMPs between cases of PET with concurrent FGR and those with PET and appropriate fetal growth (Blair JD et al., 2013)

In a study that used the Infinium HumanMethylation27 BeadChip® platform, the precursor to the 450 Beadchip, 206 term placentas identified 22 methylation loci that could act as markers for poor growth Functional pathway analysis for the 22 genes where a DMP was identified centred on placental implantation and materno-fetal interface factors (Banister C et al., 2011)..

1.17 DNA methylation in umbilical cord blood

DNA methylation has been investigated in umbilical cord blood. Using the HumanMethylation27 BeadChip a cohort of appropriately grown offspring (32-43 weeks gestation) were found to have 25 differentially methylated positions in whole cord blood DNA that were thought to be related to gestational age (Schroeder DW et al., 2011).

However on the same platform, no genome wide significant methylation differences were seen in cord blood related to birth weight in a term population (between 2.1kg and 5kg) (Adkins RM et al., 2012).

In a small cohort of term pregnancies (n=5) affected by fetal growth restriction, CD34+ cells were isolated from umbilical cord blood and methylation differences were found that were identified as being important in the pathway involving hepatocyte nuclear factor 4α (*HNF4A*) gene (Einstein F et al., 2010). As *HNF4A* is a

candidate gene for non-insulin dependent diabetes mellitus this observation opens up the possibility that differential methylation of genes important in diabetes might play a role in fetal growth.

1.18 Animal models supporting a role for DNA methylation in FGR

In an animal model of FGR induced through utero-placental insufficiency (secondary to ligation of the uterine arteries), cytosine methylation in pancreatic islet cells was different at approximately 1400 loci in male rats at 7 weeks of age, before they went to develop diabetes. Of the top 53 genes almost half of those tested were associated together in a single functional network centred on a collection of important metabolic and cellular regulators (Thompson RF et al., 2011). The majority of changes occurred in evolutionarily conserved DNA sequences with some loci in proximity to genes manifesting changes in gene expression which were enriched near genes regulating vascularization, proliferation and insulin secretion (Stoffers DA et al., 2003 & Simmons RA et al., 2001).

A rodent model of FGR induced through placental insufficiency, showed a decrease in the *PDX1* gene transcription mediated through a cascade of epigenetic modifications culminating in silencing of the *PDX1* gene (Park JH et al., 2008). Histone modifications in association with DNA methylation differences in the promoter region of the gene *PDX1* were identified in growth restricted rodent offspring, eventuating in permanent gene silencing and a diabetic phenotype (Park JH et al., 2008). In health *PDX1* regulates pancreatic beta cell differentiation, but when silenced T2DM eventuates and this study offered a new insight in epigenetic mechanisms linking FGR to T2DM development (Brissova M et al., 2005). *PDX1* has also been implicated in insulin transcription, acting to recruit histone deacetylases (HDACs) to the insulin promoter region resulting in down regulation of insulin gene transcription under conditions of low glucose (Mosley AL et al., 2004). Epigenetic regulation of genes that influence insulin transcription offers novel insights into potential mechanistic pathways that link poor fetal growth with later risk of T2DM.

1.19 Inheritance of epigenetic modifications across generations

It is believed that DNA methylation throughout the genome of the developing zygote is almost completely removed during the cleavage phase of development. In normal conditions, methylation then re-establishes itself and in normal conditions in a tissue specific fashion (Rakyan V et al., 2006). This patterning is not limited to the embryo, but also occurs in a specific fashion in the extra embryonic lineages (placenta), although the overall levels of methylation in the placenta are significantly lower than that in the somatic tissues of the fetus (Jaenisch R, 1997).

The concept that epigenetic modifications can be inherited is controversial due to supposed resetting of methylation in gametes. There is, however, emerging evidence to support the concept that some methylation marks transmitted through the gamete are maintained. The Agouti viable yellow (A^{vy}) allele in mice is responsible for coat colour and, if over-produced, other phenotypic consequences such as obesity. Depending on the methylation status of this allele, coat colour can vary from yellow (fully methylated) to brown (not methylated) (Morgan HD et al., 1999). It has been shown that A^{vy} methylation can be transmitted and maintained through the female germline, even when accounting for other factors (Morgan HD et al., 1999) supporting trans-generational inheritance

In male rats, exposure to a high-fat-diet (HFD) program has been shown to result in β-cell 'dysfunction' in rat F_1 their female offspring (Ng SF et al., 2010). Chronic HFD consumption in Sprague–Dawley fathers induced increased body weight, adiposity, impaired glucose tolerance and insulin sensitivity. Relative to controls, their female offspring had an early onset of impaired insulin secretion and glucose tolerance that worsened with time.

The offspring of these obese male rats showed an altered epigenetic expression of more than 600 pancreatic islet genes. The greatest difference in expression was found in a gene called *Il13ra2*. In the daughters of obese fathers, the level of methylation of this gene was around 25% of the level seen in the control daughters

This is the first report in mammals of intergenerational transmission of metabolic sequelae from father to offspring that might be mediated through an epigenetic mechanism. It opens up the possibility that transmission of epigenetic, rather than genetic, modifications may be responsible for the growth restricted phenotype seen in infants and the diabetic susceptibility identified in their fathers.

1.20 Genetic-epigenetic interactions in T2DM

Emerging work centres on the complex interplay between genotype and epi-genotype. It is well established that SNPs can affect methylation. Within the 450HM array, Illumina (Illumina Inc., San Diego) have published over 150,000 recognised potential SNP/ methylation interacting loci.

However, it seems far from clear whether DNA methylation is an active or passive process in relation to genetic variation. Using fibroblasts, T-cells and lymphoblastoid cells from the umbilical cords of 204 babies it was shown that inter-individual DNA methylation changes were mechanistically associated to genetic variation and gene expression in complex and context dependent ways that could in fact be of passive or active nature (Gutierrez-Arcelus et al., 2013).

Evidence supports both genetic and epigenetic factors in development of T2DM. Whilst over 40 genetic loci have been identified in T2DM their individual molecular function in disease development is on the whole unclear. In many cases it has not been possible to identify any link to the hallmarks of T2DM, namely insulin resistance and/or beta cell dysfunction. The possibility arises that methylation differences secondary to SNPs may be critical in the pathway to disease and provide a molecular explanation for the mechanism by which the T2DM risk alleles act.

It has been shown that genetic, epigenetic and non-genetic factors integrate to influence the expression of a candidate gene for T2DM in human skeletal muscle. A SNP near the gene *NDUFB6* is able to introduce a CpG site which is associated with increased DNA methylation, decreased gene expression, and insulin resistance in the muscle of elderly, but not young, individuals (Ling C et al., 2007).

A comprehensive study of T2DM SNP loci in conjunction with methylation differences has shown that 19 of 40 known SNPs associated with T2DM cause the introduction or removal of a CpG site. These CpG positions exhibit differential methylation in human pancreatic islet cells. Furthermore, in a number of instances, surrounding CpG positions, methylation was also affected by the presence of the SNP (Dayeh TA et al., 2013)

SERPINA3 (Serpin peptidase inhibitor clade A member 3), is a protease inhibitor involved in a wide range of biological processes. It is up-regulated in human placental diseases in association with a hypomethylation of the 5′ region of the gene (Chebil ST et al., 2012).

It has been shown that the promoter of *SERPINA3* is transcriptionally activated by three transcription factors (TFs) (*SP1, MZF1 and ZBTB7B*) with the level of induction being dependent on the SNP rs1884082, the T allele being consistently induced to a higher level than the G, with or without added TFs. When the *SERPINA3* promoter is methylated, the response to *ZBTB7B* was allele specific (the G allele was strongly induced, while the T allele was strongly down regulated). Interestingly in a small case/control study of pregnancies, the presence of T allele seemed to predispose to growth restriction (Chebil St et al., 2012).

This interaction between methylation and genotype which could lead to alteration of gene expression and, therefore, disease provides a novel insight into how communicating genetic and epigenetic mechanisms could be involved in placental disorders such as fetal growth restriction.

1.21 Rationale for study

Fetal growth restriction is associated with future risk of T2DM and it is hypothesised that these disorders are two phenotypes of the same genotype expressed under different environmental conditions; in utero and as an adult with abundant nutrition. Epidemiological studies show diabetic fathers have smaller babies and that smaller babies are themselves more prone to diabetes as adults. This evidence supports the inheritance of risk alleles for fetal growth and T2DM from the father. However it is also possible that genes and in utero environment interact through epigenetic modifications that result in both FGR and T2DM. The pathophysiology of these interactions is still unclear. A more complete understanding may identify interventions before or during pregnancy, or in early neonatal life that might reduce future risk of T2DM. Primary prevention of T2DM started early in life is more likely to be effective at reducing morbidity and mortality of this globally endemic condition.

1.22 Hypothesis

Fetal growth restriction secondary to placental dysfunction is mediated in part by genetic factors inherited from father to offspring.

Fetal growth restriction secondary to placental dysfunction creates an environment that modifies the fetal epigenome and which predisposes the fetus to T2DM in later life.

1.23 Thesis Aims

- To investigate the fetal insulin hypothesis in relation to a growth restricted cohort by the use of a case-control study. Specifically to identify if there is a paternal phenotype in men who father fetal growth restricted offspring that predisposes to T2DM.
- To discover, using whole exome sequencing, if novel genetic variants are present in FGR offspring, which could contribute to disease pathogenesis. Furthermore, to investigate known candidate genes (implicated in MODY and T2DM) in FGR offspring. To validate findings and establish patterns of inheritance.
- To explore the DNA methylation profile in FGR offspring in placental tissue and cord blood using a whole genome approach. Furthermore, to assess potential functional significance of methylation differences using other bioinformatics tools.
- To interrogate potential genetic-epigenetic interactions from the data gathered on these FGR offspring that provide plausible functional pathways in FGR.

Chapter 2

Research Methods and Materials

Chapter 2: Research Methods and Materials

2.1 Phenotype study

The research focused on fetal growth restriction and characterisation of parental phenotype. A case control study was designed which is described in detail in sections 2.1 to 2.6. Physical, vascular and metabolic measurements were made on mothers and fathers as described in section 2.7 to 2.9.

DNA was extracted from whole blood (parents) and cord blood and placenta (offspring) of participants in the phenotype study. The quality of extracted DNA was assessed as described in section 2.12 to 2.12.4. Exome sequencing and validation was performed on samples as described in sections 2.13 and 2.14. DNA methylation was analysed using the Illumina Human methylation 450 array as described in sections 2.15 and 2.16.

2.1.1 Case- control study design

An unmatched case control study was undertaken of babies with fetal growth restriction (cases) and babies of appropriate fetal growth (controls). The study was carried out at University College London Hospital (UCLH) between September 2009 and May 2011. Ethical approval for the study was granted by the Joint UCLH/UCL (alpha) Ethical Committee (09/H0715/28). All participants gave informed consent.

2.1.2 Statistical methods:

Multivariable logistic regression was used to analyse phenotype data with case or control as the outcome measure. Prior to study commencement, the primary explanatory variables to be explored were paternal insulin resistance [using the Homeostasis model assessment (HOMA) with glucose (mmol/l) and insulin (pmol/l)] and paternal endothelial function by flow mediated dilation (% change in diameter of

brachial artery between before and after flow i.e. 100x(after-before)/before). Ten events (cases) are required for each coefficient estimated in a logistic regression. In this analysis eight coefficients were estimated. Therefore, sample size was calculated at 80 cases and 80 controls.

2.1.3 Sample size calculations

Sample size calculations were undertaken using NQuery 6.0 to show the size of effect detectable in a case-control study with 160 participants. A total of 151 observations are required to detect an odds ratio (OR) of 0.5 (or 2.0) as significantly different from 1.0 with 80% power and 5% two-sided significance when one normally distributed covariate is added to the model after adjustment for prior covariates and it is assumed that the covariate's multiple correlation with covariates already in the model is 0.4 and that the ratio of cases to controls at the mean is 1:1.

In logistic regression an OR of 0.5 for a continuous variable implies that there is a 50% decrease in the odds of being a case for every one unit increase in the continuous variable and an OR of 2.0 for a continuous variable implies that the odds of being a case is doubled for every one unit increase in the continuous variable. Thus if a log transformation is applied to the primary explanatory variable IR so that the variable is normally distributed and IR is input as 10 x log(IR) then the study is large enough to detect a doubling in the odds of being a case for every 0.1 increase in log(IR). Or if the primary explanatory variable FMD then the study is large enough to detect a 50% decrease in the odds of being a case for every 1% increase in FMD, after adjustment for other covariates under the above assumptions, after adjustment for other covariates under the above assumptions.

2.1.4 Statistical analysis

Cases and controls were assessed for comparability with simple descriptive statistics. Baseline characteristics of cases and controls are summarised with means and standard deviations (or medians and interquartile ranges, if appropriate). 95% confidence limits are given for continuous variables, e.g. age, and frequency counts and percentages for categorical variables, e.g. sex. Statistical analysis was carried out using STATA 10 (StataCorp LP, Texas USA) and GraphPad Prism version 5. Visual

assessment of normality curves for variables was undertaken to assess for normality. When comparing the means of two variables an unpaired t-test was used for continuous variables and a 2x2 table for categorical data. Two tailed P-values <0.05 were considered statistically significant.

Multivariable logistic regression was carried out on paternal phenotype data using STATA. Results are reported as odds ratios with 95% confidence limits. Explanatory variables were assessed as significant at the 5% level.

2.2 Feasibility of recruitment

An audit performed in 2007 showed almost 300 babies whose birth weight was <10th centile were delivered at UCH (2007 delivery rate of 3,450). Delivery rates subsequently increased to nearly 5,000/ year with more cases of FGR being seen. At the start of the study a further audit was done. As customising of birth weights is not routine at UCH the audit investigated clinical information about babies born of low birth weight (defined as babies born at <2.5kg). Between August 2010 and July 2011, 120 babies were born weighing <2.5kg.Of these babies, 50% were delivered early due to maternal complications and were appropriately sized for their gestation; 30% were born early but appropriately sized for their gestation (classified as spontaneous preterm deliveries); 10% of babies born <2.5kg had fetal anomalies and in 22.5% of cases FGR was identified. 1/3 of these cases of FGR were accompanied by maternal pre-eclampsia.
Extrapolation of these figures suggested 24 babies <2.5kg would be eligible for recruitment to the study.

Assuming half of all babies <10th centile are likely to be constitutionally small and not be eligible for inclusion and of the remainder of these about a fifth (based on the audit figure) would meet the inclusion criteria, then in a unit delivering 5,000 babies/year one would expect to identify 50 babies/ year.

Hence, based on an 80% uptake for recruitment, a 2 year recruitment period for the study was felt to be justified.

2.3 Recruitment criteria

In order to minimize differences in fetal growth caused by the maternal *in utero* environment, pregnant mothers were included as cases or controls only if:

- Aged 18 to 45 years
- Body mass index (BMI) 20 - 35kg/m2,
- Non smokers
- No significant maternal medical problems
- No drug, alcohol or substance use in pregnancy

And where the pregnancy was:

- A natural conception
- Singleton pregnancy

Women whose partner smoked had an additional test for cotinine (ABS Laboratories Ltd., Birmingham, UK), a metabolite of nicotine as an indicator of passive smoking.

Couples were recruited as cases with a pregnancy affected by fetal growth restriction (FGR), as defined by an antenatal $<10^{th}$ customised birth weight centile where induction of labour or delivery by caesarean section was planned in view of fetal size.

Pregnant women and their partners who were thought to be having a normally grown baby with an estimated fetal weight between 10^{th} and 95^{th} customised centile were recruited to the study as controls. Participants were offered an additional fetal ultrasound scan at 34 weeks, which I performed, to confirm predicted size.

Women and their partners who met inclusion criteria were approached while attending antenatal clinics or fetal ultrasound sessions. Some eligible participants responded to a research poster.

2.4 Diagnosis of FGR

FGR is defined as a fetus who has failed to reach its growth potential (Pallotto E et al., 2006). The diagnosis of FGR *in-utero* remains a difficult subject. A small for gestational age (SGA) baby is commonly defined as one who's estimated fetal weight and/or abdominal circumference (AC) on ultrasound is less than the 10th centile for its gestational age (RCOG Green Top). However, this cohort is a heterogeneous group and includes constitutionally small babies, pathologically small babies and babies small due to known causes such as chromosomal abnormalities. True fetal growth restriction reveals an inability of a fetus to reach its (unique) growth potential not simply a predisposition to being small compared to a cohort of babies born at the same gestation. Babies born with true FGR have a much poorer prognosis than those who are SGA (Jones, 1986). In order to improve our prediction of true FGR, customized antenatal fetal growth charts (www.gestation.net) were used.

EFW and AC centile are customised depending on maternal height, weight, fetal sex and ethnicity, which have all been shown to have a bearing on expected growth potential. Customising has been shown to be able to reclassify 27.5% of SGA fetuses (as defined by 10th centile) back to being appropriately grown (Mongelli & Gardosi).

2.5 Exclusion Criteria

In order to exclude babies who are small due to known causes and, therefore, not likely to be secondary to placental insufficiency the following exclusion criteria were made;

1. Small babies in the presence of severe maternal medical problems such as cardiac, endocrine, renal or haematological disease
2. Known maternal drug or excess alcohol intake
3. Multiple pregnancies
4. Known fetal congenital infections such as CMV, toxoplasmosis
5. Known congenital chromosomal abnormalities
6. Known congenital structural malformations

2.6 Study Protocol

All study assessments were carried out in the Clinical Research Facility, UCLH (Figure 2.1). This facility has a dedicated temperature controlled (maintained at 24^{o}C) clinical study room.

Men were studied within 4 weeks of offspring birth; women were asked to return for a study visit six months after the birth of their child. Participants were asked to fast overnight for at least 10h before the study.

Each study visit required approximately 45 minutes to complete. All subjects were given an information leaflet and asked to sign consent forms prior to commencing the investigations.

Figure 2.1 The study room in the Clinical Research Facility containing equipment necessary to perform studies

Subjects completed a questionnaire enquiring about past medical, family and treatment history (Appendix 4 S9). Study participant's own birth weight was recorded as remembered personally or from a parent. Weight, height and waist circumference (measured twice between the top of the iliac crests) were recorded. After resting, two measures of supine blood pressure (BP) were taken 15 minutes apart.

Endothelial function was assessed by performing flow mediated dilatation (FMD) on the brachial artery and pulse wave velocity recordings were undertaken.

Feedback to participants about results was available. Clinically significant results were conveyed to the subject as a priority along with relevant support literature and advice. Where appropriate, general practitioners (GPs) were contacted on behalf of the subject.

Fasting venous insulin, glucose and lipid levels were measured. Insulin resistance was calculated using HOMA (Matthews DR et al., 1985). The blood was centrifuged within 1hr of venepuncture and plasma and serum were frozen at -80^{o}C. All blood samples were processed in the same laboratory within the Clinical Research Facility.

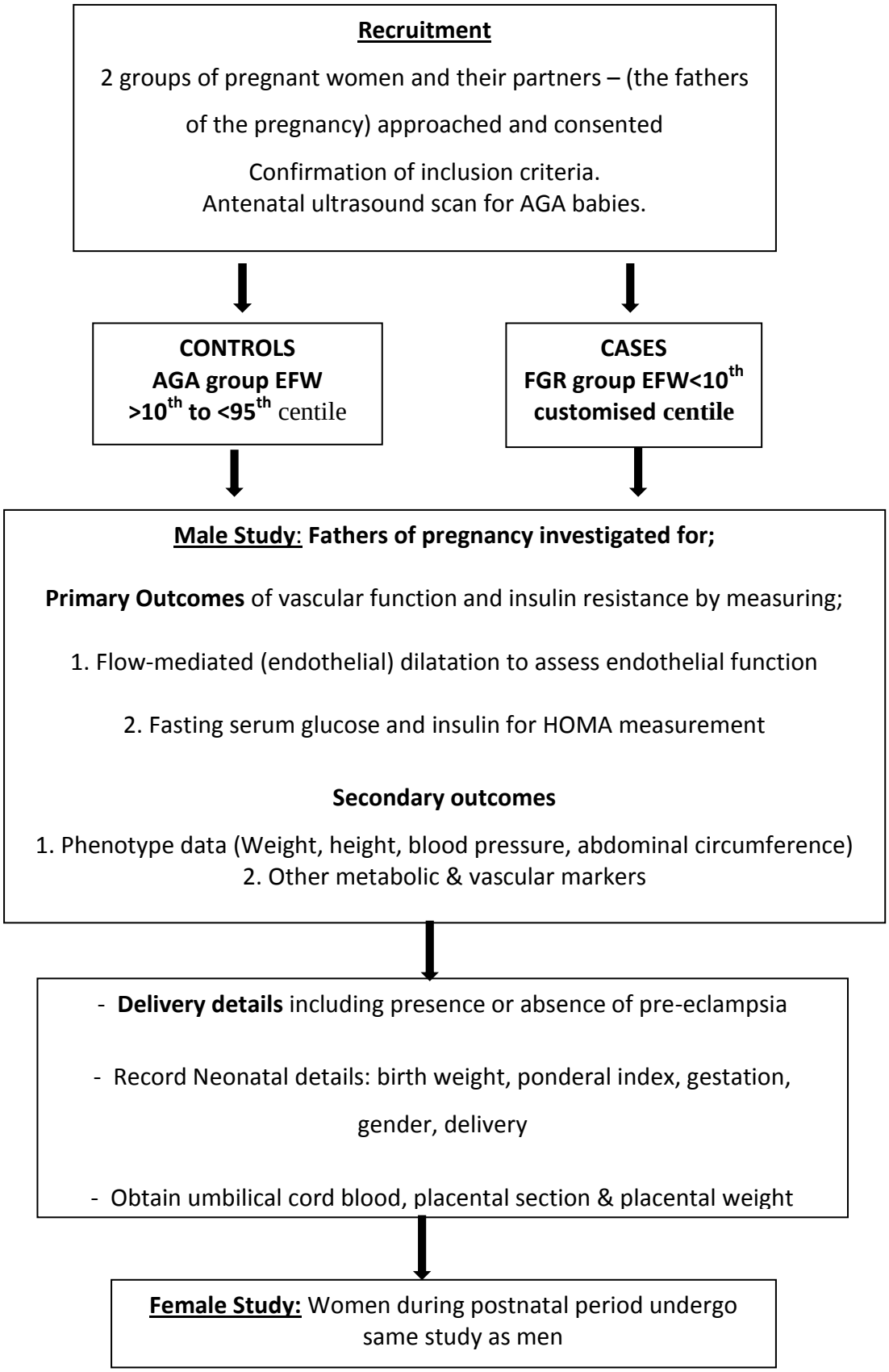


Figure 2.2 Flow diagram of recruitment pathway along with timeline for study visits and measurements being taken.

2.7 Homeostasis model assessment (HOMA) measurement

HOMA was first described in 1985 (Matthews DR et al., 1985) and it uses fasting state glucose and insulin to predict beta cell function and insulin resistance.

It has been published over 500 times and in more than half of these studies a non-diabetic population was used. The model has been validated compared to both direct and indirect methods of assessing beta cell function and insulin resistance. It is described as one of a family of "paradigm models" which use theoretical solutions adjusted to population norms in order to predict outcomes (Wallace TM et al., 2004).

The original model was improved with the understanding that there is variation in hepatic and peripheral glucose resistance. A second updated model was published in 1996. In contrast to the original model, this version allows assessment of hyperglycaemic subjects and an estimate of proinsulin secretion can be made. This is significant from a clinical laboratory perspective because it permits the use of either total radioimmunoassay or specific insulin assays.

2.7.1 Sampling

Taking three paired samples of glucose and insulin is theoretically better due to the pulsatile nature of insulin secretion. However, this has been shown to be unnecessary in practise. It has been suggested that individual sampling is an acceptable practise (Wallace TM et al., 2004). This study, which included thirty subjects, demonstrated individual paired samples yielded near perfect correlations to the mean taken from three paired samples taken at five minute intervals.

2.7.2 Validation

HOMA has been validated compared to several other methods to measure IR and beta cell function. The "gold standard" test looking at these measurements is often quoted as the hyperinsulinemic euglycemic clamp which acts as a complex stress test and is time and labour intensive. HOMA correlates well to this test and furthermore is simpler to apply to large cohorts of subjects (Bonora E et al., 2000).

The model has been used in more than 150 epidemiological studies looking at groups and their differing glucose tolerance. A study from Mexico looked at 1,449 individuals with normal and impaired glucose tolerance and showed that those with higher baseline IR were more likely to go onto develop diabetes during the 3.5 year follow up period (Haffner SM et al., 1996) .
A comparison of HOMA reproducibility and the OGTT suggested that the HOMA model reflects the most physiological reaction and differentiates well between groups with normal glucose tolerance, impaired glucose tolerance (IGT) and T2DM (Levy JC et al., 1998).

2.7.3 Evidence for impaired glucose tolerance and T2DM risk

There are numerous papers exploring impaired glucose tolerance in relation to the development of T2DM. In a meta-analysis of 28 studies that looked at the relative risk of developing diabetes in populations of subjects with and without impaired glucose tolerance there was a six fold increase in risk (95% confidence interval 4.66 to 7.38). Furthermore there was a 1.67- 3.08 relative risk of suffering from a fatal cardiovascular event in association with IGT (McMaster University Evidence Based Practice Center).

2.8 Endothelial function measurement

I measured endothelial function using brachial artery flow-mediated dilatation (FMD) in the quiet, temperature-controlled CRF study room (in accordance with previously reported protocols in Okorie et al., 2011).

The vascular endothelium responds to flow by releasing mediators that cause dilation. By measuring the diameter change of the brachial artery during flow situations it has been shown that endothelial function can be reliably assessed (Calermajer et al., 1997).
FMD should be performed in a standard way in order to achieve a high degree of accuracy and reproducibility. Using an ultrasound probe in the longitudinal plane, a segment of brachial artery is identified and studied before and after a five minute

compressive arterial occlusion. FMD is expressed as a percentage change from baseline brachial artery diameter (mm) to maximum dilation (mm) after hyperaemia.

2.8. 1 Image Acquisition

Participants were positioned on a bed and left to rest for at least ten minutes, having fasted overnight. The right arm was rested in an arm holder and the forearm was raised slightly using a foam pad (see figure 2.3). An 8.5cm wide pneumatic cuff was then placed 2cm distal to the medial epicondyle. Cuff pressure inflation was controlled by an automatic cuff inflator (Hokanson Cuff Inflator. PMS Instruments Ltd, Maidenhead, UK)

The ultrasound machine used was an Aloka SSD 5000 and the probe a 13MHz linear array transducer (Aloka Holding Europe, AG Switzerland).The image was acquired in B-mode and the probe fixed so that a 5-10cm segment of the brachial artery proximal to the antecubital fossa was in view.
Longitudinal end diastolic images, in conjunction with ECG tracing were acquired every three seconds during the 11minute recording (1minute baseline, 5minutes with cuff inflation at 300mmHg and 5minutes with cuff deflated). Blood flow velocity was continuously monitored during the scan by switching to a B/D (Doppler) mode for the test duration.

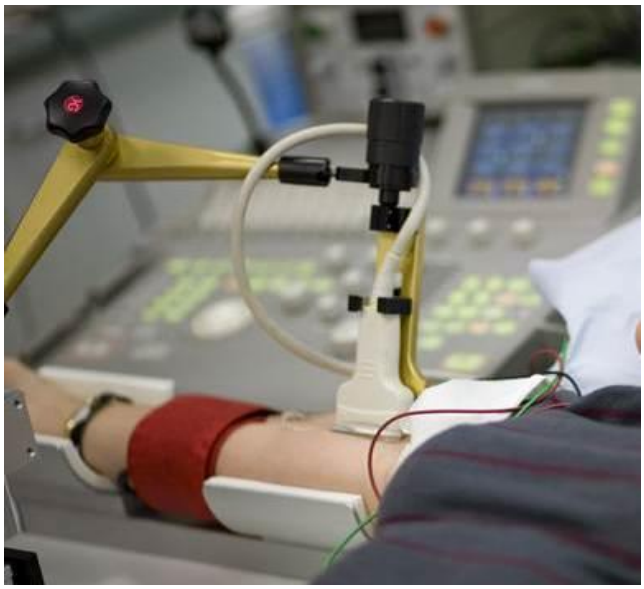


Figure 2.3 The apparatus used in order to study the brachial artery by ultrasound. Ultrasound probe is held in position over the brachial artery with subject's arm resting.

2.8.2 Analysis

Images were acquired and analysed (Figure 2.4) using automated software (Brachial Tools Medical Imaging Applications, Iowa, USA). With the initial segment recorded a smaller region of maximum clarity was selected for analysis. Internal reproducibility was performed by looking at different segments within each scan.

Analysis exclusion criteria were poor image quality, if recording was deemed to be unrepresentative of the true vessel diameter or if the image did not return to near baseline after reperfusion. Each image-recording was validated by a second operator blinded to the subject's group. This led to 22 (18.5%) scans being excluded from the final analysis due to lack of agreement.

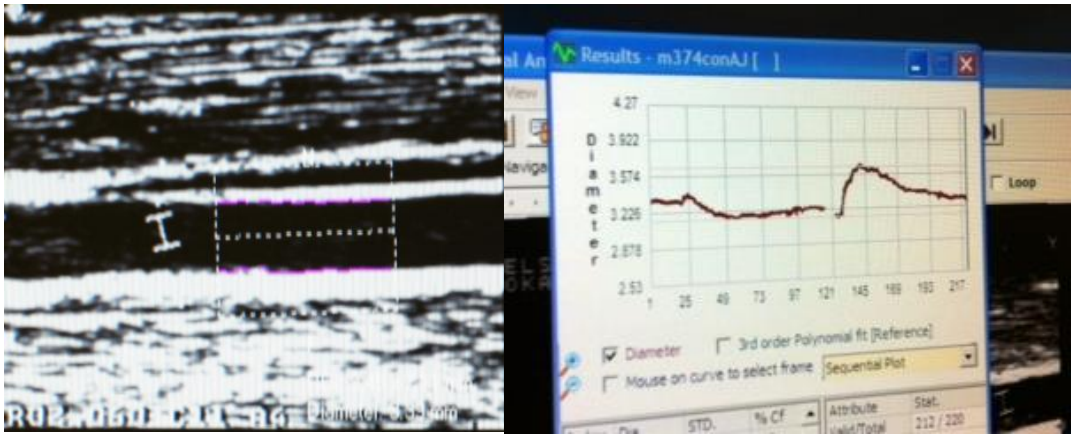


Figure 2.4 The image on left is picture obtained on the screen of the brachial artery and the box is where the analysis of flow occurs. The image obtained by the software of the dilation recorded is shown on the right)

2.9 Pulse Wave Velocity

Pulse wave velocity was recorded using the Vicorder automatic device (Smart Medical, Moreton-in-March, Gloucester)
Participants were placed in a semi-prone position with the head and shoulders raised approximately 30 degrees. The neck pad was placed around the participant's neck with the pressure pad over the carotid artery and care taken not to over tighten.

An appropriately sized cuff was placed around the right upper thigh. Both neck and thigh cuffs were attached to the Vicorder. A measurement was taken from the suprasternal notch to the middle of the thigh cuff using the umbilicus as a midpoint landmark. This was input along with the lying blood pressure reading. Cuff inflation then occurred, resulting in waveform generation. Once a steady pattern was established the image was frozen and data recorded (Figure 2.5).

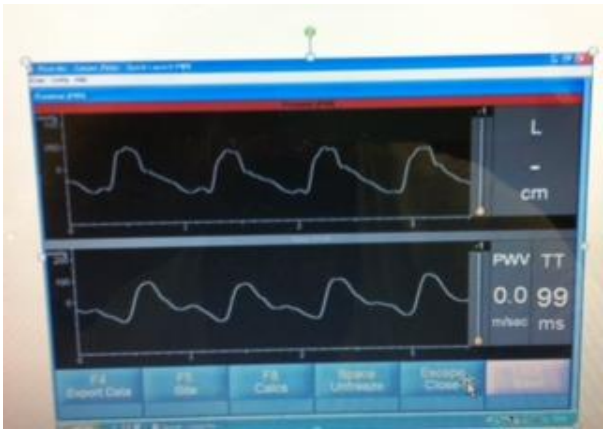


Figure 2.5 Carotid artery recording top and below is femoral. TT is transit time

2.10 Sample collection, handling and storage

2.10.1 Parental Blood Samples

Blood was taken at the time of the study visit for fathers. Mothers had blood taken when recruited antenatally and postnatally as part of their study visit.

Venepuncture was performed on each participant and 20mls of blood removed. All participants had fasted and been resting for 20minutes before sampling for study visits. All sampling occurred in the study room which was a temperature controlled area.

Following the study visit whole blood was sent to the UCLH biochemical laboratory using routine validated biochemical assays for the following measurements: fasting glucose, insulin, triglycerides (TG), total cholesterol (TC), high density lipoprotein cholesterol (HDL) & low density lipoprotein cholesterol (LDL), urea, creatinine and

C-reactive protein (CRP). LDL cholesterol was calculated using the Friedewald formula.
Homeostasis model assessment (HOMA) of insulin resistance (IR) was measured from fasting glucose and insulin values (HOMA-IR = fasting glucose x fasting insulin/ 22.5). These values were calculated using a freely available software package from the Diabetes Trial Unit, Oxford, UK (http://www.dtu.ox.ac.uk/homacalculator/download.php)

Whole blood was also stored in two aliquots and frozen immediately at -80^{o}C. Blood was centrifuged at 3000rpm for 10minutes. The resulting plasma and serum was stored in 2-4 aliquots at -80^{o}C within one hour of the sample being taken.

All samples were colour coded, logged and stored according to being a case or control. All participants were allocated unique anonymised identification number. These allowed family units to be linked (fathers, mothers and babies) and enabled one to identify if a participant was a case or control.

2.10.2 Placental like growth factor (*PLGF*) measurements

Aliquots of 250µl of plasma, extracted from paternal whole blood, were assayed for *PLGF* by Dr Anna Jeffery Smith using the Alere Triage *PLGF* test (Alere Inc., San Diego). This assay employs antibodies specific for the *PLGF-1* isoform and runs on the commercially available, point of care Alere Triage MeterPro (Alere, San Diego, USA). The MeterPro is a portable device which can be used in combination with the Triage test cartridges to analyse the concentration of a range of analytes in blood. The device has a laser that hits a test cartridge when inserted in the meter. This causes fluorescent dye in the test cartridge to emit energy. The amount of energy the fluorescent dye emits is proportional to the concentration of analyte in a sample.

2.11 Fetal samples

Immediately following delivery of the placenta, umbilical cord blood was taken from the umbilical vein or artery, centrifuged and stored at -80^{o}C as plasma and serum for later measures of metabolites. Gestational length, offspring gender, weight and length were recorded.

Plasma extracted from whole cord blood was used to assess insulin, c-peptide and glycated haemoglobin by TDL laboratories (London, UK) using standardized assays and methods.

2.12 DNA extraction

2.12.1 DNA extraction fetal cord blood and parental whole blood

Cord blood was extracted from the umbilical cord immediately after delivery of the placenta. A portion of free flowing vein or artery was used to collect up to 20mls of blood. DNA extraction was performed as soon as possible after collection. If any delay was expected in extracting DNA it was refrigerated at +4^{o}C prior to extraction.

DNA was extracted with the QIAamp DNA Mini Kit (QIAGEN Ltd, Manchester UK) using 200µl of blood. At least three aliquots were extracted for each subject. The optional step to remove RNA was used and additional steps recommended were carried out. Elution was performed using 200µl RNA/DNA free water. Water was used in preference to TE after discussion with Deborah Hughes from the Institute of Neurology due to concerns that the TE could interfere with processes required in DNA preparation for exome sequencing. DNA was stored at -20^{o}C once eluted.

2.12.2 DNA extraction of placental and cord tissue

Placental and umbilical cord tissue were stored at -80 ^{o}C, removed at time of extraction and placed onto ice. For each sample, 480µl of buffer solution and 20µl

proteinase K solution (Qiagen Ltd, UK) was put into a 1.5ml eppendorf. The buffer solution contained 50mM Tris 8.0, 10mM EDTA and 0.5% SDS and was suggested by Dr Melissa Smart, Queen Mary University, London. A 5mm slice of tissue (around 10g) was removed by scalpel and placed in the buffer.
This was left in a dry heater at 55^{o}C overnight in order to digest the tissue. Once all tissue appeared digested the solution was vortexed for 15 seconds and 200µl of solution was removed and used to continue the extraction process.
DNA was isolated from placental tissue (paired with cord blood samples) using the DNeasy Blood & Tissue Kit (Qiagen, UK) according to manufacturer's protocols. DNA samples were then further processed to ensure DNA quality using the Genomic DNA Clean & Concentrator Kit (Zymo Research, Cambridge Biosciences, UK).

2.12.3 Genomic DNA clean-up

In order to improve the quality of genomic DNA (improved 260/230 ratio), a genomic DNA clean-up kit was used (Zymo Research, through Cambridge Bioscience, UK) according to manufacturer's instructions. All samples that were being run on the arrays were cleaned up, including DNA from placentas and cord blood.

2.12.4 DNA quantification and quality

DNA concentration and purity was assessed by a Thermo Scientific NanoDrop™ 1000 spectrophotometer according to the manufacturer's protocol. DNA concentration from all individuals sequenced was also assessed by fluorometric assay using a Qubit® 2.0 Fluorometer (Life Technologies) following manufacturer's instructions. A 1% Agarose gel was run in order to assess DNA integrity.
Agarose gels were prepared by weighing out the appropriate amount of agarose in a conical flask and adding the correct volume of 1 × TBE. The flask contents were mixed by gentle swirling and heated in a microwave to dissolve the agarose. The dissolved agarose was left to cool to approximately 60^{o}C before the addition of 5 µl of 0.5µg/ml ethidium bromide. The melted agarose was then poured into a pre-

sealed gel cast containing a well comb. Agarose gels were left to set for 20-30 minutes at room temperature.

The agarose gel was placed in an electrophoresis tank (BIO-RAD Laboratories, Inc) containing sufficient 1X TBE buffer. Electrophoresis was carried out at 80 volts for 1 hour or until the required resolution was achieved. DNA fragments were visualised under UV light and the gel was photographed on a UV trans illuminator.

Stock DNA was normalised to a concentration of 45ng/µl by dilution with Molecular Biology Grade water. 1µl of normalised DNA was employed to measure the 260/280 and 260/230 ratios using a Nanodrop-1000. Samples with a 260/280 ratio >1.75 and a 260/230 ratio > 1.8 were discarded from analysis.

2.13 Whole exome sequencing

Next generation sequencing platforms are high-throughput DNA sequencing technologies based on the principle of running thousands or millions of clonally amplified or single DNA sequences in parallel (Mardis ER, 2008).
Exome sequencing is one of the ways in which NGS has been successfully used to identify novel genetic mutations in rare diseases by the targeted sequencing of the subset of the human genome that is protein coding.
Exomes are the subset of the genome that is protein coding. Whilst exomes constitutes <2% of the genome it has been shown that most alleles that underlie Mendelian disorders disrupt the protein coding region of the genome (Bamshed MJ et al., 2011). It is also known that a defect in the protein coding region is far more likely to have a functional/ deleterious defect. As such the exome is a highly enriched subset of the genome to search for variants with large effect sizes.

In exome sequencing, a DNA library is constructed from total genomic DNA and enriched for sequences corresponding to exons by aqueous phase hybridization capture (in-solution capture). Genomic DNA is initially randomly fragmented and oligonucleotide adaptors are ligated to each end. The whole exome is then captured and enriched by in-solution hybridisation to 95-mer biotinylated DNA

oligonucleotide baits (Illumina TruSeq). Illumina provide a kit for exome capture that is scalable to automation using DNA bait (Bamshed MJ et al., 2011). Amplification and parallel sequencing (described below) of the fragments can then occur. Samples are run on an Illumina next generation sequencing platform (in this study a HiSeq1000, Institute of Neurology), to generate 2x100bp read length DNA sequences.

The first step is bridge amplification of single DNA molecules (generated during the library preparation) on the surface of a flow cell by an automated device called a Cluster station. A flow cell is an 8-channel sealed glass micro fabricated device allowing amplification of fragments on its surface and containing DNA polymerase to produce multiple copies (around one million) of the molecule that initiated the cluster reaction.

The second step is the sequencing-by-synthesis approach during which DNA polymerase is added together with all four nucleotides each carrying a base-unique fluorescent label and with the 3'-OH group chemically blocked such that each incorporation becomes a unique event. After incorporation, an imaging step occurs prior to the 3' blocking group being removed to prepare the strand for the next incorporation.

By continuing for a set number of cycles discrete sequences reads of 76 or 100 bases in length can be obtained. With a base-calling algorithm assigning sequence and quality values to each read. Illumina have also incorporated a quality checking pipeline into the data run.

2.13.1 Exome Protocol

2.13.1.1 Sample Preparation

DNA samples being used for exome sequencing (n=10) were subject to the same DNA quality control checks as described previously (2.12.4). As per UCL Institute of Neurology requirements, 3.2ug of DNA at 100ng/µl according to Qubit reading were prepared in 32µl of H_2O. Only samples with an A_{260}/A_{280} ratio of 1.8-2.0 were sent. To asses DNA integrity, 1µl of diluted DNA (100ng/µl) was run on a 1% agarose gel with 1Kb ladder alongside.

2.13.1.2 Exome capture and sequencing

A total of 3.2µg of high quality DNA was sent to UCL Institute of Neurology in London. Samples were processed using Illumina TruSeq sample preparation and capture (64Mb), (Illumina). Samples were processed according to the Illumina Paired-End Sequencing Library Protocol and run on an Illumina HiSeq1000 sequencer; paired end 2x100bp read length with an aimed median coverage of 30x. Samples were multiplexed 5 per lane of a flowcell.

2.13.2 Exome analysis

The sequencing reads resulting from the HiSeq1000 were aligned by the GOSgene Bioinformatics team to the human reference genome (GRCh37 release, downloaded from the ENSEMBL database, corresponding to Hg19 NCBI built 37.1) with BWA (Burrows-Wheeler Alignment Tool) software (Li H, 2009). The GATK tool suite was used to further process the alignments (base quality score recalibration, indels realignment, duplicate removal) (McKenna A, 2010). Variant calling was also performed using GATK Unified Genotyper multi-sample SNP caller. A summary of the sequencing coverage produced by GATK DepthOfCoverage profiler was generated. SNPs and small insertions/deletions (indels) annotation and interpretation analysis were generated through the use of Ingenuity Variant Analysis™ software (www.ingenuity.com/variants) from Ingenuity Systems.

A binary format of aligned sequencing data named BAM file was generated for each sample. BAM files containing SNP and indels calls were visualised using the Integrative Genomics Viewer IGV (http://www.broadinstitute.org/igv/UserGuide). Relevant known potential target genes were checked for pathogenic variants and coverage. Variants were prioritised by predicted frequency starting from "novel" variants (i.e. not annotated in databases including dbSNP, NHLBI Exome Sequencing Project (ESP) and 1000genomes), predicted effect (starting from essential splice site, stop gained, stop lost, frameshift indels and non-synonymous variants).

The dbSNP database serves as a central repository for both known single base nucleotide substitutions and short deletion and insertion polymorphisms. The NHLBI GO Exome Sequencing Project (ESP) provides an exome variant server. This

repository currently contains exome data representing over 200,000 diseased individuals. Exome data from 6,500 individuals is publicly available and is used as a means of identifying variants and their frequencies across individual genes.

2.13.2.1 Downstream analysis

Downstream variants (an example of output is shown in figure 2.6) analysis included the assessment of the quality of the reads from the bam files, assessment of the mappability and existing structural variations in the chromosomal region, assessment of the novelty against internal GOSgene databases (for common currently not publically available variants), details of gene functions (review of databases and literature), affected domains and prediction of the biological impact (SIFT, Polyphen and Mutation Taster), amino acid conservation, expression pattern, existing animal models and any other relevant information (including phenotype indications via OMIM and HGMD).

2.13.2.2 Functional predictions of candidate mutations

Evidence for pathogenicity for missense variants was assessed using Polymorphism Phenotyping v2.0.23 (PolyPhen-2) [http://www.genetics.bwh.harvard.edu/ pph)] and SIFT (http://sift.jcvi.org/). SIFT (Sorting Intolerant From Tolerant) is a program that predicts whether an amino acid substitution affects protein function so that users can prioritize substitutions for further study. SIFT can distinguish between functionally neutral and deleterious amino acid changes in mutagenesis studies and on human polymorphisms (Ng PC et al., 2003). The PolyPhen-2 software classifies the functional effect of a missense mutation into three categories using empirically derived rules to predict whether the mutation is damaging, i.e. is supposed to affect protein function, or benign, i.e. most likely lacking any phenotypic effect (Ramensky V, 2002).

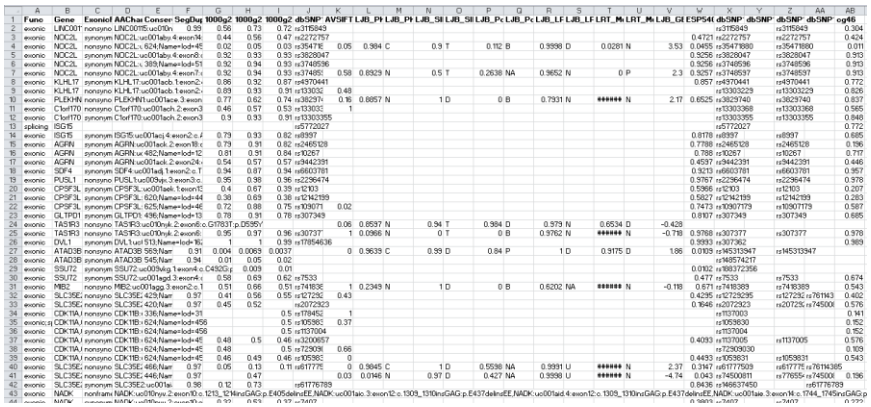


Figure 2.6.An example of the output of exome data from a single sample which includes all annotated gene variants identified across all the exomes covered.

2.13.2.3 Data filter pipeline

A data filter pipeline was developed which consisted of several steps in order to reduce the number of variants assessed. First, variants were eliminated that were not present in the actual coding sequence and synonymous variants other than those occurring at canonical splice sites. Initial analysis, therefore, concentrated on stop-gain, stop-loss, frameshift, splice site and those mutations which were undefined. Stop gain mutation are defined as those where an amino acid is substituted for a stop instructions thus terminating the sequence early whereas a stop loss mutation has the opposite effect of not terminating a sequence as expected. Frameshift mutations refer to those mutations that, through the insertion or deletion of nucleotides, lead to an altered amino acid sequence.

A frameshift mutation (also called a framing error or a reading frame shift) is a genetic mutation caused by indels (insertions or deletions) of a number of nucleotides in a DNA sequence that is not divisible by three. Due to the triplet nature of gene expression by codons, the insertion or deletion can change the reading frame (the grouping of the codons), resulting in a completely different translation from the original. The earlier in the sequence the deletion or insertion occurs, the more altered the protein (Losick R et al., 2008). Often the severity of the effect will be a result of a premature termination of the protein sequence.

Then a second filter was applied in order to identify which of these mutations were potentially novel and remove common variants. Variants were filtered out that were present and referenced in dbSNP 130 or in the ESP (exome server project) with a

frequency of >0.3 % or the 1000 Genomes phase 1 project data with a mean allele frequency (MAF) > 0.3%. Finally, the remaining variants were filtered against an in-house exome database of genetic variants from over 300 other individuals (all non-diabetic) obtained from other GOSgene exome sequencing projects.

The same variant pipeline filter was carried out for all ten samples and individual samples were then cross referenced to identify overlapping variants. Interesting rare variants related to MODY or those with known pathological associations were taken forward if found in isolated samples. For all other candidates, if the same variant was identified in three or more samples those variants were taken forward for subsequent assessment.

A second form of data filtering was performed by inputting the annotated Variant Call Format (vcf) file into Ingenuity Variant Analysis (Ingenuity, San Diego, USA). The Ingenuity software uses an Interactive Filter Cascade which allows for elimination of common and non-deleterious variants with a basic set of filters, reducing the number of variants significantly. Then, using keywords, a search can be performed on the remaining variants. This keyword search within the Ingenuity ontology provides identification and prioritization of compelling variants based on knowledge of disease, phenotypes, processes, and pathways and the pathway information available in the Ingenuity Knowledge Base. Variants in genes within 1- or 2- network "hops" of upstream or downstream genes can also be identified. A hop refers to the number of genes between the gene of interest and the condition.

2.13.3 Candidate gene and SNPs selected for *a priori* list

Candidate genes were chosen based on their association with known or putative role in monogenic diabetes and transient neonatal diabetes (Bonneford A et al., 2010) in both published literature and OMIM database.
MODY is not a single entity as at least twelve MODY subtypes with distinct genetic aetiologies have been reported in the literature: MODY1-*HNF4A;* MODY2-*GCK,; MODY3-HNF1A;* MODY*4-PDX;* MODY*5-HNF1B;* MODY6-*NEUROD1;* MODY7-*KLF11;* MODY8-*CEL;* MODY9-*PAX4;* MODY10-*INS;* MODY11-*BLK* and more recently MODY12-*ABCC8* (Bonneford A et al., 2010; Bowman P et al., 2012 and Farjans SS et al., 2011).

SNPs implicated from whole-genome-significant associations with T2DM (Voight BF et al., 2010; Zeggini E et al., 2008; Dupuis J et al., 2010; Feero WG et al., 2010 & Granup N et al., 2010), fasting glucose or birth weight (Hirokoshi et al. Nat Genet: 2013, Freathy et al., 2010 and Andersson et al., 2009) were included.

Candidate SNPs thought to have a role in beta cell function and glucose homeostasis were also included. In total 73 candidate genes were selected along with MODY candidates (Appendix 4 S3).

2.14 Validation of exome data

Exome data was validated in one of two ways; variants of interest that corresponded to a risk SNP were genotyped in the entire cohort (offspring, mothers and fathers in cases and controls). Variants thought to be novel were Sanger sequenced. In these instances parental DNA was also Sanger sequenced along with the case offspring in which the variant was identified.

2.14.1 Genotyping

Genotyping employs the use of real time polymerase chain reaction (PCR) which is able to amplify and simultaneously quantify targeted DNA molecules. A thermal cycler is used for PCR in order to rapidly heat and cool samples allowing the core processes (separation of the nucleic acids double chain, alignment of the primers with the DNA template and extension) to occur repeatedly.

Real time PCR utilizes sequence specific DNA probes containing oligonucleotides, labelled with fluorescent reporters, which allow for detection only after hybridization of the probe with its complimentary DNA sequence.
Thermal light cyclers have the ability to illuminate each sample with a beam of light of a specified wavelength and detect the amount of fluorescence emitted.

2.14.2 TaqMan Genotyping

The TaqMan method relies on the 5'-3' exonuclease activity of Taq polymerase. Taqman probes consist of a flourophore covalently attached to the 5' end of the oligonucleotide probe and a quencher at the 3' end. A quencher molecule quenches the fluorescence emitted by flourophore excited by a cycler's light source. The inhibiting action of the quencher is dependent on its proximity to the flourophore.

Taqman probes are allele specific annealing with a DNA region amplified by a specific set of primers. During the PCR cycle, the forward/reverse primer pair anneals to the template DNA, along with the respective allelic probe. When the probe is entirely complementary to the target sequence, *Taq* polymerase extends the primer and synthesises the complimentary strand from 3' to 5'. The 5' to 3' exonuclease activity of the polymerase degrades the probe that has annealed to the template.

Probe degradation releases the fluorophore which breaks and moves away from the quencher, allowing fluorescence of the fluorophore. A 7900HT Sequence Detection machine (Applied Biosystems, California, US) which is a real-time PCR thermal cycler, is able to detect the fluorescence which is directly proportional to the fluorophore released and the amount of DNA template present in the PCR. As the number of PCR cycles increases the level of fluorescence increases. The magnitude of the increase is dependent on the sample genotype; this property enables genotype identification.

2.14.2.1 TaqMan genotyping protocol

DNA concentration was standardised to 5ng/µl and a total of 10ng of DNA was used in each reaction. I was guided through the genotyping protocol with Dr Melissa Smart and reactions were performed at QMUL Genome Centre, London, UK. A Biomek 2000 robot (Beckman Coulter, High Wycombe, UK) transferred DNA from four 96-well plates into one MicroAmp® Optical 384-Well Reaction Plate with Barcode (ABI, Life Technologies, UK). The plate was left to dry out overnight at room temperature in a sterile paper bag.

A master mix comprising 2.5µl TaqMan Absolute quantitative PCR Rox Mix (ABgene, UK), 0.125µl 40x Assay(TaqMan forward and reverse oligonucleotides and their respective labelled probe pairs (Appendix 4 S5), 2.375µl Sigma ddH_2O per well was prepared and a 5µl aliquot was added to each well using an 8-channel pipette. Plates were sealed with a clear plastic lid (ABgene, UK) to prevent excess evaporation and then centrifuged at 1,000 rpm for 1 minute. Each plate underwent the following PCR program on a Tetrad 2 Thermal Cycler PCR System (Bio-Rad, USA) 384- well heated block:

95°C for 10 minutes
95°C for 15 seconds
60°C/62°C for 1 minute
} x 40 cycles

2.14.2.2 Sample analysis

The 7900HT Sequence Detection machine read all plates and different genotypes were identified by SDS v2.3 (Applied Biosystems). Output was both an allelic discrimination plot (Figure 2.2) and in a text output file. The file was then converted from 384-well to four 96-well plates in Microsoft Excel, and merged with the study database.

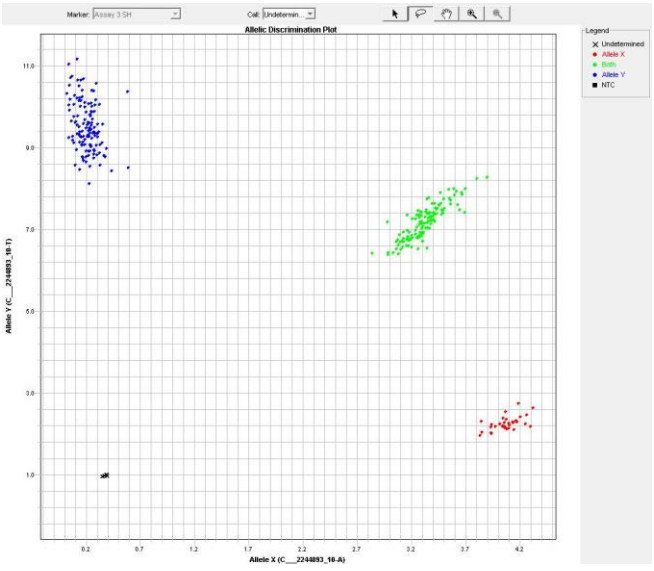


Figure 2.7: Screen capture from SDS v2.3**.** A typical allelic discrimination plot, with successful separation of the three genotype groups – homozygotes (• & •) and heterozygotes (•). Unassigned genotypes were those that did not cluster with other samples, this included NTCs-negative controls (×).

2.14.2.3 Statistical methods

Statistical analyses were performed using Intercooled Stata 10 for Windows (StataCorp LP, Texas, USA). Cohort characteristics are presented as means and standard deviation unless otherwise stated.
Hardy-Weinberg equilibrium (HWE) was examined using the Pearson's Chi-square test which tests goodness of fit. A p value of <0.05 was taken as deviation from HWE. Blood pressure and all anthropometric measures were natural log-transformed. For association studies a p value of <0.05 was taken as statistically significant. Setting a threshold of significance was the chosen method above Bonferroni corrections, since the candidate genes studied had been selected for based on an *a priori* hypothesis and biological plausibility. Haplotype association analysis was carried out using Thesias (www.ecgene.net).

2.14.3 Sanger Sequencing

Sanger's method is based on the use of dideoxynucleotides (ddNTP's) in addition to the normal nucleotides (NTP's) found in DNA. ddNTPs are essentially the same as nucleotides except they contain a hydrogen group on the 3' carbon instead of a hydroxyl group (OH).

DNA is denatured by heating and specifically designed primers (3' end is located next to the sequence of interest) anneal to the template strand. Fluorescently labelled nucleotides allow the final product to be detected on a gel (Russell, 2002)Different ddNTPs are added to the primer/DNA template solution in an automated process. This is done in a single tube containing all 4 ddNTPs each labelled with a different fluorescent dye. DNA is synthesised and nucleotides are incorporated into the chain. ddNTPs will be randomly incorporated. This causes a chain terminating event by preventing the addition of further nucleotides.

The key to the Sanger method is that all reactions start at the same nucleotide and end with a specific base. When the same chain is synthesized many times different length DNA fragments will result with the same starting and end nucleotide. The chains are then denatured again, and run on a gel to separate out the different size bands. By combining all 4 reactions on one gel the DNA sequence can be determined. In automated procedures, this reading is done by a laser which is able to identify the different wavelengths that the ddNTPs fluoresce at.

2.14.3.1 Sanger Protocol

Dr Louse Ocaka at UCL GOSgene performed sample preparation for Sanger sequencing and showed me the technique. I provided purified DNA of which 1µl underwent a polymerase chain reaction (PCR). This was performed on a PTC-100™ Peltier Thermal Cycler (MJ Research™) equipped with a heated lid thus eliminating the necessity to overlay PCR reactions with mineral oil to prevent evaporation. Amplification was carried out under the following general conditions:

Initial denaturation:		95°C	for 5 minutes
35 cycles of:		95°C	for 30 seconds
		50 - 65°C	for 30 seconds
		72°C	for 30 seconds
Final extension:		72°C	for 5 minutes
Hold:		10°C	

PCR products were visualised by loading 5µl of the reaction on an agarose gel of the appropriate concentration depending on the size of PCR product. A negative control (no DNA) was always included in the experiment.

2.14.3.2 Primer Design

PCR primers are oligonucleotides usually in the region of 18-25 bp in length. These oligonucleotides bind to a target sequence allowing a polymerase to initiate synthesis of a complementary strand. To ensure optimal results, random base distribution for each primer was designed with a similar GC content for both primer pairs, with levels of GC content between 40% and 60%. Primers were designed with an A or T nucleotide at the 5’ end and a C or G at the 3’ end to act as an anchor. Each primer

was more than 20bp in length to increase complexity and binding specificity. Primers were designed at least 40 bp away from the intron/exon boundary to ensure successful amplification and sequencing of the splice site. Primer pairs were designed to have similar annealing temperatures. All primers were synthesized by Sigma-Aldrich Ltd (Appendix 4 S4). Primers were diluted according to manufacturer's instructions in order to obtain a final stock concentration of 100µM diluted with dH_2O and working stock concentration of 10mM.

2.14.3.3 PCR Product Purification

The ExoSAP-IT enzyme (Amersham Life Science, Buckinghamshire, UK**),** which removes unwanted dNTPs and primers, was used to purify PCR products. Generally, 30µL reactions were prepared consisting of 1µL (1 unit) ExoSAP-IT enzyme, added to 100-200ng (generally 2-3µL) of PCR product and made up to 30 µL with dH_2O. The reaction was mixed thoroughly and incubated at 37°C for 15 minutes. Enzymes were inactivated by heating to 80°C for 15 minutes.

2.14.3.4 QIAquick Gel Extraction

PCR products were electrophoresed on a 0.5 % agarose gel, gel extracted, and purified. DNA in the range of 70 bp to 10 Kb was extracted and purified using the QIAquick gel extraction kit according to the manufacturer's guidelines (QIAGEN®). Upon resolving DNA on an agarose gel, the band of interest was excised from the gel and the gel slice weighed. Buffer was added to the gel and incubated at 50^oC for ten minutes A QIAquick spin column and centrifugation was then used to elute the DNA.

2.14.3.5 Automate d DNA Sequencing

DNA was sequenced using ABI BigDye terminator cycle sequencing chemistry version 3.1 on an ABI 3100 automated sequencer (Applied Biosystems) according to manufacturer's instructions. A 13µl aliquot of purified PCR product was then used for cycle sequencing. The 20µl Cycle sequencing reaction contained 13µl of purified PCR product, 1µl of BigDye, 1µl of a 0.8µM concentration primer, 5µl of ABI sequencing buffer. Cycle sequencing of sequencing reactions was conducted on a PTC-100™ Peltier Thermal Cycler (MJ Reseach™) using the following thermocycling conditions for 30 cycles of 96°C for 30 second, 50° C for 15 minutes and 60° C for 4 minutes.

2.15 DNA methylation

Epigenetic modifications are modifiable structural changes to DNA, such as histone modification and DNA methylation that result in changes to gene expression and phenotype without altering the nucleotide sequence (Bird A, 2007). DNA methylation is widely studied and occurs as a result of covalent binding of a methyl group to cytosine residues in cytosine/guanine rich regions of DNA (CpG islands). It is maintained during cell division in mammals only at the dinucleotide C-G by the enzyme DNA methyltransferase 1

There are several ways to assess methylation ranging from locus specific to genome wide. The Infinium 450 Human Methylation (450 HM) Bead chip (Illumina Inc, San Diego, CA) is one of the latest platforms able to assess methylation at a genome wide level and was the platform used for this work.

2.15.1 The Infinium HumanMethylation450 (HM450) BeadChip

The HM450 BeadChip allows epigenome wide scanning of > 485,000 CpG methylation sites per sample at a single-nucleotide resolution. The HM450 Beadchip incorporates the chemistry of two different chemical assays (Infinium I and Infinium II) carrying information about the methylation state via the number of probes.

Infinium I contains two beads per CpG locus. Each of the two bead types are attached to single stranded DNA oligonucleotide probes that differ in sequence only at the 3' free end. The 3' end of each probe is designed to match either a protected "methylated" cytosine or a thymine, bisulphite converted, "unmethylated" base. The probes are designed on the assumption that within a 50 base pair (bp) span of the genome, methylation of any CpG loci are likely to have the same methylation status.

The Infinium II assay is designed with one probe. The 3' end of the probe complements the base directly upstream of the query site. A single base extension results in the addition of a labelled G or A base complementary to either the methylated C or unmethylated T. This design has the advantage of allowing up to three CpG sites within a 50bp span, removing the assumption of concurring methylation status within the query site.

The array is run using high quality DNA that has been bisulphite converted. Bisulphite conversion converts unmethylated cytosine bases into uracil. The resulting product contains unconverted cytosine where they were previously methylated, and uracil, where cytosine was previously present but unmethylated.

Bisulphite treated DNA is subjected to whole genome amplification (WGA) amplifying uracil as complementary thymine. The products are then enzymatically fragmented, purified and applied to the chip. The bisulphite converted amplified DNA products are denatured into single strands and hybridized to the chip via allele specific annealing to either the methylation specific probe or the non-methylation probe Allele-specific single base extension of the primer incorporates a biotin nucleotide or a dinitrophenyl (DNP) labelled nucleotide. (C and G nucleotides are biotin labelled; A and T nucleotides are DNP labelled.) Signal amplification of the incorporated label further improves the overall signal-to-noise ratio of the assay. Immuno-histochemical assays are performed by repeated rounds of staining with a combination of antibodies to differentiate the two types (www.illumina.com). After staining, the chip is scanned using the Illumina HiScan SQ two-colour laser scanner, to show the intensities of unmethylated and methylated bead types.

2.15.2 Internal Quality Control Checks

Data generated is subject to internal quality control checks incorporated into the genome studio package. In order to assess the quality of specific steps in the array, sample independent checks are performed using control probes which include:

- bead-staining controls with biotin (green) and DNP (red) of DNA;
- extension controls
- hybridisation controls and subsequent target removal

Staining controls assess the efficiency of the intensity of the red and green channel. Background fluorescence failure leads to difficulty discriminating red and green channels and requires removal from further analysis. Signal intensity can also be affected by the C-G content and leads to sample removal.

Extension controls measure the efficiency of extension of the nucleotide sequences to the probes. A and T are assessed in the red channel and C and G in the green channel.

Hybridisation probes test the overall performance of the entire assay using synthetic targets (which complement the sequence on the array) not DNA. The synthetic targets are present at low, medium and high concentrations in the hybridization

buffer so should be represented as specific low, medium and high intensities, representing the concentration present. Target removal tests the efficiency of stripping after extension. The control probes are extended using the probe sequence as template (which generates labelled targets) and should result in low signal compared to hybridized controls as probe sequences are designed so that extension from the probe does not happen. Both hybridisation and target removal should be read in the green channel.

Sample dependent QCs allow for evaluation of performance across samples. The controls are designed against bisulphite converted human DNA sequence. These probes do not contain CpG dinucleotides so performance does not rely on methylation status of template DNA and include;

- a test of the bisulphite conversion efficiency,
- comparison of specificity of probe binding and primer extension,
- non-polymorphic (NP) controls
- negative controls

Bisulphite conversion (BS) controls query a C/T polymorphism created by conversion of non CpG cytosines (C) in the genome. BS I conversion use Infinium I probe design. If conversion is successful, converted C probes will match converted sequence and be extended. C1, C2 and C3 probes should be read in the green channel and C4, C5 and C6 in the red channel. BS failure results in unconverted U probes where extension fails
BS II controls use Infinium II probe design. Successful BS conversion incorporates an A nucleotide and the probe will have intensity in the red channel. If cytosine unconverted the G base will be incorporated and the probe will have intensity in the green channel.

Specificity controls monitor non-specific primer extension for Infinium I and II probes and are designed against non polymorphic T sites. Specific I probes are designed to monitor allele specific extension for Infinium I probes. Methylation of a specific cytosine is assessed using query probes for methylated and unmethylated state of each CpG locus. A PM control queries an A/T match which corresponds to an unmethylated C state; a G/C match corresponds to a methylated C so G/T and is

interrogated through an MM control probe. A G/T mismatch corresponds to non-specific detection of methylation signal over unmethylated background so should give a low signal.
Cross hybridisation can occur on all bead arrays due to regions of sequence homology. Specific probes susceptible to this cross hybridisation have been identified and are easily filtered out of downstream analysis.

The mean signal of negative control probes will define background signal. Negative controls are randomly permutated sequences that should not hybridize to the DNA template. The average background signal establishes methylation limits for the probes.
Finally, non-polymorphic regions of the bisulphite-converted genome are used as a control to test overall assay performance by designing A, C, T and G probes within these non-polymorphic regions that can be compared across different samples.

2.15.3 Sample Dependent QCs and expected intensities

Genome studio incorporates and assesses the sample dependent QC's. The package generates visual maps of signal intensities that are read within the specified channel (red, green or both). For BS conversion I (C1-C3) in the green channel and C4-C6 in the red channel, signal intensity should be high compared with probes U1-U3 in green channel and U4-U6 in red channel. Specificity probes I and II are read in the green and should have a high signal intensity for GT mismatch 1, 2, 3 (PM) but low for mismatch 1, 2, 3 (MM). Specificity II probes should be read in the red channel and have a high intensity. Non polymorphic probes for bases A and T should have high intensity in red channel and C and G should have high intensity in green channel. Negative controls can be read in either channel and should yield low signal intensity (Table 2.1 reports Genome studio expected intensities in QC probes as described).

QC	Name	Channel	Intensity
BS conversion I	C1, C2, C3	green	High
BS conversion I	U1, U2, U3	green	Background
BS conversion I	C4, C5, C6	red	High
BS conversion I	U4, U5, U6	red	Background
BS conversion II	1,2,3,4	red	High
Specificity I	GT mismatch 1,2,3 (PM)	red/green	High
Specificity I	GT mismatch 1,2,3 (MM)	red/green	Background
Specificity II	Specificity 1,2,3	red	High
Non polymorphic	NP (A), (T)	red	High
Non polymorphic	NP (C) (G)	green	High
Negative	Average	both	Background
Negative	StdDev	both	Background

Table 2.1 Expected intensities of sample dependent QC probes. Adapted from Illumina Infinium HD assay methylation protocol guide 2013

2.15.4 Probe Filtering

The following probes were filtered out of the final analysis:

1. X and Y chromosome probes were removed in all studies to prevent differential methylation being called from different numbers of X and Y chromosomes in mixed sex experimental groups
2. Cross-hybridising probes were removed (see section above)
3. P values are generated for each probe on the beadarray, describing the variance in intensity generated from the beads per probe. A high level of intensity variation implies that there is signal noise and only probes with a p value of <0.01 are included in analysis

2.15.5 Assessment of Methylation data

Illumina's custom software Genome Studio is able to generate an .idat file from raw data and methylation values (beta values) are generated as a fluorescence intensity ratio between the two bead types are calculated. A ratio value of 0 equals to non-methylation of the locus; a ratio of 1 equals to total methylation; a value of 0.5 means that one copy is methylated and the other is not, in the diploid human genome.

Classically, beta values of samples can be plotted using unsupervised cluster plots or multi-dimensional scaling (MDS). The dimensions on these plots (shown on the X and Y axes) represent a calculation based on the number of probes examined and the number of samples. To quantitatively estimate differences of MDS plots, both of these variables need to be constant. However, all being constant, clustering of samples can help to visually identify differences that may be technical i.e. due to batch effect or biological i.e. disease compared to control.

Beta values are variable or ‘ heteroscedastic’ and prevent use of statistical tests that rely on Gaussian distribution i.e. t tests. Beta values are often transformed to M values to improve normality and more accurately take into account the bimodal distribution of methylation across the genome (Du et al, BMC Bioinformatics 2010). The M value equates to the log2 ratio of the intensities of the methylated versus unmethylated probes. Once transformed, M- values do conform to statistical tests that rely on Gaussian distribution and allow for more direct analysis comparison of methylation values within cohorts.

2.15.6 Methylation data analysis

Once the processing steps described previously have been undertaken, t-test statistics can be used to generate p values, to determine the degree of significance between the means of two groups. As these array experiments incorporate multiple testing (i.e. interrogation of each array probe is one test), a standard p value is not sufficient to control for the likelihood of false positives. A correction for multiple testing should be performed, as a False Discovery Rate (FDR)-adjusted p value.

Through the use of t tests, an FDR adjusted p values can be generated between two groups (for these experiments this was the mean difference between case and controls) for each position. Those that reached an adjusted p value <0.05 were classified as differentially methylated positions (DMPs).

2.15.7 Assessment of single nucleotide polymorphism (SNP) influenced methylation

Single nucleotide polymorphisms (SNPs) can influence methylation values. In order to assess their potential influence a histogram of beta values at each probe position can be generated to visualize methylation values. DMPs that exhibit methylation (beta) values of 0%, 50% or 100% methylation are consistent with the presence of 0, 1 or 2 copies of a SNP that can create or abolish a CpG site. Illumina software that lists known interacting SNP-CpG sites (now over 150,000 positions) and in conjunction with methylation value evidence can support discovery of those DMPs that are influenced by the presence of a SNP.

2.15.8 Gene Ontology (GO) pathway analysis

The GO (The GO Consortium, 2000) database assigns biological descriptors (GO terms) to genes on the basis of the properties of their encoded products. These terms fall into three types: cellular component, biological process, and molecular function. Genes assigned the same GO term can thus be regarded as members of a category (‘‘GO category’’) of genes that are more closely related in terms of some aspect of their biology than are random sets of genes (Holmans P, 2009).

Gene Ontology (GO) analysis was performed using software that directly interacts with the Limma R package. GO terms were generated from DMPs identified in the analysis. Only those terms that reached a threshold of significance are reported and were taken forward. A cut-off p-value of <0.0001 and a fold change of 2 rounded up (expected versus actual gene count) was taken to be significant.

2.15.9 Marmal-aid analysis

‘Marmal-aid’, is the first meta-analytical tool for DNA methylation (freely available at http://marmal-aid.org). In Marmal-aid, all publicly available Illumina450K data (8,531 different arrays as of April 15, 2013) is incorporated into a single repository allowing for re-processing of data including normalisation and imputation of missing values. This database can be easily queried to gain insight into the functionality of

certain CpG sites of interest without having to rely on gene centric approaches such as Gene Ontology.

2.16 Illumina 450HM array protocol

2.16.1 Bisulphite conversion and QC

With direct guidance from Dr Melissa Smart at the Blizard Institute QMUL, London; 500ng of high-quality genomic DNA was bisulphite converted using the EZ-96 DNA Methylation Deep-well Kit (Zymo Research, Cambridge Biosciences, UK) according to manufacturer's protocols for the Illumina Infinium Assay.
In order to check bisulphite conversion success a qPCR was performed (Appendix 4; S7 and S8). This experiment uses two qPCR reactions designed against parts of the MLH and GAPDH gene. The MLH primer (Sigma Aldrich, UK) is designed against the expected sequence after bisulphite conversion and the GAPDH (Sigma Aldrich, UK) for the non-converted sequence. If conversion has been successful the MLH primer pair product will be seen. If unsuccessful, then GAPDH primer sequence will match template and produce product. The ratio of each of these products in a given sample can then be used to determine the conversion efficiency.

MLH1	F(5'>3')	GGAGTGAAGGAGGTTACGGGTAAGT
	R(5'>3')	AAAAACGATAAAACCCTATACCTAATCTATC
GAPDH	F(5'>3')	CGCCCCCGGTTTCTATAAAT
	R(5'>3')	CAAAAGAAGATGCGGCTGAC

Table 2.2 MLH1 and GAPDH forward and reverse primers as supplied by Sigma Aldrich (UK).

2.16.2 Allocation to array

4ul of bisulphite-converted samples were randomly assigned to the Illumina HumanMethylation 450Beadchip (Sandoval et al. 2011). Samples were randomized at the point of bisulphiting and before being plated out. Samples that initially failed the bisulphite QC (did not meet threshold of >97% conversion) were removed and replaced with further aliquots of the sample that did pass a subsequent QC.

2.16.3 Data generation

An .idat file was generated and using Limma (Smyth 2004), an R-based package available via Bioconductor 2.11 and beta methylation values were generated from raw array data. Data quality checks were performed using in-built control probes to examine bisulphite conversion and array performance as described.

Chapter 3

Parental Phenotype of FGR Offspring

Chapter 3: Parental Phenotype of FGR Offspring

3.1 Introduction

Fetal growth is influenced by maternal *in utero* environment and genetic factors inherited from both parents.

Studies of monogenic disease, in particular MODY and *GCK* gene variants, support a paternal genetic influence in reduced fetal growth. The correlation between father and offspring of phenotypic measures (such as similarly low birth weight) support a paternal genetic influence (Godfrey KM et al., 1997)

Men who fathered small for gestational age infants were also more likely to be obese and have larger waist circumferences (McCowan LM et al., 2011). Epidemiological studies have identified that men who develop diabetes in later life were more likely to have fathered low birth weight offspring (Davey Smith G et al., 2004; Wannamethee SG et al., 2004, and Hypponen E et al., 2003). Furthermore, that poor fetal growth and low birth weight are risk factors for the future development of impaired glucose tolerance metabolic syndrome and T2DM (Hales CN et al., 1991& Barker DJP et al., 1993). These findings strongly support a paternal effect on both fetal growth and birth weight and transmission of metabolic risk.

In this first experiment I aimed to identify if a paternal phenotype could be recognised amongst men who fathered pregnancies affected by fetal growth restriction. I recruited and studied a cohort of otherwise healthy couples thought to be having a growth restricted offspring and compared them with couples having an appropriately grown baby. In order to categorize if a specific metabolic and vascular phenotype could be identified, parental anthropomorphic measurements were taken along with fasting glucose and insulin. A surrogate marker of insulin resistance (HOMA-IR) was calculated and compared between men who fathered a growth restricted offspring and those who fathered an appropriately sized offspring.

Although poor growth results from several different mechanisms, placental dysfunction is thought to be the underlying reason behind most unexplained or idiopathic growth restriction. As the placenta is half paternal in origin, it seemed likely that a paternal influence would be most apparent in cases of growth restriction thought to secondary to placental dysfunction (as opposed to maternal disease or poor nutritional status). Therefore, only such cases of growth restriction were recruited. In order to enhance any underlying paternal effect, extensive maternal exclusion criteria were employed with the aim of recruiting non-smoking mothers of similar age, weight, parity and ethnicity in both groups.

3.2 Hypothesis

Insulin resistance, identifiable in fathers at the time of the pregnancy and inherited by a fetus from its father, will result in a growth restricted phenotype *in-utero* and predispose both offspring and father to later life development of diabetes.

3.3 Specific aims

1. To identify if there is a paternal phenotype associated with fathering a growth restricted offspring at the time of the index pregnancy.
2. To specifically identify whether men who fathered a growth restricted pregnancy had an insulin resistant phenotype as assessed by the HOMA-IR model. This would support the hypothesis that insulin resistance has a dual effect to cause growth restriction and diabetes.
3. To assess vascular function of men who father growth restricted pregnancies and compare them to men who father appropriately grown babies through the use of flow mediated dilation of the brachial artery.
4. To study female partners in the same way as fathers in order to assess the extent of maternal influence in these cases.
5. To characterise the phenotype of FGR offspring and perform measures of insulin and c-peptide hormones on both FGR and AGA babies in order to support the hypothesis of insulin resistance being present in FGR offspring.

3.4 Methods

3.4.1 Study Design and Population

A case control study was undertaken at University College London Hospital (UCLH) between September 2009 and May 2011. Ethical approval for the study was granted by the Joint UCLH/UCL (alpha) Ethical Committee (09/H0715/28). All participants gave informed consent.

Couples were recruited who were thought to be having a growth restricted or appropriately grown baby as described in the detailed methods section 2.3 to 2.5. Briefly to recap, growth restriction was defined as an antenatal <10th customised birth weight centile as assessed by ultrasound where induction of labour or delivery by caesarean section was planned in view of fetal size.
Pregnant women and their partners were recruited as controls that were thought to be having a normally grown baby with an estimated fetal weight between 10th and 95th customised centile as assessed by an antenatal ultrasound.

Study design methods are described in detail in methods section 2.6 to 2.11. Briefly, key measurements for each participant were:

3.4.2 Father Studies

Each father completed a questionnaire enquiring about past medical, family and treatment history. Own birth weight was recorded as remembered personally or from a parent. All study assessments were carried out in the Clinical Research Facility, UCLH. Men were studied within 4 weeks of offspring birth. The study room was temperature controlled at 24^{o}C. Participants were asked to fast overnight for at least 10h before study. Weight, height and abdominal circumference were recorded. After resting, two measures of supine blood pressure (BP) were taken 15 minutes apart. Fasting venous insulin, glucose and lipid levels were measured. Insulin resistance was calculated using the HOMA model (as described in section 2.7). The blood was spun within 1hr of venepuncture and plasma and serum were frozen at -80^{o}C. All blood samples were processed in the same laboratory.

I measured endothelial function using brachial artery flow-mediated dilatation in a quiet, temperature-controlled room in accordance with protocol described in section 2.8.

3.4.3 Fetal sample collection and measurements

At the time of childbirth, umbilical cord blood was taken from the umbilical vein or artery, centrifuged and stored at -80^{o}C as plasma and serum for later measures of fetal insulin and c-peptide levels. Gestational length, offspring gender, weight and length were recorded.

3.4.4 Statistical Analysis

An initial calculation (using STATA) determined that 151 observations would be sufficient to detect a doubling in the odds ratio of being a case for every unit increase in log HOMA insulin resistance after adjusting for other co-variables, at 0.80 power and 0.05 significance.
Further statistical analysis of the data was performed using the STATA 10 package with guidance from a UCL statistician (Mrs Pauline Rogers). HOMA insulin resistance was log^{e} transformed to improve normality. All results are recorded as mean (standard deviation) unless otherwise stated. Paternal data was initially analysed by univariable logistic regression. The four coefficients with the lowest p value were used to generate a multivariable model (further information provided in methods section 2.1.2 to 2.1.4).

3.4.5 Statistical revision of numbers

In all, after recruitment, 119 paternal observations were made. Using the actual HOMA IR results, we were able to re- calculate using our sample size of 119 observations which showed we were powered at 0.96 to detect a doubling in the odds of being a case for every unit increase in log HOMA IR. Hence we were able to stop recruited at the intended time end-point despite failure to recruitment the initial sample size.

3.5 Results

3.5.1 Recruitment numbers to study groups

Recruitment occurred over a 15month time period (September 2009- May 2011). Of couples approached, 38 families declined to participate, 10 potential cases and 28 controls. Of the 10 cases, women experienced stillbirths in two instances, thought to be as a consequence of growth restriction. Families where in-utero demise was a possibility were approached in this study as it would have been useful to be able to include this population. However, of three families that were approached in the antenatal monitoring period, in two cases the family declined to take part in the study. In the third case, the family wanted to be involved in the study, but delivered unexpectedly at another hospital and could not be included.

Forty four couples were included as cases with a pregnancy affected by fetal growth restriction (FGR), $<10^{th}$ customised centile where induction of labour or delivery by caesarean section was planned in view of fetal size. Of these cases, 2 women went into spontaneous labour before their planned induction date but were still included in the study.

Eighty five pregnant women and their partners who were thought to be having a normally grown baby with an estimated fetal weight between 10^{th} and 95^{th} customised centile as were included as 'controls'. All participants had an additional fetal ultrasound scan at 34 weeks to confirm predicted size (unless a third trimester scan had already been done n=10). After delivery, neonatal measures that did not match predicted fetal growth led to the exclusion of 8 families (6 controls and 2 cases) and two further families withdrew consent to study participation after delivery. Final study analysis was between 42 cases and 77 controls (Figure 3.1).

In terms of recruitment the vast majority of participants were recruited from within the maternal fetal assessment unit at University College Hospital. A minority were approached directly in the fetal medicine unit (where complex fetal growth restriction cases are managed) or through antenatal classes or in antenatal clinic. Fourteen couples expressed interest through email correspondence after seeing posters about the study. Four of these were classified as cases and 10 as controls. All couples who instigated email contact agreed to participate in the study when further information was provided to them.

Seven couples, thought to be having a growth restricted offspring, agreed to donate fetal samples but did not want to undergo any studies themselves.

3.5.2 Baseline characteristics

Baseline characteristics of offspring confirmed that case and control subjects met the study criteria (Table 3.1). Data was available for all 77 control offspring and 42 case offspring that were being studied as part of a family.

	CASES				CONTROLS			
BABY								
	Mean (n)	**SD**	**25th**	**75th**	**Mean (n)**	**SD**	**25th**	**75th**
Customised birth centile (%)	1.8	2.2	0.1	3.2	49.6	27.5	26	73.9
Gestation (days)	256	29	236	279	283	9.3	274	291
Birth weight (g)*	2019	752	1378	2650	3517	367	3290	3800

Table 3.1 Birth characteristics of offspring cases (<10th customised birth weight centile: n=42) and controls (10th - 95th customised birth weight centiles; n=77). Customised birth centiles are gestation specific and detect fetal growth restriction despite reduced gestation

3.5.3 Offspring Data

There were significant differences between case and control birth weight and gestational age at delivery which was to be expected by the study design. The mean customised birth centile in controls was 49.6th ±27.5 centile and in cases was 1.8th ± 2.2 centile.

Of the seven cases in which permission was given for collection of just fetal samples the average birth weight was 2.29±0.29 kg and gestation 273±12 days (equivalent to 39 weeks gestation).

3.6 Maternal phenotype data

Female pre-pregnancy age, weight (kg), heights (cm) and parity were recorded through interview and from antenatal notes (Table 3.2) and revealed no statistically

significant results. Antenatal data was available on 76 controls and all 42 cases. Smoking status was assessed through antenatal records and confirmed that no women recruited smoked.

MOTHERS	CASES				CONTROLS			
	Mean	SD	Min	Max	Mean	SD	Min	Max
Age (years)	33.8	4.52	25	44	32.3	3.57	20	39
Weight (kg)	62.7	9.2	48.4	84.1	64.4	10.5	44	94
Height (cm)	163	5.7	152	175	166	5.8	155	185
Birthweight (g)	3007 (20)	576	1000	3800	3322 (45)	425	2300	4300
BMI (kg/m^2)	23.3	3.5	19	32	23.4	3.4	18	32
Nulliparous	31 (73.8%)				57 (74.0%)			
Multiparous	11 (26.2%)				20 (26.0%)			

Table 3.2 Maternal baseline characteristics. Antenatal statistics gathered from antenatal notes and interview (for own birthweight). Confirmed all women met inclusion criteria.

3.6.1 Routine Antenatal Maternal Glucose Data

As part of routine antenatal care at University College Hospital, London all pregnant women are offered a random glucose at booking and 28 weeks of pregnancy and these results were extracted and compared between control and case subjects.

Maternal random glucose levels at the booking antenatal visit (usually between 8 and 12 weeks of pregnancy) were similar [case subjects 4.45 ±0.52 mmol/L and control subjects 4.54± 0.54 mmol/L, P = 0.44].

3.6.2 Antenatal vitamin and micronutrient data

A protocol amendment meant that at the time of recruitment to the study (34 weeks of pregnancy) an extra sample of blood could be taken from female participants at the same time as their routine antenatal blood test. Whole blood was stored for DNA extraction but samples were also centrifuged and serum stored. Aliquots of female antenatal serum were processed by The Doctors Laboratory Ltd, London, UK for vitamin D, vitamin B12, folate and homocysteine levels using routine validated clinical biochemical assays.

As these micronutrients have been shown to influence fetal outcomes and deficiencies are associated with poor fetal growth it was important to investigate maternal nutritional levels during pregnancy in order to identify their potential role in this cohort. Due to the later addition of this sampling time point, 45 samples (29 control and 18 cases) were processed.

Women with FGR had lower antenatal folate, vitamin D and homocysteine levels but this did not reach statistical significance (Table 3.3).

	CASES (mean)	SD	CONTROLS (mean)	SD	P value (95% CI)
No samples	18		29		
B12 (pmol/L)	384.5	188.2	332.2	137.5	0.29 (-44.8 to 149.3)
Folate (nmol/L)	10.1	4.33	11.3	4.6	0.37 (-1.5 to 4.01)
Vitamin D (ng/mL)	44.2	29.5	50.6	31.4	0.49 (-25.4 to 12.4)
Homocysteine (umol/L)	6.9	2	17.1	26.4	0.12 (-2.75 to 22.6)

Table 3.3 Measurement of female participants' antenatal (34weeks gestation) micronutrient and vitamin levels for B12, folate, vitamin D and homocysteine.

3.6.3 Antenatal maternal plasma cotinine levels

Women who smoked were excluded from the study due to the known association of smoking with fetal growth restriction. However, we not exclude fathers who smoked. In order to investigate the potential *in-utero* environmental effect of smoking within a household (i.e. by a father) cotinine (a metabolite of nicotine) levels were measured in women whose partners smoked. Only 3 women (2 cases and 1 control) out of a sample of 17 women who had a partner, who smoked, had detectable serum cotinine levels. These levels were compatible with passive smoking (15.7, 45.9, 59.3ng/mL).

3.6.4 Postnatal maternal phenotype study

All women were recruited antenatally, but studied postnatally due to the confounding effect of pregnancy. Visits took place 6 months after delivery +/- 1 month, except if women were exclusively breastfeeding (n=9), in which case they were seen between 6months and 1 year (when infant weaning commenced).

Women were invited for follow up, by letter, email and telephone and in total 60 (51.2%) female participants (37 controls and 23 cases) returned to be studied. Several of the initial cohort could not be contacted due to out-of-date contact details (n=10). 8 women were pregnant again at the time of call-back and hence not eligible to be studied and due to the time lapsed were not called back after the subsequent pregnancy. 10 women declined to come back having moved area and 15 women declined due to work or family commitments. Due to time constraints 9 women recruited at the end of the study were not invited back. All women were asked about their on-going consent to the study when contact was established and all were happy for data to be used in situations where they were unable to attend postnatal follow-up.

Women who returned underwent exactly the same study as their partners in the same room within the Clinical Research Facility. Physical measurements revealed very little difference in weight (kg), abdominal circumference (cm), blood pressure (mmHg) between control and case women (Table 3.4)

	CASES	**CONTROLS**		
	Mean (SD)	Mean (SD)	p-value	95% CI
Number (n)	23	37		
Weight (kg)	63.2 (10.6)	62.9 (8.4)	0.9	(-6.6 to 5.87)
Abdo Circumference (cm)	82.3 (8.6)	81.8 (6.4)	0.84	(-4.63 to 5.68)
Systolic BP (mmHg)	103.6 (15.2)	109.9 (7.8)	0.08	(-0.72 to 13.2)
Diastolic BP (mmHg)	64.1 (7.9)	65.2 (8.0)	0.64	(-3.53 to 5.65)
	Post pregnancy	**Pre pregnancy**		
Case weight (kg)	63.2 (10.6)	63.2 (9.6)	0.98	(-6.75 to 6.62)
Control weight (kg)	62.9 (8.4)	64.1 (10.4)	0.55	(-2.82 to 5.21)

Table 3.4 Female postnatal study. Physical measurements of the 37 control women originally recruited and 23 case women including weight (kg), height (cm), abdominal circumference (cm) and blood pressure (mmHg) which revealed no statistical differences.

It is striking to note that pre and post-pregnancy weights were similar in the female cohort. Assuming that weight is one surrogate marker, it suggests that this post-pregnancy study was a reasonable surrogate marker of pre-pregnancy female status (which is the time point I was interested in trying to recreate).

3.6.5 Postnatal maternal blood results (glucose homeostasis)

As shown in table 3.5, blood glucose results remained similar postpartum (control subjects 4.57± 0.46 mmol/L and case subjects 4.53± 0.37 mmol/L, P = 0.78). Postpartum, maternal insulin levels were also similar between control (3.92 ±2.15 mIU/L) and case (4.38 ±2.41 mIU/ L) subjects; P = 0.57. Postnatal maternal HOMA index was similar between control (0.56±0.24) and case (0.58 ±0.29) subjects; P = 0.57.

	CASES Mean	SD	CONTROLS Mean	SD	P-value (95% CI)
Fasting glucose (mmol/L)	4.53	0.37	4.57	0.46	0.78 (-0.21 to 0.26)
Fasting insulin (mIu/L)	4.38	2.41	3.92	2.15	0.57 (-2.27 to 0.46)
IR (HOMA)	0.58	0.29	0.56	0.24	0.57 (-0.19 to 0.34)

Table 3.5 Postnatal maternal glucose homeostasis including glucose (mmol/L), insulin (mIu/L) and HOMA IR in control (n=37) and in case women (n= 23).

3.6.6 Postnatal maternal blood results (metabolic status)

Blood from participants was processed at University College Hospital biochemistry laboratory, using standard validated biochemical assays for; urea, creatinine, c-reactive protein and lipid profiles along with glucose, insulin and glycated haemoglobin (HbA1C).

As shown in table 3.6, results revealed no significant differences between controls and cases particularly in the lipid profile of triglycerides (control 0.60 ±0.23 mmol/L and case 0.61±0.26 mmol/L) and HDL cholesterol (control 1.81±0.38mmol/L and case 1.80±0.33 mmol).

	Cases mean (SD)	Controls mean (SD)	p-value	95% CI
Number samples	23	37		
Urea	4.55 (1.15)	4.77 (1.02)	0.56	(-0.44 to 0.8)
Creatinine	61.5 (9.6)	61.7(13)	0.95	(-6.56 to 7.05)
Triglycerides (mmol/L)	0.61 (0.26)	0.6 (0.23)	0.98	(-0.14 to 0.14)
Cholesterol (mmol/L)	4.79 (1.1)	4.49 (0.69)	0.26	(-0.78 to 0.22)
HDL (mmol/L)	1.8 (0.33)	1.81 (0.38)	0.98	(-0.21 to 0.21)
LDL (mmol/L)	2.55 (0.80)	2.41 (0.67)	0.53	(-0.57 to0.3)

Table 3.6 Female postnatal biochemical results. Urea, creatinine and lipid profile all reveal no statistically significant differences between controls and cases.

3.6.7 Postnatal serum vitamin and micronutrient levels

Postnatal serum vitamin B12, folate and vitamin D were assessed in 18 control and 11 case women. In both cases and controls levels were higher than seen during pregnancy. However, no statistical difference was seen between cases and controls for serum B12, folate or vitamin D levels postnatally (Table 3.7).

	CASES (mean)	SD	CONTROLS (mean)	SD
No samples	11		18	
B12 (pmol/L)	462.9	198.1	469.1	157.9
Folate (nmol/L)	11.6	5.1	12.3	4.5
Vitamin D (ng/mL)	51.1	28.2	45	16.7

Table 3.7 Measurement of female participants' postnatal micronutrient and vitamin levels for B12, folate, and vitamin D in 18 controls and 11 cases which revealed no statistically significant results.

3.6.8 Postnatal maternal vascular results

Women were also studied using both flow mediated dilatation (FMD) and pulse wave velocity (PWV).

Flow mediated dilatation in women postnatally revealed some interesting differences between case and control women. Of the 33 postnatal controls where FMD was performed (4 women returned on a day where FMD was not possible), 29 of the

scans were eligible for inclusion after quality control checks and 18 of 21 scans performed on case women were eligible for comparison.
Control women postnatally had higher FMD 10.26 ±2.44% than case women FMD 6.98±0.97% which was statistically significant p=0.0003 (95% CI 1.61-5.06). In this analysis smoking did not need to be adjusted for as no women included in the study smoked. Of case women studied, 4 had experienced early onset pre-eclampsia in conjunction with growth restriction. Postnatal FMD in these women revealed a mean FMD of 6.68±1.5% versus 7±0.68%) in case women with growth restriction only.

Pulse wave velocity revealed that despite no statistical difference there was a trend for case women to have higher PWV [controls 6.54±0.57 m/s and cases 6.95± 0.10 m/s] with a p-value= 0.1 (95% Confidence Interval -0.86 to 0.77).

3.7 Paternal phenotype data

Phenotype data was available on all male participants (n=119). In contrast to mothers, fathers of growth-restricted offspring (case subjects) had greater waist circumference and blood pressure and were more likely to smoke than fathers of normal grown offspring (Table 3.8).

	CASES				CONTROLS				
	Mean (n)	SD	Min	Max	Mean (n)	SD	Min	Max	p value
Age (years)	34.8	5.71	23	47	33.3	4.69	22	50	0.158
Weight (kg)	83.4	11.8	57	106.9	80.2	10.2	55.2	112	0.13
Height (cm)	177	7.0	166	193	179	6	161	195	0.28
Birthweight (g)	3127 (20)	597	1000	3970	3506 (37)	380	2780	4540	0.005
Waist circumference (cm)	94.1	8.23	71	114	89.6	8	60.5	113	0.005
BMI (kg/m^2)	26.2	3.25	18.7	31.5	25.2	2.93	16.8	32.2	0.08
Systolic BP (mmHg)	121.5	8.73	102	137	117.7	5.86	105	130	0.006
Diastolic BP (mmHg)	70.5	7.16	57	85	66.9	6.31	55	83	0.005
Fasting glucose (mmol/L)	4.87	0.41	4.2	6.4	4.72	0.32	3.9	5.4	0.037
Fasting insulin (mIU/L)	7.1	4.29	2	17.5	4.81	2.61	1.5	11	0.001
Insulin Resistance (HOMA)	0.93	0.51	0.4	2.2	0.63	0.31	0.4	2	<0.001
Cholesterol (mmol/L)	4.85	0.89	2.8	7.2	4.65	0.86	3.1	7.8	0.27
Triglycerides (mmol/L)	1.06	0.48	0.5	2.5	0.9	0.38	0.4	2.2	0.06
HDLCholesterol (mmol/L)	1.29	0.32	0.7	2.3	1.41	0.34	0.9	2.7	0.09
LDL Cholesterol (mmol/L)	3.01	0.92	1.2	5.7	2.82	0.83	1.2	6.1	0.28
FMD (%)	6.45	3.5	1.15	15.6	8.12	3.08	3.65	17.17	0.017
Smokers	17 (40.5%)				15 (19.5%)				0.008

Table 3.8 Baseline paternal phenotype including; physical, vascular, metabolic and smoking status

3.7.1 Paternal blood results- glucose homeostasis

Blood was available to analyse and a result was available in all cases (in three situations the test was repeated due to failure of the initial sample). Fasting glucose and insulin levels were both higher in case subjects, which resulted in an elevated HOMA-IR (Figure 3.1).

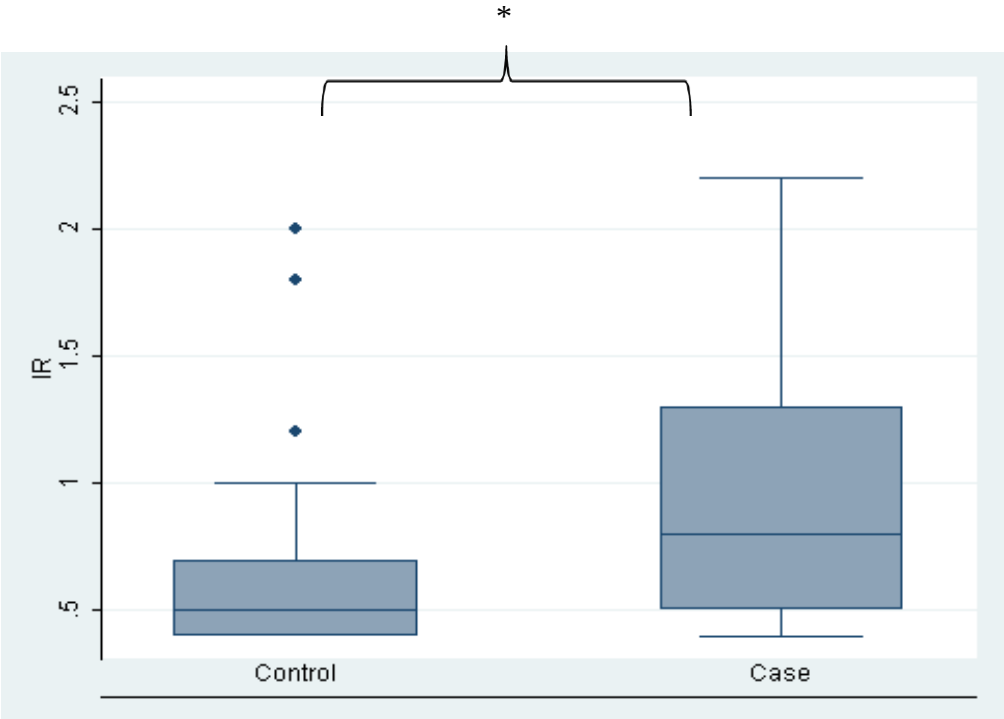


Figure 3.1Box Plot of IR values which were statistically significantly different between case (42) and control (77) men (* p<0.0001)

3.7.2 HOMA beta cell function

In contrast to HOMA-IR beta cell function (as measured through HOMA-β%) was seen to be increased in cases versus controls. Male controls had a HOMA- β of 73.4±21.6 with cases mean 94.7± 40 and a significant p-value of 0.0007 (-33.28 to -9.14).

Mid- way through the trial we also began to assess glycated haemoglobin (HbA1C) in some of the study men but the difference between cases and controls did not reach statistical significance (p=0.6 95% CI -0.361 to 0.220) in part due to the limited

numbers (cases n=11 and controls n=14) . Case men had an HbA1C of 5.31 ± 0.34% versus control men who had a slightly lower value of 5.24± 0.34%

3.8 Paternal vascular status

3.8.1 Flow mediated dilation

In analysis, exclusion criteria of recordings included: poor image quality; if recordings were deemed to be unrepresentative of the true vessel diameter or if images did not return to near baseline after reperfusion (Figure 3.3a). Each image-recording was validated by a second operator unaware of the subjects' group. Of 119 scans assessed in this fashion 22 (18.5%) were excluded from the final analysis due to above criteria. The remainder were included in analysis (example in Figure 3.3b)

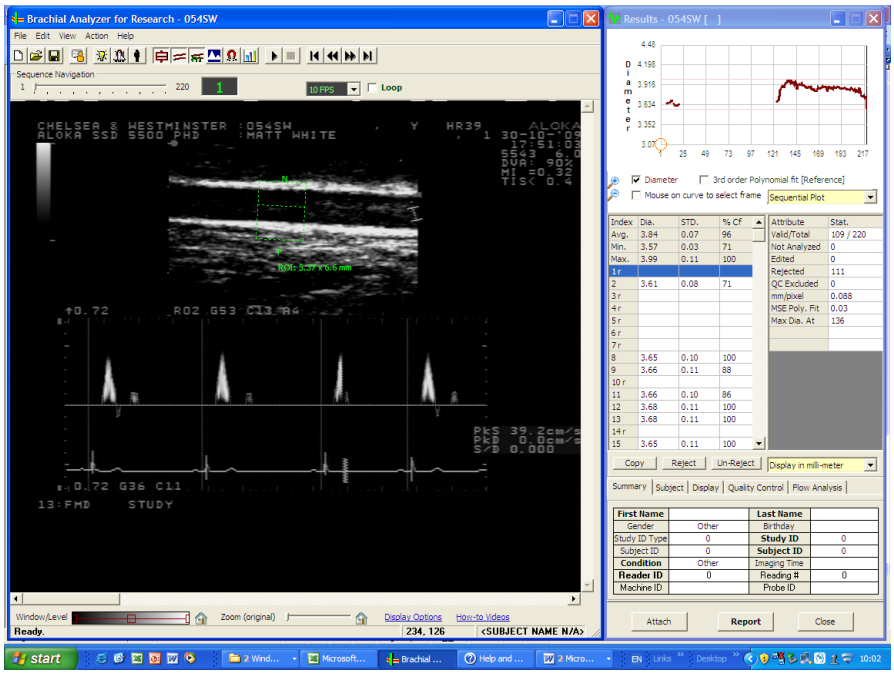


Figure 3.2a. FMD scan (above) excluded from analysis due to failure of image to return to baseline

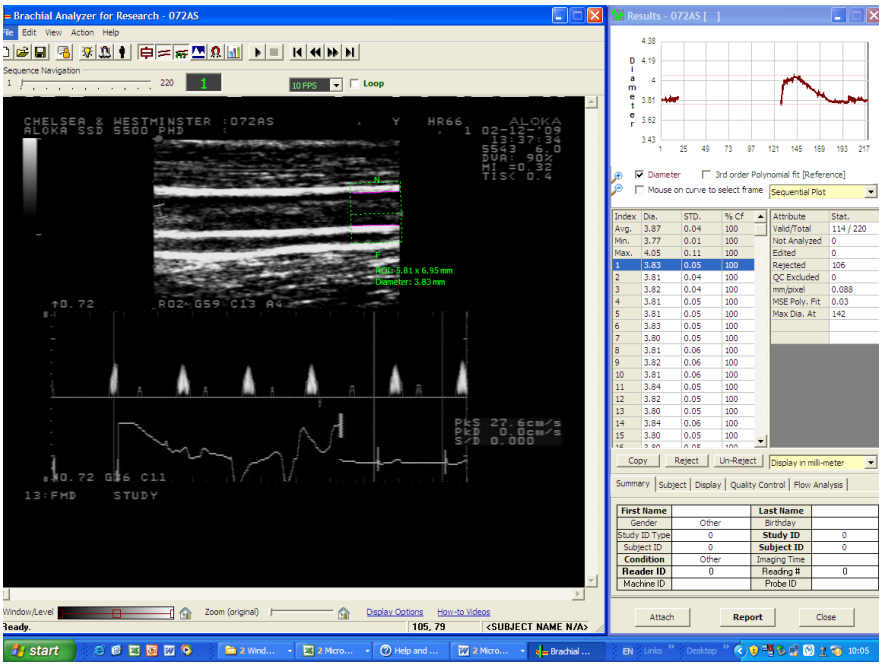


Figure3.2b. Suitable FMD scan for analysis : flat baseline, clean peak and return to baseline

As an independent variable, FMD in case men was significantly lower [6.45±3.45%] than control men [8.12±3.08%], suggesting more endothelial dysfunction. However, once adjustment had been made for smoking status then this difference was removed. Smoking is known to affect endothelial function (Hirofumi T, 2008) and as such should be included in any analysis; however, it does not negate the increased cardiovascular risk identified in these men.

3.8.2 Pulse Wave Velocity

Pulse wave velocity was only performed on 50 controls and 30 cases and 4 of these recordings were excluded due to technical issues after analysis. The results (Table 3.9) showed that case men had slightly higher PWV than control men but this was not significant (p=0.6 95 CI% -0.56 to 0.37). I believe that this technique would have yielded information consistent with FMD given larger numbers of participants.

	CASES	CONTROLS
Mean (m/s)	7.21	7.12
SD	0.98	0.99
SEM	0.18	0.14
N	29	47

Table 3.9 PWV data for control (n=50) and case (n=30) men. Case men had a higher mean PWV (m/s) but this did not reach statistical significance

3.9 Phenotype statistical analysis

Initial univariable logistical regression analysis confirmed that paternal insulin resistance, blood pressure, and waist circumference were higher in case compared with control subjects, while flow mediated dilatation was reduced and case fathers were more likely to smoke cigarettes (Table 3.9). The four most statistically significant paternal coefficients were log HOMA-IR, smoking, waist circumference, and diastolic blood pressure, which were analysed in the multivariable analysis. Paternal insulin resistance and smoking remained different after multivariable analysis, and therefore the final model was run with these two variables (Table 3.10). Compared with fathers of normal grown offspring, men who fathered pregnancies affected by fetal growth restriction had an OR 7.68 (95% CI 2.63–22.4, P , 0.0001) of having a 1-unit higher log HOMA-IR value and 3.39 (1.26–9.16, P = 0.016) of being a smoker (Table 3.10).

Coefficients	**Univariable analysis**			**Multivariable analysis**		
	Odds ratio	*95% Confidence Interval (CI)*	*p*	*Odds ratio*	*CI*	*p*
logIR (HOMA)	5.99	2.25-15.91	<0.0001	**7.68**	2.63-22.4	**<0.0001**
Smoker	3.09	1.10-8.22	0.01	**3.39**	1.26-9.13	**0.016**
Diastolic Blood Pressure (mmHg)	1.09	1.02-1.16	0.006			
Systolic Blood Pressure (mmHg)	1.08	1.02-1.14	0.007			
Waist circumference (cm)	1.08	1.02-1.14	0.007			
FMD (%)	0.84	0.73-0.97	0.021			
BMI (kg/m^2)	1.12	0.99-1.28	0.081			
Weight (kg)	1.03	0.99-1.07	0.133			
Age (years)	1.06	0.98-1.14	0.158			

Table 3.10 Logistical regression of paternal variables. Compared with fathers of normal grown offspring, men who fathered pregnancies affected by fetal growth restriction had an OR 7.68 (2.63-22.40; p<0.0001) of having a 1 unit higher log HOMA IR value and an OR 3.39 (1.26-9.16; p=0.016) of being a smoker. (OR, odds ratio: HOMA, homeostasis model assessment; FMD, flow-mediated dilatation; BMI, body mass index).

In order to assess maternal confounding sensitivity analyses was performed using forward stepwise regression, first including all paternal explanatory variables followed by addition of maternal explanatory variables (maternal age and BMI) to the model, With this use of step-wise regression, no maternal variable affected the differences identified in paternal insulin resistance or smoking.

HOMA beta cell function was assessed in a similar fashion. Significance was lost in a model including adjustment for logIR (data not shown). However, it did retain significance when adjusted for age, smoking status and abdominal circumference. Men who fathered a growth restricted offspring had an OR 1.03 (p=0.03) of having a one unit increase in beta cell function compared to men who fathered appropriately sized babies.

cas_con	Odds Ratio	Std. Err.	z	P>|z|	[95% CI]	
Homa β	1.015724	.0073583	2.15	0.031	1.001403	1.030249
age	1.037631	.0510038	0.75	0.452	.9423298	1.14257
smoker	3.516152	1.793553	2.46	0.014	1.293834	9.555575
Abdo	1.033914	.0339851	1.01	0.310	.9694048	1.102716

Table 3.11- Logistical regression of HOMA-β with age, smoking status and abdominal circumference (cm) including odds ratio, standard error (Std Err), p-value and 95% confidence interval

3.9.1 Secondary Analysis

Secondary analysis was performed assessing the data in a more clinical context. Using the criteria for the insulin resistance syndrome defined by the European Group for the study of insulin resistance in non-diabetic individuals the majority of men in the highest quartile for insulin resistance (n=30) were cases (19/42 cases (45%) compared with 11/77 controls (14%) (Figure 9a).

Further risk factors as defined The European Group for the Study of Insulin Resistance are:

(1) Central obesity: waist circumference 94 cm or more

(2) Dyslipidaemia: triglycerides (TG) 2.0 mmol/l or more and/or HDL-C less than 1.0 mmol/l or treated for dyslipidaemia

(3) Hypertension: blood pressure 140/90mmHg or more

(4) Fasting plasma glucose at least 6.1 mmol/l.

With use of these criteria for the insulin resistance syndrome, the majority of men in the highest quartile for insulin resistance (n = 30 of 119) were case subjects (19 of 42 case subjects (45%) compared with 11 of 77 control subjects [14%] (Table 3.12). Most of these case subjects (10 of 19 [53%]) also had an increased abdominal circumference >94 cm (Figure 3.3b), which was only noted in 2 of 11 (18%) control subjects (Figure3.3a). This information was then used to generate a second model limited to fathers in the top quartile for insulin resistance and their related insulin resistance traits (Figure 3.3b).

	n=	n=
Abdominal circumference (cm)	10	2
Triglycerides (mmol/L)	2	1
HDL-Cholesterol (mmol/L)	4	2
Blood Pressure (mmHg)	0	0
Fasting glucose (mmol/L)	1	0
Top 25% (n=30) fasting insulin value of our population(n=119)	19 (45.2%)	11 (14.3%)
Of these: no further risk factors n(% of 19)	5 (26.3%)	8 (72.3%)
1 further risk factor	**10 (52.6%)**	1 (9%)
2 further risk factors or IR syndrome	**4 (21%)**	1 (9%)

Table 3.12 Fathers with insulin levels in the highest quartile were selected (n=30). This included 19/42 (45%) of cases and 11/77 (14%) controls. These men were assessed for the presence of other risk factors for the insulin resistance syndrome as a % of the number of men identified in each group. The majority of men in this top quartile for insulin resistance who fathered growth restricted offspring had 1 or 2 further risk factors for the insulin resistance syndrome (14/19; 74%) versus only 18% (2 of 11) of the controls identified.

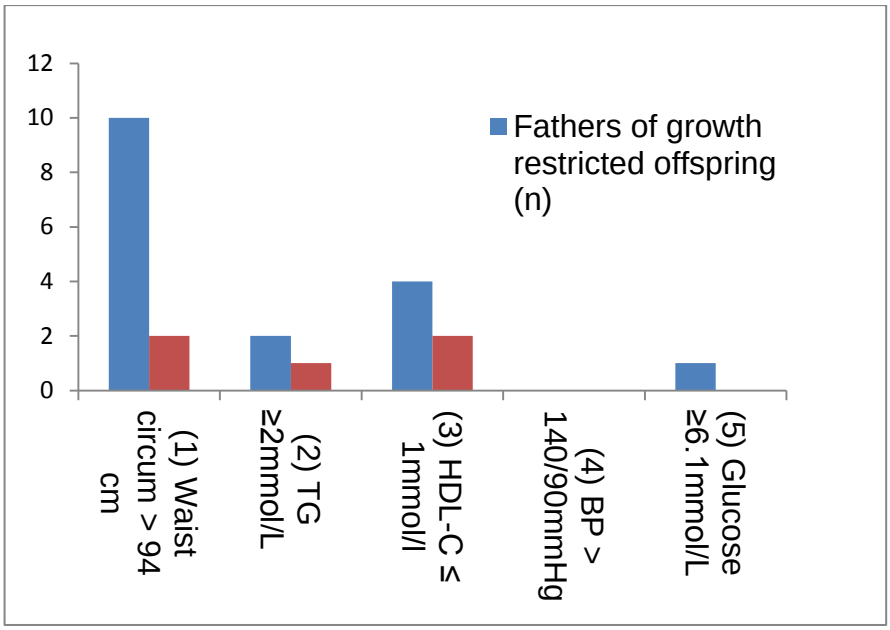


Figure 3.3a. Of individuals who were within top 25% insulin, the number of subjects who met other criteria for insulin resistant syndrome including waist circumference >94cm, TG ≥ 2mmol/l, HDL-C≤ 1mmol/l, BP >140/90 mmHg or blood glucose≥ 6.1mmol/l. Case men are in blue and controls in red.

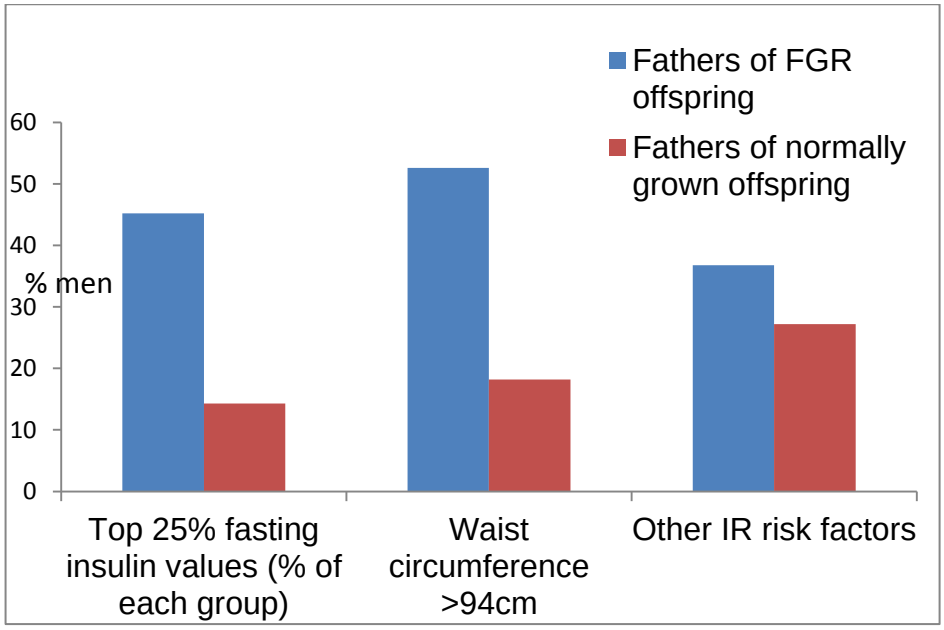


Figure 3.3b - Association between paternal insulin resistance syndrome (with individual criteria represented) and fathering a pregnancy affected by fetal growth restriction. Case fathers are in blue and controls in red.

Graph shows % of men in top quartile of insulin and of those, who had further risk factors for IR syndrome. Men who fathered growth restricted offspring were more than three times more likely to fall into the highest insulin resistance quartile, with more than half of them having a waist circumference >94cm.

This information could be used to generate a second model looking specifically at the risks for those in the highest quartile of insulin levels.

After multivariable analysis, men in the top quartile for insulin resistance who fathered a growth-restricted fetus had an OR 6.72 (95% CI 2.43–18.58; P ,0.0001) of having further risk factors for the insulin resistance syndrome and OR 3.36 (1.28–8.28; P-value 0.013) of being a smoker compared with fathers of normal grown offspring.

	Univariable			Multivariable		
	OR	P value	95% CI	OR	p value	95% CI
IR syndrome risk factor	5.86	<0.0001	2.25-15.27	**6.72**	<0.0001	2.43-18.58
Smoking	3.09	0.01	1.10-8.22	**3.26**	0.013	1.28-8.28

Table 3.13 Secondary multivariable analysis of men in top quartile of insulin levels in cohort, including smoking status then men who fathered growth restricted baby nearly seven times more likely to have other risk factors for insulin resistance syndrome.

3.10 Paternal Placental like growth factor *(Plgf)*

63 male control and 32 case serum samples were suitable for analysis. In 8 of the 32 cases mothers had PET. No controls had PET. As shown in Figure 3.4, *PlGF* levels in male controls; 11.80± 0.51 ng/L (mean±SEM) and *PlGF* levels in cases; 10.86±0.67 ng/L (no difference in cases with PET 10.8ng/L) A student t-test was performed on the data which revealed the difference did not reach significance [p value 0.3 (CI -0.8- 2.6)]

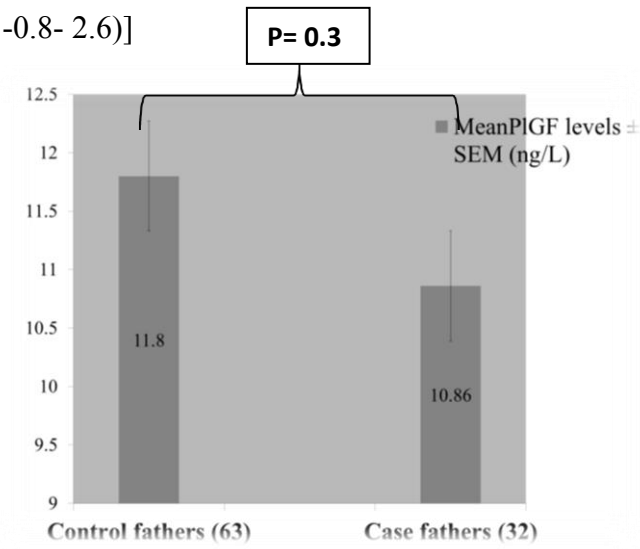


Figure 3.4 Bar chart of PLGF levels (ng/L) in case and control fathers

3.11 Fetal phenotype measurements

Weight and gestation was recorded for all babies and this data is presented already in section 3.5.2. It proved to be far more difficult recording length due to many of the FGR babies requiring admission to the Special Care unit where this measurement was not an immediate priority. It was, therefore, decided not to continue to record ponderal index due to the variability of data collection.
Overall baby gender was well matched in the study with 39 boys and 38 girls being born in the control group versus 20 boys and 22 girls in case offspring. In the control group boys were 228g heavier on average than girls. There was no difference in weights seen between boys and girls born in the case group (Table 3.1).

3.11.1 Fetal sample collection

As babies were being born throughout the study, it was not possible to achieve 100% collection of cord blood and placental samples. Of 119 deliveries, samples were retrieved in 58 controls (75% of trios) and 30 cases (3 where consent just for samples had been given) so 64% of trios) which resulted in an overall collection rate of above 70% for offspring where parental data was also available). The slightly lower number of cases also reflected situations where it was not possible to collect umbilical cord blood or where placentas needed to be sent before any sampling (total n=5). In two control cases where sampling was missed at the time of delivery, buccal swabs were used on offspring during postnatal visits at the request of parents.

3.11.2 Glucose homeostasis in offspring

Due to the nature of collections and prioritisation of use, it was not always possible to collect enough cord blood in order to obtain plasma and serum. Only samples where there was enough cord blood and the collection had been taken within one hour of delivery were centrifuged to achieve plasma or serum from cord blood.

This meant that 12 plasma samples were available from fetal growth–restricted offspring and 21 from normal grown offspring in order to assess for insulin and c-peptide level (Table 3.14). In 2 controls and 2 cases a c-peptide result was unavailable. Both control and case offspring had similar insulin levels (5.41±3.52

and 7.12±8.66 mIU/L, P = 0.43) respectively, and similar c-peptide levels (1.04±0.42 and 1.21±0.63 ug/L, P = 0.38).

	CASES (mean)	SD	CONTROLS (mean)	SD	P value
No samples	11		18		
Insulin (mIu/l)	7.12	8.66	5.41	3.52	0.43
No samples	10		19		
C-peptide (ug/l)	1.21	0.63	1.04	0.42	0.38

Table 3.14 Offspring plasma levels of c-peptide (ug/l) and insulin (mIU/l) in growth restricted offspring (cases) and appropriately grown offspring (controls).

3.11.3 Family Glucose Homeostasis Comparison

Individual analysis was undertaken to assess correlation between fetal cord blood insulin and paternal insulin levels (Table 3.15).
Firstly, for cord blood plasma insulin levels available, paternal log IR was interrogated to assess whether the small sample size could reflect the larger dataset. This initial analysis seemed promising and revealed the mean paternal IR (with offspring data available) was significantly different (p=0.008 95% CI 4.13-7.3) with control mean serum insulin 4.36±2.8 mIU/L and cases 8.69±5.8 mIU/L.
In concordance with the above data paternal logIR between controls (-0.27±0.04) and cases (-0.03±0.09) remained significantly different (p=0.006 95% CI -0.28 to -0.1) However, comparison of all samples revealed no correlation seen between paternal logIR and offspring cord blood serum c-peptide or insulin levels in our cohort (p=0.57) (Figure 3.5)

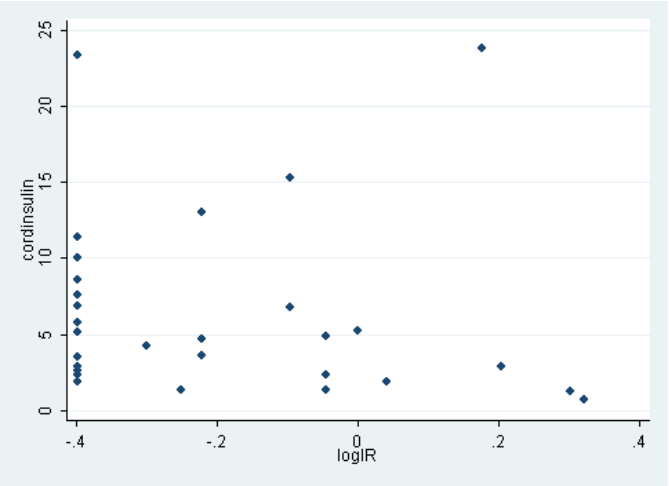


Figure 3.5. Graphical representation of correlation between paternal log HOMA-IR and offspring serum cord insulin (mIU/l)

Family member	Group	Random glucose (mmol/L) AN	Fasting glucose (mmol/L)	Fasting insulin* (mIU/L)	IR (HOMA)	C-peptide* (ug/L)
BABIES	Cases Controls p value			7.12 (8.66) 5.41 (3.52) 0.43		1.21 (0.63) 1.04 (0.42) 0.38
MOTHERS	Cases Controls p value	4.45 (0.52) 4.54 (0.54) 0.44	4.53 (0.37) 4.57 (0.46) 0.78	4.38 (2.41) 3.97 (2.15) 0.57	0.58 (0.29) 0.56 (0.24) 0.57	
FATHERS	Cases Controls p value		4.87 (0.41) 4.72 (0.32) 0.037	7.1 (4.29) 4.81 (2.61) 0.001	0.93 (0.52) 0.63 (0.31) 0.0003	

Table 3.15 Family glucose homeostasis in controls and cases. In babies only glucose and c-peptide recorded. In mothers, antenatal random glucose (mmol/l) along with postnatal study recorded. Fathers fasting values recorded.

3.12 Remembered parental birth weight

As part of the questionnaire, both parents were asked if they knew of their own birth weights. This information was self -reported often through another family member. Parental birth weight was known in 57 (48%) fathers and 65 (55%) mothers. Case fathers, were lighter than control fathers (birth-weight 3127±597g, versus 3506±380g, respectively (p=0.0045). Case mothers, were lighter than control mothers (birth-weight 3007±576g versus 3322±425g, respectively (p=0.012).

CONTROLS					**CASES**				***p value***
	n	*Mean (grams)*	*SD*	*Range*	*N*	*Mean (grams)*	*SD*	*Range*	
Baby (total)	77	3517	367	2760-4400	42	2019	752	560-3280	<0.001
Male baby	39	3629			20	2019	903		
Female baby	38	3401			22	2019	627		
Father*	37	3506	380	2780-4540	20	3127	597	1000-3970	0.0045
Mother*	45	3322	425	2300-4300	20	3007	576	1000-3800	0.012

Table 3.16 Family birth weights in grams (as remembered) in relation to offspring birth weight.

3.13 Discussion

This case-control study identified women with pregnancies affected by fetal growth restriction and showed that their partners were more insulin resistant and more likely to smoke compared with fathers of normal grown offspring. Fathers of growth-restricted offspring also had other elements of the insulin resistant syndrome including; high blood pressure, endothelial dysfunction, upper body fat redistribution and a more atherogenic lipid profile. These observations support epidemiological studies that have consistently observed an increased incidence of type 2-diabetes and cardiovascular disease amongst men who previously fathered low-birth weight offspring (Davey Smith G et al., 2004; Wannamethee SG et al., 2004, and Hypponen E et al., 2003).

Although the study provides objective evidence of sub-clinical insulin resistance at the time of fathering a growth-restricted offspring, the Exeter Family Study of Childhood Health (EFSOCH), which studied almost 1000 normal grown offspring and their fathers, did not find an association between offspring birth weight and paternal insulin resistance (Knight et al., 2006). Unlike the cohort I recruited, however, the EFSOCH study only investigated offspring with a normal birth weight (2.95kg - 3.98kg). Under these circumstances paternally inherited insulin resistance may be compensated by increases in fetal insulin production. This suggestion was supported by their observation that paternal insulin resistance was inversely correlated with fetal insulin concentrations (Shields et al., 2007). In a sub-set of our study population, I was unable to detect such a relationship between paternal insulin resistance and fetal cord blood insulin or c-peptide levels although fetal insulin levels were higher in FGR offspring but this did not reach statistical significance. It is possible that the secondary analysis did not have statistical power to detect such a correlation. Another explanation is that, in this study, cases may have included a mixture of growth restricted offspring, some with insulin resistance and high fetal insulin levels and others with reduced beta cell function and low fetal insulin secretion, so that a clear difference could not be detected in overall insulin levels between cases and controls. It is possible that other heritable factors that are passed from father to offspring influence both paternal phenotype and fetal growth.

Men who fathered growth restricted offspring were themselves smaller at birth. This observation could be explained by either the inheritance of genes that limit fetal growth, such as described in the fetal insulin hypothesis or fetal adaptations in response to relative malnutrition in utero, the thrifty phenotype (Hales CN et al., 1991). Fetal growth restriction is associated with an increased risk of perinatal death (Cruz-Lemini M et al., 2012). It is also possible that survivors of fetal growth restriction, like all the cases in this study, have inherited paternal characteristics that predispose to weight gain and cardiovascular risk factors. During the study, cases of fetal growth restriction resulted in intrauterine death, but the families were unwilling to participate in the study. Whether families of fetal growth restricted offspring that result in perinatal death are different from families of survivors remains a challenging question to answer.

In the study the effect of maternal environment was minimised by only including healthy pregnant women within pre-specified phenotypic limits and excluding fetal growth restriction due to recognized maternal or fetal diseases. This allowed study pregnancies predominantly affected by placental disease. These predetermined maternal inclusion criteria are likely to have strengthened the effects of paternally inherited factors. Others have found that men who father small for gestational age (SGA) offspring are more likely to be overweight and to have a greater waist circumference than fathers of normal grown offspring (McCowan et al., 2011). This study adds objective measures of paternal insulin resistance, endothelial function and lipid profile to these phenotypic characteristics.

Differences in paternal blood pressure and endothelial function between the groups were no longer evident following adjustment for paternal smoking. Cigarette smoking is known to independently raise BP and impair flow-mediated dilatation (Hirofumi T et al., 2008). Maternal smoking is a recognized risk factor for fetal growth restriction (Reeves S et al., 2008). Women who smoked were therefore excluded from the study. However, the partners of some women smoked during pregnancy. Paternal smoking has previously been associated with fetal growth restriction and correlates with levels of maternal cotinine (Suter M et al., 2010). In this study, maternal cotinine was only detectable in serum of 3/17 women whose partner smoked, compatible with low-level passive smoking. Although fetal cotinine levels from umbilical cord blood were not checked, it is unlikely that maternal

passive smoking contributed to fetal growth restriction. Maternal smoking can cause epigenetic change to human placental genes (Suter M et al., 2010). It is currently unknown whether paternal smoking can cause epigenetic change that is inherited by the fetus and placenta.

For pragmatic reasons, insulin resistance was measured using the HOMA model derived from fasting insulin and glucose. A more robust, but invasive technique is the euglycemic insulin clamp (Matthews et al., 1985). HOMA correlates well with clamp-derived methods (Matthews et al., 1985) and, in this cohort, is supported by other parameters associated with the insulin resistant syndrome.

The metabolic pattern in case men was of insulin resistance but accompanying beta cell dysfunction was not evident. Whilst T2DM classically presents with both features, studies have shown that in isolation insulin resistance can represent one of the earliest phases of progression towards disease development several years later (Skarfors ET et al., 2010). This early pre-diabetic stage can be accompanied by other signs of diabetes risk including hypertension and obesity which is consistent with the findings in this study.

It has been suggested that T2DM disease is triggered by different genetic defects; one defect affects beta cell function and is expressed later in life and another, present from early life, is responsible for the insulin resistant phenotype particularly through the impaired stimulation of glycogen formation in skeletal muscle (Beck-Nielsen H et al., 1994). This sequence of events would be consistent with the paternal phenotype identified by this study.

GCK mutations are thought to be responsible for a significant proportion of diagnosed MODY cases but classically present from a young age with impaired glucose stimulated insulin secretion and mild beta cell impairment (Hattersley A et al., 1999). This presentation gives rise to the possibility that the metabolic pattern seen in these men who fathered growth restricted offspring may share limited aetiology with current known genetic causes of MODY where the primary defect is beta cell impairment rather than insulin resistance.

Only 4 fathers of growth-restricted pregnancies (9.5% of cases) fulfilled a European definition of the insulin resistant syndrome (Balkau B et al., 1999) but this does not negate the early pre-diabetic state I identified. The majority of cases in the top 25% for insulin resistance also had central obesity as defined by a waist circumference >

94cm. Insulin resistance is closely linked with central obesity, which in turn precedes other elements of the metabolic syndrome (Cameron AJ et al., 2008). Sub-clinical insulin resistance at the time of fathering a growth restricted offspring not only explains the association with future paternal T2DM and cardiovascular disease, but identifies a group of men with a reversible risk factor for future metabolic and vascular disease (Reddy KJ et al., 2010), just as gestational diabetes identifies women at risk of future diabetes (Bellamy L et al., 2009).

In men with established T2DM, a similar log unit increase in HOMA insulin resistance to that found in fathers of growth restricted offspring, has been associated with a 31% increased risk of cardiovascular disease (Bonora E et al., 2002). Insulin resistance in men without diabetes is an independent risk factor for future cardiovascular disease (Nakamura K et al., 2010 and Hanley AJ et al., 2002). Dietary and lifestyle measures can reverse insulin resistance and reduce future cardiovascular risk (Eriksson J et al., 1999 and Caterson ID et al., 2012). The observations from this study suggest at the time of fathering a growth restricted offspring that these men may benefit from advice on a healthy life-style as part of primary prevention of diabetes and cardiovascular disease.

Inheritance of common insulin control genes (Freathy et al., 2010), rarer monogenic disorders of glucose metabolism whose primary defect is not beta cell defects (Hattersley A et al., 1999), or other as yet unidentified heritable factors, may explain this observed link between paternal insulin resistance and fathering a growth restricted offspring. It is possible that paternal life-style leading to obesity and smoking, may influence change that leads to insulin resistance, which is inherited by offspring and manifest as fetal growth restriction (Ng SF et al., 2010).

Despite other anthropomorphic parameters revealing no significant differences between case and control women antenatally or postnatally, endothelial dysfunction appeared to differ between the two groups. Postnatal FMD in women, whose pregnancy had been affected by growth restriction, suggested significantly more endothelial dysfunction (6.98 ±0.97%) than women who had an appropriately grown baby (10.26 ±2.4%). This pattern of endothelial function has been reported before in relation to postnatal FMD studies on women whose pregnancies have been affected by pre-eclampsia (Noori M et al., 2010 and Hamad RR et al., 2007).

In fact postnatal FMD measurements by Noori et al. on pre-eclamptic subjects yielded very similar results (FMD controls 9.64±0.6% versus cases 7.3±2.4 %) to the ones seen in this cohort. When women who had experienced pre-eclampsia were removed from the growth restriction analysis then results remained significantly different between FGR case women and controls.

A previous study (Paradisi G et al., 2011) has reported similar differences in reduced FMD in women who had experienced a small for gestational age baby but in this small cohort, these women had higher blood pressure and a strong family history of heart disease, making it difficult to interpret whether pre-existing risk factors or a pregnancy affected by FGR was the cause of the endothelial dysfunction.

Given that antenatally these women exhibit no other signs of increased cardiovascular risk then the possibility arises that the pregnancy itself causes some underlying damage to the mother. A large retrospective study has shown that maternal risk of ischaemic heart disease (IHD) admission or death is associated with delivering a baby in the lowest birthweight quintile for gestational age (adjusted hazard ratio 1·9 [95% CI 1·5–2·4]) as well as preterm delivery (1·8 [1·3–2·5]) and pre-eclampsia (2·0 [1·5–2·5] but that these associations are additive (Smith GC et al., 2001). In this study maternal smoking was not factored in and the presence of underlying maternal disease (although screened for) was not known.

Recently, risk of maternal IHD was evaluated in a population based cross-sectional study of 6,608 women with a prior live term SGA birth. Delivery of a SGA infant was strongly associated with greater maternal risk for IHD (age adjusted OR; 95% CI: 1.8; 1.2, 2.9; p=0.012), independent of the family history of IHD, stroke, hypertension and diabetes (family history-adjusted OR; 95% CI: 1.9; 1.2, 3.0; p=0.011) as well as other risk factors for IHD (risk factor-adjusted OR; 95% CI: 1.7; 1.1, 2.7; p=0.025). (Bukowski R et al., 2012). However, despite decreased FMD, it remains to be seen if cardiovascular disease will develop in the current cohort of women who had a growth restricted pregnancy complicated specifically by placental dysfunction.

3.14 Summary

This case-control study identified that men who fathered growth restricted offspring (mean 1.8±2.2th customised birth centile) were more insulin resistant and more likely to smoke compared with fathers of normal grown offspring (49.6±27.5th) customised birth centile. In the study all offspring met the inclusion criteria.
Fathers of growth-restricted offspring also had other elements of the insulin resistant syndrome, including high blood pressure, endothelial dysfunction, and upper body fat redistribution and tended to a more atherogenic lipid profile.

Of those fathers who smoked, plasma cotinine levels in their partners during the pregnancy revealed no evidence that women were actively smoking or exposed to high levels of passive smoke.

In contrast to fathers, mothers recruited and shown to similar in weight, age, height and parity at the time of their pregnancy. Postnatally they again were phenotypically similar with no differences identified between blood pressure, abdominal circumference or weight. Metabolically; postnatal lipid profiles and glucose homeostasis were similar between mothers. Endothelial function was, however, found to be impaired in case women.

Although insulin and c-peptide levels were found to be higher in FGR offspring there was no statistically significant difference between case and control offspring plasma cord blood values but sample size was small.

This study identified that men who fathered a growth restricted pregnancy had a pre-diabetic phenotype at the time of the pregnancy that could have contributed to fetal phenotype. Furthermore, that their metabolic dysfunction was one primarily of insulin resistance in conjunction with higher abdominal circumferences.

Chapter 4

The Genetics of Fetal Growth Restriction

Chapter 4: The Genetics of Fetal Growth Restriction

4.1 Introduction

The fetal insulin hypothesis states that low birth weight and susceptibility to T2DM could both be two phenotypes of the same genotype. Fetal insulin mediated growth can reflect maternal glycaemia and fetal genetic factors which regulate the ability of the fetus to secrete insulin and the degree of insulin resistance in fetal tissues. A genetic defect in the sensing of maternal glucose by the fetal pancreas, fetal pancreatic insulin secretion, or the action of the insulin on the insulin-dependent tissues, would all result in reduced fetal growth (Hattersley and Tooke, 1999). T2DM is characterized by insulin resistance and/or beta-cell dysfunction. Therefore, genetic variants impacting on insulin secretion and action could alter both fetal growth and susceptibility to T2DM.

Studies of rare monogenic diabetes support this idea. Mutations in a single gene, such as *GCK* or *KCNJ11*, are associated with diabetes and reduced birth weight (Hattersley A et al., 1998). Furthermore, studies of the *GCK* gene, has shown that a mutation present in a father and inherited by his offspring, results in a lower weight baby, than if the offspring is born without the mutation or if the mutation was present in mother and offspring (Hattersley A et al., 1998).
Genome wide association studies of birth weight have identified variants in genes known to predispose to diabetes, which if present in the fetus are associated with reduced birth weight (Freathy et al., 2010).

Together these data provide evidence for a genetic contribution to T2DM and low birth weight which is most evident when inherited from father to offspring.

In this chapter I investigated family genotypes in order to identify whether there was any evidence of a specific, paternally inherited, fetal genotype in FGR offspring. As genes associated with monogenic diabetes have been implicated in low birth weight with evidence of inheritance from the father, these genes were of particular interest.

Whole exome sequencing (WES) of FGR offspring was the platform selected as it provided the opportunity to capture genes of interest, but also investigate FGR offspring for novel genetic variants.

Validation of exome sequencing was undertaken with Sanger sequencing when novel and very rare variants were identified. Taqman genotyping was used to validate other MODY variants of interest in FGR offspring. Parental DNA was also genotyped in order to assess patterns of inheritance.

A panel of risk SNPs associated with T2DM, insulin resistance syndromes and elevated fasting glucose were genotyped if they were present in 3 or more of the exome samples. The cohort was underpowered, but associations identified with FGR may provide support for a larger study to detect differences. Parental DNA, allowed for assessment of inheritance of risk SNPs and their possible role in FGR in support of the fetal insulin hypothesis.

4.2 Hypothesis

Genetic inheritance of insulin resistance by a fetus from its father will result in a growth restricted phenotype *in-utero* and predispose both offspring and father to T2DM in later life.

4.3 Specific Aims

1. To discover novel genetic variants in growth restricted offspring through the use of whole exome sequencing.
2. To identify whether variants in MODY candidate genes are associated with growth restriction and identify whether there is evidence of transmission from father to offspring.
3. To identify the presence of T2DM SNPs in the cohort of growth restricted and appropriately grown offspring and parents.

4.4 Methods

4.4.1 Study population and design

Ten Caucasian FGR offspring were selected from the original cohort of trios collected that were felt to best represent a growth restricted phenotype. Cases were selected with different gestational ages at delivery. WES required 3.3ug of DNA from DNA extracted from FGR offspring cord blood (as described in methods section 2.12) which was used as it had yielded the highest DNA concentrations between cord blood, cord and placenta. After preparation and enrichment (as described in methods section 2.13) samples were run on the HiSeq1000 genome analyzer.

Data was aligned and annotated as described previously in section 2.13.1. Filtering of the data was performed manually and with the use of Ingenuity Variant Analysis software as detailed in methods section 2.13.2.

Candidate genes for MODY, T2DM, insulin resistance (IR) and elevated fasting glucose were identified through WES and genotyped using TaqMan technology (Applied Biosciences, ABI, Warrington UK) as described in detailed methods section 2.14.1. Reactions were performed on 384-well microplates and analysed using ABI TaqMan 7900HT software. Sanger sequencing was performed with novel/rare variants from FGR offspring, identified through WES (along with parental DNA). Primer design and PCR conditions were followed as described in methods section 2.14.2.

4.4.2 Statistical Analysis

Statistical analyses of genotyping were performed using Intercooled Stata 10 for Windows (StataCorp LP) according to detailed methods in section 2.14.1.3. Data is presented as means± SD unless otherwise stated.

4.5 Results

4.5.1 FGR offspring birth phenotype characteristics

Ten cases were selected from the original FGR offspring cohort as being most representative of a cross-section of pre-term and term growth restricted offspring secondary to placental dysfunction (Table 4.1). Four of the 10 cases were delivered at 34 weeks or less. In 2 of these cases the primary indication for delivery was maternal pre-eclampsia in association with growth restriction. In the other two cases, delivery was indicated purely for the fetal condition of severe *in-utero* growth restriction.

Of the remaining 6 cases all were delivered after 34 weeks and none of these pregnancies were complicated by pre-eclamspsia. All women were induced or delivered by elective caesarean section because of the presence of fetal growth restriction as evidenced by ultrasound.

All babies were born alive with an average birth weight of 1854g (range 779-2730g) but separating babies born ≤34 weeks, the average birth weight for those 4 infants was 965g versus 2446g (range 1900-2730) for the 6 babies born after 34 weeks. Customising birth centiles which is calculated from birth weight, gestation, maternal ethnicity, parity, weight and height has the advantage of adjusting birth weight for gestation.. The average birth centile for these 10 offspring was the 1.1th centile (range 0-3.8%). Four of the offspring were boys and 6 were girls (Table 4.1).

ID	Baby gender	Gest (weeks)	Birth Weight (g)	Birth centile %	≤34 weeks
331MBV	M	37+2	1900	0.1	
337ECb	F	29+4	850	0	y
386SSa	F	39+3	2650	3.8	
39NK	F	38+1	2410	0.7	
413BTa	M	38+4	2480	1.1	
53ALa	M	38+2	2510	3	
328LB	M	28+4	855	0.1	y
227FBa	F	30+1	779	0	y
156BTc	F	40+4	2730	2.1	
334MVb	F	34	1378	0.1	y

Table 4.1 The phenotypic characteristics of the FGR offspring samples including baby gender, gestation born at (weeks), birth weight (grams), and customised % birth centile

4.5.3 Parental Phenotype

Maternal and paternal phenotype data was also available for these offspring and is described in table 4.2. Mean maternal age was 34.2±5.5 years versus average paternal age of 35.4±4.8 years. For most couples this was their first child but in two instances they had one child already. For case mothers the mean BMI was normal at 22.3±3.1 kg/m^2 but for fathers it was raised at 27.62±.9 kg/m^2, as was their HOMA IR (mean 1.1). No mothers smoked, but two fathers did.

	Mother					Father					
ID	Mat Age	BMI	Parity	PET		Pat Age	BP	BMI	IR	Abdo	Smoker
331MBV	33	27.8	0			34	117/68	30.7	1.4	113.5	
337ECb	27	20	0			34	129/75	30.6	0.8	92.5	
386SSa	40	20.4	0			34	131/85	31.5	1.9	90	y
39NK	34	22.9	1			33	127/74	27	0.4	96	y
413BTa	33	20.4	1			35	119/69	24	0.6	92	
53ALa	26	26	0			27	114/61	25.1	1.2	90.5	
328LB	44	25.9	0	y		44	123/76	30.1	1.5	95.5	
227FBa	37	20	0	y		40	120/73	27.3	0.9	107	
156BTc	32	20	0			37	11/737	23.5	0.8		
334MVb	36	20	0			36	132/76	26.6	1.6	98.5	

Table 4.2 Phenotypic characteristics of parents of FGR offspring used including; maternal age, BMI in kg/m^2 (calculated weight/height^2), parity and presence of pre-eclampsia (PET) and paternal age (years), blood pressure (mmHg), BMI, HOMA insulin resistance (IR), abdominal circumference and smoking status

4.5.2 DNA Quality Control (QC)

The 10 FGR offspring DNA samples that were prepared for analysis and all passed initial QC and concentration checks as described previously (Table 4.3). Samples were run on 1% Agarose gel according to methods as described previously in section 2.12.4 and the DNA was found not to be degraded (Figure 4.1).

Sample ID	Concentration (ng/ul)	Volume (ul)	Total ng	260/280	260/230
331MBV	100	32	3200	1.79	2.27
337ECb	100	32	3200	1.9	2.39
386SSa	100	32	3200	1.84	2.27
39NK	100	32	3200	1.86	2.35
413BTa	100	32	3200	1.84	2.39
53ALa	100	32	3200	1.84	2.29
328LB	70.60	43	3035.8	1.84	2.21
227FBa	100	32	3200	1.84	2.36
156BTc	74.50	42	3129	1.87	2.27
334MVb	81.40	38	3093.2	1.82	2.26

Table 4.3 Offspring cord blood DNA samples used in exome sequencing. Sample ID, concentration (ng/ul), volume and total amount present (ng) along with Nanodrop 260/280 and 260/230 purity readings.

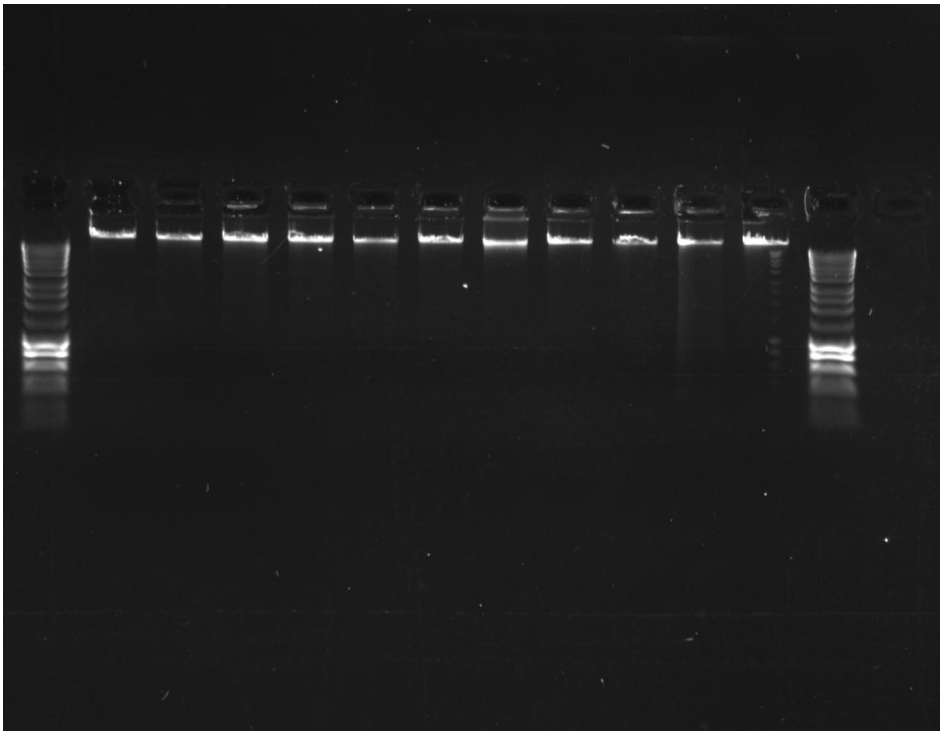


Figure 4.1 Agarose gel (1%) of genomic DNA (331,337, 385, 39, 413, 53, 328, 227, 156 and 334) with 1kb ladder at ends of gel

4.6 Analysis of variants by manual filtering

4.6.1 Loss/gain of function variant call

Using the following filters; stop gain, stop loss, frameshift insertion/deletion, splice site and not defined then 75 variants were identified that were present in 3 or more of the 10 samples.

4.6.2 Non synonymous & non frameshift insertion/deletions variant call

Non-synonymous (NS) variants result in the substitution of a nucleotide and subsequent introduction of a different amino acid in a sequence and can have functional significance. Non frameshift mutations were annotated separately, but in effect are similar to NS variants. These variants result in insertions or deletions of nucleotides in a number divisible by three which means the amino acid sequence is altered by an insertion or deletion but does not give rise to a shift in the triplet reading frame of the gene.
Assessment of non-synonymous variants yielded a further 158 variants with a frequency in ESP database of <0.05 or blank (meaning no frequency in ESP database). The same filter was applied for non frameshift variants which revealed a further 82 variants in three or more samples.

Overall 75 novel loss/gain of function mutations and 240 non synonymous or non-frameshift variants were identified in three or more FGR offspring samples. A second approach using Ingenuity was employed to try and identify which of the novel candidates were most likely to have a biologically significant effect.

4.6.3 Ingenuity variant analysis call incorporating a biological filter

The 10 FGR ‘vcf’ files were input into Ingenuity Variant Analysis along with 5 control samples from the GOSgene database.
From all samples uploaded to Ingenuity (including the 5 control samples) there were 414304 variants identified on 20496 genes. Ingenuity filters were then set using similar terms as employed in the manual assessment approach. For common variants, Ingenuity was set to exclude all referenced dbSNP variants. However, as Ingenuity

used the latest version of dbSNP (137 at the time of analysis) this resulted in a reduced set of variants than the manual filtering process. Other filters were set to exclude all referenced variants that had a frequency <3% according to the 1000 genomes project and <1% of the NHLBI ESP database. Those variants not thought to be pathogenic (according to Polyphen and SIFT) were excluded and finally a biological significance filter was applied to the remaining variants (Figure 4.2).

Figure 4.2 Filter cascade applied to genetic variants identified in 10 FGR samples

The following biological terms were applied as a filter: fetal growth restriction; maturity onset diabetes of the young; type 2 diabetes and pre-diabetes. Lists of variants were then generated for genes directly associated or 1 or 2 hops away to compare sensitivity. When only those genes directly associated with the terms were used this resulted in only 2 potential variants being identified.

When two hops were applied this resulted in 174 variants of interest and when one hop was applied this resulted in 71 variants (Appendix 4 S6) being identified and on further assessment interesting candidate genes were present so the final analysis was therefore carried forward on these variants.
None of the 71 variants were seen in the 5 controls (which were set as a separate filter) and the next step in the analysis was again to assess whether they could be identified in any of the current exome data held by GOSgene.

This was done through the use of a software package PLINK/SEQ (a library for the analysis of genetic variation data) and the assistance of Dr Chiara Bacchelli who had access to the data.

Of the 71 variants, 10 could not be identified in any other GOSgene samples and a *TBP* variant was included as it was a tri-nucleotide repeat insertion which showed similarities to another pathological condition (spino-cerebellar ataxia) and needed further sequencing for clarity.

Of the 10 variants identified (Table 4.4), I then sought to stratify those with greatest biological plausibility in order to take them forward to validate. Of these 10, *KRT8, TBP, TDG* and *KDM4B* were connected in Ingenuity to fetal growth restriction with *KRT8* also connected to T2DM and MODY. *SAFB2* was connected to T2DM and MODY.
KRT8 and *SAFB2* are predicted by SIFT and Polyphen to be pathogenic which supports the concept that these genetic variants could have a functional role. However, when BAM files were checked, poor coverage was found in *KRT8* and *SAFB2* raising the possibility of a false positive.

Sanger sequencing was, therefore, performed in order to validate the exome finding. In total, 6 variants (highlighted in bold in table 4.4) were taken forward to validate by Sanger sequencing in each of the FGR offspring DNA in which they were present as well as parental DNA.

These variants in *KRT8, SAFB2, TDG, TBP, RPL14 and KDM4B* and the two variants identified in *HNF1a* and *HNF4a* were Sanger sequenced.

No samples	Ch	Position	Ref Allele	Sample Allele	Variation Type	Gene Symbol	Transcript Variant	Protein Variant	Translation Impact	SIFT Function Prediction
2	12	1.21E+08	C	A	SNV	*HNF1A*		p.P291Q	missense	
1	20	43043289	C	T	SNV	*HNF4A*	c.C569T:p.P190L	p.P212L	missense	
3	2	1.32E+08	G	T	SNV	*TUBA3C*	c.1040G>T	p.C347F	missense	Damaging
4	11	47857258	G	A	SNV	*NUP160*	c.1046C>T	p.S349F	missense	Damaging
7	12	53343231	G	C	SNV	***KRT8***	c.274G>C; c.-44734+268C>G	p.A92P	missense	Damaging
7	19	5587270	A	G	SNV	***SAFB2***	c.2846T>C	p.F949S	missense	Damaging
3	12	1.04E+08	G	C	SNV	***TDG***	c.23+1G>C			
3	6	1.71E+08		CAGCAGCAG	Insertion	***TBP****	c.189_190insCAGCAGCAG	p.43_44insQQQ	in-frame	
9	6	1.71E+08		GCA	Insertion	*TBP**	c.155_156insGCA	p.52_53insQ	in-frame	
3	19	50308921	C	A	SNV	*AP2A1*	c.2472C>A	p.Y824*	stop gain	
5	2	1.79E+08	A	C	SNV	*TTN*	c.61653T>G	p.I20676M	missense	
8	3	40503520		***	Insertion	***RPL14***		p.149_150insAAAAA	in-frame	
	19	5144411				*KDM4B*	**	p.963_964ins ESITVSCGVGQGAGGGWEHSDNLYP		

** *KDM4B* Transcript variant c2889_2890ins

* *TBP* transcript suggested varying length of insertions in samples (treated as one variant)

*** *RPL14* Transcript variant c.445_446insCTGCTGCTGCTGCTG
GAGAGCATCACGGTGAGCTGTGGGGTGGGGCAGGGGGCGGGGGGAGGCTGGGAGCACAGCGACAACCTGTACCCC
KDM4B Protein variant p.963_964insESITVSCGVGQGAGGGWEHSDNLYP *** *RPL14* Transcript variant c.445_446insCTGCTGCTGCTGCTG

Table 4.4 The 2 MODY variants and 10 novel variants identified in Ingenuity according to filter settings previously described and also not recognised as being present in in-house database of exomes. Number of offspring present in, chromosome and position, reference and observed allele (where applicable), transcript variant and gene name, impact and protein transcript along with SIFT function prediction reported. Highlighted in bold are variants identified to be taken forward for validation

4.7 Novel/ rare variant in MODY genes

I screened for currently known MODY genes in the exome data and included the following genes; *HNF1α; HNF1β, HNF4α, GCK, KCNJ11, CEL, NEUROD1, PDX1, WFS1, PPARG, NOTCH2, MADD* and *GLIS3.* Ingenuity was used for this process and a gene candidate text file was uploaded and used as a separate filter.

Displayed in table 4.5, are the 6 rare genetic variants, according to ESP (underlined) within MODY genes identified or not recorded in dbSNP130 and predicted to be pathogenic in SIFT and Polyphen. In total, 8 variants in 7 genes across the 10 samples were identified, all of which were non synonymous mutations (see next section for remaining two variants).

Ch	Position	Obs	Ref	Gene	Si	ESP	dbSNP (130)	
17	36059141	T	C	*HNF1β*	D	0.00028		het
1	120478125	A	C	*NOTCH2*	D	0.00223		het
4	6302816	C	G	*WFS1*	D	0.00437	rs35031397	het
12	12143527	G	A	*HNF1α*	D	0.24642	rs2464196	het
20	43034783	C	T	*HNF4α*	D	0.04211	rs736823	het
11	17408831	G	C	*KCNJ11*	D	0.03365	rs1800467	hom

Table 4.5 MODY variants of interest identified, gene name, predicted dangerous by SIFT, ESP frequency as recorded in Ingenuity and rs number where available as taken from Ingenuity and location on chromosome, variant observed and whether heterozygous or homozygous.

It was possible using custom assays (designed by Sigma Aldrich) to perform Taqman genotyping in order to validate findings in the variants.

A manual search was also undertaken through the data, which revealed two very rare variants in *HNF1α*and *HNF4α* (in the same fetal sample-227FB) not predicted to be dangerous by Sift (therefore not identified by Ingenuity).

Gene		Ch	Position	Obs	Ref	
HNF1α	-	12	12143125	T	C	het
HNF4α	-	20	43043289	T	C	het

These two rare MODY variants were Sanger sequenced in the one FGR offspring in which they were identified. Corresponding maternal and paternal DNA in this

offspring were also Sanger sequenced in order to identify the inheritance pattern for the variants. Sanger sequencing was performed as these variants had only been reported once previously in ESP.

4.8 Assessment of exome coverage in relation to MODY genes

As part of the assessment of WES capture, overall read depths were calculated for all exome coverage (Table 4.6). The actual coverage of the exomes within each of the individual candidate genes of interest was also investigated in order to assess whether coverage was adequate to capture the entire region of interest (Table 4.7)**.**

Read depth	**1X**	**5X**	**10X**	**15X**	**30X**
HNF1α	99.63	95.74	90.65	86.05	64.24
HNF1β	100.00	98.88	97.22	92.19	71.35
HNF4α	97.93	95.00	93.16	86.11	64.46
GCK	90.97	85.83	83.68	82.23	69.29
KCNJ11	96.69	87.90	83.32	81.37	71.22
PDX	79.74	62.08	48.61	40.03	25.76
NOTCH2	99.32	95.74	92.50	86.18	54.40
WFS1	94.60	86.13	79.03	73.27	52.23

Table 4.6 represents the mean coverage of all individual exomes together within each of the MODY genes of interest where variants were identified. Values are a % of coverage at 1, 5, 10, 15 or 30 read depth. For *HNF1α, HNF1β, HNF4 α* and *NOTCH2* coverage is above 90% at 10X. *GCK* and *KCNJ11* are coverage is above 80% at 10X but *PDX* is not well covered above 5X.

Sample id	total coverage	Mean (X)	%_bases_ above_1	%_bases above_5	%_bases_ above_10	%_bases_ above_15	%_bases _above_30	Median (X)
Internal								
328LB	1840720090	29.65	94.7	89.2	82.4	73.4	40.5	26.11
39NK	2101494548	33.85	91.3	75.2	65.2	58.2	42.3	25.2
53AL	2073917523	33.4	94.6	89.6	83.9	76.7	48.6	31.26
156BT	1905955206	30.7	94.2	89.1	82.6	73.9	43	27.07
227FB	1746333617	28.13	93.3	83.4	73.6	64.4	38.5	24.1
331MVB	2959079719	47.66	90.1	69.2	59.8	54.9	45.4	30.19
334MV	3135911493	50.51	92.9	84	77.7	72.7	59.1	43.88
337EC	1920511772	30.93	94.8	89.8	83.9	76.1	44.8	29.2
385SS	4982261824	80.25	94.6	82.5	75.3	70.9	62.2	69.4
413BT	1801203846	29.0	94.3	83.3	71.5	62.3	40.4	27.74

Table 4.7 Coverage of all exomes across samples. Total coverage of DNA bases and mean coverage depth of read (X). Along with % of bases above 1,5,10,15 and 30 covering areas

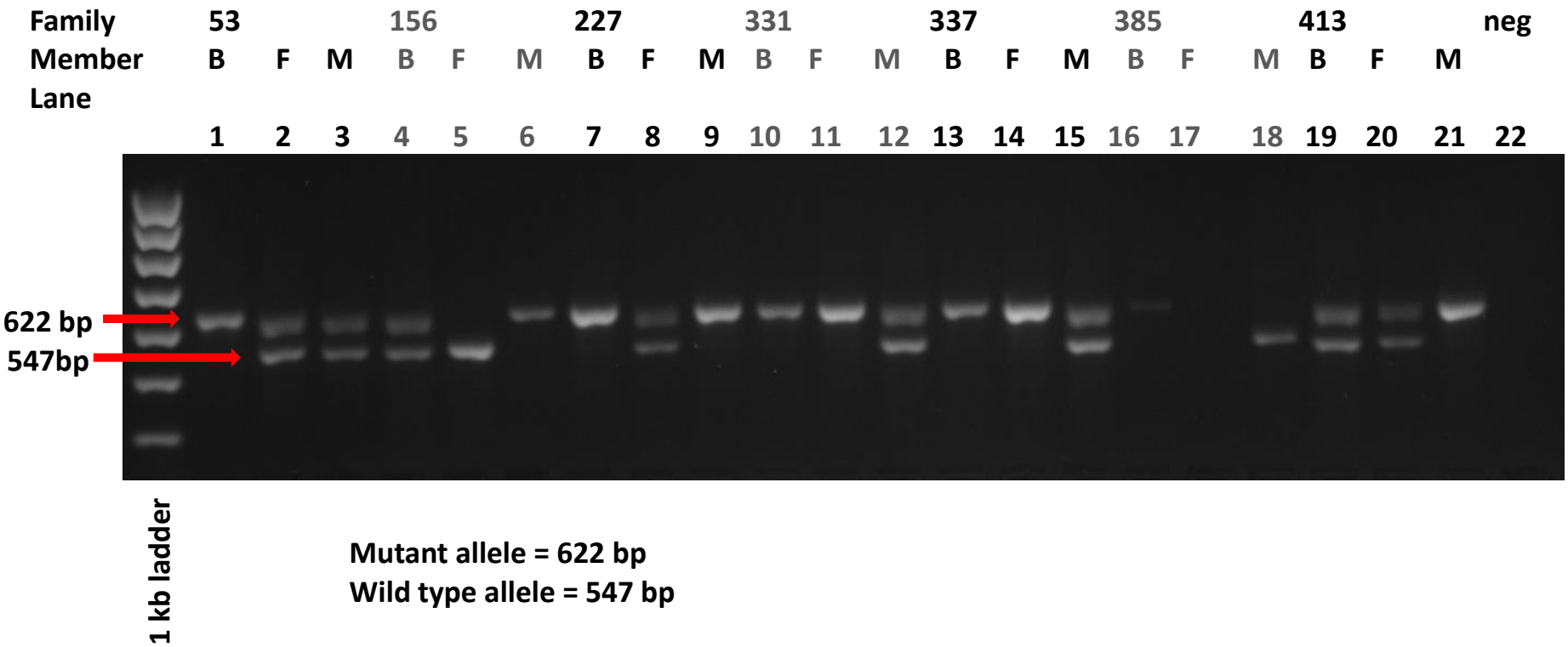
4.9 Sanger sequencing validation of novel variants

4.9.1 Investigation of exome predicted novel *KDM4B* variant

In order to confirm that the insertion c.2889_2890ins in *KDM4B* was present, Sanger sequencing was attempted on the samples. However, it was not possible to perform Sanger sequencing to validate this variant due to the length (75 bases) of the insertion. Instead PCR amplification and visualization of the PCR products on a 1% agarose gel was performed on all family members (Table 4.8). The two PCR products resulting from the different size alleles can be clearly distinguished on the gel (Figure 4.3) confirming an insertion was present.

Lane	Family	Member	
1	53	Baby	Homozygous
2		Father	Heterozygous
3		Mother	Heterozygous
4	156	Baby	Heterozygous
5		Father	Wild type
6		Mother	Heterozygous
7	227	Baby	Homozygous
8		Father	Heterozygous
9		Mother	Homozygous
10	334	Baby	Homozygous
11		Father	Homozygous
12		Mother	Heterozygous
13	337	Baby	Homozygous
14		Father	Homozygous
15		Mother	Heterozygous
16	385	Baby	? Homozygous
17		Father	Fail
18		Mother	?Wild type
19	413	Baby	Heterozygous
20		Father	Heterozygous
21		Mother	Homozygous
22		control	negative

Table 4.8 Family members and presence of *KDM4B* insertion



Variant Validated:

Sequence: GAGAGCATCACGGTGAGCTGTGGGGTGGGGCAGGGGGCGGGGGGAGGCTGGGAGCACAGCGACAACCTGTACCCC;

Figure 4.3 Agarose gel of KDM4B insertion fragment (mutant allele 622bp) as visualised on gel in families with insertion

B baby, F father, M mother on gel

4.9.2 Validation of location of *KDM4B* insertion

Whilst Sanger sequencing had proved difficult, it was possible from the gel to amplify the DNA fragments present and therefore confirm the location and sequence of the insertion. The sequence predicted from the exome data was validated but not the location of the insertion. Instead the sequence was found and confirmed to be an intronic insertion observed between c.2901+63 and c.2901+64 with c.2901 being the last coding nucleotide of exon 20 of the *KDM4B* transcript. The change was actually observed in intron 20-21.The correct description of the insertion is as follows: c.2901+63_2901+64insGAGAGCATCACGGTGAGCTGTGGGGTGGGGCAGGG GGCGGGGGGAGGCTGGGAGCACAGCGACAACCTGTACCCC

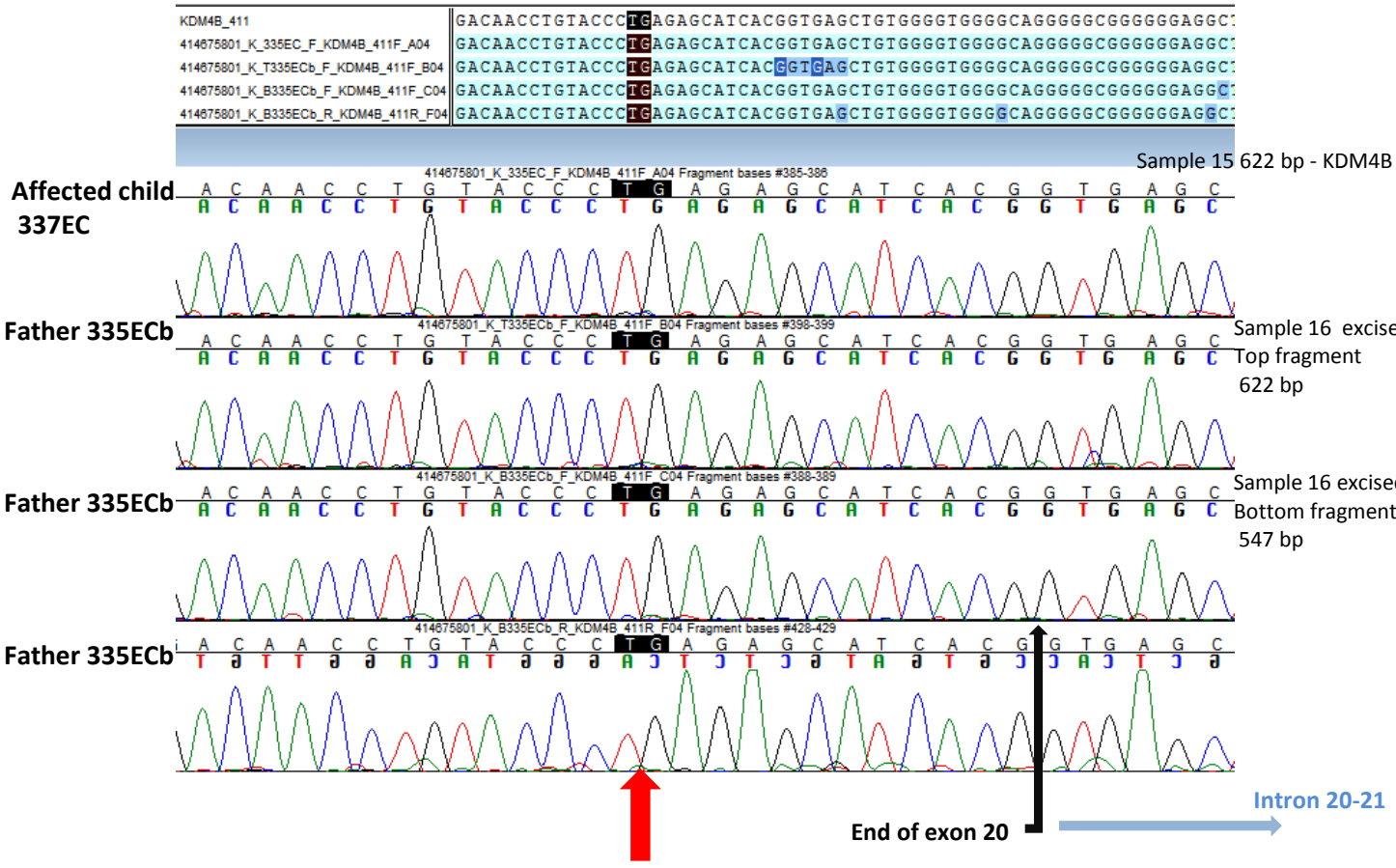


Figure 4.4 Confirms the insertion c.2889_2890ins p.963_964insESITVSCGVGQGAGGGWEHSDNLYP **IS NOT PRESENT IN EXON** in Child (337EC) in the homozygous state, **NOT PRESENT IN EXON** in the father (335ECb) in the heterozygous state (excised fragment of 622 bp)

4.9.3 Presence of *KDM4B* insertion in control populations

This intronic insertion was still not found to be present in the 1000 genome database, however, further validation was sought and the control population recruited for the study were then assessed for the presence or absence of the insertion.

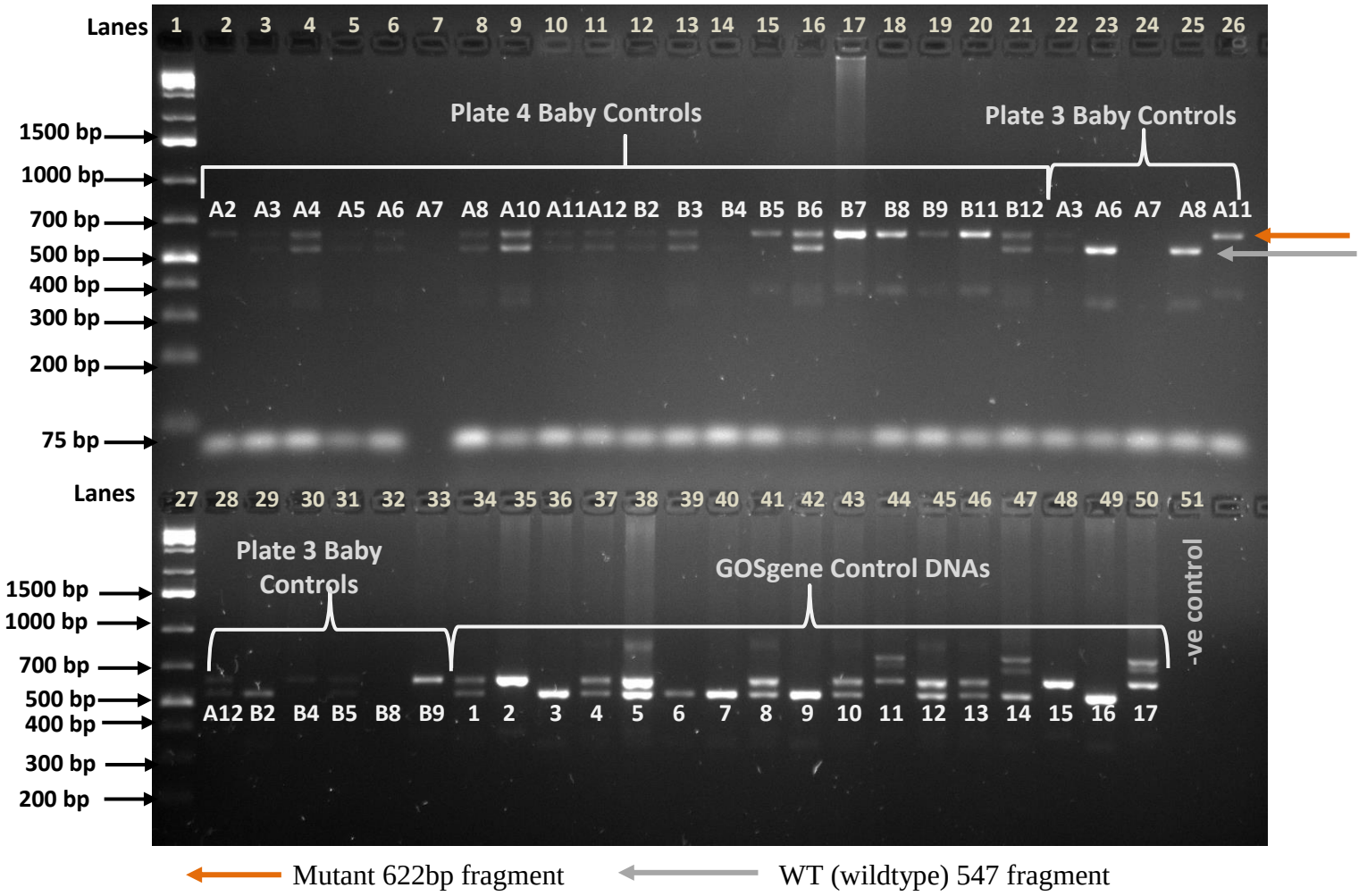


Figure 4.5 shows WT (547bp) and mutant fragments run on 1% agarose gel from 30 control samples from cohort as well as 18 GOSgene controls.1kb ladder used in lane 1 and lane 27

Figure 4.5 shows that, despite its absence in the 1000 genome database, the insertion identified was also present in control samples from the same cohort. In total 48 samples were run of which 3 failed. Of the remaining 45, 9 did not have the insertion but 22 were heterozygous and 14 were homozygous for the insertion. As the insertion was present in cases and controls I concluded that what appeared to be a novel finding in FGR families was also common in appropriately grown offspring.

4.9.4 Investigation of other potentially novel variants identified

The *TBP* insertion was found to be a variant of a recognised single triplet base CAG repeat (coding for glutamine). A single triplet base insertion had been reported in the in-house exomes analysed. An increased number of repeat CAG sequences have been noted before in relation to neurodegenerative diseases, specifically spinocerebellar ataxia (Nolte D et al., 2010).

In the FGR individuals, where a variant in *TBP* was identified by exome sequencing, insertion and deletion of the CAG repeat sequences was variable in number. In two cases there were 8 and 9 deleted glutamines identified. The remainder of samples had just a one insertion increase or were neutral in relation to the total number of CAG repeats within the stretch (Table 4.8).

The variant in *RPL14* validated but was found to be a polymorphic insertion that varied between individuals. Controls were also found to be a polymorphic in this region. ESP recognises a similar polymorphism close to our interrogated site and it, was therefore felt that we had identified a region already reported.

4.10 Sanger sequencing of rare variants in MODY genes

The two rare variants in one FGR offspring validated with Sanger sequencing. Inheritance was found to be maternal in *HNF4α* (Figure 4.6) and paternal in *HNF1α* (Figure 4.7).

Sample ID	c.189_190ins CAGCAGCAG; p.63_64ins QQQ	c.215_216insGCA; p.72_73ins Q	c.223_225del CAG; p.72_73del Q	c.223_228del CAGCAG;p.75_76del QQ	c.223_231del CAGCAGCAG;p.75_76del QQQ	c.223_259delCAGCAG CAGCAGCAGCAGCAGCAGCAG;p.75_83del QQQQQQQQQ	c.223_237delCAG CAGCAGCAGCAGCAGCAGCAG;p.75_79del QQQQQQQQQ
39NK	Hom	Clear	Clear	Clear	Hom	Clear	Clear
53AL	Clear	Hom	Clear	Clear	Clear	Clear	Clear
156BT	Clear	Hom	Clear	Clear	Clear	Clear	Clear
227FB	Clear	Hom	Hom	Clear	Clear	Clear	Clear
328LB_p3	Clear	Hom	Clear	Clear	Clear	Hom	Clear
331MVB	Clear	Clear	Hom	clear	Clear	Clear	Clear
334MVa	Clear	Hom	Hom	clear	Clear	Clear	Clear
385SS	Clear	Hom	Clear	Clear	Clear	Clear	Hom
413BT	Hom	Clear	Clear	Clear	Hom	Clear	Clear

Table 4.9 *TBP* insertion/deletion variant of triplet CAG/GCA repeat leading to variable polyglutamine repeat sequence

Hom- homozygous for allele Het- heterozygous for allele

HNF4A – Family 227

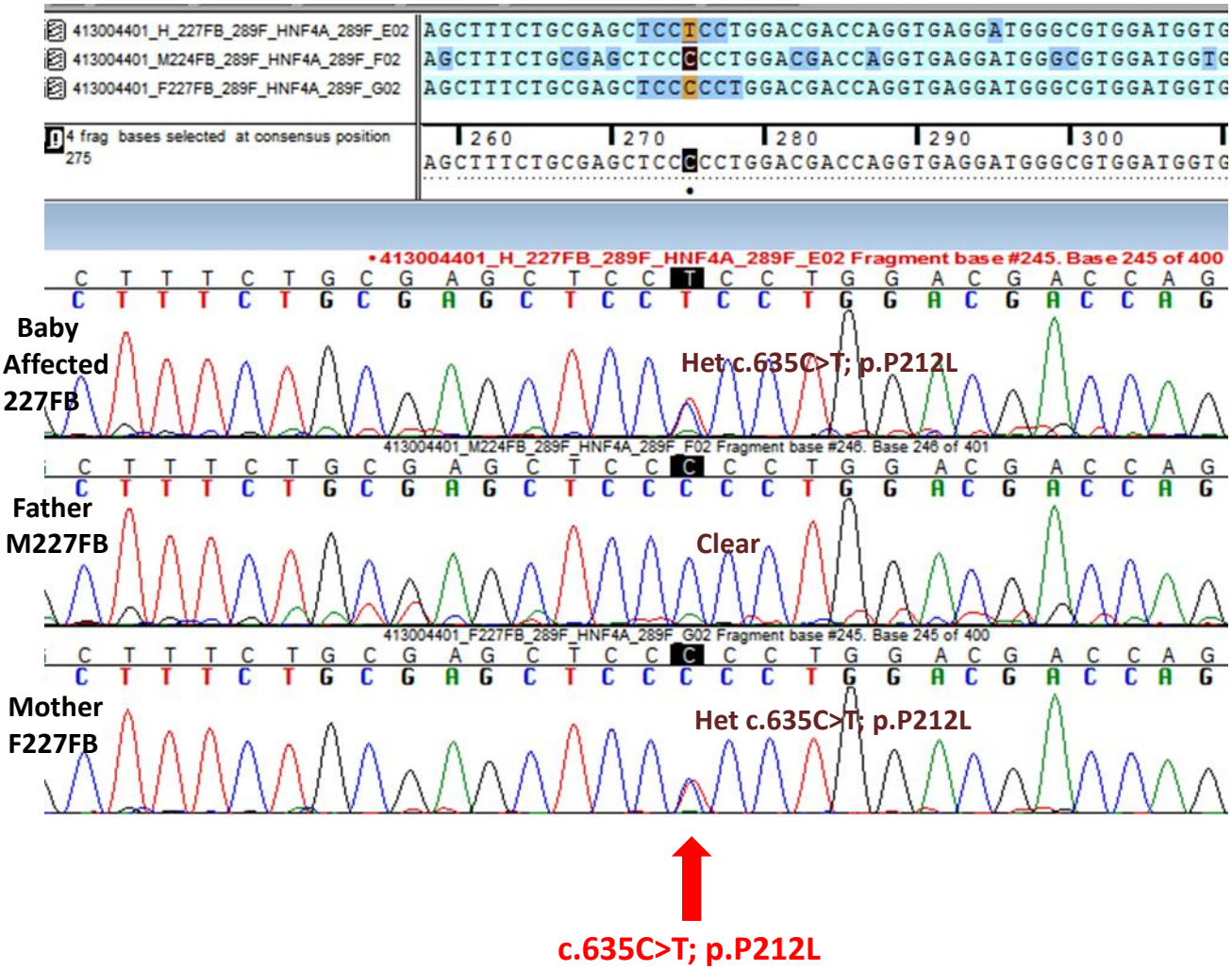


Figure 4.6 Sanger sequencing of HNF4α variant in offspring, mother and father confirming segregation with mother. Red arrow is position of change

HNF1α – Family 227

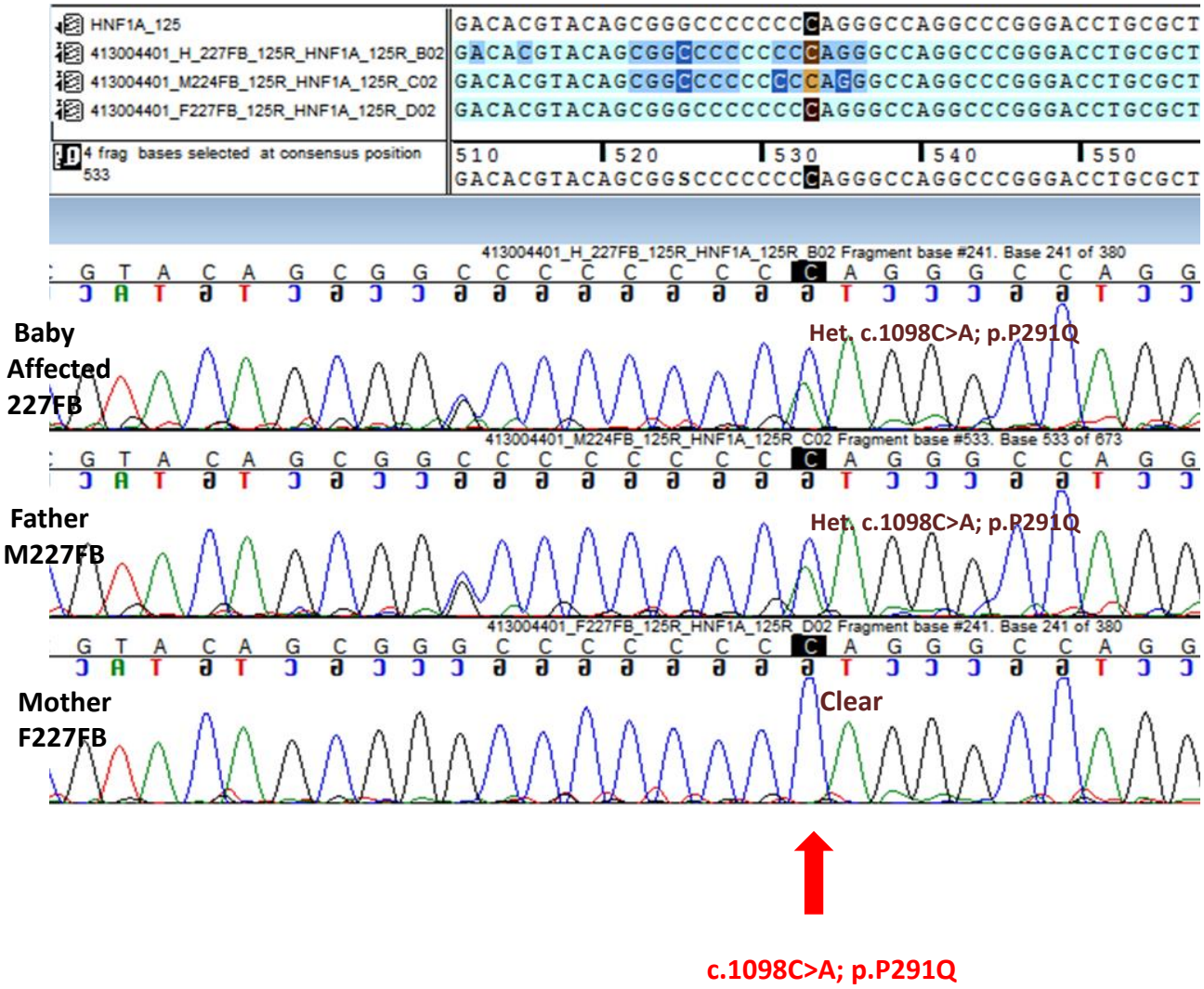


Figure 4.7 Sanger sequencing of HNF1α variant in offspring, mother and father confirming segregation with father. Red arrow is position of change

4.11 Genotyping of risk SNPs identified from exome data

From the a priori SNP list (Appendix 4 S3) those candidates present in three or more of the FGR offspring exome data were taken forward for genotyping. Of the original list, 16 SNPs (Table 4.10) present in three or more cases and this panel, along with the 2 variants in MODY associated genes (discussed in section 4.7), were taken forward to examine in the entire cohort of offspring, mothers and fathers in both cases and controls.

Assay	Gene	rs
1	*ABCC8*	757110
2	*ALDHIL1*	1127717
3	*BMP2*	235768
4	*CEL*	62576769
5	*GCKR*	1260326
6	*KCNJ11*	5215
7	*KCNJ11*	5218
8	*KCNJ11*	5219
9	*NEUROD1*	1801262
10	*PPARG*	1801282
11	*SLC2A2*	5400
12	*SLC30A8*	1326663
13	*THADA*	7578597
14	*WFS1*	734312
15	*WFS1*	1801214
16	*INSR*	1799817

Table 4.10 Panel of 16 T2 risk SNPs identified through exome sequencing of FGR offspring

4.11.1 Sample numbers

All available DNA passed QC checks as described in methods section 2.12.4. All assays worked and all SNPs provided genotypic data for analysis. Each SNP had a call rate above 97% and all were within HWE (except rs5400 on *SLC2A2* where HWE was affected by the calls in the fathers). Due to failures and original availability of DNA samples, sample numbers varied between offspring, mothers and fathers. Forty-three control offspring and 42 case offspring were genotyped; 47 control mothers and 29 case mothers were genotyped and 62 control fathers and 41 case fathers were assessed.

Baseline characteristics of the cohorts were similar to previously reported values (results chapter 3) but due to sample variation, exact values are different. Gestation,

birth weight and birth centiles were significantly different between FGR case offspring (1.7±2.2%) and controls (53.2±28.4%) when all offspring- all genders and ethnicities are compared. Offspring gender ratios were significantly different between case and controls and there was a difference in male/female birthweight within groups. However, offspring birthweight remained significantly different when adjusted for ethnicity and gender. There was no significant difference in maternal age or parity and all women in the study were nonsmokers. Although, pregnancies were identified on the presence or not of FGR, case fathers were significantly different than control fathers in several phenotypic respects including; insulin resistance, abdominal circumference and smoking when t-test statistics were applied (Table 4.11). When paternal data was adjusted for ethnicity, age and smoking status and run in a logistical regression model, phenotypic anthropomorphic measurements (blood pressure, BMI, abdominal circumference and insulin resistance) remained significantly different between male cases and controls.

	CASE	**CONTROL**	
Offspring	Mean (SD)	Mean (SD)	p value
No. samples	42	43	
Birth weight (g)	2192 (174)	3658 (125)	<0.0001
Birth centile (%)	1.7 (2.2)	53.2 (28.3)	<0.0001
Gender	17 male 25 female	28 male 15 female	0.002
Mothers			
No. samples	29	47	
Age (years)	35.8 (4.3)	35.6 (8.8)	0.4
Parity	0.26	0.2	0.53
Smoker (y)	0	0	
Fathers			
No. samples	41	62	
Age (years)	34.7 (5.2)	33.5 (5)	0.06
Smoker (y)	14/41	13/62	0.04
SBP (mmHg)	121	117	0.008
DBP (mmHg)	71	67	0.019
BMI (kg/m^2)	26.4	25.1	0.05
Abdo (cm)	94	89.4	0.01
HOMA IR	0.78	0.58	0.003

Table 4.11 Phenotypic characteristics of offspring, mothers and fathers in genotyping cohort

4.11.2 Allele Frequencies

No significant differences were identified in allele frequencies when the SNPs were assessed in all individuals with a chi^2 case versus control test (Fisher's exact in relevant samples). Frequencies were then assessed in separate case versus control mothers, fathers and offspring. In the majority of cases we were underpowered but in 2 instances a significant result was seen. With rs757110 rarer homozygote (C/C) mothers were identified as being mothers of growth restricted babies and for rs62576769 less rare homozygotes (T/T) were identified as being growth restricted offspring.
The SNP rs757110 on *ABCC8* and rs62576769 on *CEL* are both implicated in MODY, however, frequencies in case and control parents (in rs757110) and in case/control offspring frequencies for rs62576769 did not did not show any differences.

4.11.3 Regression of offspring birth weight in relation to fetal genotype

Using offspring birth weight as a continuous variable and adjusting for baby gender, SNPs identified through the candidate panel were then assessed in a regression model. No significant p values were identified (primarily as the study was underpowered) but non-significant trends with birth weight could be seen and are reported. Only individuals of Caucasian ethnicity were used for this analysis.

For the rare allele of SNPs rs62576769, rs5218, rs5400, rs734312 and rs1801214 the trend was increased birth weight.

For the rare allele of SNPs rs757110, rs1127717, rs235768, rs1260326, rs5215, rs5219, rs1801282, and rs7578597 the trend was decreased birth weight.

4.11.4 Regression of paternal logIR and BMI in relation to paternal genotype

Regression analysis was also performed for paternal phenotype in relation to HOMA insulin resistance (logIR) when adjusted for BMI, smoking status and age. One significant result was identified with carriers of the rare allele of rs5400 having a lower IR value ($p=0.04$). This argues against this SNP having an effect on an insulin resistant mechanism in FGR fathers.

Trends could be seen between paternal insulin resistance and genotype:-

- Higher IR values were observed in carriers of the rare allele of; rs1127717, rs235768, rs1260326, rs1801282, rs734312 and rs1801214.
- Lower IR value were observed in carriers of the rare allele of; rs 62576769, rs5215, rs5218, rs5219, rs1801262, rs5400 and rs7578597.

BMI was assessed in a similar fashion to insulin resistance using logistical regression and adjusted for age and smoking status. Carriers of the rare alleles of rs757110, rs5215 and rs5219 were found to be significantly associated (p=0.05) with an increased BMI. Non significant trends are also reported:-

- Increased BMI (kg/m^2) was observed in carriers of the rare allele of: rs1127717 and rs5218.
- Decreased BMI (kg/m^2) was observed in carriers of the rare allele of: rs1801262; rs7578597; rs734312 and rs1801214.

Abdominal circumference was also assessed but there were no significant associations. However, increased abdominal circumference trends were found to correlate well with increased BMI results and conversely the same allele trends were identified with decreased BMI and lower abdominal circumference.

- Higher abdominal circumference (cm): rs757110, rs62576769, rs5215, rs5219, rs5400 and rs1799817
- Lower abdominal circumference (cm): rs1801262, rs7578597, rs734312 and rs1801214

4.11.5 Evidence of a paternal-fetal genetic effect through T2DM

The same SNP on *ALDH1L1* (rs1127717) was associated with a decreasing offspring birth weight trend along with an increased BMI and IR trend in fathers. Furthermore, decreased offspring birthweight and paternal BMI was raised with SNP rs5215 and rs5219 on *KCNJ11*.

4.12 Genetic variants in MODY genes replicated in Taqman genotyping

4.12.1 Novel variants in FGR offspring in relation to rest of cohort

The two rare variants in separate offspring were identified through exome sequencing on *HNF1β* and *NOTCH 2* and were replicated through Taqman genotyping. Both offspring were heterozygotes. The parents of these offspring were the only other carriers of the same variant. No other carriers of these variants were identified in the remaining cohort.

The *NOTCH2* variant (rare "A" allele) seen in one FGR offspring (156BT). The father of this offspring was homozygous for the common "T" allele and the mother was the heterozygous carrier of the "A" allele. However, in the instance of the *HNF1β* mutation in a FGR offspring (53AL carrying the rare "T" allele), this variant had been inherited from the father (who was also heterozygous the "T" allele) with the mother being homozygous for the common "C" allele.

4.12.2 Variants present in controls as well as cases

The heterozygous variant on *HNF4α* (rs736823) was found in only one case offspring in the cohort compared with 5 control offspring. In the FGR offspring the variant was inherited from the mother, versus in the controls 4 of the 5 offspring inherited the allele from their father. This argues against this allele being influential in a father to offspring genetic risk predisposing to FGR

The variant on *KCNJ11* (rs1800467) found through exome sequencing was validated with genotyping and again found to be homozygous in one FGR offspring (39NK). Both parents were found to carry one copy of the rare "C" allele. In three further cases, this "C" allele was identified in FGR offspring who were found to be heterozygotes for this variant. In 2 of the 3 cases the variant was present in the father and not the mother. In one FGR case the parental data was not available. Two control offspring were heterozygotes for the variant, in one control this variant was present in the father and the second control father failed to genotype.

The variant on *WFS1* also validated with Taqman genotyping. The SNP (rs35031397) was present in 2 case offspring who were both heterozygotes. In one case this variant was also present in the father and in the other case parental DNA was not available. In the controls, one offspring was a heterozygote for the rare variant and in this instance transmission was through the mother.

The variant identified in *HNF1α* (rs2464196) from exome sequencing was identified in several case and control individuals in the genotyping with no discernible significant difference in allele frequency in offspring or fathers.

4.13 Discussion

4.13.1 Novel variants/ polymorphisms discovered

The nature of the variation in the variants called in *TBP* and *RPL14*, in themselves, do not represent a likely novel or known pathological change, rather are found to be polymorphisms of insertion sequences. However, the variability of the CAG repeat sequence in *TBP* requires further work to elicit whether any conformational or functional change is occurring depending on the number of consecutive glutamines present.

These variants are discussed in more detail in chapter 7 in relation to potential role in the methylation differences identified in the FGR cohort.

4.13.2 The novel insertion variant on *KDM4B* did not validate

The insertion identified on *KDM4B* was initially an interesting finding. *KDM4B* is a Jumonj domain containing lysine K specific demethylase involved in the histone code, specifically able to demethylate trimethylated lysine 9 on H3.
Along with other members of the Jumonji family it has been identified as a histone demethylase involved in transcriptional repression or activation, thereby able to have an impact on important cellular processes. However, its presence in control samples within our cohort of well grown fetuses does not support a unique functional role for this particular variant on *KDM4B* despite the fact that it has not been reported before in the 1000 genome project. It is unclear why this *KDM4B* variant was not previously reported.

4.13.3 Discovery of rare MODY variants

This work has identified rare non synonymous genetic variants in FGR offspring in genes associated with MODY. Of these variants, 2 (one in *HNF1β* and one in *NOTCH2*) are predicted to be pathogenic by Sift and Polyphen. In the case of the *HNF1β* variant, transmission was from father to offspring. Two other variants found in the same FGR offspring, (in *HNF1α* and *HNF4α*) whilst not predicted to be pathogenic have only been reported once in ESP as documented below (Table 4.12).

ch	Position	Gene	Obs allele	Geno type	other	SIFT	ESP (all) minor allele frequency	Genotype frequency	Inh
12	121432125	***HNF1α****	T	TC	het		T=1, C=13000	TT=0, **TC= 1**, CC=6500	pat
17	36059141	***HNF1β***	C	CT	het	D	C=4, T=13002	CC=0, CT=4, TT=6499	pat
20	43043289	*HNF4α**	T	TC	het		T=1, C=13005	TT=0, **TC= 1,** CC=6502	mat
11	17408831	*KCNJ11*	C	CC	hom	D	C=434, G=12552	CC=5, CG=424, GG=6064	both
1	120478125	*NOTCH2*	C	CA	het	D	C=27, A=12979	CC=0, CA=27, AA=6476	mat
4	6302816	*WFS1*	G	GC	het	D	G=49, C=12957	GG=1, GC= 424, GG=6064	pat

* inherited in same offspring

Table 4.12 Frequency of genetic variants reported as according to Exome Server Project including; chromosome, position, observed and referenced allele and MAF along with parent-of-origin (pat or mat)

Whilst WES has been used to identify novel variants in subjects with clinical symptoms of MODY before, this study is one of the first to use WES on subjects with the phenotype of FGR. This work offers novel insights into potential genetic mutations that could be influencing the development of growth restriction in a fetus. Paternal inheritance was shown to be present in one case, in *HNF1β*, a gene recognised in MODY and supports the fetal insulin hypothesis

With the evolution of whole exome sequencing, it has become possible to widen the search for candidates implicated in disease causation in relation to MODY. Recent studies have yielded the identification of the 13th MODY gene variant in *KCNJ11* through the use of WES (Bonnefond A, 2012), along with the identification of rare variants on genes (in *MADD,NOTCH2, WFS1* and *PPARG*) that have been highly implicated in MODY (Johansson S et al., 2012).

In this study, 2 FGR offspring were heterozygotes for the two rare variants identified (*HNF1β and NOTCH2*) consistent with the autosomal dominant pattern of

inheritance seen in MODY. According to recent diagnostic criteria for MODY (in relation to variants in *GCK, HNF1α* and *HNF4α* genes) novel variants must be reported as novel in the context of a relevant population and include the evidence in support of pathogenicity. This might include the absence from a large series of ethnically matched controls or MODY patients (testing of 210 normal chromosomes is necessary to achieve at least 80% power to detect a polymorphism present in 1% of the population (Ellard S et al, 2008).
Based on both predictive modelling programmes SIFT and Polyphen the variants present in *HNF1β* and *NOTCH2* have been classified as likely to have a pathogenic function. Furthermore, in the remaining genotyped cohort of DNA from 101 fathers, 83 offspring and 76 mothers (along with 21 DNA samples from families with a pregnancy affected by gestational diabetes- data not shown) these two variants were absent. These variants were also found to be absent from over 300 subjects in the UCL GOSgene exome database and in positions captured by exome sequencing of these individuals reported in ESP (Table 4.12).

Whilst *HNF1β* mutations are well established in MODY presentations, they are more commonly identified in association with renal cysts. It remains to be seen whether or not offspring go on to develop the clinical features of MODY later in life or if growth restriction might be the clinical phenotype of this variant. MODY is known to present as a heterogeneous condition with varying degrees of clinical manifestations and it may be that FGR should be considered within the classifying spectrum of physical presentations of MODY.

NOTCH2 is more controversial as a MODY gene, having previously been linked to T2DM risk (Zeggini E et al., 2008). It has recently been put forward as a MODY candidate with a rare variant having been identified through WES performed on a subject with a clinical diagnosis of MODY (Johansson S et al., 2012).

Interestingly, the same variant identified in the MODY patient is the one identified in our FGR offspring. Together these two studies provide evidence that links a genetic variant directly to both FGR and MODY.

The maternal inheritance raises the question about the mechanism through which this *NOTCH2* mutation may be acting. Other presentations of maternally inherited MODY mutations normally result in bigger babies due to the maternal

hyperglycaemic environment. Antenatal maternal glucose measurements in the FGR case revealed no such evidence of hyperglycaemia. Functional work is warranted in order to provide further evidence of the influence that this variant has on gestational glucose levels or another mechanism through which is might be acting. However, this variant in *NOTCH2* does fit the criteria to be considered a potential MODY candidate being both rare and predicted to be pathogenic

In offspring 227FB, two very rare genetic variants (one inherited from the mother and the other from the father) were identified and validated with Sanger sequencing. In isolation, they are not predicted to be pathogenic, however, it has been established that HNF4α regulates the expression of HNF1α (Fajans SS et al., 2001) and variants in both these genes could be acting synergistically to negatively affect fetal growth. Maternal antenatal and postnatal glucose levels were normal with little other evidence of glucose intolerance. The clinical presentation in this offspring was of severe growth restriction in the presence of increased paternal IR and increased abdominal circumference. This FGR phenotype contrasted with the other cases of FGR where a potential MODY variant was identified which less severe presentations than 227FB as shown previously in table 4.2.

When investigated in the entire cohort, the remaining two variants in *KCNJ11* and *WFS1* identified by WES were not isolated to FGR offspring and parents; however, this does not exclude an influence on fetal growth from these variants.

In the current study I identified a rare variant in *KCNJ11* (rs1800467), predicted to be dangerous by SIFT. One FGR offspring was found to be homozygous for this variant, confirmed through genotyping. Three further FGR cases were heterozygous for the mutation but two controls were also found to be heterozygotes. The homozygous genotype for this variant is very rare and supports a potential role in an FGR phenotype.

KCNJ11 encodes the Kir6.2 subunit of the potassium ATP channel within the pancreatic beta cell. This channel is essential in glucose-stimulated insulin secretion and is the target of sulfonylurea drugs, which enhances insulin secretion.

This gene has been suggested as a candidate potentially involved in both monogenic and T2DM diabetes pathogenesis. Activating mutations are known to cause neonatal

diabetes and a recent WES study identified a rare heterozygous mutation in *KCNJ11* (known to cause neonatal diabetes) as being causative in a clinical case of MODY (Bonnefond A et al., 2012). This finding supports the notion that *KCNJ11* mutations can be associated with a large spectrum of diabetes phenotypes. *KCNJ11* SNPs have also been one of the most robustly associated alleles with increased T2DM risk. The E23K variant of *KCNJ11* (rs5219) has been shown to be associated with T2DM with an odds ratio of 1.23 (Zeggini et al., 2008).

Heterozygous mutations in *KCNJ11* cause neonatal diabetes which is different to the finding in the FGR subject. In contrast, mutations in *ABCC8* (next to *KCNJ11*) which is also known to cause neonatal diabetes (Edghill EL et al., 2010) presents with both recessive and dominant mutations with subsequent varying phenotype. It is possible that different inheritance patterns of a mutation in *KCNJ11* could lead to different phenotypic presentations which may include growth restriction.

4.13.4 Birth Weight Associations with T2DM SNPs

Trends were identified for both increased and decreased birth weight in relation to the panel of T2DM risk SNPs. The dual association with increased and decreased birth weight is not surprising as T2DM risk alleles in mothers could also be influencing birth weight. There is strong evidence that gestational diabetes in mothers is associated with increased birth weight of offspring, along with an increased maternal risk of T2DM later in life (Bellamy L et al., 2009). The association of mothers having an increased risk of the rare allele of SNP rs757110 on *ABCC8* in association with bigger babies (although not pathologically big) would be consistent with this theory.

4.13.5 Limited evidence for T2DM risk SNPs influencing FGR pathology

My investigations have provided very limited support for T2DM risk SNPs having a pluripotent role in FGR and T2DM development, particularly in relation to a paternal inheritance pattern. The absence of substantive findings in FGR offspring in relation to these SNPs may be an issue of power but could also suggest that the influence of these SNPs may be less direct and part of a complex pathway eventuating in altered

fetal growth. Interpretations made using the birth weight regression analysis do warrant further discussion in three of the genotyped SNPs.

When the more common "T" T2DM risk allele on *KCNJ11* (rs5219) was assessed in the cohort, a trend towards lower birth weight and higher paternal BMI was observed. This would be consistent with the hypothesis that variants in *KCNJ11* could have a role to play in fetal growth and potentially in conjunction with obesity mechanisms, increase the risk of T2DM.

The T allele of *KCNJ11* SNP rs5219 has been found to be associated with an SGA cohort in New Zealand (Morgan AR et al., 2010), however, two previous studies which evaluated this variant in relation to birth weight found no association (Weedon MN et al., 2003 & Bennett AJ et al., 2008).

The findings that both a rare variant and more common T2DM risk SNP on *KCNJ11* are present in this current FGR cohort means that this gene remains of interest as a candidate in FGR. However, the small sample number genotyped for the SNP means any results need to be interpreted with caution and repeated in a bigger cohort.

The common proline (Pro) allele at position 12 on the *PPARG* gene is associated with T2DM risk (Altshuler D et al., 2000). In this study I demonstrated that the presence of the protective alanine (Ala) allele was associated with a trend of lower birth weight in offspring with 21.4% of FGR offspring having a Pro-Ala/ Ala-Ala genotype versus 11.6% of appropriately grown offspring (Table 4.13). Despite the Ala allele being associated with decreased risk of diabetes it is associated with increased BMI in overweight individuals (Masud S et al., 2003). Furthermore, this allele has previously been associated with lower birth weight (Eriksson JG et al., 2002) and with higher fasting insulin levels and insulin-to-glucose ratios in a group of young adults born SGA (Jaquet D et al., 2002).

	Pro-Pro genotype (n,%)	**Pro-Ala**	**Ala-Ala**
Baby case	33 (78.6%)	8 (19%)	1 (2.4%)
Baby control	38 (88.4%)	5 (11.6%)	
Male case	32 (80%)	7 (17.5%)	1 (2.5%)
Male control	50(82%)	11 (18%)	
Female case	23 (79.3%)	6 (21.7%)	
Female control	38 82.6%)	8 (17.4%)	

Table 4.13 Allele frequency for SNP rs 1801282 Pro12Ala polymorphism on *PPARG* for offspring, mothers and fathers in the cohort

Subjects who carried the Ala allele who were born SGA were found to have significantly increased insulin excursion under OGTT and fasting insulin-to-glucose ratio which was surprising, given the known increase in insulin sensitivity with this allele (Ek J et al., 2001). Furthermore, in this group of adults, increased BMI strongly potentiated IR in the presence of the polymorphism. This study supports the idea that genetic polymorphisms in *PPARG* can modulate insulin resistance parameters in adult subjects born SGA and provides a biological pathway to disease development. The question arises about its role in the fetus and whether environmental cues may be modulating fetal metabolic responses and influencing growth.

Common variants in the gene *WFS1* including the SNP rs734312 have been shown to confer risk of T2DM (Sandhu MS et al., 2007). The minor rare "G" allele in this cohort associated with a lower offspring birth weight trend and higher paternal insulin resistance. However, this is in contrast to the reduced risk of T2DM diabetes reported in association with the minor "G" allele. A study by Florez et al. did find that baseline minor allele homozygotes at SNP rs734312 had a higher insulinogenic index ($p = 0.02$), but interpreted this as an appropriate compensatory response to their nominally lower insulin sensitivity (Florez JC et al., 2008).

It has been shown that this allele can confer differing risk depending on the glucose tolerance status of the individual. This may go some way to explain why no clear effect was shown in the case fathers – those with abnormal glucose regulation responding differently than glucose-tolerant individuals (Sparsø T, 2008).

4.14 Summary

Rare genetic variants in genes associated with Maturity onset diabetes of the young (MODY) were discovered in fetal growth restricted offspring. Different variants, identified through WES, were present in 4 separate offspring affected by fetal growth restriction. Whilst proof of cause is still required they are highly implicated in the pathogenesis of growth restriction through their predicted pathogenicity, rareness and known biological mechanisms affecting insulin secretion and beta cell function.

The panel of risk SNPs implicated in T2DM and insulin resistance risk were not shown to have a strong influence in this analysis, but interpretations of this data is limited by the small sample numbers. Lower birth weight trends were implicated in SNPs on *KCNJ11, PPARG* and *WFS1* but alone are unlikely to be causal or have a strong direct effect on FGR pathogenesis. It is more likely these variants modulate fetal growth within normal parameters.

Support for the fetal insulin hypothesis is derived from the findings of rare variants in MODY genes. The genes described affect insulin sensitivity and beta cell function and inheritance of this trait in a fetus may lead to growth restriction. A paternal role was identified with respect to some of these cases but the overall pattern of transmission of genetic risk factors also included mothers.

Whilst genes seem to play a significant role in FGR pathogenesis this does not exclude other influences, such as DNA methylation, being involved.

Chapter 5

DNA Methylation in Growth Restricted Offspring

Chapter 5: DNA Methylation in Growth Restricted Offspring

5.1 Introduction

Fetal growth restriction (FGR) and low birth weight are a major cause of neonatal morbidity and mortality, associated with metabolic and cardiovascular disease in later life (Barker DJP et al., 1993 and Hales CN et al., 1991).
It has been proposed that epigenetic reprogramming of fetal genes could cause both growth restriction and T2DM if they persist into later life (Simmons RA et al., 2007 & Cutfield WS et al., 2007). Direct evidence that epigenetic processes operating in human pre-natal life might influence postpartum phenotype, have come from studies of altered methylation status of selected candidate genes (Godfrey et al., 2007 and Heijmans BT et al., 2006). However these findings have not been reproducible in prospectively collected umbilical cord blood samples of babies born with growth restriction or across a range of birth weights (Tobi EW et al., 2011 and Adkins RM et al., 2012).

To fulfil its growth potential, a fetus must assimilate an adequate nutritional and oxygen supply from its mother across the placenta. This process is therefore governed by a complex interplay between the feto-placental genome and its maternal environment (Brodsky D et al., 2004). In low-income nations, environmental influences on fetal growth are predominantly related to poor maternal nutrition, chronic disease or infections (Black RE et al., 2013). In high-income nations, fetal growth is more commonly compromised by poor placental development.
Inadequate invasion of the placenta with increased resistance within utero-placental arteries can be identified at the end of the first trimester and therefore, pre-dates the growth decline in the fetus. This sequence of events would suggest that poor placental development must precede the manifestations of growth-restriction in a fetus and hence investigation of the epigenome of both the placenta and fetus is warranted.

Most human studies that have investigated the role of epigenetic events in FGR pregnancies have been limited to a small subset of the thousands of genes expressed in the FGR placentae (Piedrahita et al., 2011). New technologies, such as the Infinium Human Methylation 450 Beadchip®, provide an epigenome-wide assessment of methylation on more than 450,000 CpG sites on 99% of known genes. Placentas from pre-eclamptic pregnancies have been investigated using this technology and found to have a distinct methylation profile compared with normotensive pregnancies (Blair JD et al., 2013) but within this PET cohort, pregnancies also affected by FGR could not be differentiated. A genome wide approach has shown methylation differences throughout the genome on haematopoietic CD34+ cells from umbilical cord blood of FGR (Einstein F et al., 2010). However, this study was limited to comparison of only 5 FGR cases and observed low levels (6%) of methylation difference.

Therefore, in this chapter I investigated whether genome-wide DNA methylation might be altered in umbilical cord blood and placental tissue of growth-restricted offspring. In order to discover novel epigenetic events associated with placental insufficiency; a case-control study was conducted that identified otherwise healthy women with a pregnancy affected by FGR or appropriately grown offspring (AGA). An unbiased approach was taken by investigating the whole methylome of umbilical cord blood and placentae using the Infinium Human Methylation 450 Beadchip®. Findings were validated with a previously published placenta dataset and objective bioinformatic tools were used to identify functional pathways likely to be affected by the most differentially methylated positions.

5.2 Hypothesis

Aberrant DNA methylation may predispose a fetus to growth restriction *in utero* and lead to T2DM in adulthood.

5.3 Specific Aims

1. The interrogation of DNA methylation in both placenta and cord blood in order to identify differentially methylated positions (DMPs) between growth restricted and appropriately grown offspring.
2. To assess the similarities and differences between cord blood and placenta in FGR offspring.
3. To use bioinformatics tools (GO and Marmal-aid) to provide functional evidence for DMPs identified

5.4 Methods

5.4.1 Study population

Offspring placental and cord blood samples from the case/control study undertaken at University College London Hospital (UCLH) were used. All parents gave informed consent.

At child-birth, umbilical cord blood was taken from the umbilical vein, and stored at -80^{o}C. Placental samples were taken from the surface closest to the fetus in 1cm blocks of tissue from four quadrants around the umbilical cord insertion and stored at -80°C. Gestational age, offspring gender, weight, length and delivery details were recorded to confirm antenatal diagnosis of growth restriction.

This experiment was designed to investigate DNA methylation in FGR offspring thought to be secondary to placental disorder. Women in both groups recruited were phenotypically similar in order to minimise in-utero environmental differences between the groups.

5.4.2 Methodology

High-quality genomic DNA (500ng) from paired cord blood and placenta samples was extracted and cleaned up (as described in methods section 2.12) from FGR and control AGA offspring. DNA was bisulphite converted using the EZ -96 DNA Methylation Deep-well Kit (Zymo Research, Cambridge Biosciences, UK). Samples that showed >97% bi-sulphite conversion were included in the array.

Bisulphite-converted samples were hybridised to the Illumina HumanMethylation 450Beadchip (Illumina, Inc. CA, USA) (Sandoval et al. 2011) using standard protocols (described in methods section 2.16) and arrays were scanned using the Illumina iScan SQ two-colour laser scanner.

Data produced was subject to quality control measures and probes were filtered (as described in methods section 2.15).

5.4.3 Statistical Analysis

Differentially methylated positions (DMPs) were identified using t test-based statistics in the Limma package provided by Bioconductor 2.11 (Smyth GK, 2004). Adjustment for genome-wide significance was performed using the Benjamini and Hochburg False Discovery Rate (FDR) with a threshold p value for DMPs of <0.05 (Benjamini and Hochburg, 1995).

A case/ control model was first run with all samples and then in a sub-cohort of term cases and controls only. Term samples were defined by a GA >255 days, the equivalent to a full term gestation.

The influence of gestational age on DMPs between FGR and AGA was assessed using a continuous linear regression model in the Limma package, incorporating all samples and gestational age. Results were validated with an external dataset placenta tissue from the Illumina 27k methylation array (Novakovic B et al., 2011).

To further characterize the DMPs, we utilized Marmal-aid (http://marmal-aid.org) a meta-analysis tool recently developed. In Marmal-aid, all publicly available Illumina450K data (8,531 different arrays as of April 15, 2013) is incorporated into a single repository and re-processed to enable customizable meta-analyses (Marmal-aid manuscript currently under review).

Gene Ontology (GO) pathway analysis was also performed using these DMPs. A cut off p value <0.0001 and over representation genes (from expected) of a 2 fold enrichment was taken as the criteria for GO pathways to be reported.

5.5 Results

5.5.1 Quality Control and sample removal

Data generated was subject to quality control checks as described in the methods section previously (section 2.15). Sample independent QCs of control probes (staining, extension, density) showed good performance (data not shown). Sample dependent QC's (bisulphite conversion, specificity, negative controls and non-polymorphic controls) also showed good performance but did result in removal of samples from final analysis according to whether the probes performed according to expected intensities.

5.5.1.2 Bisulphite Conversion Failure

Bisulphite converted DNA performance of Infinium I probes was assessed through signal intensity. Control probes C1, C2 and C3 assess bisulphite- converted DNA performance of the Infinium I probes in the green channel. Due to difference in extension of sequences, control probes (C1, C2 and C3) had higher signal intensity than unconverted control samples (U1, U2 and U3). The same pattern was observed in C4, C5 and C6 compared to U4, U5 and U6 in the red channel (see figure 5.1).

Due to the Infinium II design, converted and unconverted DNA is hybridised to a single probe and bisulphite conversion performance is represented by higher intensities across all control probes (II1, II2, II3 and II4) in the red channel (data not shown).

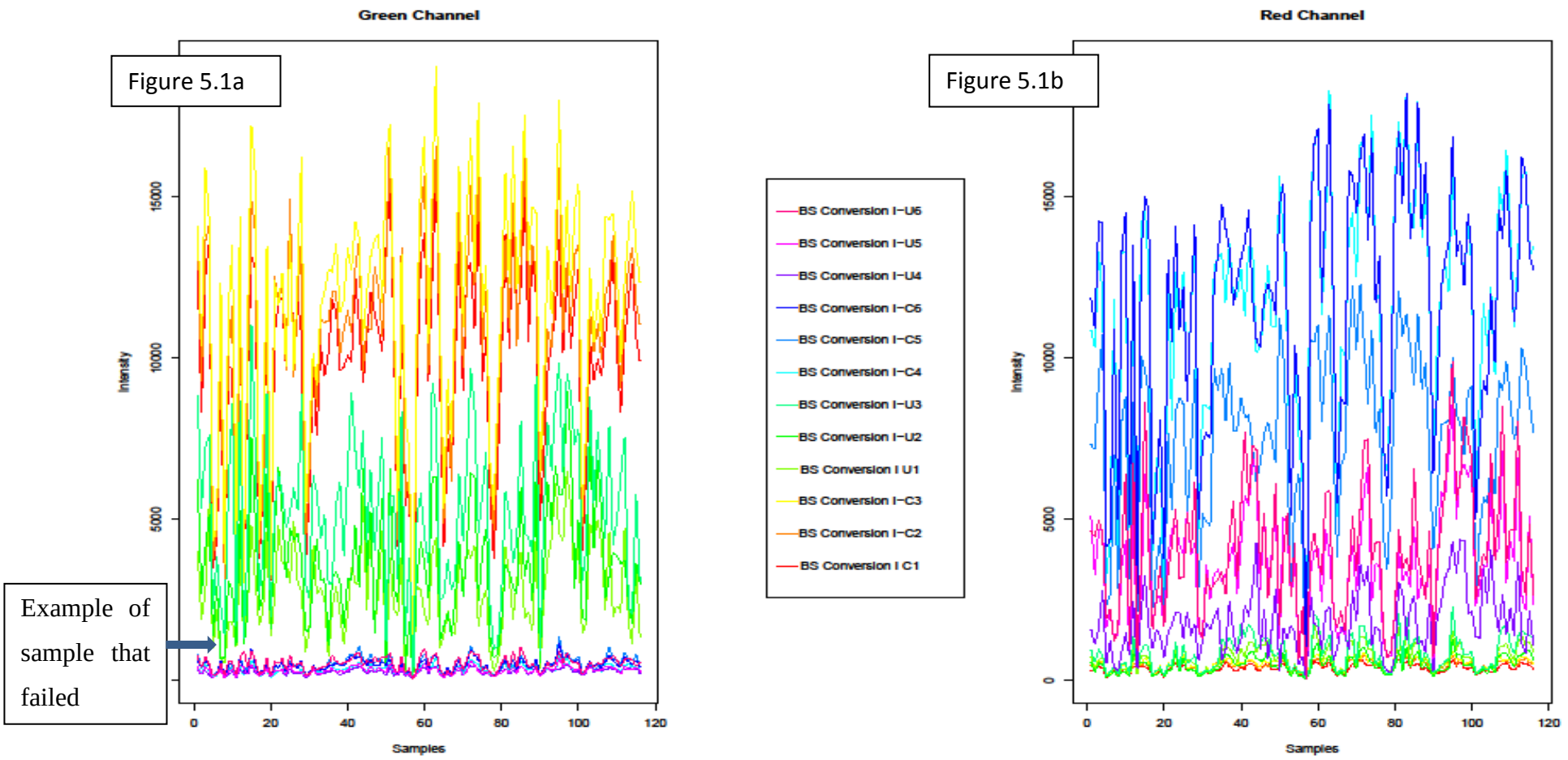


Figure 5.1 Bisulphite control probes (C1, C2 and C3) should have higher signal intensity than unconverted control samples (U1, U2 and U3) in green channel. The same high signal intensity should be seen in C4, C5 and C6 compared to U4, U5 and U6 in the red channel.

5.5.1.3 Specificity Controls - Specificity of probes I (Figure 5.2) which reflect high signal intensity of unmethylated interrogated C using PM control (reflecting a perfect A/T match) and G/T mismatch (MM control) at background signal intensity in the green channel and in Specificity II 1,2,3 should be of high intensity in the red channel (data not shown).

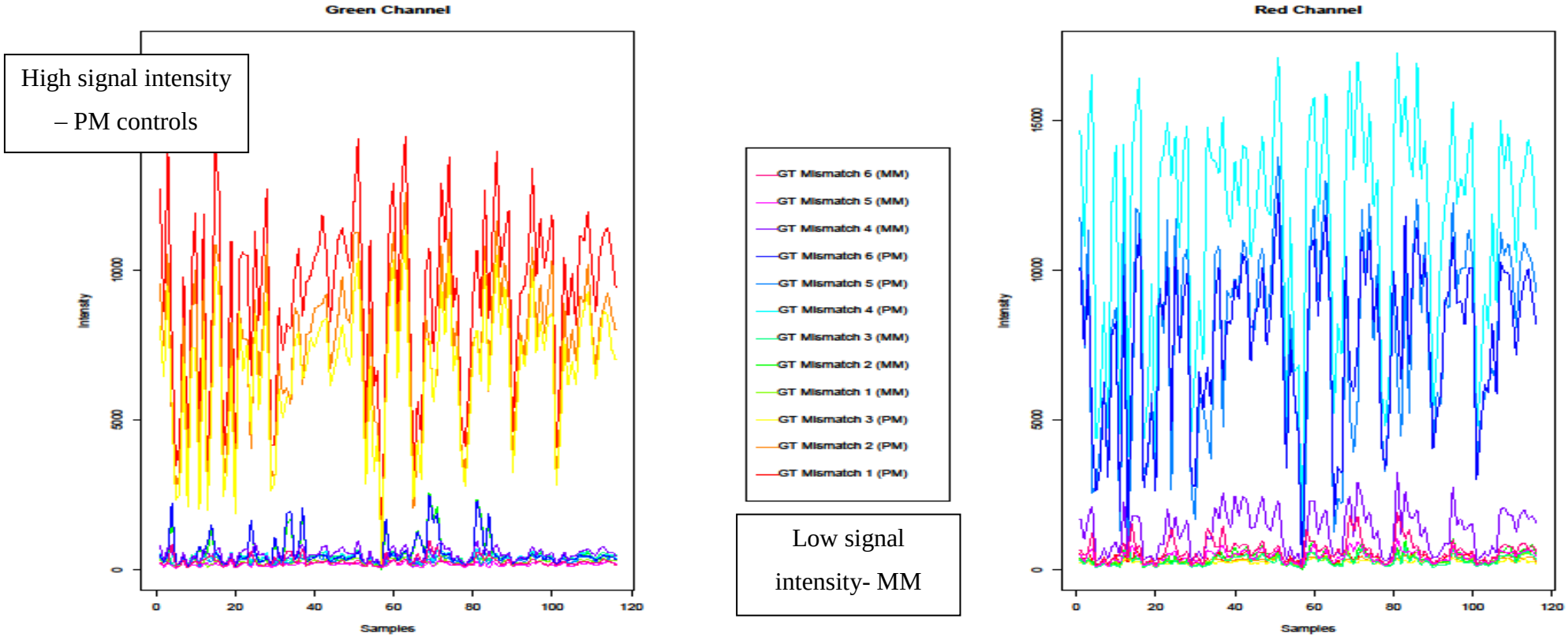


Figure 5.2 Specificity of probe I design- high signal intensity of PM controls and low intensity of MM controls passes QC.

5.5.1.4 Non polymorphic Controls

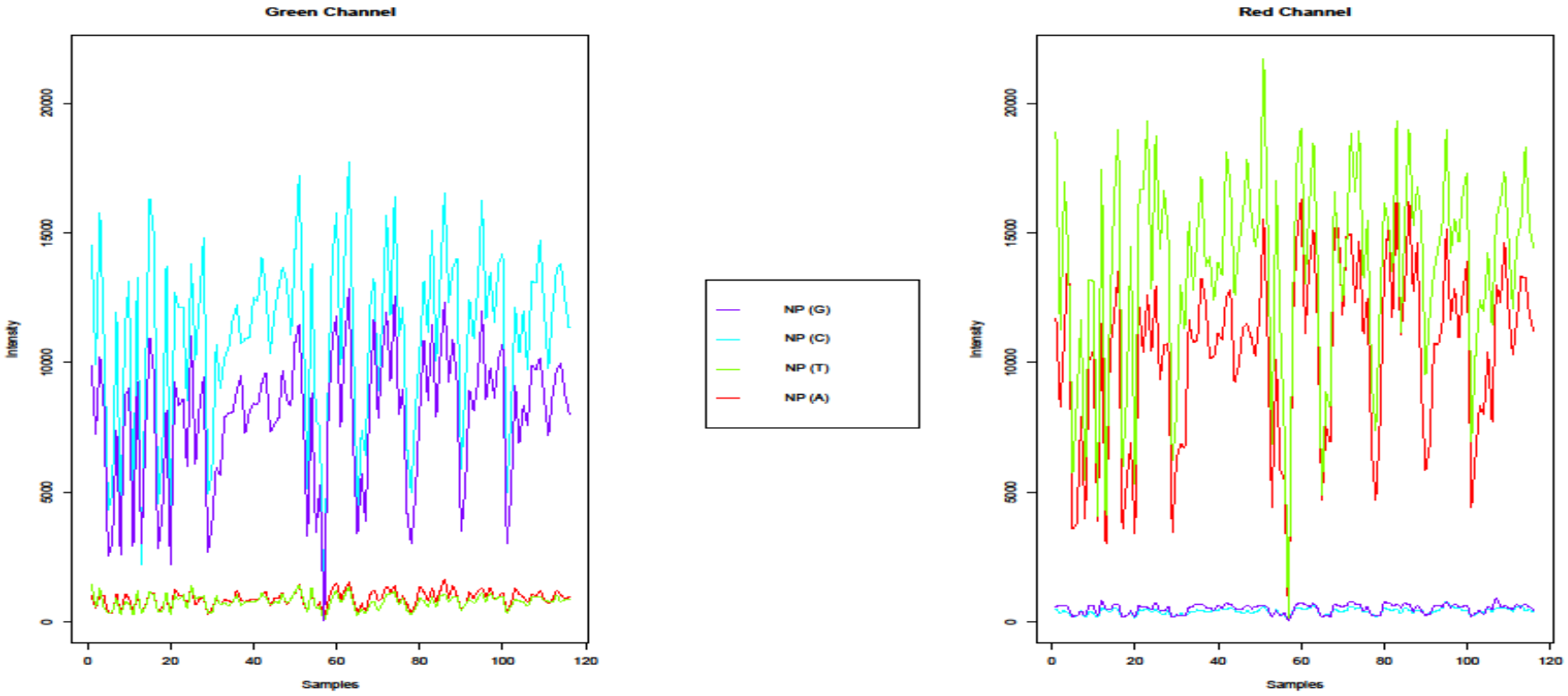


Figure 5.3 Non polymorphic controls -Non-polymorphic regions of the bisulphite-converted genome are used as a control to test assay performance by designing A, C, T and G probes within these non-polymorphic regions that can be compared across different samples. Non polymorphic probe performance of samples which should have high signal intensities (again the same samples are shown to have failed).

5.5.1.5 Negative Controls

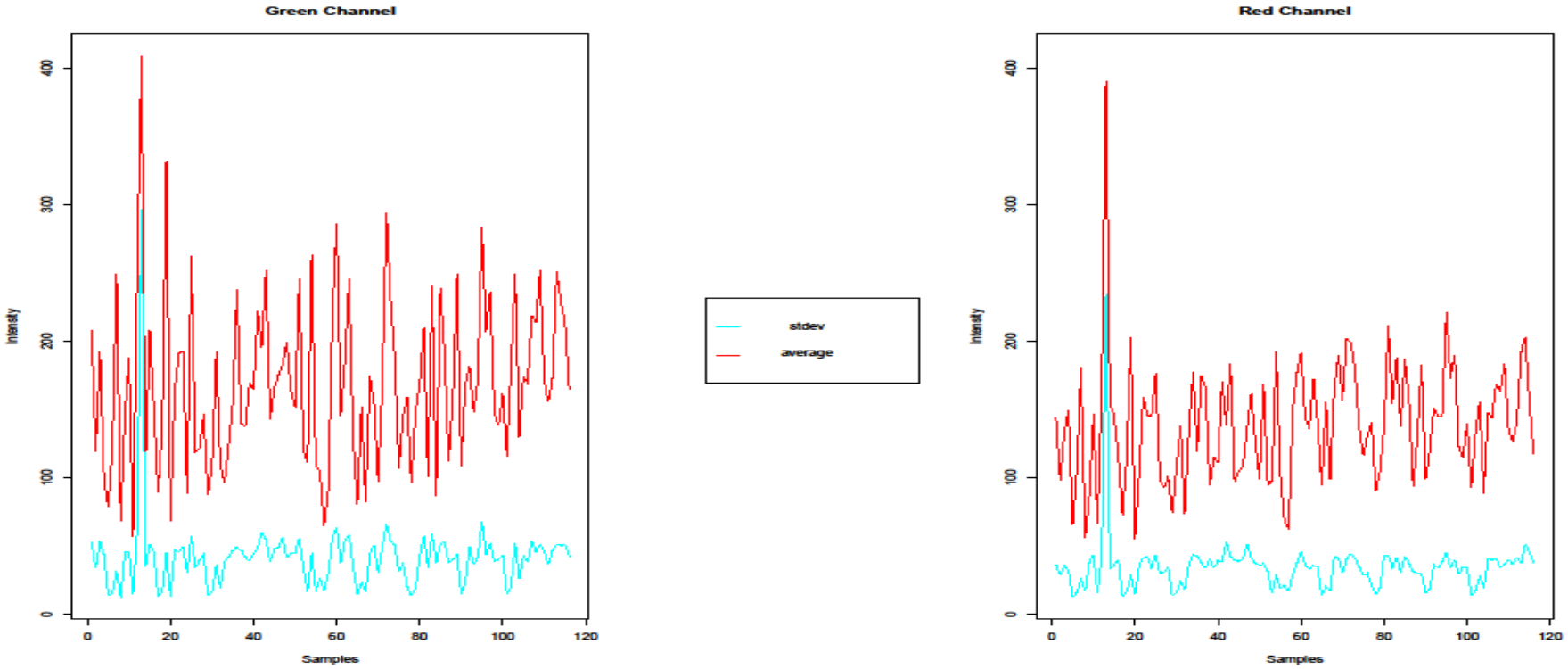


Figure 5.4 Negative Controls- Negative control probes are based on comparison with the average and standard deviation from 600 control samples available in the genome studio platform. These controls can be read in both the green and red channels

5.5.2 Intra-Array Normalisation

Initially quantile normalisation was performed separately for the red and green channels (U unmethylated and M methylated) and the outlying tissue sample was removed from further analysis. However, it was quickly identified that within the red (Figure 5.5a) and green (Figure 5.5b) channels, tissues were clustering out separately (see colour figure, cord blood in green and placenta in red) so raw data was then quantile normalised separately for each tissue (cord blood and placenta) in order to adjust for inter-array variations. Final normalisation beta value patterns were generated for each sample and are represented visually (Figure 5.6a placenta and Figure 5.6b cord blood).

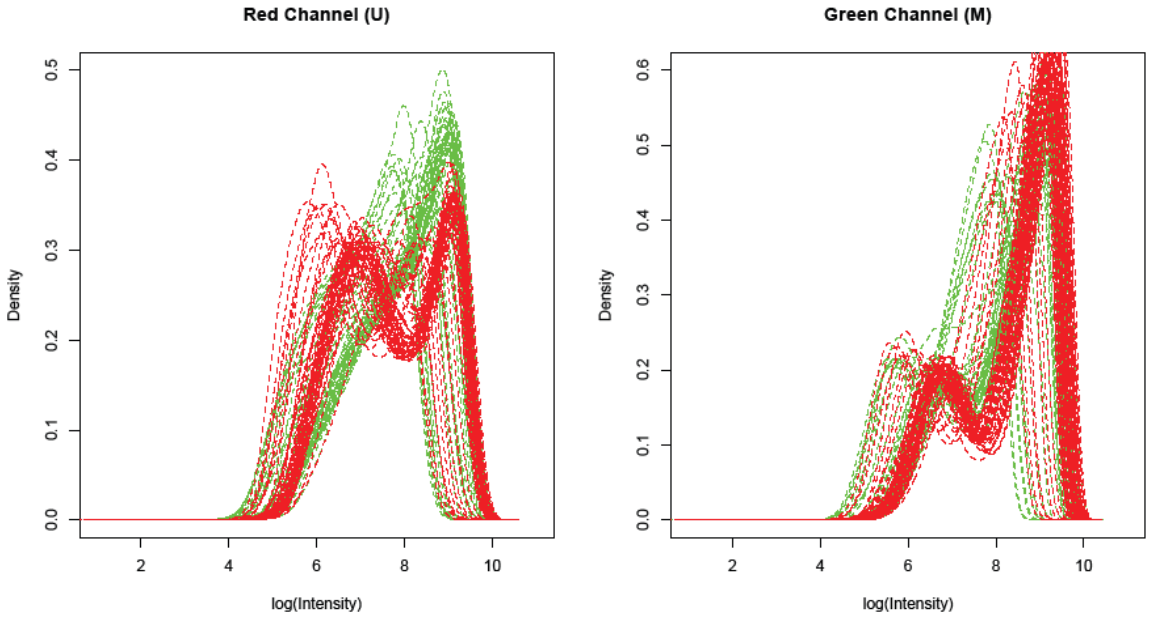
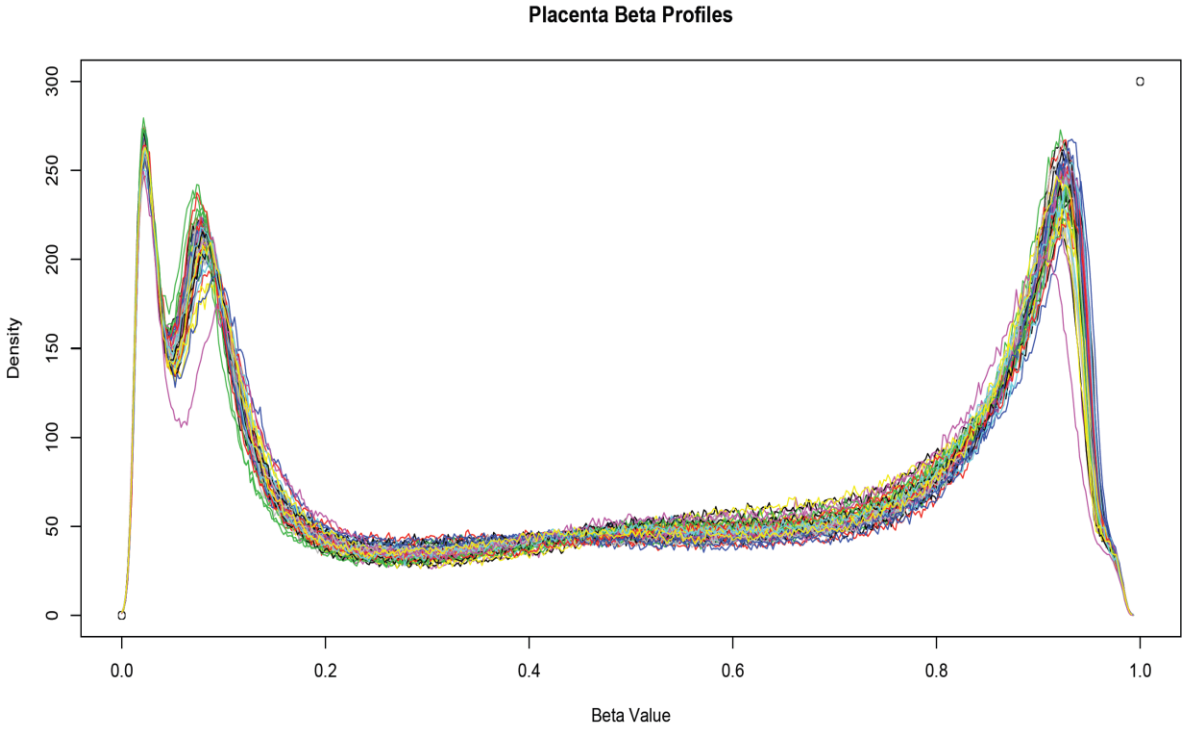


Figure 5.5 Samples represented in colour in red (5.5a) and green (5.5b) channels (cord blood green colour and placenta in red colour) it can be seen that tissues cluster out distinctly.

5.5.3 Post normalisation beta profiles

Post separate quantile normalisation a beta profile of samples was generated in order to confirm homogeneity between samples analysed (Figure 5.6a and 5.6b).



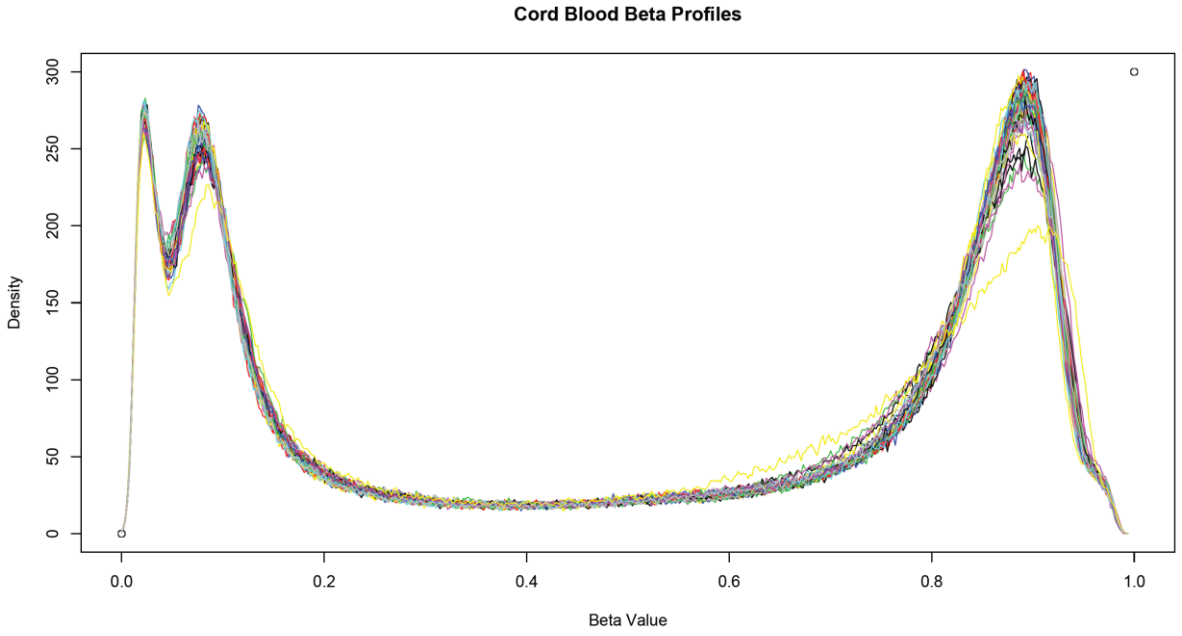


Figure 5.6a is the post normalisation beta profile in placenta and cord blood (in figure 5.6b) of each of the individual samples analysed (individually coloured).

5.5.4 Probe Filtering

Of 486,428 probes present on the array, 11,648 X and Y chromosome probes were removed; 1,927 were removed for not meeting the detection P-value cut-off (P=0.05) and 43,097 were removed due to probe cross-hybridisation. This resulted in a total of 429,756 probes, representing individual positions of methylation available for downstream analysis.

5.5.5 Sample Identification

Sample identification was confirmed using a combination of Y-chromosome methylation profiles in order to generate gender matches and interrogation of predicted single nucleotide polymorphism (SNP) probes. SNP sites were identified by selecting positions where 3 individuals had a beta value difference >0.3 from the median in both tissues.

5.5.5.1 Gender Matches

Gender assessment through Y chromosome methylation profiling (Figure 5.7) was undertaken for pairs of cord blood and placental samples in order to establish concordance which was then cross referenced with known phenotype data of fetal sex at birth. In one case a gender mismatch was identified; sample 90 (female) did not match expected gender (which was paired with 32- male) and as a result samples 32 and 90 were removed from analysis. Upon further investigation of initial DNA used it was identified that maternal DNA of the offspring had been used (labelling of tube was insufficient in this case).

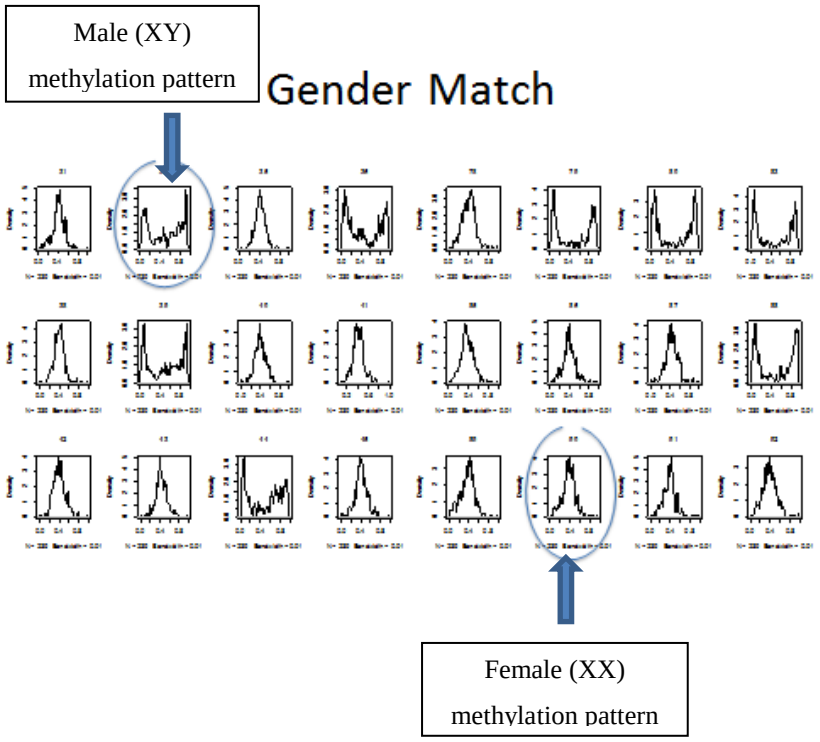


Figure 5.7 Beta methylation profile for sex chromosomes (XX or XY) successfully identifying gender. Sample 32 and 90 are highlighted showing discordant gender- in sample at top of figure the XY methylation pattern is revealed but in the bottom profile the XX female methylation pattern is present.

5.5.5.2 Genetic Cluster Analysis

Genetic cluster analysis of paired (umbilical cord blood and placental tissue) samples assessed the homogeneity of pairs in order to establish mismatches.

Identification of a SNP methylation profile led to the discovery of around 1,800 probes. These probes were then used to identify concordance of all placental and cord blood samples using a genetic clustering map (data not shown).

Cluster analysis revealed all but three sample pairs clustered out and the rest matched and identified all pairs correctly with single samples forming out-groups. These outliers were correctly identified as the non- Caucasian samples run on the array. When plotted on an MDS plot (as described in the methods section 2.15.5) the three outliers can be seen in green within the samples labelled in orange (Figure 5.8). These samples were not included in downstream analysis for this experiment.

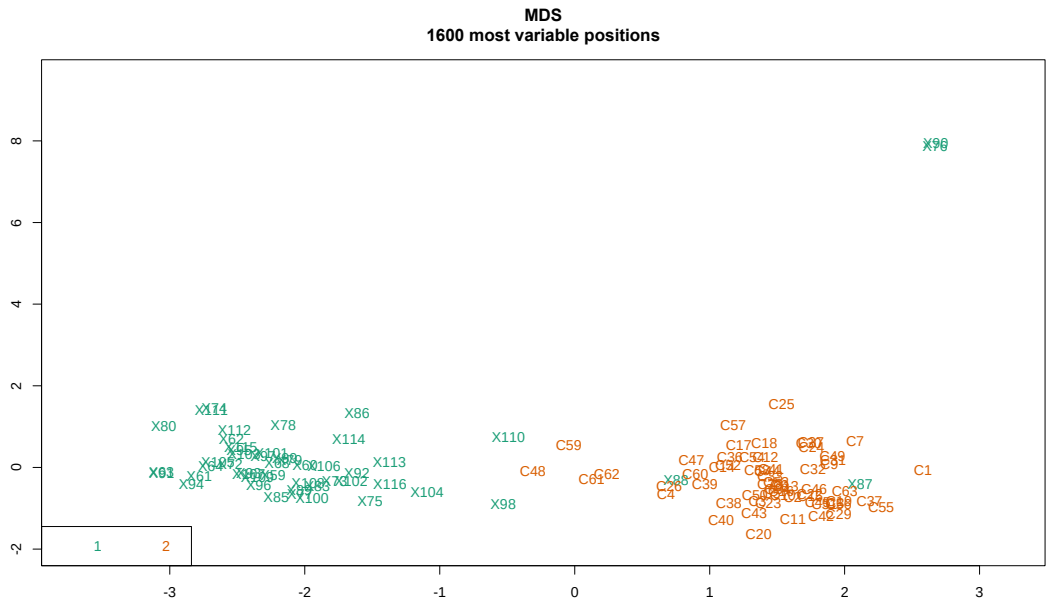


Figure 5.8 An MDS plot of the 1800 positions identified as having a SNP methylation profile. The orange samples correspond to a separate Asian (Hitman group unpublished data) cohort and the three outliers in green from the FGR cohort correspond to the three Asian cases run additionally on the array.

5.6 Final sample number analysis

5.6.1 Maternal Characteristics

Baseline maternal characteristics are described in table 5.1, for all mothers who had either placental or umbilical cord blood samples analysed (AGA mothers n=24 and FGR mothers n= 29, excluding 3 mothers of different ethnicity) which confirms that case and control maternal phenotype was similar.

The majority of women had not given birth before and this was their first child (77% primiparous mothers of AGA offspring versus 86% FGR offspring). There were no significant differences between maternal age [AGA 31.9±3.98 years versus FGR 32.8±2.98 years] or BMI [23.7±3.5 kg/m^2 AGA versus 23.4±3.6 kg/m^2 FGR].

Mode of delivery did vary for cases and controls with more FGR pregnancies ending in delivery by caesarean section (62% versus 25% in control group) but this was expected as, particularly at earlier gestational ages, elective caesarean delivery may

be the most appropriate mode of delivery. When term deliveries were assessed then the rate of caesarean section in the FGR group of mothers fell to 37% (n=7 of 19 term cord blood samples)

	FGR mothers (mean ± SD)	**Control AGA mothers (mean ± SD)**
	n=29	n= 24
Maternal age (years)	32.8 ± 2.98	31.9 ± 3.98
BMI (kgm^{-2})	23.4 ± 3.6	23.7 ± 3.5
Maternal parity	Primiparous, n=25 (86%)	Primiparous, n=21 (77%)
Mode of delivery	Vaginal, n=11 (38%) Caesarean section, n=18	Vaginal, n=18 (75%) Caesarean section, n=6

Table 5.1 Maternal baseline characteristics of appropriately grown for gestational age (AGA) offspring and growth restricted (FGR) offspring.

5.6.2 Offspring Samples

A total of 56 cord-blood and 54 placenta samples were run on the array, including 49 paired samples and 12 unpaired tissue samples (7 cord blood and 5 placental samples).

After QC and removal of three samples of different ethnicity, data from 45 umbilical blood samples (FGR n= 27 and AGA n=18) and 45 placental samples (FGR n=22 and AGA n= 23) were available for further analysis which included 44 paired samples (19 paired in cord blood and 15 in placenta) and left 7 cord blood and 12 placental samples unpaired.

Excluded samples meant that numbers and characteristics between offspring cord blood and placental samples did vary slightly (Table 5.2).

CORD BLOOD	All AGA offspring (mean ± SD)	All FGR offspring (mean ± SD)	Term FGR offspring (mean ± SD)
Number of samples	18	27	19
Birth Weight (grams)	3634 ± 369	2065 ± 770	2487 ± 420
Customised centile (%)	51.8 ± 30.5	1.9 ± 2.3	2.6 ± 2.3
Gestation (range) days	285 (272-299)	258 (202-293)	272 (256-293)
Baby Gender	Male, n=9 Female, n = 9	Male, n= 9 Female, n = 18	Male n=7 Female n=12
PLACENTA			
Number of samples	23	22	15
Birth Weight (grams)	3614 ± 361	2011 ± 710	2428 ± 342
Customised centile (%)	47.6 ± 28.6	1.8 ± 2.3	2.3 ± 2.0
Gestation (range) days	285 (266-300)	257 (202-286)	272 (256-286)
Baby Gender	Male, n=11 Female, n = 12	Male, n= 9 Female, n = 13	Male, n=6 Female, n = 9

Table 5.2 Offspring characteristics of all cord blood samples and placental samples analysed as well as comparison with the term FGR samples analysed (by inclusion criteria all AGA samples term gestation so no difference in values than the figures presented for All AGA samples)

In all cord blood samples, AGA offspring samples (n=18) had a mean customised birth centile of 51.8±30.5% and birth weight of 3634±369g versus FGR offspring (n=27) whose mean customised birth centile was 1.9±2.3% with a birth weight mean of 2065±770g. All offspring met the inclusion criteria for the study and AGA and FGR offspring customised centiles were statistically significantly different (p=0.001). There were more growth restricted females (67%) than in the AGA offspring (50%).

Of umbilical cord blood samples 19 of 27 FGR offspring were born at a gestation of 255 days or more which we defined as term gestation. Recruitment criteria for control offspring meant that all 18 AGA cord blood samples were 255 days or more gestation. For the term comparison these cord blood samples were used.
The cut-off of more than 255 days was taken based on a pervious study of methylation differences seen in a term cohort (Banister et al., 2011) who defined a term cohort as more than 36 weeks and also from investigation of methylation profiles which supported a natural cut off around this gestational age. In term cord blood samples, data was the same for AGA offspring but for FGR offspring (n= 19) their mean customised birth centile increased to 2.6±2.3 % and birth weight increased to 2487±420g. Again in growth restricted offspring there were more females (63%) than AGA offspring. All statistical differences were maintained between AGA and FGR offspring in the term samples (data not shown).

In all placental samples AGA offspring samples (n=23) had a mean customised birth centile of 47.6±28.6% and birth weight of 3614±361g versus FGR offspring (n=22) whose mean customised birth centile was 1.8±2.3% with a birth weight mean of 2011±710g. All offspring met the inclusion criteria for the study and AGA and FGR offspring customised centiles were statistically significantly different (p=0.001). There were more growth restricted females (59%) than in the AGA offspring (52%).

Of placental samples 15 of 22 FGR offspring were born at more than 255 days. Again in term placental samples, data was the same for AGA offspring but for FGR offspring (n= 15) their mean customised birth centile increased to 2.3±2.0% and birth weight increased to 2428±342g. Again in growth restricted offspring there were more females (60%) than AGA offspring. All statistical differences were maintained between AGA and FGR offspring in the term samples (data not shown).

5.7 Data Analysis

Firstly an MDS plot was plotted as a means of visualising the data. This was done for all the probes available (429,756). Cord blood samples can be seen clustering separately to placental samples (Figure 5.9).

MDS plot were then generated for both cord blood and placental tissue with AGA and FGR samples annotated separately (Figure 5.10a & Figure 5.10b). When tissues plotted separately then it becomes more obvious that cases (green) and controls (orange) are behaving differently in cord blood.

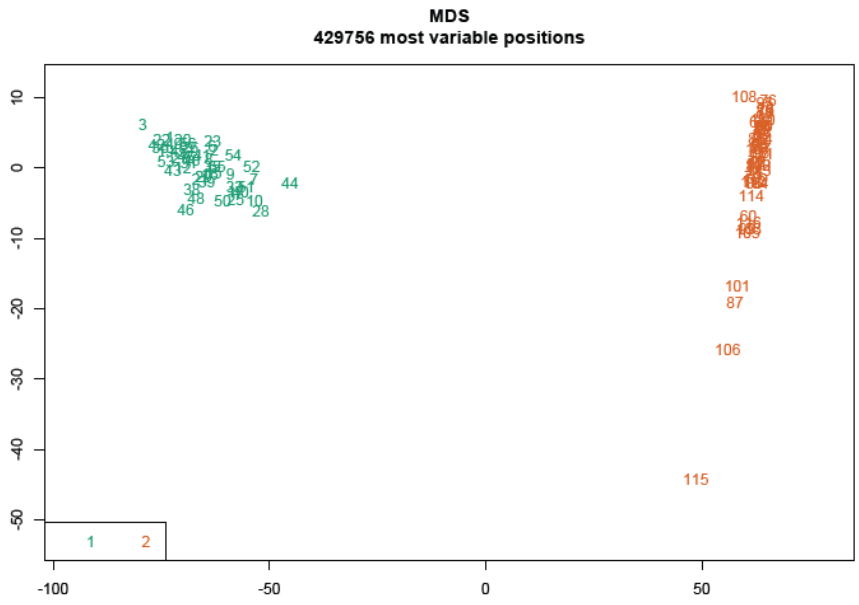


Figure 5.9 MDS plot of cord blood in green versus placental samples in orange which show two tissues cluster out separately.

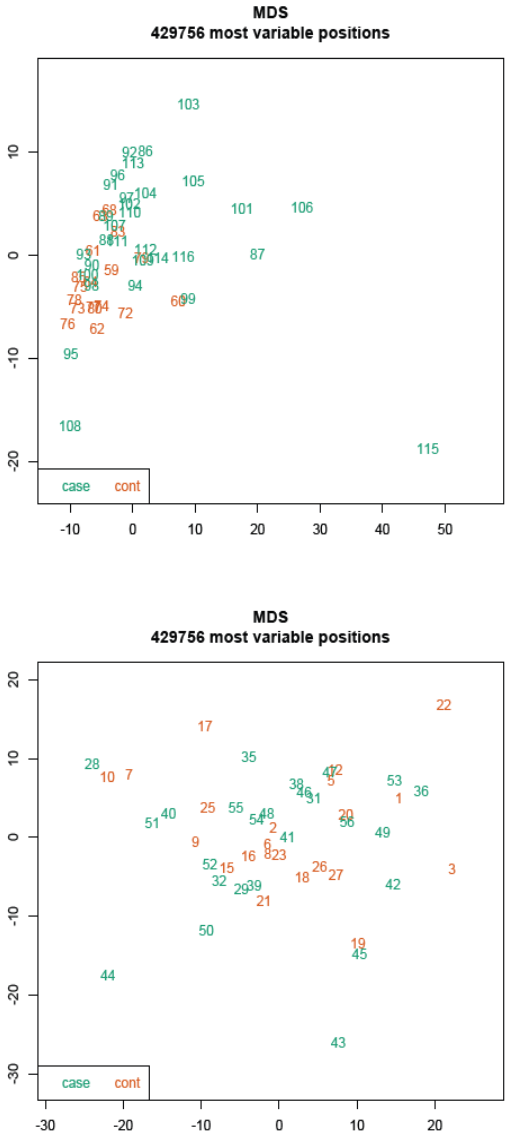


Figure 5.10 When tissues separated out in cord blood samples AGA (orange) and FGR (green) samples are seen to separate out (Figure 5.10a above). This pattern is not as obvious in placental tissue (Figure 5.10b).

5.7.1 Analysis pipeline

As discussed previously there were pre-term FGR offspring within the samples processed. All AGA offspring were born at term and hence did not include a pre-term cohort. The approach to the analysis, therefore, took into account the possible effect of gestational age on methylation change.

Analysis was performed separately for the tissues in the following ways:

- A case/control call between all FGR and AGA offspring
- A term (>255 days) case/control call between FGR and AGA offspring
- Age matched term FGR offspring versus AGA offspring in cord blood to validate findings
- A gestational age model using all samples
- Validation of placental tissue gestational age methylation
- Overlap between FGR and AGA call and gestational age
- Assessment of paternal IR and offspring methylation

5.8 Differentially methylated positions (DMPs) in growth restricted versus appropriately grown offspring

In all FGR offspring, 553 differently methylated positions (DMPs) in cord blood and 453 DMPs in placenta reached genome significance (adj. p <0.05) compared with all AGA offspring. Of these hits, in placental tissue, differences were predominantly hypomethylated in FGR samples (Figure 5.11 is an example) but hypermethylated in FGR cord blood samples.

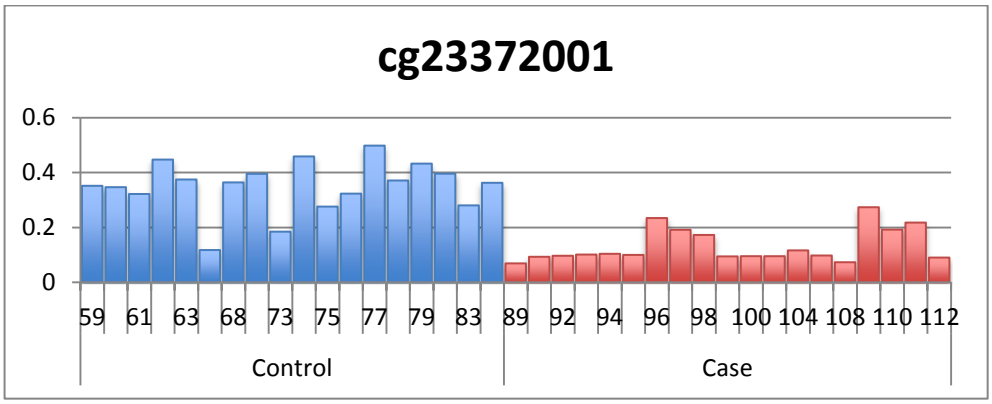


Figure 5.11 Example of a DMP called as hypomethylated in case FGR offspring in placenta.

5.9 Differentially Methylated Positions in umbilical cord blood from fetal growth restricted offspring at term

From baseline data it was clear that gestational age varied in the FGR cohort with the presence of pre-term offspring in this group. As no AGA offspring were born pre-term (<255 days), gestational age was not controlled for in this original case/control call. In order to minimise the confounding effect of gestational age in the next analysis, just FGR cases (n=19) and AGA controls (n=18) that delivered at term (≥ 255 days) were compared.

This analysis of genome-wide methylation differences yielded 839 DMPs in the cord blood of FGR offspring, with the majority (86%) being hypermethylated (Figure 5.12) in FGR compared to AGA offspring. Only 67% were common to the DMPs identified in the case/control call between all samples, suggesting that gestational age has independent effects on methylation in the initial call.

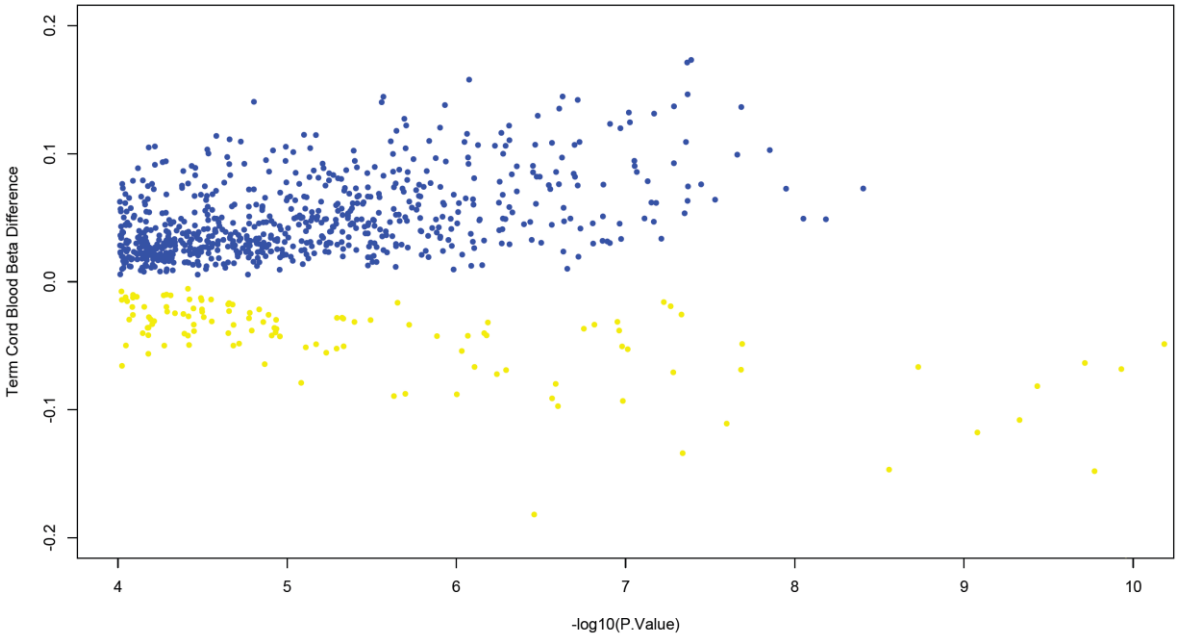


Figure 5.12 Volcano plot showing differentially methylated positions in cord blood between growth restricted and appropriately grown offspring at term gestation. Hypermethylated positions are shown in blue and hypomethylated positions in yellow.

In order to identify the DMPs with a high likelihood of a functional effect, these 839 DMPs were then stratified according to the degree of methylation difference that

they exhibited. Of the 839 DMPs, 304 had a beta methylation difference of >5% and 53 DMPs had a beta methylation value difference of >10% (Table 5.3). The majority of DMPs with >10% beta methylation difference were hypermethylated (41/53) (Table 5.3). Of the 839 DMPs that were identified on Ref-seq annotated genes, 658 DMPs were within 1000bp of each other and 25 DMPs co-methylated in the same direction. Co-methylation in the same direction was exhibited in 8/53 of the most differentially methylated hits including 4 co-methylated DMPs within the gene *C5orf39* and 3 in *FOXP1* and *NRN1*. Two DMPs were within 5 base pairs of each other on *RIOK3,* with beta methylation values of 11 and 17%.

Hypomethylated positions

ID	CHR	MAPINFO	UCSC_RefGene	Beta value	p value
cg23372001	6	27791640	*HIST1H4J*	-0.22	1.10E-10
cg26758857	22	36649135	*APOL1*	-0.18	3.47E-07
cg05533953	4	331685	*ZNF141*	-0.15	1.69E-10
cg26542283	5	43040505	*C5orf39*	-0.15	2.77E-09
cg07875360	5	1801344	*NDUFS6;MRPL36*	-0.13	4.60E-08
ch.20.327316F	20	15628644	*MACROD2*	-0.12	8.34E-10
cg01458605	3	147128679	*ZIC1*	-0.11	4.69E-10
cg04410715	19	39283334		-0.11	2.53E-08

Table 5.3a Hypomethylated positions with beta value >10% in cord blood from FGR compared with AGA offspring, according to cg identifying site, gene identifier, chromosome number, chromosome position and beta value

ID	CHR	MAPINFO	UCSC_RefGene	Beta value	P.Value
cg25418748	5	178977236	*RUFY1*	0.17	4.10E-08
cg20796556	18	21032774	*RIOK3*	0.17	4.32E-08
cg17724814	1	155881358	*RIT1*	0.16	8.41E-07
cg07979357	19	14142353	*IL27RA*	0.15	4.30E-08
cg01769037	6	15246613	*JARID2*	0.14	2.36E-07
cg22256960	15	77711686		0.14	2.70E-06
cg19857457	18	47018947	*RPL17;SNORD58A*	0.14	1.92E-07
cg27069263	2	55278118	*RTN4*	0.14	1.57E-05
cg18261223	5	32711517	*NPR3*	0.14	2.76E-06
cg09435090	4	17513980	*QDPR*	0.14	1.17E-06
cg20684973	10	105127632	*TAF5*	0.14	5.18E-08
cg00565558	7	6120396		0.14	2.07E-08
cg23802518	10	80827482	*ZMIZ1*	0.14	2.47E-07
cg18036763	22	45404910	*PHF21B*	0.13	9.57E-08
cg08965527	16	84178213	*HSDL1;LRRC50*	0.13	6.78E-08
cg10861135	9	14345348		0.13	3.31E-07
cg00510787	6	151772946	*C6orf211;RMND1*	0.13	2.03E-06
cg09747578	1	46769076	*LRRC41;UQCRH*	0.12	9.43E-08
cg18022926	10	70093071	*PBLD;HNRNPH3*	0.12	1.24E-07
cg10013169	6	5997028		0.12	1.98E-06
cg22276571	20	17549599	*DSTN;DSTN*	0.12	4.88E-07
cg16595484	3	122512170	*HSPBAP1*	0.12	1.25E-06
cg03393426	10	102986686		0.12	1.07E-07
cg04921814	1	145575587	*PIAS3*	0.12	2.26E-06
cg21935083	5	131892314	*RAD50*	0.12	5.42E-07
cg23653187	22	44319257	*PNPLA3*	0.12	8.65E-07
cg17344340	22	46044084		0.11	7.95E-06
cg07772516	15	52107742	*TMOD2;TMOD2*	0.11	6.76E-06
cg08779777	7	106505772	*PIK3CG*	0.11	2.62E-05
cg21860429	6	105389544		0.11	2.19E-05
cg14555127	7	35841578	*SEPT7;*	0.11	4.86E-07
cg04609859	17	46655736	*HOXB4*	0.11	1.45E-06
cg21312090	8	116681727	*TRPS1*	0.11	2.44E-06
cg17373554	5	72594735		0.11	1.88E-05
cg05489143	3	40498640	*RPL14*	0.11	8.96E-07
cg04005701	17	1733641	*RPA1;SMYD4*	0.11	1.87E-07
cg20760063	17	41277580	*NBR2;BRCA1*	0.11	4.40E-08
cg06361531	16	68057779	*DUS2L;DDX28*	0.11	2.72E-07
cg01013600	19	49141177	*DBP;SEC1*	0.11	3.42E-07
cg25845597	6	27841122	*HIST1H4L;HIST1H3I*	0.11	2.00E-07
cg14036868	2	38604442	*ATL2*	0.11	7.45E-07
cg26220528	12	57623348	*SHMT2*	0.11	5.92E-07
cg01147107	1	47696505	*TAL1*	0.11	5.12E-07
cg16942681	18	21032779	*RIOK3*	0.11	6.05E-05
cg08215925	3	71633214	*FOXP1*	0.11	1.02E-05

Table 5.3b Hypermethylated positions with beta value >10% in cord blood from FGR compared with AGA offspring

5.10 Term Placenta FGR versus AGA

When the same term analysis was performed on placental tissue (AGA n= 23, FGR n=15) genome wide significant DMPs were not identified. The top 1000 non-significant hits in the placenta were then compared with the significant term cord blood DMPs in order to make an assessment of whether the placenta was simply underpowered to identify differences. There was zero overlap between positions, suggesting that the phenomenon seen in cord blood was not present in placental tissue in term growth restricted offspring rather than the placental samples being underpowered to identify it.

When p value distribution for the two different tissues is called in cord blood enrichment can be seen but this is not the case for placental tissue where DMPs are not enriched above background noise (Figures 5.13a and 5.13b)

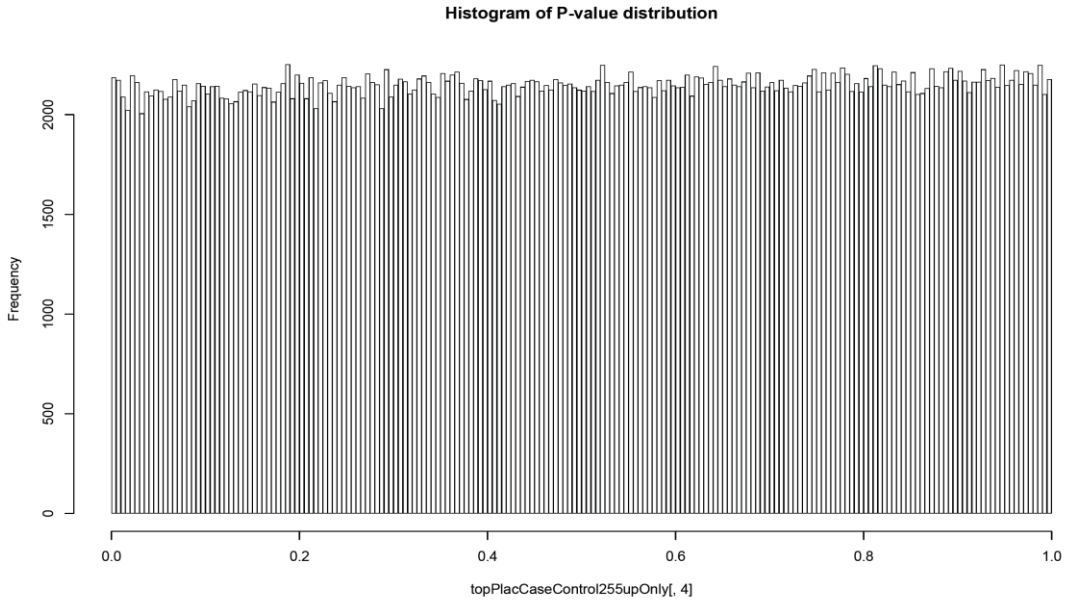


Figure 5.13a. Representation of p value distribution across placental term samples

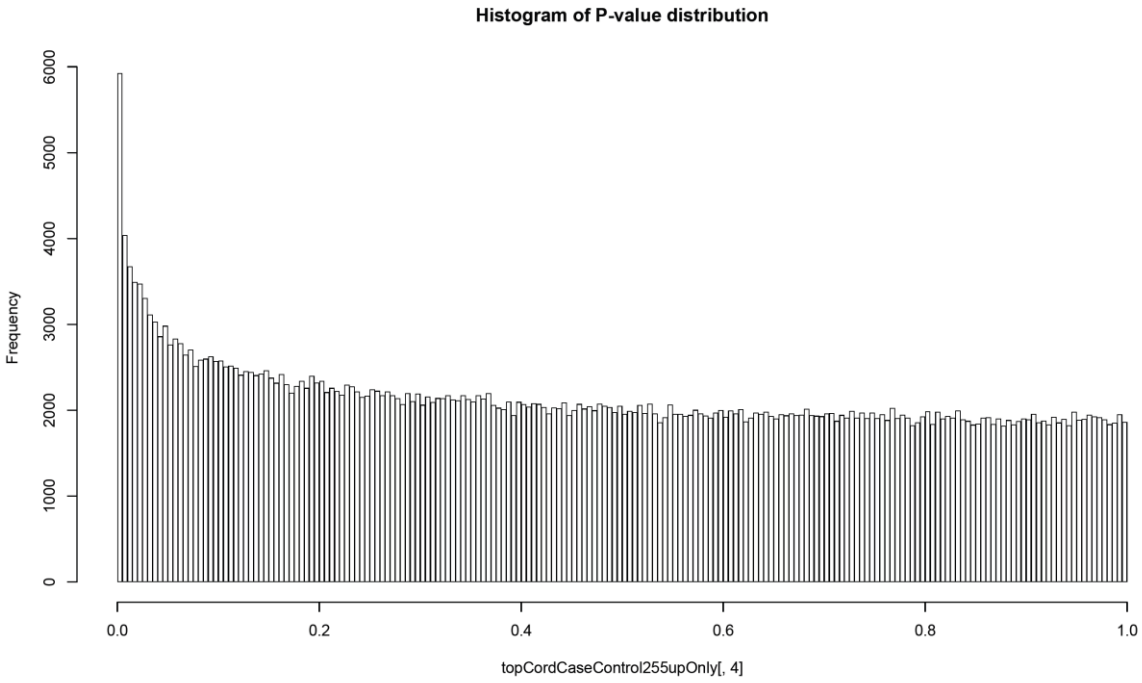


Figure 5.13b Representation of p value distribution across cord blood term samples shows enrichment

Before the linear model using gestational age was run, models using birth weight centiles were also run. Using this variable failed to yield genome wide significant DMPs, as clustering together of centiles meant significance was lost. However, of the top 2000 non-significant results from the birth centile call when placenta and cord blood were compared only 21 probes are common between the two - just above what would be expected by chance. Again this supports the idea that methylation profiles between cord blood and placenta are different in growth restriction.

5.11 Differentially methylated positions driven by gestational age

To elucidate whether the 453 placenta DMPs identified in the first analysis were driven by differences in gestational age between experimental groups, a linear model was run on all samples using gestational age as a continuous variable. Across the gestational age range (189 days to 300 days) in the placenta there were 20,893 DMPs. In cord blood the same linear model yielded 76,249 DMPs (adjusted p<

0.05). Of these DMPs, 7,041 were common to both placenta and cord blood, which was enriched above background probability. By chance you would expect a distribution with mean ~3500 and range from 3300-3700 DMPs.

DMPs in placenta became more methylated with advancing gestation which was also a characteristic of cord blood DMPs. However, the majority of DMPs, irrespective of tissue type, are hypomethylated (93% and 68% in placenta and cord blood, respectively) as the difference is likely to be driven by the preterm samples.

Individual positions were plotted for all samples against gestational age. The top placental DMPs were plotted in this way. With advancing gestation an increasing trend in methylation can be seen in the top 20 positions (Figure 5.14). The clustering of samples towards term gestation is due to the numbers of control samples (all term gestation) along with a significant proportion of FGR cases being >255days in placental samples (n=15). However, despite the small number of preterm samples the trend seems consistent with increasing methylation associated with advancing gestation across the most significant DMPs.

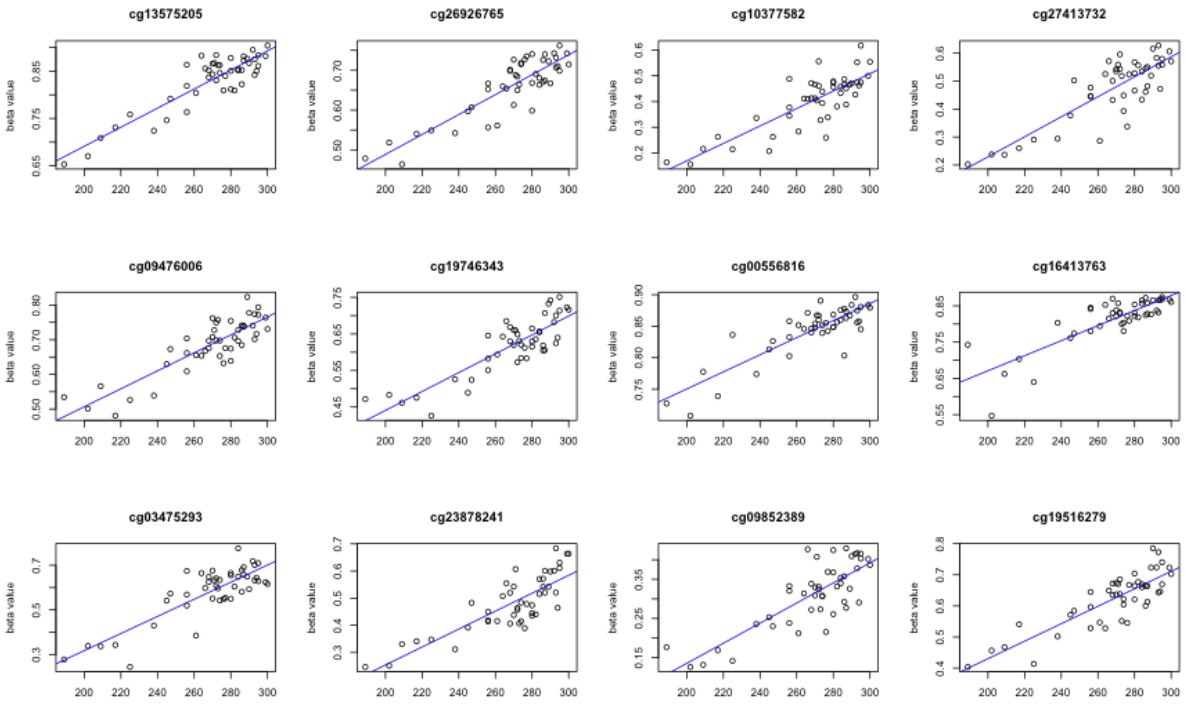


Figure 5.14 Top 12 placental hits plotted according to gestation (days) on x-axis and beta methylation value showing increasing methylation status with advancing gestational age

5.12 External validation of DMPs in placental tissue

Placental results were validated using methylation data from placental tissue run on an Illumina HumanMethylation27 BeadChip, a precursor to the 450k array (Novakovic B et al., 2011)

Using the same cutoff for statistical significance as our study (FDR-adjusted $p<0.05$) placental DMPs were identified between the second trimester and term groups in the published dataset (Novakovic B et al., 2011) and similar placental DMPs were searched for between the second and third trimesters in our own cohort.

There were 23,677 positions available for direct comparison between the 27k and 450k platforms where probe type and design were identical. 825 DMPs were called as significant from these positions on the 450k platform and of these, 541 (65%) were also identified as being significant in the independent dataset run on the 27k platform (Novakovic et al). There were 519 similarly hypomethylated and 13 similarly hypermethylated positions in both datasets. In only 9 probes was the direction of methylation opposite between the two datasets (Figure 5.15).

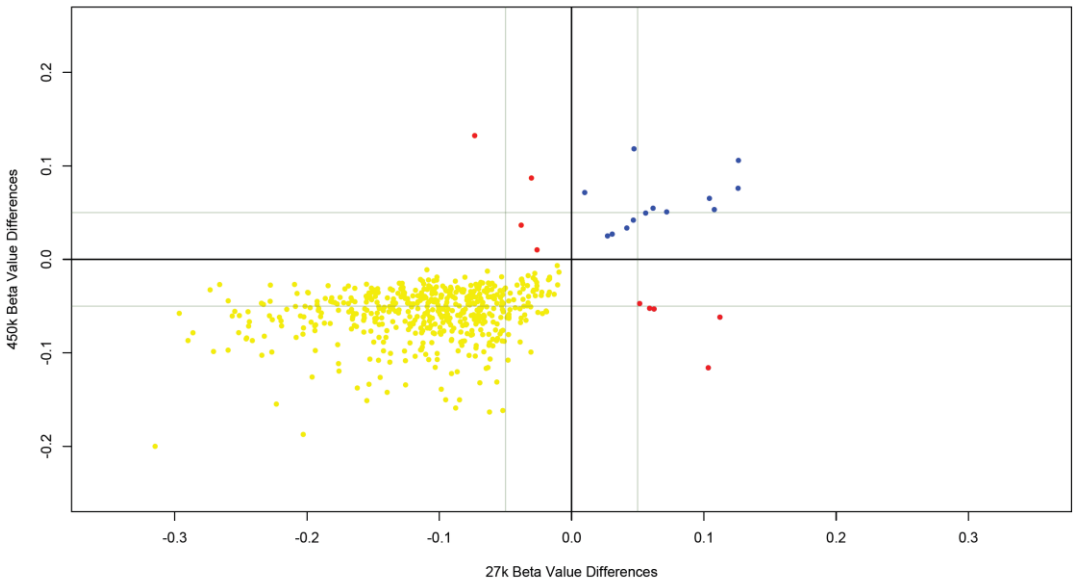


Figure 5.15 Differentially methylated positions present in placenta from our data set (450K) and from the data of Novakovic (27K). Hypomethylated regions common to both platforms (519 DMPs; yellow), hypermethylated regions (13 DMPs; blue) and mismatched direction (9 DMPs; red).

Individual positions were also correlated according to gestational age and visually represented in order to assess whether differences in our FGR and AGA offspring could be interpreted.

This confirmed that in a significant number of positions the trend identified in the Novakovic preterm cohort was mirrored in FGR and AGA samples from our experiments (Figures 5.16c and 5.16d below). However, anti-correlations were also identified where there were clear differences in FGR offspring (in green) versus AGA offspring (red) which fell into the Novakovic controls (in black) represented by figure 5.16a and to some degree 5.16b. The limitation in this assessment is, however, the lack mid trimester samples in our cohort that would be needed to identify the methylation trend during this missed period of gestation (as it is conceivable that methylation status changed in a non-linear trajectory despite the start and end point making an increasing linear course most likely).

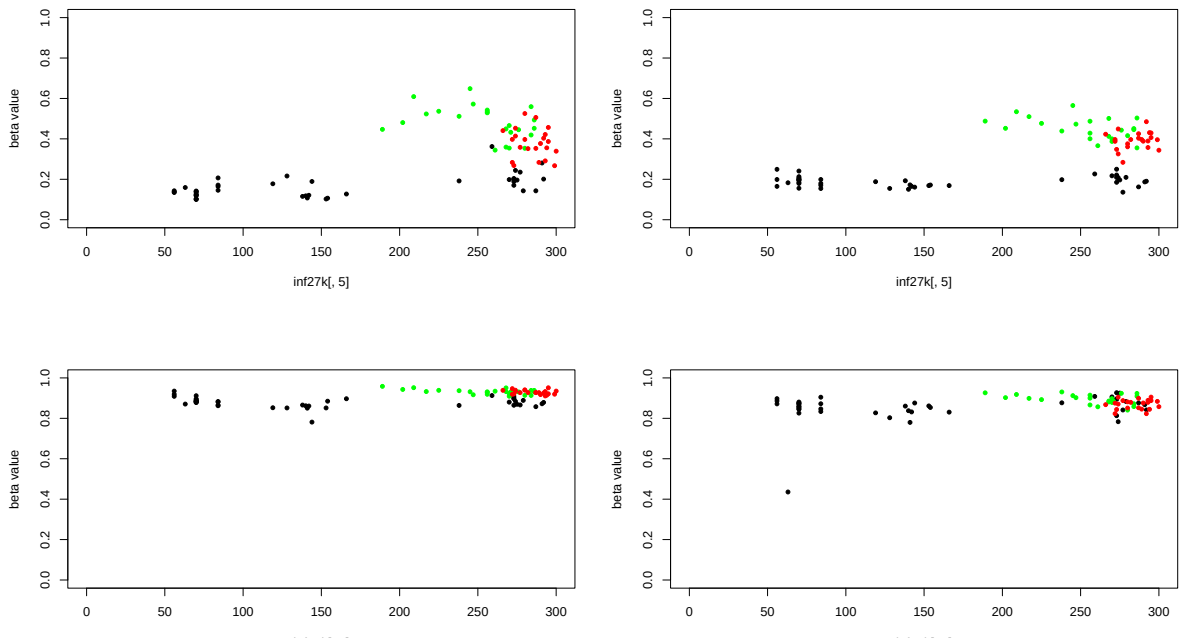


Figure 5.16 (a top left, b top right, c bottom left and d bottom right) representing individual DMPs in placental tissue and trends of case (green), control (red) or Novakovic samples (black). In bottom two figures there is a correlated trend suggesting methylation is similar in all cohorts. In the top two examples there seems to be a hypermethylated trend in FGR cases as opposed to either of the other two cohorts.

5.13 Placental DMPs associated with FGR independently of gestational age

By identifying 541/835 DMPs common to both our dataset and the appropriately grown fetal dataset by Novakovic, it is more likely that the remaining 284 DMPs (data not shown) unique to our FGR group, have a role independent of gestational age that influences fetal growth. Out of these 284 placental DMPs, the top DMP stratified according to adjusted p values was on *FGFR1* (fibroblast growth factor receptor 1**)** which is a growth factor involved in angiogenesis and embryonic development (Bottcher RT et al., 2005).

Placental DMPs identified in the original case/control model were also compared with the DMPs identified through the linear gestational age model. Of the 443 DMPs, 39 (Table 5.4) were significant in the original FGR versus AGA comparison but did not come up in the gestational age model. Again, through exclusion, is seems possible that these 39 DMPs might have a role in the placenta in growth restricted pregnancies that is independent of gestational age. The most significant placental DMP of the remaining 39 DMPs was located on latent transforming growth factor beta binding protein 3 *(LTBP3*) which encodes for a protein which forms a complex with transforming growth factor beta (TGF-beta) proteins and it may play a critical role in controlling and directing the activity of *TGFB1.*

It is likely that the pre-term placental samples power the initial case/control as evidenced by the lack of DMPs in the term cohort. In combination with the placental DMPs identified as discordant from Novakovic, these DMPs offer novel placental targets to investigate and suggest key methylation differences exist between placenta and cord in FGR samples that are also related to gestational age and likely to be specific to the developing placenta as the term FGR versus AGA placentae did not yield significant DMPs.

cg identifier	Chr	Position	Gene	adj. p value
cg04089426	11	65307791	*LTBP3*	1.34E-06
cg02696763	1	156252642	*SMG5;TMEM79*	1.40E-06
cg19883905	18	12254024	*CIDEA*	2.18E-06
cg03795316	10	12110955	*DHTKD1*	2.56E-06
cg23844623	11	47290730	*MADD*	4.50E-06
cg05827732	3	96533699	*EPHA6*	5.81E-06
cg13549444	12	7060474	*PTPN6*	6.16E-06
cg13682375	5	16828148	*MYO10*	6.35E-06
cg22416002	22	39713062	*RPL3;RNU86*	7.82E-06
cg11031567	17	79378207	*BAHCC1*	9.69E-06
cg25063733	1	43996686	*PTPRFRF*	1.04E-05
cg05284727	4	1874366	*WHSC1*	1.27E-05
cg03437741	1	117491966	*PTGFRN*	1.52E-05
cg07137519	10	131688736	*EBF3*	1.57E-05
cg00210210	6	99798738	*C6orf168*	2.27E-05
cg11682350	6	33240284	*VPS52;RPS18*	2.35E-05
cg14528147	1	113001201	*CTTNBP2NL*	3.00E-05
cg03427058	1	1565025	*MIB2*	3.28E-05
cg23496755	7	20865522	*RPL23P8*	3.29E-05
cg02747671	13	39612086	*C13orf23;NHLRC3*	3.53E-05
cg06915053	12	7245056	*C1R*	3.65E-05
cg07130266	3	50375624	*RASSF1*	3.65E-05
ch.7.2667284R	7	128843678	*SMO*	3.82E-05
cg06534427	22	39713022	*RPL3;RNU86;RPL3*	3.99E-05
cg17755460	1	117491905	*PTGFRN*	4.20E-05
cg09005580	3	184636467	*VPS8;VPS8*	4.59E-05
cg17444015	1	2322981	*MORN1;RER1;MORN1*	4.62E-05
cg21113768	6	144386848	*PLAGL1*	4.85E-05
cg26286901	8	22595307	*PEBP4*	4.92E-05
cg08920071	1	117544781	*CD101*	5.18E-05
cg09993092	6	33167743	*SLC39A7;RXRB;SLC39A7*	5.21E-05
cg24780479	19	14581142	*PKN1*	5.26E-05

Table 5.4 39 DMPs not present in gestational placenta age model but remaining from the original case/control call where a Ref-seq gene annotation has been made

5.14 Cord blood FGR DMPs independent of gestational age

To determine whether advancing gestation at 'term', between 256 - 300 days, was still able to influence cord blood DMPs, DMPs were identified that were present both in FGR offspring cord blood at term and the gestational age cord blood linear model. Of 839 DMPs identified in FGR cord blood at term, 121 DMPs were also called in the gestational age cord blood analysis. These 121 DMPs were matched for gestational age between 12 FGR (mean 279.1 ± 8.3 days) and 12 AGA individuals (mean 280.3± 7.2 days). The direction of methylation is 100% correct with the overall case / control call from samples> 255 days (Figure 5.17) wwhich shows that the methylation differences seen in the term model persist when gestational age is confounded for through a matching model. This supports the idea that the 121 DMPs called in both were likely to be driven by fetal growth restriction rather than gestational age.

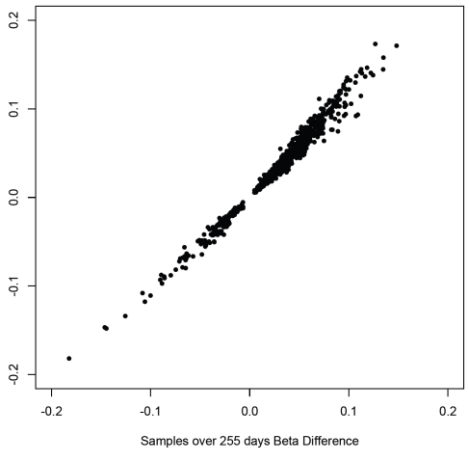


Figure 5.17 Correlation of direction methylation between the 121 DMPs in both models when compared in gestational age matched FGR and AGA offspring cord blood samples

It is likely that these 121 DMPs have a similar influence in growth restriction to the other 718 DMPs in term cord blood. When these 121 are stratified according to beta values they reveal similar results to the overall 839 DMPs in term cord blood. 54 of the 121 have a beta difference greater than 5% (24% of all 839 hits which is a similar proportion to overall number of DMPs in term call with beta difference>5% at

26.8%) and 4 have a beta difference of >10% (7.5%) which is similar to the overall proportion of DMPs with beta difference >10% (6.3%) – see table 5.5 for split according to hyper and hypomethylated state.

	5% beta value difference (n)	GA associated (n/%)	10% beta value difference (n)	GA associated (n/%)
Hypermethylated	162	48 (30%)	45	4 (8.9%)
Hypomethylated	63	6 (9.5%)	8	0 (0%)

Table 5.5Table of significance of term cord blood DMP at >5% and >10% cutoff levels and number of each of these that are present in the gestational age (GA) model.

Term cord blood DMPs were also compared to results from a study performed on whole cord blood of a separate cohort of AGA offspring between 32-43 weeks gestation (Schroeder DW et al., 2011). Using the HumanMethylation27 BeadChip, 25 DMPs were identified in whole umbilical cord blood, related to gestational age. None of these 25 DMPs were present in the term FGR cord blood, further supporting the fact that methylation differences have been identified unique to term growth restricted offspring, not driven by gestational age.

5.15 Functional assessment tools

5.15.1 Term cord blood DMPs in relation to Marmal-aid database

The Marmal-aid tool was used to gain insight into the functionality of the 839 differentially methylated CpG sites that we discovered in cord blood from FGR offspring born at term. The Marmal-aid output suggested that these DMPs are found in normally very stable unmethylated or lowly methylated gene loci consistent across all blood cell types (Figure 5.20). Also, these same positions vary little in methylation status between tissue type and through ageing and are located in highly conserved, invariate regions of the genome (Figure 5.19). In contrast, the majority of growth restricted DMPs was hypermethylated, forming a distinct profile in FGR offspring.

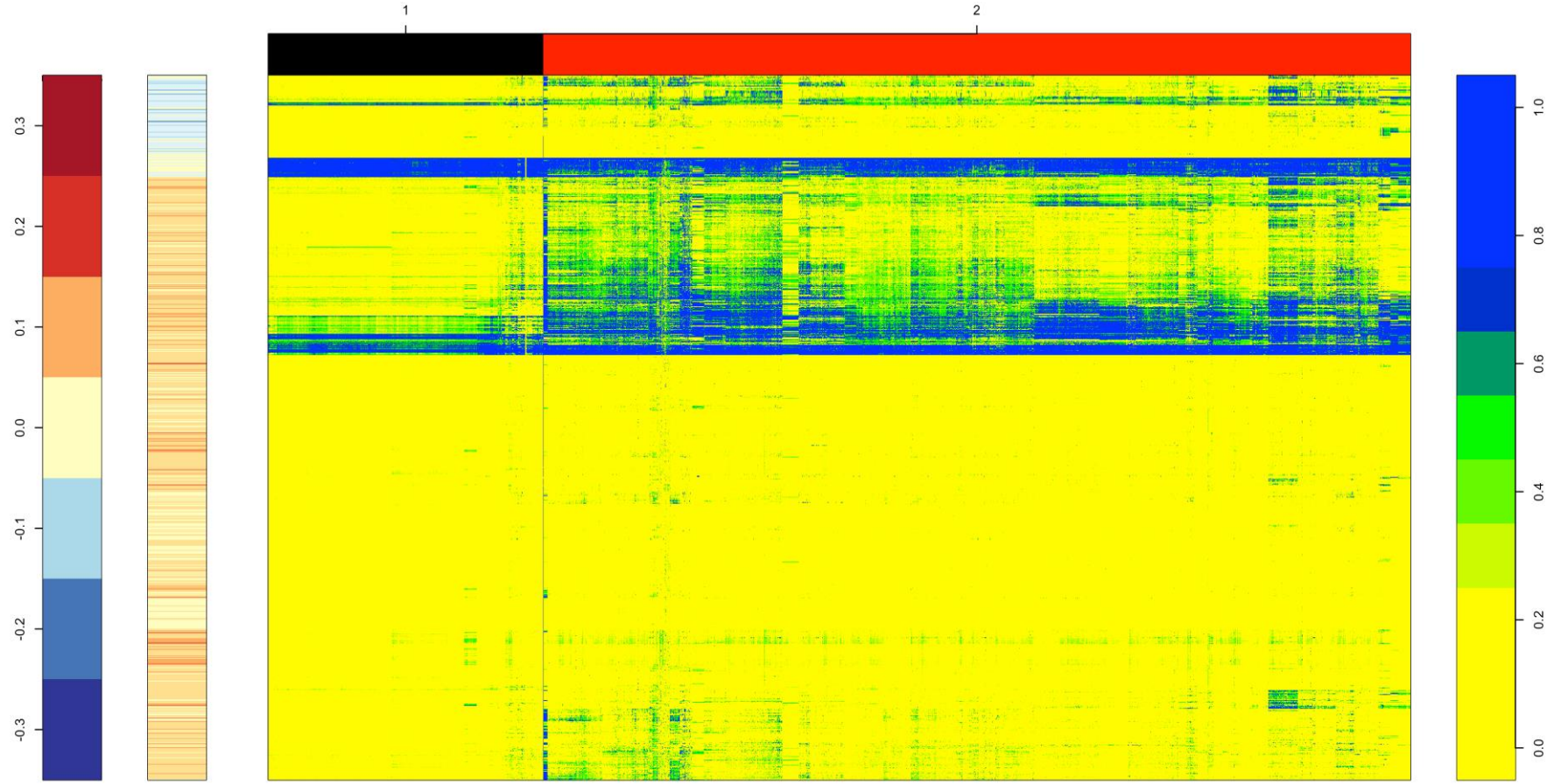


Figure 5.18. Heat map incorporating the 839 DMPs from term cord blood. Cluster 1 (black) for the all samples plotted are the blood samples including AML. Cluster 2(red) is everything else in Marmal-aid. The big blocks of yellow suggest fairly consistent methylation across a largenumber of different tissues and disease states. Directionality of the differences observed coloured from red-blue

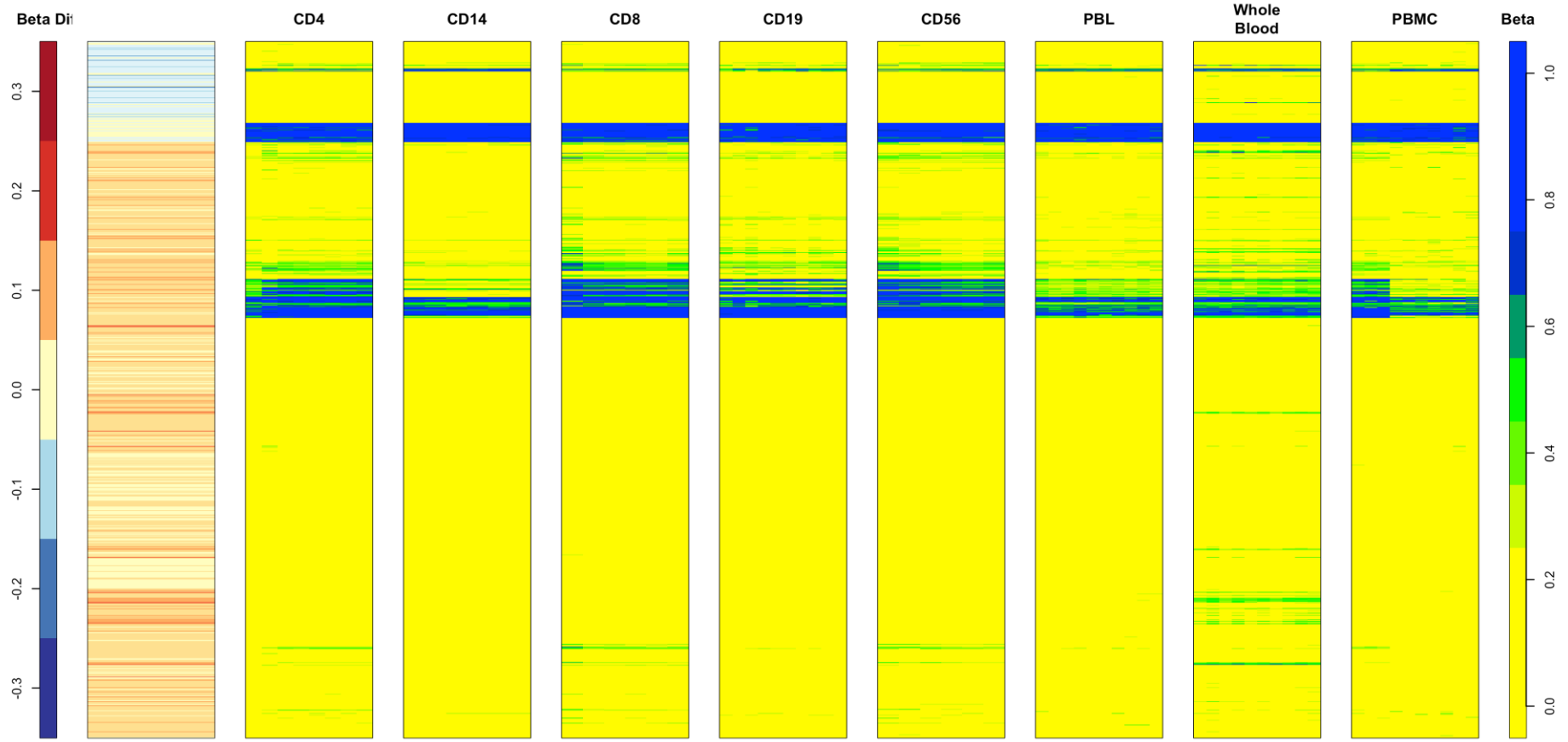


Figure 5.19 Heat map incorporating the 839 DMPs from term cord blood. Clusters of different cell types are recorded showing methylation status in relation to different cells of whole blood. The big blocks of yellow suggest fairly consistent methylation across a large number of different cell types.

5.15.2 Investigation of Gene function Associated with Epigenetic Dysregulation

5.15.2.1 Term Cord Blood

To investigate whether genes associated with the loci exhibiting epigenetic dysregulation in FGR cohort had potential functional roles, we used GO pathway analysis on the 839 term cod blood DMPs, which were linked to RefSeq-annotated genes. Using thresholds described in the methods section, 23 GO pathways were identified (Table 5.6) many of which contained DMPs identified with >10% beta value differences, strengthening and supporting a potential functional role for these DMPs. Several genes were differentially methylated in more than one position at CpG sites close together. The majority of these 23 pathways relate to regulation of gene transcription and expression and control of metabolic processes. Those with the greatest fold change suggest influence in regulation of DNA recombination and cellular processes. Transcription and regulation of RNA polymerase II is also overrepresented with fold change above 2.2. Embryonic development is present as well as CNS development and organ morphogenesis which are consistent with a fetal growth phenotype. Negative and positive regulation of metabolic processes (especially nucelobase containing and nitrogen compound metabolic processes) and regulation of gene expression are also represented in the most significant GO pathways.

GOBPID	Term	Exp Count	Count	Size	Odds Ratio
GO:0006310	DNA recombination	6	18	183	3.05
GO:0009987	cellular process	410	440	11554	2.6
GO:0006357	regulation of transcription from RNA polymerase II promoter	35	71	983	2.32
GO:0006366	transcription from RNA polymerase II promoter	42	82	1189	2.23
GO:0009790	embryo development	27	53	768	2.15
GO:0007417	central nervous system development	21	40	581	2.15
GO:0045944	positive regulation of transcription from RNA polymerase II promoter	20	39	570	2.09
GO:0019219	regulation of nucleobase-containing compound metabolic process	107	172	3012	2.03
GO:0045934	negative regulation of nucleobase-containing compound metabolic process	28	51	781	2.01
GO:0051171	regulation of nitrogen compound metabolic process	110	174	3087	2.00
GO:0032774	RNA biosynthetic process	103	165	2890	2.00
GO:0010629	negative regulation of gene expression	27	50	771	1.99
GO:2000113	negative regulation of cellular macromolecule biosynthetic process	27	50	771	1.99
GO:2000112	regulation of cellular macromolecule biosynthetic process	99	159	2793	1.98
GO:0006355	regulation of transcription, DNA-dependent	92	149	2580	1.98
GO:0051172	negative regulation of nitrogen compound metabolic process	28	51	790	1.98
GO:0010468	regulation of gene expression	106	168	2987	1.97
GO:0006351	transcription, DNA-dependent	100	160	2822	1.97
GO:2001141	regulation of RNA biosynthetic process	92	149	2596	1.97
GO:0009887	organ morphogenesis	24	44	682	1.97
GO:0090304	nucleic acid metabolic process	139	208	3905	1.96
GO:0010556	regulation of macromolecule biosynthetic process	101	161	2851	1.96
GO:0010558	negative regulation of macromolecule biosynthetic process	28	51	798	1.96
GO:0051252	regulation of RNA metabolic process	94	151	2651	1.96

Table 5.6 GO pathways identified through DMPs in cord blood

When individual assessment of the genes identified in the GO pathways was made it was apparent that several of the most significant DMPs according to beta value differences were also frequently present in the pathways identified through GO analysis. This was particularly true for the hypermethylated positions present on *TAF5; JARID2, HOXB4; TRPS1; FOXP1* as well as the hypomethylated positions on *ZNF141* and *ZIC1* (Table 5.7) and some of these will be discussed further in the discussion section.

GO term	**Gene in pathway >10% beta value difference**	
	Hypermethylated	**Hypomethylated**
DNA recombination	*IL27RA, FOXB1, RAD50*	
cellular process	*RIOK3; IL27RA; RIT1; JARID2; TAF5; RAD50; HOXB4; TRPS1; RPL14; ATL2; FOXP1*	*APOL1; ZNF141; MACROD2; ZIC1*
regulation of transcription from RNA polymerase II promoter	*JARID2; HOXB4;TRPS1; FOXP1*	
transcription from RNA polymerase II promoter	*JARID2; TAF5; HOXB4; TRPS1; FOXP1*	*ZNF141*
embryo development	*HOXB4; FOXP1*	*ZIC1*
central nervous system development	*JARID2; FOXP1*	*ZIC1*
positive regulation of transcription from RNA polymerase II promoter	*FOXP1*	
regulation of nucleobase-containing compound metabolic process	*IL27RA; JARID2; TAF5; RAD50; HOXB4; FOXP1*	*ZNF141; ZIC1*
negative regulation of nucleobase-containing metabolic process	*JARID2; HOXB4; FOXP1*	
regulation of nitrogen compound metabolic process	*IL27RA; JARID2; TAF5; HOXB4; TRPS1; FOXP1*	*ZNF141; ZIC1*
RNA biosynthetic process	*TRPS1; FOXP1*	*ZIC1*
negative regulation of gene expression	*JARID2; HOXB4; TRPS1; FOXP1*	
negative regulation of cellular macromolecule biosynthetic process	*JARID2; HOXB4; TRPS1; FOXP1*	
regulation of cellular macromolecule biosynthetic process	*JARID2; HOXB4; TRPS1; FOXP1*	*ZNF141; ZIC1*
regulation of transcription, DNA-dependent	*JARID2; TAF5; HOXB4; TRPS1; FOXP1*	*ZNF141; ZIC1*
negative regulation of nitrogen compound metabolic process	*IL27RA; JARID2; TAF5; HOXB4;TRPS1; FOXP1*	*ZNF141; ZIC1*
regulation of gene expression	*ZIC1; ZNF141; HOXB4; FOXP1; JARID2; TAF5*	*ZNF141; ZIC1*
transcription, DNA-dependent	*JARID2; TAF5; HOXB4; FOXP1*	*ZNF141; ZIC1*
regulation of RNA biosynthetic process	*JARID2; TAF5*	
organ morphogenesis	*HOXB4; FOXP1*	*ZIC1*
nucleic acid metabolic process	*IL27RA; JARID2; TAF5; HOXB4; FOXP1*	*ZNF141; ZIC1*
regulation of macromolecule biosynthetic process	*JARID2; TAF5; HOXB4; FOXP1*	*ZNF141; ZIC1*
negative regulation of macromolecule biosynthetic process	*JARID2; HOXB4; TRPS1; FOXP1*	
regulation of RNA metabolic process	*JARID2; TAF5 ; HOXB4; FOXP1*	*ZNF141; ZIC1*

Table 5.7 53 DMPs with beta methylation values above 10% and their presence in GO pathways identified

5.15.2.2 Placental DMPs in FGR not AGA pre-term samples

GO pathway analysis was used on the 284 placental DMPs identified in in the validation step that associated only with the FGR cohort. This yielded 16 pathways met the criteria for reporting. The majority of the pathways involved regulation of autophagy, response to oxidative stress and hormone stimuli (Table 5.8).

GOBPID	Odds Ratio	Exp count	Actual count	Term
GO:0016239	31.4	0.14	3	positive regulation of macroautophagy
GO:0090026	31.4	0.14	3	positive regulation of monocyte chemotaxis
GO:0010508	16.8	0.3	4	positive regulation of autophagy
GO:0032312	12	0.4	4	regulation of ARF GTPase activity
GO:0010506	9.78	0.72	6	regulation of autophagy
GO:0006914	5.12	1.49	7	autophagy
GO:0043627	4.24	2.31	9	response to oestrogen stimulus
GO:0006979	4.08	3.5	13	response to oxidative stress
GO:0031667	3.06	5.31	15	response to nutrient levels
GO:0032870	3.05	6.07	17	cellular response to hormone stimulus
GO:0043434	3.05	5.71	16	response to peptide hormone stimulus
GO:0051270	2.85	6.07	16	regulation of cellular component movement
GO:0009991	2.85	5.68	15	response to extracellular stimulus
GO:0009725	2.49	10.6	24	response to hormone stimulus
GO:0071495	2.46	8.35	19	cellular response to endogenous stimulus
GO:0032879	2.04	18.4	34	regulation of localization

Table 5.8 16 GO gene pathways identified in relation to DMPs in placental tissue

5.16 Offspring methylation and Paternal IR

This experiment was not powered to identify differences in the father through methylation status of offspring. None the less as paternal phenotype data was available for all offspring, analysis was run comparing paternal HOMA insulin resistance (IR) (as described previously in chapter 3) with methylation in offspring. Whilst no DMPs reached genome wide significance (when adjusted) for paternal IR range, 24 probes were found to be in common between placenta and cord blood of the top 100 hits. These are not necessarily changes in the same direction but most of

these 24 IR hits could be seen to represent a methylation profile influenced by a SNP (Figure 5.20) in the FGR offspring (not present in AGA controls).

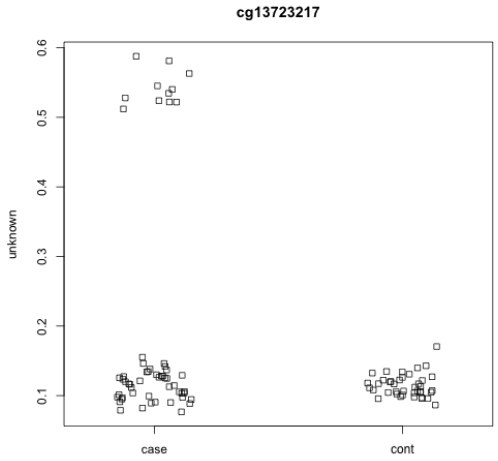


Figure 5.20 CpG profile in case and control offspring, called based on paternal IR that correlates between cord blood and placenta and shows apparent SNP profile (0,50% methylation) cg13723217 is on the *ELAC1* gene and was one of the 24 DMPs identified in analysis

The most striking examples were seen in methylation profiles for the five sites on *PAX8* that were common to placenta and cord blood (Figure 5.21).

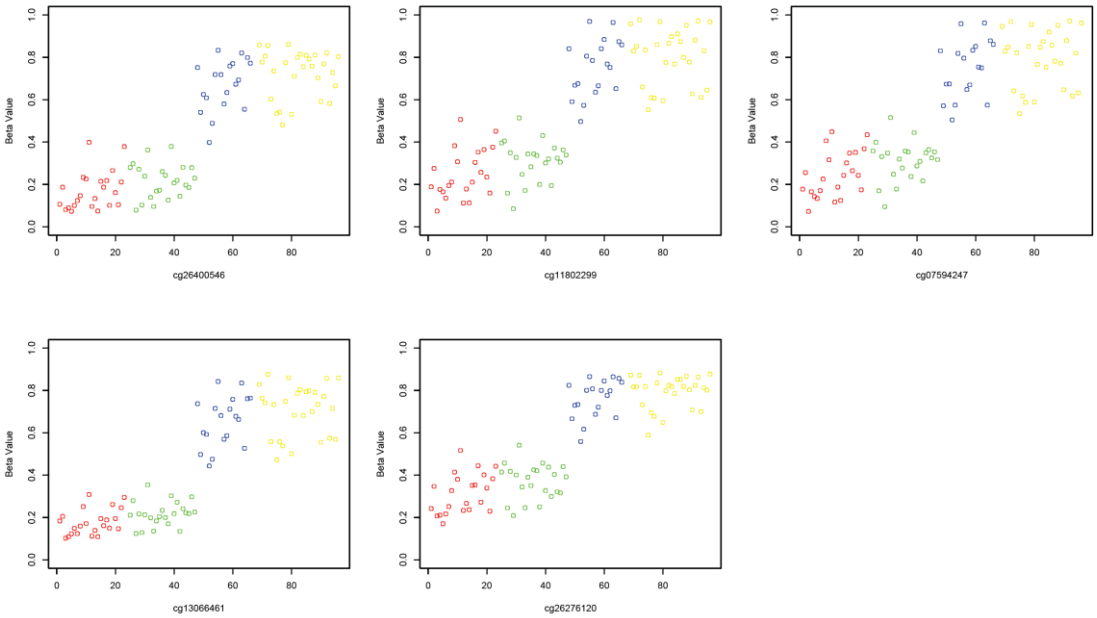


Figure 5.21 represents overlapping DMPs in cord blood and placenta in case (blue/yellow) versus control (red/green) offspring showing SNP methylation profile (0, 50%, and 100%) in FGR offspring in *PAX5*.

5.17 Influence of SNPs on methylation profiles

Finally an assessment was made of the overall influence of SNP profiles in the data. Obviously from the example above, particularly where similarities between tissues were identified, then there was the potential for a SNP to influence the methylation profile seen. Position cg13723217 (Figure 5.21) is in extremely close proximity (1bp away) to the SNP rs60618945 on the *ELAC2* gene with a minor allele frequency of 0.5 (C/G) making the difference seen highly likely to be SNP driven.

This (0%, 50%, 100% methylation) pattern of SNP profiles was not as evident in the 839 term cord blood hits and there was no tissue overlap between cord and placenta again arguing against SNPs directly driving the methylation differences seen.

To more comprehensively assess SNP influence, the next stage of analysis compared the DMPs identified in the 839 term cod blood hits with the current Illumina database of SNPs known to be able to influence methylation. This database is freely available from Illumina (http://support.illumina.com/downloads/infinium_hd_methylation_snp_list.ilmn) and contains more than 150,000 SNPs that have the potential to influence methylation. When this database was compared with the 839 DMPs, 221 positions were present in the Illumina SNP database (supplementary information) of which of the 53 positions with a beta value difference of >10%, 13 hypermethylated positions and hypomethylated positions were represented (Table 5.9).

ID	Ch	Position	UCSCRef Gene	Beta value	p value	SNP rs	Dist. (bp)	Minor Allele Frequency
cg09747578	1	46769076	*LRRC41; UQCRH*	0.12	9.43E-08	728834261 / 115665648	25/3 5	0.33/ 0.008
cg01147107	1	47696505	*TAL1*	0.11	5.12E-07	57869712	1	0.08
cg27069263	2	55278118	*RTN4*	0.14	1.57E-05	1348528	50	0.1
cg05533953	4	331685	*ZNF141*	-0.15	1.69E-10	3749520	29	0.4
cg21935083	5	131892314	*RAD50*	0.12	5.42E-07	2706335	43	0.04
cg26542283	5	43040505	*C5orf39*	-0.15	2.77E-09	77647328/ 116073875	7/24	0.05/0.00 8
cg21312090	8	116681727	*TRPS1*	0.11	2.44E-06	76321771	5	0.05
cg20684973	10	105127632	*TAF5*	0.14	5.18E-08	10883857	47	0.4
cg23802518	10	80827482	*ZMIZ1*	0.14	2.47E-07	7087678	24	0.45
cg26220528	12	57623348	*SHMT2*	0.11	5.92E-07	28365862	15	0.09
cg08965527	16	84178213	*HSDL1; LRRC50*	0.13	6.78E-08	78309971	25	0.03
cg19857457	18	47018947	*RPL17; SNORD58A*	0.14	1.92E-07	4402665	12	0.41
ch.20.327316	20	15628644	*MACROD2*	-0.12	8.34E-10	111615106	9	0.5

Table 5.9 13 DMPs with beta methylation >10%, identified in close proximity to a SNP – with information on distance away from and minor allele frequency as provided by Illumina (bp- base pairs away)

5.18 Discussion

This study identified significant DNA methylation differences in offspring affected by fetal growth restriction secondary to placental dysfunction. In term cord blood the DMPs were found to be associated with pathways crucial to key cellular processes, particularly regulation of gene expression and transcription. Within the 53 DMPs with the greatest methylation differences several genes were co-methylated at more than one CpG site.

Examples of genes that these DMPs were identified on include: *FOXP1*, whose gene in FGR offspring was hypermethylated in 3 places (beta methylation difference up to 11%). *FOXP1* is expressed in many tissues and is important for organogenesis, metabolism and immunity (Koon HB et al., 2007). *RIOK3* was hypermethylated on 2 CpG sites just 5 base pairs apart (beta values 11% and 17%). RIOK3 is an atypical protein kinase that interacts with caspase-10 and negatively regulates the transcription factor, NF-Kappa B (Shan J et al., 2009). Inhibition of NF-Kappa B reduces angiogenesis, but permits apoptosis amongst other key immune responses with the potential to compromise fetal growth and development. *TAF5* was also hypermethylated (14%) in our FGR cases. *TAF5* encodes transcription initiation factor TFIID subunit 5 (TFIID), which controls the initiation of transcription of RNA polymerase II and is crucial to the assembly and transcription initiation of DNA into mRNA. RNA polymerase II pathways featured heavily in the GO analysis of term hits.

Another similarly hypermethylated position was Jumonji (*JARID2),* which is an important regulator of cardiovascular development (Toyoda M et al., 2000; Kim TG et al, 2004). Jumonji inhibits atrial natriuretic peptide (ANP) gene expression by inhibiting the transcriptional activities of GATA4 (Kim TG et al., 2004), which was also hypermethylated (7%) in our FGR cord blood. In pregnancy, ANP is essential for physiological changes at the maternal-fetal interface. Mice deficient in ANP show poor trophoblast invasion and impaired remodelling of the maternal spiral arteries (Cui Y et al., 2012), a characteristic of pregnancies affected by FGR.

Some genes were hypomethylated for example *C5orf39* (Beta value -15%), which is unique to humans and was co-methylated in 4 positions. *C5orf39* encodes Annexin II receptor, which is a novel activator of apoptosis (Xiong Y et al Apoptosis 2013; 18: 925-939). Excessive cellular expression of Annexin II receptor, as may be the case when *C5orf39* is hypomethylated, leads to apoptosis (Whitehead CL et al., 2013). Apoptosis and autophagy has been recognised in low birth weight offspring (Hung et al., 2012). In particular, low birth weight offspring that gain excessive weight have lymphocyte apoptosis, which interestingly correlates with insulin resistance (Barg E et al., 2012).

Similar findings to this human study have been identified in an animal model of fetal growth restriction, (Thompson RF et al., 2010). Of the top genes, found to be differentially methylated in pancreatic islet cells of male rats at 7 weeks of age born growth restricted, almost half of those tested were associated together in a single functional network centred on important metabolic and cellular regulators. Furthermore, the majority of changes in the animal model occurred in evolutionarily conserved DNA sequences in proximity to genes regulating vascularization and insulin secretion (Thompson RF et al., 2010).

Within the 839 DMPs, low-level methylation differences in 2 imprinted genes *PLAGL1* and *KCNQ1* were identified, but no evidence of *IGF2/H19* involvement was found. Fetal growth restriction secondary to poor placental development appears not to be primarily controlled by imprinted genes. Indeed, this whole genome approach has led to the discovery of novel differentially methylated genes associated with FGR secondary to poor placental development in humans.

Fetal growth restriction often results in iatrogenic premature child-birth due to poor fetal development in utero. This leads to the dual problems of low birth weight and prematurity. In this study we attempted to dissect apart the influence of gestational age from impaired fetal growth by comparing placental data with other published datasets.

The placenta is a dynamic system maturing in-utero with multiple roles critical to normal fetal development. When I compared DMPs from placentae in our cohort,

with the AGA offspring studied by Novakovic, 65% of DMPs correlated, supporting their role in normal feto-placental development. The 35% of DMPs that did not correlate between datasets are likely to be associated with changes in a growth restricted placenta. GO analysis of these DMPs identified differential methylation on genes implicated in pathways important in autophagy and response to oxidative stress and hormone stimulus. These pathways are crucial in normal placental development (Jauniaux E et al., 2006) and the presence of aberrant methylation provides evidence for a fundamental role for epigenetic change in placental dysfunction response specifically in relation to oxygen and hormonal stimuli. A more targeted approach may be able to determine the exact role of these differentially methylated regions in an FGR placenta.

We confirmed previous reports of genome-wide hypomethylation in the placenta (Gama-Sosa MA., 1983 & Macaulay EC., 2011) and, as reported before that as pregnancy progresses, placental DNA becomes increasingly methylated (Schroeder DI et al ., 2013). It is plausible that certain placental genes need to methylate towards the end of pregnancy to prevent disproportionate placental and fetal growth and obstructed labour. At term, we did not observe the placental methylation differences in FGR and AGA seen in earlier pregnancy, which supports the notion that given time, pathological influences give way to gestational influences.

In a study that used the Infinium HumanMethylation27 BeadChip® platform, 206 term placentas identified 22 methylation loci that could act as markers for poor growth (Banister CE et al., 2011). We did not replicate these methylation loci on the Infinium HumanMethylation450BeadChip®. It is possible that these DMPs might be influenced by gestational age. Of the 22 DMPs, 6 were present in the recalled second and third trimester Novakovic placental dataset, supporting an influence of gestational age on these loci.

One of the limitations when assessing whole cord blood methylation is the change in fetal circulating cell types as pregnancy progresses. As we studied whole blood, it is possible that some of the variation seen in cord blood DMPs was due to different cell type populations between individuals. Using Marmal-aid, it was seen that the 839

DMPs (identified in whole cord blood) are normally unmethylated across cell types within whole blood. This supports the idea that cell type is not the underlying reason for the differences seen in growth restricted whole cord blood. Furthermore, at term, gestational ages were closely matched between cases and controls and changing cell composition due to gestation is unlikely to have an impact. However, cell composition is likely to play an important role in the 76,249 DMPs from cord blood of pre-term FGR offspring.

It is unknown whether methylation differences are a cause or a consequence of fetal growth restriction. Having identified DMPs that are associated with FGR irrespective of gestational age, a more targeted approach may be able to determine the role of these methylated regions. The permanence of these DMPs has not been addressed in this study. However, DMPs in newborn cord blood at the moment of early independent life, provides evidence that at least some of these marks have the potential for persistence into later life. Follow up studies that reassess the children born with FGR would determine if DMPs identified at birth have endured into later life and influence vulnerability to adult disease.

5.19 Summary

Using the HumanMethylation450 BeadChip, 839, predominantly hypermethylated DMPs were identified in DNA from umbilical cord blood of fetal growth restricted offspring at term. These DMPs were shown to be predominantly driven by fetal growth rather than gestational age and that these differentially methylated genes have fundamental roles in controlling gene regulation and transcription related to important metabolic and vascular functions. Altered expression of these genes in utero that endures in to later life might explain the predisposition of growth restricted offspring to adult disease. These observations provide a robust platform for further studies into the functional consequences of DMPs in fetus and placenta of growth restricted offspring.

Chapter 6

Genetic-Epigenetic interactions in Growth Restricted Offspring

Chapter 6: Genetic-Epigenetic Interactions in Growth Restricted Offspring

6.1 Introduction

Recent studies have begun to explore the interplay between genetic and epigenetic modifications. For example, SNPs that create or destroy CpG sites clearly have the ability to influence methylation status (Schalkwyk LC et al., 2010). Methylation variation as a direct consequence of genetic variation could be the route through which pathology emerges. The mechanism by which many of the T2DM risk SNPs cause disease is unknown but differential methylation on these genes is one possibility which may also provide new insight into how MODY gene mutations influence the methylome; ultimately resulting in increased risk of T2DM and cardiovascular disease.

PDX-1 is a gene that has been recognised as being associated with MODY (Naylor R et al., 2007). An animal model has shown that *PDX-1* silencing through differential methylation leads to growth restriction and T2DM (Park JH et al., 2008). It is possible that genetic variation could be driving epigenetic modifications in key regulating genes, the result of which being growth restriction and later T2DM in the individual.

In this chapter I firstly identified SNPs associated with T2DM which have the potential to influence methylation and T2DM risk. The aim was to correlate methylation change to the presence or absence in FGR offspring sampled.

Novel/ rare genetic variants present in FGR offspring are more likely to directly affect function, or influence methylation, which in turn may lead to functional change. Therefore, the variants identified in Chapter 4 in growth restricted offspring are of particular interest in the search for underlying mechanisms that may be important in the pathogenesis of FGR.

In this chapter, therefore, I also sought to identify whether there was any evidence of genetic-epigenetic interactions between genetic variants identified in FGR offspring and DNA methylation obtained from the same FGR cohort.

6.2 Hypothesis

Genetic variants influence the state of DNA methylation in areas of the genome involved in control of transcription and expression of genes and regulation of metabolic processes. As a result, aberrant interaction can lead to poor fetal growth.

6.3 Specific aims

1. To investigate the degree of influence of all known SNPs on differential methylation identified in FGR offspring. Within this aim, further characterise influence of risk SNPs in metabolic processes previously investigated in this thesis.
2. Investigate any epigenetic-genetic interactions with novel/rare candidate genetic variants identified of interest in FGR offspring.

6.4 Methods

Firstly, overlap between T2DM risk SNPs/ MODY genes and DNA methylation was explored using the 839 term cord blood DMPs identified.
In order to assess the role of genetics and the influence of SNPs on methylation status, interaction between known SNPs in the Illumina SNP database and the 839 term cord blood DMPs identified in table 5.10 was assessed. Exome data was used where applicable to identify whether the relevant SNPs were present in the individual term offspring.
Novel and rare variant calls from the 10 whole exomes of FGR offspring (identified as previously described in Chapter 4) were analysed using the 839 term cord blood DMPs as a filter in the Ingenuity Pathway Analysis. Methylation was introduced in the form of a 5 line bed file incorporating chromosome position as a unique identifier. This file was uploaded to Ingenuity software and amalgamated into the output.

6.5 Results

6.5.1 SNP data in relation to exome

Of the 839 cord blood hits in FGR offspring, 221 SNPs by Illumina were identified that are in close proximity and could influence methylation (data not shown). A more structured approach was sought to assess genetic-epigenetic interactions that might be important in FGR. Only those methylation differences that had the greatest beta value difference (as identified in section 5.9) and hence had the most functional potential.

Of the 53 DMPs with a beta value difference of more than 10%, 13 were present in the Illumina SNP list (Table 6.1).

Next, using the exome data from individual samples, the presence of the relevant SNP was investigated. Exome data from all term offspring (that were also included in the methylation array) were assessed. This meant that 6 of the exomes could be used.

ID	CHR	MAPINFO	UCSCRef	B value	p value	SNP rs	Distance From SNP (bp)	MAF
cg09747578	1	46769076	*LRRC41; UQCRH*	0.12	9.43E-08	728834261/ 115665648	25/35	0.33/ 0.008
cg01147107	1	47696505	*TAL1*	0.11	5.12E-07	57869712	1	0.08
cg27069263	2	55278118	*RTN4*	0.14	1.57E-05	1348528	50	0.1
cg05533953	4	331685	*ZNF141*	-0.15	1.69E-10	3749520	29	0.4
cg21935083	5	131892314	*RAD50*	0.12	5.42E-07	2706335	43	0.04
cg26542283	5	43040505	*C5orf39*	-0.15	2.77E-09	77647328/ 116073875	7/ 24	0.05/ 0.008
cg21312090	8	116681727	*TRPS1*	0.11	2.44E-06	76321771	5	0.05
cg20684973	10	105127632	*TAF5*	0.14	5.18E-08	10883857	47	0.4
cg23802518	10	80827482	*ZMIZ1*	0.14	2.47E-07	7087678	24	0.45
cg26220528	12	57623348	*SHMT2*	0.11	5.92E-07	28365862	15	0.09
cg08965527	16	84178213	*HSDL1; LRRC50*	0.13	6.78E-08	78309971	25	0.03
cg19857457	18	47018947	*RPL17; SNORD58A*	0.14	1.92E-07	4402665	12	0.41
ch.20.327316F	20	15628644	*MACROD2*	-0.12	8.34E-10	111615106	9	0.5

Table 6.1 Thirteen DMPs within top beta value table (n=53) that were present on Illumina SNP list. Table includes distance of the methylation position from the SNP as well as the MAF- minor allele frequency of the SNP

However, of these 13 DMPs only 4 on the genes *ZN141, SHMT2, RPL17* and *C5orf39* were located within exomes or sequence that was captured within the exome dataset.

All 6 exome samples were assessed for recognition of these 4 SNPs. None of the samples contained evidence of these particular SNPs. Whilst this provides some evidence that these SNPs are not driving methylation on these particular positions, further interpretation of the influence of other SNPs on the methylation profile cannot be made. The majority of the 839 DMPs (including a majority 44/53 of top hits), however, either do not have recognised SNPs within them or have been ruled out as being SNP driven through exome data. This provides further support that the methylation differences seen in term cord blood are not due to SNPs.

6.5.2 T2DM risk SNPs identified in FGR cohort and influence on methylation status

In order to assess any methylation influenced change secondary to the presence of T2DM risk SNPs, differential methylation in FGR offspring in relation to the 16 T2DM SNPs identified through the exome data was investigated. None of the SNPs identified through exome data overlapped (within 1000bp) with the 839 DMPs identified in term cord blood. There was no evidence in this cohort of T2DM risk SNPs influencing methylation, however, a number of T2DM risk SNPs are intronic and it was not possible in this study to investigate these interactions.

6.5.3 MODY genes and methylation status

Within the 839 term cord blood DMPs, there was no overlap with MODY gene variants identified in the exome data. Of the 839 DMPs, positions on *PDX-1* (beta value difference in FGR offspring 3%), *PLAGL1* (identified on two positions at 8% and 4% difference) and *KCNQ1* (8% difference) were identified in FGR offspring but no new variants of interest were located on these genes. *PLAGL1* and *KCNQ1* have been implicated in presentations of neonatal diabetes and are known to be imprinted genes.

6.5.4 Novel genetic variants and differential methylation

Having explored the interaction between T2DM risk SNPs and MODY gene variants with DNA methylation interactions and discovering no convincing associations, the next stage in the analysis involved assessing potentially novel variants and the presence of differential methylation.

Two approaches were undertaken with input from Ingenuity software analyst Tim Morewood, Dr Chiara Bacchelli (UCL) and Chris Mathews (QMUL).

The gene on which the novel/rare variant was identified through Ingenuity Pathways Analysis was assessed against the 839 DMPs identified in term cord blood.
From the first analysis within the novel gene variants, 2 differentially methylated positions were identified on the same gene in two instances; *KDM4B* and *RPL14*.

The methylation position on *KDM4B* was located more than 100,000 bases away from the novel insertion but the methylation position on *RPL14* was 4880bp away from the polymorphism identified from exome data.

The 839 DMPs were then input as a text file into Ingenuity and used as a filter. Firstly, a window of 10000bp around each of the 839 DMPs was also generated and uploaded as a 5 line .BED file. This methylation window was connected to one of the 71 gene variants (identified by Ingenuity in Chapter 4) if its chromosome position was contained within the 10000bp window.

The 839 methylation positions were uploaded to Ingenuity. In this analysis one different filter was set in order to incorporate regions around the exome that had been captured (as the methylation array had more extensive coverage across the whole genome than just exomes) which meant that 214 variants were assessed against the 839 methylation positions. A window of 10000bp around the methylation position was first used which yielded 9 hits (Table 6.2 below). The majority of these positions were found to be in promoter and 3'UTR regions of the genome. When a more stringent window of 1000bp around the variant was used then only 2 of these methylation positions were found to overlap a variant (*CLASP1* and *BCL11A*) which

were both in promoter regions and hence not identified in the original variant analysis.

Position	Gene Region	Gene Symbol	Regulatory Site	File 2: Methyl 10000	bp distance	File 3: Methyl 1000
33286389	3'UTR; Promoter	*ZBTB22; DAXX*	Promoter Loss	33292028	5639	
33286402	3'UTR; Promoter	*ZBTB22; DAXX*	Promoter Loss	33292028	5626	
40503520	Exonic	*RPL14*		40498640	4880	
57301820	Promoter	*PAICS; PPAT*	Promoter Loss	57303156	1336	
60780662	Promoter	***BCL11A***	Promoter Loss	60781128	466	60781128
69599088	Promoter	*NFAT5*	Promoter Loss	69600792	1704	
69599097	Promoter	*NFAT5*	Promoter Loss	69600792	1695	
1.21E+08	3'UTR	*INHBB*	microRNA Binding Site	121116384	7410	
1.22E+08	Promoter	***CLASP1***	Promoter Loss	122407280	225	60781128

Table 6.2. Potentially interacting genetic variants and DMPs from cohort of FGR offspring. Table includes positions of gene variants and DMP and distance apart- file 2 refers to distance of <10000bp and file 3 refers to a distance <1000bp.

6.5.5 Manual annotation of 53 DMPs in FGR cord blood with 71 novel gene variants

A final approach was undertaken incorporating the list of 53 most differentially methylated positions and 71 novel variants and performing a manual literature search in order to identify any potential interactions between a genetic variant and methylation change not previously identified through pathway analysis.

The TATA-binding protein (*TBP*) gene variants were identified as of interest in relation to the DMP identified in term cord blood *TBP* associated factor 5 (*TAF5*). *TBP* is part of the transcription factor IID (*TFIID*) protein, required to coordinate initiation of transcription by RNA polymerase II (RNAPII) by binding to the core promoter to position the polymerase correctly. *TFIID* is composed of the (*TBP*) and a group of evolutionarily conserved proteins known as *TBP*-associated factors or *TAF*s.

6.6 Discussion

Epigenetic modifications secondary to genetic variations offer a novel pathway for disease development. In this study methylation differences seen in our FGR offspring term cord blood do not seem to be driven by common SNPs. There was no evidence of a correlation between T2DM risk SNPs or MODY gene variants with methylation differences. Caution must be applied to negative findings, however, due to the lack of whole genome comparison with the methylation array data.

Novel genetic mutations identified in FGR offspring yielded limited useful insights, correlating loosely with some of the DMPs identified in the cord blood of FGR offspring. The *TBP* variant present in FGR offspring has the potential to be interacting with *TAF5* and provides a plausible argument for genetic variation having the potential to impact methylation status. Given the important role that *TAFs* play in the control of gene transcription and expression this opens up the possibility of a functional effect. *TAFs* participate in basal transcription, serve as co-activators, function in promoter recognition or modify general transcription factors (GTFs) to facilitate complex assembly and transcription initiation (Hampsey M et al., 1998). The structure of GTFs is dynamic and differences in subunit composition affect conformation (Coleman RA et al., 2008).

Extensive work has been undertaken to understand the interplay between *TAF* and other components in order to provide insight into how transcriptional regulation might help direct growth and development of organisms (Tora L et al., 2002).

Transcription of protein coding genes and most small nuclear RNA (snRNA) genes is performed by RNA polymerase II (RNAPII). Synthesis of pre mRNA by RNAPII involves assembly of a transcription pre initiation complex (PIC) at the promoter. The PIC is made of RNAPII and general transcription factors (GTFs) including TFIID. Recently, *TAF5* has been identified as the only *TAF* present on the promoter of the snRNA genes suggesting a unique role for *TAF5* in transcriptional regulation of snRNAs (Zaborowska J et al., 2012).

Conformational change in *TBP* may be affecting the ability of *TAFs* to interact thereby influencing the ability of GTF complexes to regulate RNA polymerase II initiated transcription. In relation to our FGR offspring, a specific interaction with conformational *TBP* change, *TAF5* methylation difference and snRNA transcription warrants further investigation.

6.7 Summary

This study has provided little evidence of SNP driven methylation differences influencing FGR phenotype. However, limited coverage of genome by using WES means further study is needed in order to provide convincing evidence that SNP driven change is not relevant in the distinct methylation profile identified in FGR offspring.

Genetic variation in *TBP* and the *TAF5* methylation change identified in the FGR offspring provide an interesting avenue of investigation in relation to individuals affected by FGR.

Chapter 7

General Discussion and Future Aims

Chapter 7: General Discussion and Future Aims

7.1 Key findings of the research and conclusions

This work has identified that men who father growth restricted offspring have a distinct phenotype at the time of the index pregnancy that will predispose them to T2DM. These findings are consistent with previous epidemiological evidence that shows that men with T2DM are more likely to have fathered low birth weight offspring (Davey Smith G, 2004 et al.; Wannamethee SG et al., 2004, and Hypponen E et al., 2003).

Identification that men have a metabolic phenotype, characterised by insulin resistance and central adiposity, at the time of an affected pregnancy, provides new avenues of investigation into the way in which a paternal phenotype, expressed through a genotype may lead to FGR in his offspring.

A paternal genetic influence on fetal growth is strongly suggested by the fetal insulin hypothesis. In my thesis, I provide direct evidence for a genetic influence on FGR. Rare genetic variants were identified in genes important in Maturity Onset Diabetes of the Young (MODY) in FGR offspring. However, transmission was not exclusively paternal in origin. A combined parental genetic influence along with the additive effects of more than one gene variant appears likely.
The identification of MODY gene variants, did not explain the majority of fetal growth restriction in the cohort studied. Furthermore, investigation of the SNPs common to T2DM and other insulin resistant syndromes did not provide support for their direct involvement in FGR; a finding consistent with other reported work.

Epigenetic modifications to the genome were investigated in FGR offspring cord blood and placenta. FGR offspring were phenotypically homogenous, with placental dysfunction thought to underlie the cases recruited. Study of both the placenta and

fetus was, therefore, important in order to identify specific methylation patterns within each tissue that could result in a growth restricted fetus.
DNA methylation in cord blood of term growth restricted fetuses revealed a distinct pattern, different to the methylation in the placenta of growth restricted individuals (across a range of gestational ages). The findings in my thesis support both the importance of epigenetic dysregulation in growth restriction and the fact that this dysregulation has a separate and distinct pattern in placenta and fetus.

It is likely that gestation specific DNA methylation occurs in the placenta in order to facilitate normal fetal growth. Several reports support a recognised reproducible placental methylation pattern that is gestation specific (Gama Sosa et al., 1993 and McCauley et al., 2011), which was confirmed in my work. Deviation from a healthy gestation specific methylation profile may lead to a hostile placental environment and sub-optimal fetal growth and development. Conversely a sub-optimal in-utero environment might drive the placental methylation change. The design and strength of my study meant that mothers in both the FGR and AGA groups were selected, and shown, to have similar phenotypic, metabolic and vascular characteristics with no underlying maternal reason for FGR development, other than placental dysfunction. This makes it less likely that a hostile *in-utero* environment (per se) is the primary driving force behind the pathogenesis of the FGR in this cohort. This does not exclude the *in-utero* environment from playing an important role; it just supports the concept that in this particular presentation of FGR it is more of a consequence of primary genetic/epigenetic influences.
The work in my thesis has sought to provide evidence for novel gene variants that could be driving the differential methylation identified in FGR offspring.

The insertion on *KDM4B* was shown to be present in a majority of the FGR offspring sampled, however, further investigation revealed it was also found in controls. Hence this particular variant is not likely to be of significance in my FGR cohort. However, the *KDM4B* gene is part of the JmJ complex of genes, recently identified as having histone demethylase activity. It is normally up-regulated in the presence of activated hypoxia inducible factor 1 *(HIF1).* Activation of *HIF1* occurs in the presence of hypoxia in the tissues and so, during normal early placentation it is up-regulated in response to the relatively hypoxic *in-utero* environment. This function

means this family of genes are still interesting candidates in FGR and may play a role in fetal growth in conjunction with other pathways.

In normal pregnancy, extra villous trophoblast (EVT) cells invade spiral arteries and replace the vascular endothelial cells (Pijnenborg et al., 1980). This remodelling dilates spiral arteries and increases utero-placental blood flow. Despite the utero-placental oxygen concentration being as low as 2% in early pregnancy (Jauniaux E et al., 2001), remodelling can occur with the appropriate response from *HIF1*. Failed remodelling and poor EVT invasion, which is the classic pathology seen in placental pregnancy syndromes, is therefore thought not to be directly due to oxygen availability, but rather the ability of the placenta to adapt to the low oxygen conditions.

A placenta, unable to respond appropriately to hypoxia will develop oxidative stress (Burton G & Jauniaux E, 2011). Part of the response to oxidative stress involves an abnormal autophagy response (Hung et al., 2012). In my thesis I have shown placental methylation differences in genes associated with the autophagy response, along with response to oxidative stress and hormonal stimuli. It is plausible that genetic variants in the hypoxia-signalling pathway may impair the ability of a fetus (and placenta) to compensate for hypoxia. DNA methylation is thus altered in genes responsible for autophagy in the placenta thereby influencing the development of placental dysfunction with subsequent compromise of fetal growth. Increased autophagy proteins have been identified in FGR placentas (Hung et al., 2012). My findings of altered placental DNA methylation of autophagy genes in FGR placentas and the evidence of altered autophagy protein expression shown in other studies, provides a novel connection between methylation change and functional effect that could explain the dysregulation seen in FGR placentas.

It is conceivable, that the methylation profile seen in FGR offspring cord blood is essentially the downstream product of the effects on the placenta during gestation. Persistence of cord blood (fetal) methylation differences in these offspring may provide a mechanism by which growth restricted individuals are more susceptible to adult cardio-metabolic disease.

The identification of methylation change in *TAF5* along with insertion of CAG sequences in *TBP* again opens up the possibility that genetic variants may be driving

methylation change with subsequent functional consequences. It remains to be proved that the methylation difference seen in term cord blood has a functional effect. However, if variation in *TBP* is the cause of the methylation difference in *TAF* seen in FGR cord blood then this may provide both the reason for the methylation difference and a pathway that confers increased T2DM risk in adults, affected by growth restriction.

FGR cord blood yielded a different methylation profile to that of the placenta. It can be argued that the epigenetic processes occurring in the placenta are critical to placental function and *in-utero* fetal growth but that the cord blood changes seen in the fetus at term are the adaptive response to a poor *in-utero* environment and are more likely to influence the consequent risk of adult cardio-metabolic disease. The finding that cord blood methylation differences correspond to pathways important in gene regulation, transcription and expression, supports the concept that they have the potential to have a profound functional effect.

7.2 Future Work

This thesis has yielded many new avenues for investigation in fetal growth restriction (FGR). FGR is an important clinical problem to tackle; its consequences are significant both for short and long term outcome of affected offspring.

This work identifies paternity of FGR offspring as a novel risk factor in male diabetes and heart disease. Through the proxy marker of low birth weight offspring, fathers were identified who are at risk of developing cardio-metabolic disease later in life. Primary prophylaxis through altered diet and lifestyle may protect these men from progression towards T2DM disease development particularly if they are also overweight (Chae JS et al., 2012).

It would be useful to establish whether the findings in this work hold true if the inclusion criteria are widened. A well-designed prospective study investigating all couples who have a baby born at a $<10^{th}$ customised birthweight centile would

identify whether fathers of these pregnancies have a reproducible phenotype. The potential for primary prevention is substantially increased if these findings can be applied generically to a male population who father a small baby. It would also pave the way for a postnatal program of referral for men, similar to that provided for women who suffer from gestational diabetes during their pregnancy due to their increased risk of subsequent diabetes development (Bellamy L et al., 2009).

It remains to be known whether reproductive outcomes can be improved if a father loses weight before conceiving a second pregnancy. Optimising periconceptual paternal health is a novel concept worthy of investigation in order to minimise the risk of an adverse pregnancy outcome. A study investigating reproductive outcomes in men post interventions that reduce their risk for diabetes may provide evidence. Clearly, a purely genetic risk transmission from father to offspring is not likely to be affected by weight loss but potential trans-generational epigenetic mechanisms may be affected by paternal weight loss and a phenotypic study such as proposed would support this concept.

If it proves to be the case that moderating paternal behaviour prior to conception, or in between pregnancies, has a positive effect on reproductive success this has far reaching implications in pre-conception advice that should be provided to couples. The finding in this thesis that men who smoked were more likely to have a growth restricted offspring is one example of a mechanism for which further investigation is warranted in order to identify if moderating this behaviour (i.e. stopping smoking) improves offspring birthweight. If it proves to be the case then investigation, of how paternal smoking translates to poor fetal growth, is warranted.

The identification of rare genetic variants in FGR offspring has the potential to strongly support a genetic component in FGR. This finding needs further functional validation in order to definitively prove a role for MODY gene variants in growth restriction. However, the success of using WES in order to capture these genetic variants suggests a future role for next generation platforms in uncovering new avenues of investigation in FGR.

Young adults born SGA have a different metabolic profile than those born AGA (Longo S et al., 2013) but MODY is not often considered. A follow up study of FGR

children is needed, specifically designed to provide evidence of whether these children go on to develop MODY. This could provide a novel screening opportunity in children who have been affected by FGR. Furthermore, identifying FGR offspring with MODY gene variants and following them up as adults, whilst time consuming, would provide comprehensive evidence of a pluripotent phenotype in MODY which includes growth restriction.

The DNA methylation data generated through the use of the 450K array strongly supports the use of next generation platforms in investigating epigenetic modifications in the genome. This thesis identified novel methylation patterns in term cord blood and placenta from growth restricted offspring. Aberrant hypermethylation in the cord blood corresponded to regions of the genome implicated in gene expression and transcription regulation. Placental hits centred on response to oxygen, nutrients and hormones and regulation of autophagy. These methylation findings need to be directly connected to functional significance. Knowing the pathways affected means that a functional study can be planned to better assess the consequences of aberrant methylation. Such a study may provide novel insights into ways in which FGR prevention can be achieved. If methylation is part of a pathway responding to genetic and/or environmental cues then interrupting the sequence of events that leads to aberrant methylation may affect subsequent outcome.

Within my investigation of DNA methylation, changes within histones were apparent. It would be very useful to study histone modifications in FGR in conjunction with DNA methylation patterns. It is likely that the histone modifications together with aberrant methylation may result in FGR.

Genetic variants in hypoxia signalling genes could be causing a functional defect, which influences the placental response to oxygen. Impaired sensing leads to altered DNA methylation and histone modification in placental genes related to autophagy and oxidative stress. Consequent fetal growth restriction occurs and is reflected in aberrant DNA methylation in genes regulating key cellular processes in cord blood seen in the neonate at birth.

Functional studies need to be performed to assess this hypothesis and to provide evidence that the downstream result will be impaired fetal growth.

Other genetic variants in the hypoxia signalling pathway, particularly those responsible for the placental response to hypoxia, need to be investigated for their potential influence on epigenetic change within a placenta and subsequent effect on fetal growth.

7.3 In summary

This thesis provides evidence of a distinct paternal phenotype in FGR related to placental dysfunction. There is evidence for a paternal genetic effect in FGR offspring but this does not explain the majority of cases.

Rare gene variants were identified in growth restricted offspring through whole exome sequencing. Distinct DNA methylation profiles were also identified in placenta and cord blood of FGR offspring. These findings support the possibility that genetic variants in glucose sensing and/or hypoxia signalling pathways are driving epigenetic change at the level of the placenta which is affecting fetal growth.

Placental methylation patterns correspond to pathways related to hypoxia response and term cord blood methylation differences were seen on genes critical in the regulation of gene expression and transcription. This suggests that DNA methylation has a crucial role to play in fetal growth. Persistence of these epigenetic marks in FGR offspring would support their role in susceptibility to future T2DM.

APPENDIX 1

PUBLISHED ARTICLES FROM THIS WORK

Hillman S, Peebles DM, Williams DJ.

Paternal metabolic and cardiovascular risk factors for fetal growth restriction: a case-control study. Diabetes Care. 2013 Jun; 36(6):1675-80

APPENDIX 2

PUBLISHED ABSTRACTS FROM THIS WORK

Hillman SL, Whitten M, Peebles D and Williams D

The metabolic and vascular phenotype of men who father pregnancies affected by fetal growth restriction. Arch Dis Child Fetal Neonatal Ed 2011;96.

APPENDIX 3

DATABASES AND BIOINFORMATICS USED IN ASSESSMENT

Various bioinformatic tools and databases were used in this study in order to identify polymorphic markers and characterise genes within the region. The main sites used for this study were as follows:

BLAST (www.ncbi.nlm.nih.gov/blast/)
Ensembl Genome Browser (http://www.ensembl.org/)
ExPASy (http://www.expasy.org).
GDB Human Genome Database (http://www.gdb.org/)
Gene ontology(http://www.geneontology.org)
Mammalian Genotyping Service (http://research.marshfieldclinic.org/)
Marmal-aid (http://marmal-aid.org)
NCBI database (http://www.ncbi.nlm.nih.gov/)
NCBI SNP database (http://ncbi.nlm.nih.gov/SNP/)
OMIM (http://www.ncbi.nlm.nih.gov/sites/entrez?db=OMIM)
UCSC Genome Browser (http://genome.ucsc.edu/)
UniGene (NCBI, http://www.ncbi.nlm.nih.gov/)

APPENDIX 4

SUPPLEMENTARY INFORMATION

Sample ID	Qubit ng/ul	Nanodrop ng/ul	A260	A280	260/280	260/230	Vole ul left	Total ng
187EHa	482	534.32	10.686	5.873	1.82	2.33	30	14460
190LFa	543	578.31	11.566	6.42	1.8	2.31	30	16290
255OAa	333	373.38	7.468	4.012	1.86	2.47	30	9990
331MBV	341	381.68	7.634	4.256	1.79	2.27	47	16027
337ECb	115	128.31	2.566	1.352	1.9	2.39	47	5405
359KCb	389	413.52	8.27	4.58	1.81	2.33	47	18283
386SSa	204	226.87	4.537	2.462	1.84	2.27	47	9588
39NK	126	131.94	2.639	1.418	1.86	2.35	47	5922
392DGa	255	280.26	5.605	3.047	1.84	2.36	47	11985
395MHa	190	199.01	3.98	2.155	1.85	2.37	47	8930
413BTa	264	274.74	5.495	2.99	1.84	2.39	47	12408
421YDc	92.30	100.52	2.01	1.076	1.87	2.3	47	4338.1
447VXaa	54.90	51.95	1.039	0.563	1.85	1.96	47	2580.3
47KWaa	61.80	64.98	1.3	0.7	1.86	2.22	47	2904.6
50STa	177	164.52	3.29	1.747	1.88	2.3	47	8319
53ALa	119	137.22	2.744	1.492	1.84	2.29	47	5593
56SWa	175	200.15	4.003	2.164	1.85	2.38	47	8225
009REaa	70.60	84.7	1.694	0.921	1.84	2.21	47	3318.2
227FBa	153.00	170.44	3.409	1.851	1.84	2.36	47	7191
3IH	114.00	133.18	2.664	1.46	1.82	2.12	47	5358
398IRbb	58.70	62.25	1.245	0.677	1.84	1.98	47	2758.9
156BTa	58.30	64.19	1.284	0.701	1.83	2.23	47	2740.1

156BTc	74.50	82.9	1.658	0.886	1.87	2.27	47	3501.5
178JSa	21.10	25.12	0.502	0.29	1.73	1.97	47	991.7
178JSb	59.60	67.8	1.356	0.732	1.85	2.06	47	2801.2
328LBa	83.80	82.46	1.649	0.91	1.81	2.13	47	3938.6
328LBb_2	79.10	81.92	1.638	0.896	1.83	2.23	47	3717.7
334MVa	75.10	86.28	1.726	0.925	1.87	2.34	47	3529.7
334MVb	81.40	90.39	1.808	0.991	1.82	2.26	47	3825.8
95EGb	105	110.77	2.215	1.196	1.85	2.33	47	4935
95EGc	69	69.7	1.394	0.744	1.87	2.32	47	3243

S1 DNA Quality Check readings on case offspring samples

Assay	Gene	rs
1	ABCC8	757110
2	ALDHIL1	1127717
3	BMP2	235768
4	CEL	62576769
5	GCKR	1260326
7	HNF1A	2464196
9	HNF1B	141193981
10	HNF4A	736823
11	KCNJ11	5215
12	KCNJ11	5218
13	KCNJ11	5219
14	KCNJ11	1800467
15	NEUROD1	1801262
16	NOTCH2	147223770
17	PPARG	1801282
18	SLC2A2	5400
19	SLC30A8	1326663
20	THADA	7578597
21	WFS1	734312
22	WFS1	1801214
23	WFS1	35031397
24	INSR	1799817

S2 Genotyping assays designed and tested. Gene name and rs number as identified through dbSNP 130 and 135.

SNP	*Gene*	Population	Trait	Risk Allele Freq in Controls	OR/Beta [95% CI]	First Author. Journal; Published Year
	ABCC8		NDM			Johansson et al. Plos One; 2013
rs4607103	*ADAMTS9*	European	T2D GWAS	C[0.76]	1.09[1.06–1.12]	Zeggini et al. Nat Genet; 2008
	ADCY5		Birth weight GWAS			Horikoshi M et al. Nat Genet; 2013
	ADRB1		Birth weight GWAS			Horikoshi M et al. Nat Genet; 2013
	AGPAT2		Syndrome			Johansson et al. Plos One; 2013
rs1127717	*ALDHIL1*		Syndrome			Johansson et al. Plos One; 2013
	ALMS1		Syndrome			Johansson et al. Plos One; 2013
	APPL1		Candidate			Johansson et al. Plos One; 2013
rs1552224	*ARAP1*	European	T2D GWAS	A[NR]	1.14[1.11–1.17]	Voight et al. Nat Genet; 2010
rs243021	*BCL11A*	European	T2D GWAS	A[NR]	1.08[1.06–1.10]	Voight et al. Nat Genet; 2010
rs235768	*BMP2*		Syndrome			Johansson et al. Plos One; 2013
	CAV1		Syndrome			Johansson et al. Plos One; 2013
	CCNL1		Birth weight GWAS			Horikoshi M et al. Nat Genet; 2013
rs12779790	*CDC123, CAMK1D*	European	T2D GWAS	G[0.18]	1.11[1.07–1.14]	Zeggini et al. Nat Genet; 2008
rs7754840	*CDKAL1*	European	T2D GWAS	C[0.36]	1.12[1.08–1.16]	Horikoshi M et al. Nat Genet; 2013
rs564398	*CDKN2A, CD-KN2B*	European	T2D GWAS	T[0.56]	1.13[1.08–1.19]	Zeggini et al. Science; 2007
rs7018475	*CDKN2B*	European	T2D GWAS	?[NR]	1.35[1.18–1.56]	Huang J et al. Eur J Hum Genet; 2012
	CEL	62576769				
rs13292136	*CHCHD9*	European	T2D GWAS	C[NR]	1.11[1.07–1.15]	Voight et al. Nat Genet; 2010
	CISD2/WFS2		Syndrome			Johansson et al. Plos One; 2013

rs5945326	*DUSP9*	European	T2D GWAS	G[NR]	1.27[1.18–1.37]	Voight et al. Nat Genet; 2010
	EIF2AK3		Syndrome			Johansson et al. Plos One; 2013
	FOXA1		Candidate			Johansson et al. Plos One; 2013
	FOXA2		Candidate			Johansson et al. Plos One; 2013
	FOXA3		Candidate			Johansson et al. Plos One; 2013
rs8050136	*FTO*	European	T2D GWAS	A[0.4]	1.23[1.18–1.32]	Zeggini et al. Science; 2007
	FXN		Syndrome			Johansson et al. Plos One; 2013
	GATA4		Candidate			Johansson et al. Plos One; 2013
	GATA6		Candidate			Johansson et al. Plos One; 2013
	HFE		Syndrome			Johansson et al. Plos One; 2013
rs1111875	*HHEX*	European	T2D GWAS	C[0.52]	1.13[1.09–1.17]	Scott et al. Science; 2007
rs1531343	*HMGA2*	European	T2D GWAS	C[NR]	1.1[1.07–1.14]	Horikoshi M et al. Nat Genet; 2013
rs7957197	*HNF1A*	European	T2D GWAS	T[NR]	1.07[1.05–1.10]	Voight et al. Nat Genet; 2010
rs4430796	*HNF1B*	European	T2D GWAS	G[NR]	1.14[1.08–1.20]	Voight et al. Nat Genet; 2010
rs4402960	*IGF2BP2*	European	T2D GWAS	T[0.3]	1.14[1.11–1.18]	Scott et al. Science; 2007
rs1799817	*INSR*		Candidate			Johansson et al. Plos One; 2013
rs7578326	*IRS1*	European	T2D GWAS	A[NR]	1.11[1.08–1.13]	Voight et al. Nat Genet; 2010
	ISL1		Candidate			Johansson et al. Plos One; 2013
rs864745	*JAZF1*	European	T2D GWAS	T[0.5]	1.1[1.07–1.13]	Zeggini et al. Nat Genet; 2008
rs5219/ 5218/5215	*KCNJ11*	European	T2D GWAS	T[0.46]	1.14[1.10–1.19]	Scott et al. Science; 2007
rs231362	*KCNQ1*	European	T2D GWAS	G[NR]	1.08[1.06–1.10]	Voight et al. Nat Genet; 2010
rs972283	*KLF14*	European	T2D GWAS	G[NR]	1.07[1.05–1.10]	Voight et al. Nat Genet; 2010
	LCORL		Birth weight GWAS			Horikoshi M et al. Nat Genet; 2013
	LMNB2		Syndrome			Johansson et al. Plos One; 2013
	LMX1A		Candidate			Johansson et al. Plos One; 2013

rs2943641	*LOC64673, IRS1*	European	T2D GWAS	C[0.63]	1.19[1.13–1.25]	Rung et al. Nat Genet; 2009
rs1387153	*MTNR1B*	European	T2D GWAS	T[NR]	1.09[1.06–1.11]	Voight et al. Nat Genet; 2010
	NKX2-2		Candidate			Johansson et al. Plos One; 2013
	NKX6-1		Candidate			Johansson et al. Plos One; 2013
rs10923931	*NOTCH2, ADAM30*	European	T2D GWAS	T[0.11]	1.13[1.08–1.17]	Zeggini et al. Nat Genet; 2008
	ONECUT1/HNF6		Candidate			Johansson et al. Plos One; 2013
	PAX6		Candidate			Johansson et al. Plos One; 2013
	PBX1		Candidate			Johansson et al. Plos One; 2013
rs1801282	*PPARG*	European	T2D GWAS	C[0.82]	1.14[1.08–1.20]	Scott et al. Science; 2007
rs8042680	*PRC1*	European	T2D GWAS	A[NR]	1.07[1.05–1.09]	Voight et al. Nat Genet; 2010
	PTF1A		Candidate			Johansson et al. Plos One; 2016
rs7593730	*RBMS1, ITGB6*	European	T2D GWAS	?[0.78]	1.11[1.08–1.16]	Qi et al. Hum Mol Genet; 2010
rs13266634	*SLC30A8*	European	T2D GWAS	C[0.61]	1.12[1.07–1.16]	Scott et al. Science; 2007
rs5400	*SLC2A2*		Candidate			Johansson et al. Plos One; 2013
	SOX9		Candidate			Johansson et al. Plos One; 2013
	SREBF1		Candidate			Johansson et al. Plos One; 2013
rs17036101	*SYN2, PPARG*	European	T2D GWAS	G[0.93]	1.15[1.10–1.21]	Zeggini et al. Nat Genet; 2008
	SYT9		Candidate			Johansson et al. Plos One; 2013
rs7903146	*TCF7L2*	European	T2D GWAS	T[0.26]	1.38[1.31–1.46]	Saxena et al. Science; 2007
rs7578597	*THADA*	European	T2D GWAS	T[0.9]	1.15[1.10–1.20]	Zeggini et al. Nat Genet; 2008
rs896854	*TP53INP1*	European	T2D GWAS	T[NR]	1.06[1.04–1.09]	Voight et al. Nat Genet; 2010
rs7961581	*TSPAN8, LGR5*	European	T2D GWAS	C[0.27]	1.09[1.06–1.12]	Zeggini et al. Nat Genet; 2008
rs9472138	*VEGFA*	European	T2D GWAS	T[0.28]	1.06[1.04–1.09]	Zeggini et al. Nat Genet; 2008

rs1801214 / 734312	*WFS1*	European	T2D GWAS	T[NR]	1.13[1.08–1.18]	Voight et al. Nat Genet; 2010
rs4689388	*WFS1, PPP2R2C*	European	T2D GWAS	T[0.57]	1.16[1.10–1.21]	Rung et al. Nat Genet; 2009
	ZAC		Syndrome			Johansson et al. Plos One; 2013
rs4457053	*ZBED3*	European	T2D GWAS	G[NR]	1.08[1.06–1.11]	Voight et al. Nat Genet; 2010
rs11634397	*ZFAND6*	European	T2D GWAS	G[NR]	1.06[1.04–1.08]	Voight et al. Nat Genet; 2010

S3 List of a priori candidate genes and/or SNPs implicated in FGR and references

Oligo name	Sequence	Synthesis scale	Scale	Tm	Oligo length	Product size
HNF1A_125F	TTCTCAGAACCCTCCCCTTC	0.025	DESALT	60.6	20.0	412.0
HNF1A_125R	AAAGAGGTTTAGGTGACTGCTG	0.025	DESALT	57.7	22.0	
HNF4A_289F	CAGCTGGAGGGCACCCACTAT	0.025	DESALT	65.8	21.0	426.0
HNF4A_289R	TTCAGACACGTGTGTGCACTA	0.025	DESALT	58.9	21.0	
KRT8_231F	AGACAAGGCTGGGAGTTGAGG	0.025	DESALT	63.0	21.0	438.0
KRT8_231R	CTCCACCTTCTCCACCAACTA	0.025	DESALT	59.2	21.0	
SAFB2_270F	GCGGGACCGTTACTGAAC	0.025	DESALT	59.1	18.0	313.0
SAFB2_270R	CCCAAGTTCGAGGGAACC	0.025	DESALT	60.4	18.0	
TDG_839F	CCACGCGGTACTACAGAGAC	0.025	DESALT	58.4	20.0	187.0
TDG_839R	TTAAAAGACCCCTGCGTGC	0.025	DESALT	61.2	19.0	
TBP_013F	ACCTGCTGTTCCACCAAGAA	0.025	DESALT	60.7	20.0	595.0
TBP_013R	TCTGAAAACAGAGCAGGAACAT	0.025	DESALT	59.0	21.0	
KDM4B_411F	ATGCCTCCCTTGAAGGCTG	0.025	DESALT	63.0	19.0	664.0
KDM4B_411R	CAGGCCACTCCTAGCTTCC	0.025	DESALT	60.0	19.0	
RPL14_520F	CTTCAGATGTAGTTGTAGGGA	0.025	DESALT	52.0	21.0	517.0
RPL14_520R	CTATCAATCTGCCTCCTAACT	0.025	DESALT	53.0	21.0	

S4 Primer sequence ordered for Sanger sequencing

	Assay ID	Forward Primer Seq.	Reverse Primer Seq.	Reporter 1 Sequence (VIC)	Reporter 2 Sequence (FAM)	Context Sequence
Assay 1	C____600632_20					CGTGCTCTGACCTTCTGTCCAGGGG[A/C]GATGAGGGCATTGACGTGCTTCAGC
Assay 2	C__11908221_20					GGCTATGAACATGTGGTCTTCCACG[C/T]CTGTGAAAACAGTTGGCTCAAAGAA
Assay 3	C___2244893_10					CCAACTCGAAATTCCCCGTGACCAG[A/T]CTTTTGGACACCAGGTTGGTGAATC
		CCCATCCCAG	CGGTGG	AGGACATC	AGGACATC	
Assay 4	AHVJIP0	GAGAATAAGAAGAAG	GCACCAGGAA	GGTCTCAAA	GATCTCAAA	
Assay 5	C___2862880_1_					CACAGCACCGTGGGTCAGACCTTGC[C/T]GGTGAGAGTCCAGCCGTGACAAAGG
Assay 7	C___1263617_10					CCTGTGCAGAGCCATGTGACCCAGA[A/G]CCCCTTCATGGCCACCATGGCTCAG
		CTGATGCTG	TGTACGCAC	TTTCCATCT	CATCTGCA	
Assay 9	AHS1MDK	CTGGTATCTGTGA	ACAAGCAGGAA	GCAGTGGTG	ATGGTG	
Assay 10	C__25618321_10					TGTGTGCCATCTGCGGGGACCGGGC[C/T]ACGGGCAAACACTACGGTGCCTCGA
Assay 11	C___2991148_10					GTGCAGAGTGGTGTGGGCACTTTGA[C/T]GGTGTTGCCAAACTTGGAGTAGTCC
Assay 12	C___2991149_20					TGAAGCAGAGGCGGCCGTGGCGCAG[A/G]GCGATCACCGCATGCTTGCTGAAGA
Assay 13	C__11654065_10					CGCTGGCGGGCACGGTACCTGGGCT[C/T]GGCAGGGTCCTCTGCCAGGCGTGTC
Assay 14	C__11654166_10					TGGTGGTGCAGGTCGCTGGGTGCCA[C/G]GTCGTAGAGTGGGCTGTTGGCATCA
Assay 15	C__11951412_10					AGTGAGTCCTCCTCTGCGTTCATGG[C/T]TTCGAGGTCGTCCTCCTTCTTGTCT
		CCCTGGATGG	AGAATGGAGG	CTTGTGAAC	TGTGAACCATG	
Assay 16	AHUAKJS	AAAATGGATAAGGAT	CACCTGTATTGAC	CATTTCAAGTG	TCAAGTG	
Assay 17	C___1129864_10					AACTCTGGGAGATTCTCCTATTGAC[C/G]CAGAAAGCGATTCCTTCAC
Assay 18	C___3142148_10					AAGCCACCCACCAAAGAATGATGCA[A/G]TCATTCCACCAACTGCAAAGCTGGA
Assay 19	C___8899319_10					AGCTGGTAGGCGTAGAGATGAGTGG[A/G]CATTTTGATCCAGTGAGGGGTATTG
Assay 20	C__32653841_10					GCCAAAGAGATTAACTCTTTCATTG[C/T]TATTTTCAACAAATCCATTCTGCCT
Assay 21	C___2401729_1_					GTCTGTAGTGTGCCCCTGCTGTTGC[A/G]CTGGTGGACCAAGGCCAGCTTCTCT
Assay 22	C___8841076_10					CTGTCGGCCACCTGGTCGTCCTCAA[C/T]GTCAGCGTCCCGTGCCTGCTCTATG
Assay 23	C__25617947_20					CAAGGACTGCATCCCCTGCTCGGAG[C/G]TGGCTGTCATCACCGGCTTCTTTAC
Assay 24	C___8356128_1_					GTCCCACCCCCACTGGACTCACCAC[A/G]TGATGGCAGGTGAAGCCCTTCATGA

S5 ABI primers and probes for genotyping. Where customised assays were designed then ABI provide only context sequences

Chr	Position	Gene Region	Gene Symbol	Protein Variant	
1	144923730	Exonic	PDE4DIP	p.E243fs; p.E309fs; p.E406fs	frameshift
1	152189055	Exonic	HRNR	p.R1684C	missense
1	152189310	Exonic	HRNR	p.G1599S	missense
1	152190257	Exonic	HRNR	p.V1283A	missense
2	132238306	Exonic	TUBA3C/TUBA3D	p.C347F	missense
2	179419226	Exonic	TTN	p.I20676M	missense
3	10088407	Exonic	FANCD2	p.L426fs	frameshift
3	40503520	Exonic	RPL14	p.149_150insAAAAA	in-frame
3	52027854	Exonic	RPL29	p.130_131insAK	in-frame
3	73111482	Exonic; Intronic	EBLN2; PPP4R2	p.R84fs	frameshift
3	73111506	Intronic	EBLN2; PPP4R2	p.91_92insV	in-frame
3	124951185	Exonic	ZNF148	p.*795fs	frameshift
3	124951186	Exonic	ZNF148	p.*795fs	frameshift
3	195505910	Exonic; Intronic	MUC4	p.S4181G	missense
3	195505925	Exonic; Intronic	MUC4	p.P4176S	missense
3	195506051	Exonic; Intronic	MUC4	p.L4134F	missense
3	195506076	Exonic; Intronic	MUC4	p.Q4125H	missense
3	195506089	Exonic; Intronic	MUC4	p.A4121V	missense
3	195507011	Exonic; Intronic	MUC4	p.L3814F	missense
3	195507324	Exonic; Intronic	MUC4	p.H3709Q	missense
3	195507605	Exonic; Intronic	MUC4	p.R3616C	missense
3	195507941	Exonic; Intronic	MUC4	p.T3504P	missense
3	195508499	Exonic; Intronic	MUC4	p.A3318P	missense
3	195508510	Exonic; Intronic	MUC4	p.L3314H	missense
3	195509861	Exonic; Intronic	MUC4	p.S2864P	missense
3	195514524	Exonic; Intronic	MUC4	p.H1309Q	missense
3	195515134	Intronic	MUC4	p.P1106H	missense
4	1388625	Exonic	CRIPAK	p.C110fs	frameshift
4	88537069	Exonic	DSPP	p.1085_1086insDSS	in-frame
5	139931629	Intronic; Exonic;	SRA1	p.109_110insR	in-frame

		ncRNA			
6	16327915	Exonic	ATXN1	p.208_209insQQQ	in-frame
6	16327915	Exonic	ATXN1	p.208_209insQ	in-frame
6	29912030	Exonic	HLA-A	p.D251fs	frameshift
6	31380162	Exonic; ncRNA	MICA	p.G318fs; p.A317fs	frameshift
6	31380162	Exonic; ncRNA	MICA	p.G318fs; p.A317fs	frameshift
6	31380162	Exonic; ncRNA	MICA	p.G318fs; p.A317fs	frameshift
6	170871013	Exonic	TBP	p.43_44insQQQ; p.63_64insQQQ	in-frame
6	170871039	Exonic	TBP	p.52_53insQ; p.72_73insQ	in-frame
6	170871047	Exonic	TBP	p.55delQ; p.75delQ	in-frame
7	15725824	Exonic	MEOX2	p.68delH	in-frame
7	22190123	Splice Site	RAPGEF5		
7	142460779	Exonic	PRSS1	p.D218Y	missense
8	101730036	Exonic	PABPC1	p.M158fs	frameshift
9	40703811	Exonic	SPATA31A6 (includes others)	p.R490W	missense
9	140637837	Exonic	EHMT1	p.V280F	missense
11	47857258	Exonic	NUP160	p.S349F	missense
11	112832341	Exonic	NCAM1	p.Q3fs	frameshift
12	53343231	Intronic; Exonic	KRT8; KRT18	p.A92P	missense
12	104359839	Splice Site	TDG		
12	104373734	Exonic	TDG	p.I98fs	frameshift
13	73409508	Splice Site	PIBF1		
14	23548797	Exonic	ACIN1	p.640_641insSR; p.600_601insSR	in-frame
14	92537354	Exonic; ncRNA	ATXN3	p.92_93insAAAAAAAAA;	in-frame
15	100252710	Exonic	MEF2A	p.410_411delQQ;	in-frame
16	1574650	Exonic	IFT140	p.H1015P	missense
16	3119304	Exonic	IL32	p.D115fs; p.D172fs; p.D152fs; p.D163fs	frameshift
17	7750177	Exonic	KDM6B	p.251_252insP	in-frame
17	45234350	Exonic	CDC27	p.Q257Q	synonymous
17	72541045	Exonic	CD300C	p.V35L	missense
19	501701	Intronic; Exonic	MADCAM1	p.234_235insTTSPEPPN	in-frame
19	501701	Intronic; Exonic	MADCAM1	p.234_235insTTSPEPPD	in-frame
19	5144411	Exonic	KDM4B	p.963_964insESITVSCGVGQ GAGGGWEHSDNLYP	in-frame
19	5587270	Exonic	SAFB2	p.F949S	missense
19	9015382	Exonic	MUC16	p.G12736S	missense

19	50308921	Exonic	AP2A1	p.Y824*; p.Y846* p.173_174insT;	stop gain
19	50881824	Exonic	NR1H2	p.76_77insT	in-frame
22	28194933	Exonic	MN1	p.533_534insQ	in-frame
22	38120180	Exonic	TRIOBP	p.540delS	in-frame
X	135956575	Exonic; Intronic; ncRNA Exonic;	RBMX	p.P301fs	frameshift
X	155239824	3'UTR	IL9R	p.N439S	missense
X	155239827	Exonic; 3'UTR	IL9R	p.N440S	missense

S6 71 gene variants identified through Ingenuity with chromosome, position, region, gene name, protein transcript, type of variant

ID	450k ID	DNA (ul)	Water (ul)	PLACENTA				
				Qubit ng/ul	Volume ng/ul	260/280	260/230	% Conversion
42FF	1	**12.8**	**32.2**	39.1	97	1.74	2.25	97.36%
68NA	2	**3.6**	**41.4**	139	97	1.8	2.31	98.64%
80CM	3	**6.2**	**38.8**	80.8	97	1.85	2.36	98.69%
83AG	4	**9.1**	**35.9**	55.2	97	1.8	2.52	90.22%
98LH	5	**3.1**	**41.9**	159	97	1.85	2.5	95.51%
101HA	6	**5.0**	**40.0**	101	97	1.85	2.66	96.81%
104NF	7	**5.7**	**39.3**	88.4	97	1.86	2.5	98.63%
131JB	8	**5.9**	**39.1**	85.2	97	1.83	2.47	98.89%
143MB	9	**10.1**	**34.9**	49.4	97	1.81	2.21	97.76%
165EK	10	**4.2**	**40.8**	118	97	1.85	2.66	98.89%
174NS	11	**12.6**	**32.4**	39.8	97	1.78	2.57	93.50%
196ST	12	**3.0**	**42.0**	165	97	1.82	2.46	97.68%
202AD	13	**3.5**	**41.5**	142	97	1.82	2.51	84.84%
209DK	14	**12.9**	**32.1**	38.7	97	1.8	2.69	80.92%
230DP	15	**2.5**	**42.5**	198	97	1.85	2.51	96.76%
248IA	16	**12.3**	**32.7**	40.6	97	1.89	2.41	98.35%
285IR	17	**11.4**	**33.6**	43.8	97	1.83	2.4	95.99%
300PB	18	**3.7**	**41.3**	135	97	1.83	2.42	97.19%
309FC	19	**4.2**	**40.8**	118	97	1.88	2.55	96.40%
320AMF	20	**8.9**	**36.1**	55.9	97	1.96	2.56	96.49%
343EE	21	**19.6**	**25.4**	25.5	97	1.94	2.8	98.51%
359LO	22	**7.7**	**37.3**	65.1	97	1.82	2.33	98.07%

376AJ	23	**19.6**	**25.4**	25.5	97	1.82	2.48	97.68%
379VMB	24	**20.6**	**24.4**	24.3	97	2.07	2.33	90.10%
442EN	25	**21.2**	**23.8**	23.6	97	1.85	2.06	98.88%
455MT	26	**13.8**	**31.2**	36.3	97	1.84	2.66	94.23%
801AG	27	**9.1**	**35.9**	55	97	1.88	1.8	99.08%
215ET	28	**2.2**	**42.8**	230	97	1.89	2.63	97.47%
235RA	29	**3.0**	**42.0**	165	47	1.84	2.47	95.51%
319RB	30	**4.7**	**40.3**	107	43	1.79	2.23	84.49%
3IH	31	**14.0**	**31.0**	35.7	97	1.85	2.31	95.45%
9RE	32	**14.7**	**30.3**	34.1	97	1.81	2.12	95.89%
39NK	33	**15.2**	**29.8**	33	97	1.84	2.41	94.61%
47KW	34	**1.5**	**43.5**	329	97	1.84	2.41	96.05%
50ST	35	**3.4**	**41.6**	149	97	1.82	2.43	99.28%
53AL	36	**5.0**	**40.0**	100	97	1.9	2.37	99.33%
56SW	37	**14.1**	**30.9**	35.4	97	1.86	2.35	93.51%
156BT	38	**3.5**	**41.5**	143	97	1.88	2.38	98.16%
161SJ	39	**3.2**	**41.8**	155	97	1.87	2.4	98.76%
184EH	40	**3.4**	**41.6**	147	97	1.84	2.58	97.81%
190LF	41	**3.0**	**42.0**	164	97	1.82	2.48	93.37%
254OA	42	**11.1**	**33.9**	45.2	97	1.79	2.57	98.28%
328LBp3	43	**6.7**	**38.3**	74.2	97	1.9	2.37	96.95%
331MVB	44	**12.3**	**32.7**	40.6	97	1.88	2.17	92.99%
334MVa	45	**13.9**	**31.1**	36	97	1.76	2.16	99.14%
337EC	46	**9.9**	**35.1**	50.7	97	1.86	2.77	97.37%
359KC	47	**4.2**	**40.8**	118	97	1.88	2.5	98.90%
385SS	48	**24.4**	**20.6**	20.5	97	1.84	2.43	97.65%
392DG	49	**6.5**	**38.5**	77	97	1.86	2.58	96.95%
395MH	50	**45.0**	**0.0**	7.83	97	1.8	2.27	99.95%
398IR	51	**13.9**	**31.1**	36	97	1.84	2.09	98.60%
413BT	52	**4.5**	**40.5**	111	97	1.84	2.55	96.99%
421YD	53	**4.8**	**40.2**	104	97	1.86	2.58	94.13%
427LN	54	**14.2**	**30.8**	35.2	97	1.89	2.39	97.76%
431MZ	55	**27.9**	**17.1**	17.9	97	1.85	2.73	99.27%
445VX	56	**12.0**	**33.0**	41.6	97	1.82	2.65	98.98%
1008JB	57	**12.8**	**32.2**	39.1	97	1.85	2.48	92.41%

433AS	**13**							**94.39%**
439ZH	**14**							**99.24%**
95EG	**30**							94.38%
164RH	**11**							93%

S7 DNA quality check and bisulphite conversion efficiency (% conversion) in placental samples run on 450K array

ID	450k ID	DNA (ul)	water (ul)			CORD BLOOD		
				Qubit ng/ul	Volume ul	260/280	260/230	% Conversion
42FF	58	**4.5**	**40.5**	112	97	1.84	2.34	92.83%
68NA	59	**5.3**	**39.7**	94	97	1.85	2.35	98.32%
80CM	60	**5.7**	**39.3**	87.3	97	1.82	2.26	98.93%
83AG	61	**5.1**	**39.9**	97.8	97	1.83	2.27	99.40%
98LH	62	**8.6**	**36.4**	58.1	97	1.8	2.21	98.17%
101HA	63	**10.2**	**34.8**	48.8	97	1.81	2.16	98.06%
104NF	64	**10.2**	**34.8**	48.8	97	1.81	2.16	98.55%
131JB	65	**8.1**	**36.9**	61.6	97	1.82	2.24	91.26%
143MB	66	**15.9**	**29.1**	31.5	97	1.8	2.1	96.48%
165EK	67	**7.6**	**37.4**	66.1	97	1.9	2.42	97.72%
174NS	68	**4.3**	**40.7**	116	97	1.82	2.32	99.09%
196ST	69	**17.9**	**27.1**	28	97	1.97	3.43	0.09%
202AD	70	**3.1**	**41.9**	163	97	1.84	2.41	96.58%
209DK	71	**4.4**	**40.6**	114	97	1.84	2.43	96.23%
230DP	72	**6.1**	**38.9**	82.3	97	1.87	2.42	97.62%
248IA	73	**2.8**	**42.2**	176	50	1.86	2.55	96.21%
285IR	74	**3.3**	**41.7**	152	97	1.84	2.35	99.48%
300PB	75	**2.5**	**42.5**	204	50	1.84	2.45	97.91%
309FC	76	**4.3**	**40.7**	117	97	1.85	2.29	98.63%
320AMF	77	**8.9**	**36.1**	55.9	97	1.94	2.56	99.03%
343EE	78	**2.6**	**42.4**	196	97	1.88	2.44	99.68%
359LO	79	**4.3**	**40.7**	117	96	1.85	2.4	99.70%
376AJ	80	**2.9**	**42.1**	174	96	1.85	2.39	99.66%
379VMB	81	**2.0**	**43.0**	249	96	1.86	2.37	99.42%
442EN	82	**13.3**	**31.7**	37.6	96	1.93	1.74	96.53%
455MT	83	**4.9**	**40.1**	103	96	1.91	2.34	99.31%
801AG	84	**14.9**	**30.1**	33.5	97	1.78	1.81	1.03%
77KP	85	**8.1**	**36.9**	61.4	97	1.83	2.3	98.47%
215Et	86	**15.8**	**29.2**	31.6	97	1.83	2.15	99.02%
235RA	87	**26.3**	**18.7**	19	97	1.82	1.86	98.95%
319RB	88	**5.9**	**39.1**	85.4	97	1.81	2.3	99.70%
3IH	89	**4.4**	**40.6**	114	18	1.82	2.12	94.28%
9RE	90	**25.5**	**19.5**	19.6	97	2	2.01	99.93%
39NK	91	**4.0**	**41.0**	126	21	1.86	2.35	93.14%
47KW	92	**8.1**	**36.9**	61.8	47	1.86	2.22	94.94%
50ST	93	**2.8**	**42.2**	177	28	1.88	2.33	95.71%
53AL	94	**4.2**	**40.8**	119	19	1.84	2.29	94.26%
56SW	95	**2.9**	**42.1**	175	28	1.85	2.38	94.34%
95EGc	96	**7.2**	**37.8**	69	47	1.87	2.32	97.24%
156BT	97	**8.6**	**36.4**	58.3	47	1.83	2.23	94.11%
178JSB	98	**8.4**	**36.6**	59.6	47	1.85	2.06	94.72%

184EH	99	**1.0**	**44.0**	482	23	1.82	2.33	96.26%
190LF	100	**0.9**	**44.1**	542	24	1.8	2.31	96.82%
227FB	101	**3.3**	**41.7**	153	25	1.84	2.36	93.70%
254OA	102	**1.5**	**43.5**	333	20	1.86	2.47	96.63%
328LBb2	103	**6.3**	**38.7**	79.1	47	1.83	2.23	94.86%
331MVB	104	**1.5**	**43.5**	341	37	1.79	2.27	94.81%
334MVa	105	**6.7**	**38.3**	75.1	47	1.87	2.34	91.91%
337EC	106	**4.3**	**40.7**	115	18	1.9	2.39	96.90%
359KC	107	**1.3**	**43.7**	389	38	1.81	2.33	96.96%
385SS	108	**2.5**	**42.5**	204	40	1.84	2.27	95.81%
392DG	109	**2.0**	**43.0**	255	34	1.84	2.36	99.41%
395MH	110	**2.6**	**42.4**	190	29	1.85	2.37	98.33%
398IR	111	**8.5**	**36.5**	58.7	47	1.84	1.98	93.66%
413BT	112	**1.9**	**43.1**	264	44.5	1.84	2.39	94.75%
421YD	113	**61.4**	**-16.4**	8.14	97	1.77	2.02	99.13%
431MZ	114	**30.7**	**14.3**	16.3	97	1.78	2.02	99.39%
445VX	115	**9.1**	**35.9**	54.9	47	1.85	1.96	96.70%
1008JB	116	**45.0**	**0.0**	7.13	147	1.77	1.58	99.73%

S8 DNA quality check and bisulphite conversion efficiency (% conversion) in cord blood samples run on 450K array

University College London Hospitals NHS
NHS Foundation Trust

A Study looking at how parents contribute to growth of babies in the womb

FATHER'S QUESTIONNAIRE

This questionnaire asks you some questions about your health and your background.

All answers you give are confidential and will only be seen by the research team

We would be grateful if you would help us by answering all the questions. If you have any queries about any of the questions or would like some help in completing this questionnaire, please contact Dr Sara Hillman on **07826 631112**

THANK YOU VERY MUCH FOR YOUR HELP

Date today: | | | |

DETAILS (number/address)

Preferred method of contact: Mobile telephone ☐
……...

Home/ Work (please delete) telephone ☐
…… ☐

Home/ Work (please delete) email

……..……..

Name: …… Date of birth:

(Partner's name: ……)

Q1. How old are you? ☐ ☐☐☐

Q2. Do you know what your weight is approximately?..... And your height?.....

Q3. Is this your first child with your current partner? ☐ no ☐ yes

Q3a. If no, how many children do you already have together? ☐

MEDICAL HISTORY

Q3. Do you have any medical problems? ☐ no ☐ yes

If yes:

i. Diabetes ☐ no ☐ yes Requiring insul ☐ no ☐ yes

ii. Heart diseas ☐ no ☐ yes

iii. Kidney diseas ☐ no ☐ yes

iv. Haematological problems ☐ no ☐ yes

v. Other - please state:
…… ☐ no ☐ yes ……
…….

Q4. Are you taking any medications?
What:……..

FAMILY HISTORY

Q5a. Do you know how much you weighed when born?
……

Q5b. Were you called a 'small' baby?

no	yes	N --

Q5c. Were you born early? If so, when?.....

no	yes	N --

Q5d. Were you a twin?

no	yes	N --

Q6. Do you recall if your mother had high blood pressure in her pregnancy /pregnancies?

no	yes

Q7. Did / do either of your parents have significant health problems?

no	yes

If yes, can you tell us whom and what?
……

.....
.....

LIFESTYLE

Q8. Have you ever smoked?

no	yes

If stopped for how long and how many were you smoking?.....

If yes, How many / day

0-5	5-	10-	20

Q9. Do you drink alcohol?

no	yes

If yes, approximately how many glasses*/ week

1 glass	1-6 glasses	7-20 glasse	21-69 glasse	70 +

*By glass we mean a pub measure of spirits, half a pint of lager or cider, a wine glass of wine, etc

Q10. If this is not your first child with this partner, has there been any significant change in your lifestyle between children i.e. smoking, exercise, development of disease or significant weight gain or loss?

……
……
……

Q11. Anything else you would like to share that may be relevant?

……
……
……

……
…… **THANK YOU**

S9 SUBJECT QUESTIONNAIRE

REFERENCES

Adkins RM, Tylavsky FA, Krushkal J. Newborn umbilical cord blood DNA methylation and gene expression levels exhibit limited association with birth weight. Chem Biodivers. 2012 May;9(5):888-99

Altshuler D, Hirschhorn JN, Klannemark M et al. The common PPARgamma Pro12Ala polymorphism is associated with decreased risk of type 2 diabetes. Nat Genet 2000: 26: 76–80

Andersson EA, Pilgaard K, Pisinger C, Harder MN, Grarup N, Færch K, Sandholt C, Poulsen P, Witte DR, Jørgensen T, Vaag A, Pedersen O, Hansen T. Do gene variants influencing adult adiposity affect birth weight? A population-based study of 24 loci in 24 744 Danish individuals. PloS One 2010; 5:e14190.

Andersson EA, Pilgaard K, Pisinger C, et al. Do gene variants influencing adult Balkau B, Charles MA. Comment on the provisional report from the WHO consultation. European Group for the Study of Insulin Resistance (EGIR). Diabetic Medicine 1999; 16: 442-3.

Bamshad MJ, Ng SB, Bigham AW, Tabor HK, Emond MJ, Nickerson DA and Shendure J. Exome Sequencing as a tool for MEndelian disease gene discovery. Nature Genetics 2011; 1 2: 745-755)

Bannister CE, Koestler DC, Maccani MA, Padbury JF, Houseman EA, Marsit CJ. 2011. Infant growth restriction is associated with distinct patterns of DNA methylation in human placentas. Epigenetics 6:920–927.

Barg E, Szopa J, Pesz KA, Gąsiorowski K. Indices of insulin resistance and dyslipidemia are correlated with lymphocyte proneness to apoptosis in obese or overweight low birth weight children. Horm Res Paediatr. 2013;79(5):293-9

Barker DJP, Hales CN, Fall CHD, Osmond C, Phipps K, Clark PMS. Type 2 diabetes mellitus, hypertension and hyperlipidaemia: relation to reduced fetal growth. Diabetologia 1993;36: 62-7

Bartoov-Shifman R, Hertz R, Wang HY, Wollheim CB, BarTana J, Walker MD (2002) Activation of the insulin gene promoter through a direct effect of hepatocyte nuclear factor 4alpha. J Biol Chem 277:25914-25919

Beck-Nielsen H, Groop LC.Metabolic and genetic characterization of prediabetic states. Sequence of events leading to non-insulin-dependent diabetes mellitus. J Clin Invest. 1994 Nov;94(5):1714-21

Bejamini and Hochburg, 1995 Controlling the False Discovery Rate: a practical and powerful approach to multiple testing. Journal of the Royal Statistical Scoiety, Series B57 (1): 289-300)

Bellamy L, Casas JP, Hingorani AD, Williams D. Type 2 diabetes mellitus after gestational diabetes: a systematic review and meta-analysis. Lancet 2009;

Bennett AJ, Sovio U, Ruokonen A, Martikainen H, Pouta A, Hartikainen AL,Franks S, Elliott P, Järvelin MR, McCarthy MI: No evidence that established type 2 diabetes susceptibility variants in the *PPARG* and *KCNJ11* genes have pleiotropic effects on early growth. Diabetologia 2008, 51(1):82-85.

Beyer S, Maag Kristensen M, Steen Jensen K, Vilstrup Johansen J and Staller P. The Histone Demethylases JMJD1A and JMJD2B are Transcriptional Targets of Hypoxia-inducible Factor HIF. *J. Biol. Chem. 2008, 283:36542-36552*

Bird A. Perceptions of epigenetics. Nature 2007; 447:396–398.

Bodsky D and Christou H. Current Concepts in Intrauterine Growth Restriction. Journal of Intensive Care Medicine 2004: 19; 307-319

Bonnefond A, Froguel P, Vaxillaire M (2010) The emerging genetics of type 2 diabetes. Trends Mol Med 16: 407–416.

Bonnefond A, Philippe J, Durand E, Dechaume A, Huyvaert M, Montagne L, Marre M, Balkau B, Fajardy I, Vambergue A, Vatin V, Delplanque J, Le Guilcher D, De Graeve F, Lecoeur C, Sand O, Vaxillaire M, Froguel P. Whole-exome sequencing and high throughput genotyping identified KCNJ11 as the thirteenth MODY gene. PLoS One. 2012;7(6):e37423

Bonora E, Formentini G, Calcaterra F, Lonbardi S et al. HOMA-estimated insulin resistance is an independent predictor of cardiovascular disease in type 2 diabetic subjects:prospective data from the Verona Diabetes Complication Study. Diabetes Care 2002; 25: 1134-41

Bonora E, Targher G, Alberichie M, Bonadonna RC, Saggianni F, Zenere MB, Monauni T, Muggeo M (2000): Homeostasis model assessment closely mirrors the glucose clamp technique in the assessment of insulin sensitivity. Diabetes Care 23: 57-63

Böttcher RT, Niehrs C. Fibroblast growth factor signaling during early vertebrate development. Endocr Rev. 2005 Feb;26(1):63-77

Bourque DK, Avila L, Penaherrera M, von Dadelszen P and Robinson WP Decreased placental methylation at the H19/IGF2 imprinting control region is associated with normotensive intrauterine growth restriction but not preeclampsia. Placenta 31: 197-202, 2010.

Bowman P, Flanagan SE, Edghill EL, Damhuis A, Shepherd MH, et al. (2012) Heterozygous ABCC8 mutations are a cause of MODY. Diabetologia 55: 123–127.

Brissova M, Blaha M, Spear C, Nicholson W, Radhika A, Shiota M, Charron MJ, Wright CV, Powers AC. Reduced Pdx1 expression impairs islet response to insulin resistance and worsens glucose homeostasis. Am J Physiol Endo Met 2005; 288: E707-E714

Bukowski R, Davis KE, Wilson PW. Delivery of a small for gestational age infant and greater maternal risk of ischemic heart disease. PLoS One. 2012;7(3):e33047

Burton GJ & Jauniaux E. Oxidative stress. Best Pract Res Clin Obstet Gynaecol. 2011 Jun;25(3):287-99

Cameron AJ, Boyko EJ, Sicree RA et al. Central obesity as a precursor to the metabolic syndrome in the AudDiab study and Mauritius. Obesity 2008;16: 2707-2716.

Carter AM. Placental oxygen consumption. Part I: in vivo studies— a review. Placenta 2000;21:S31– S37

Caterson ID, Finer N, Coutinho W, Van Gaal LF et al Maintained intentional weight loss reduces cardiovascular outcomes: results from the Sibutramine Cardiovascular OUTcomes (SCOUT) trial. Diabetes Obes Metab. 2012; 14: 523-530

Celemajer D, Sorensen KE, Gooch VM, Spiegelhalter DJ, Miller OI, Sullivan ID. Non invasive detection of endothelial dysfunction in children and adults at risk of atherosclerosis. Lancet 1992; 340: 1111-1115

Chae JS, Kang R, Kwak JH, Paik JK, Kim OY, Kim M, Park JW, Jeon JY, Lee JH. Supervised exercise program, BMI, and risk of type 2 diabetes in subjects with normal or impaired fasting glucose. Diabetes Care. 2012 Aug;35(8):1680-5

Chaunn and Magnn 2006 Chauhan SP, Magann EF. Screening for fetal growth restriction. Clin Obstet Gynecol. 2006 Jun;49(2):284-94.

Chelbi1Sonia T, Melissa L. Wilson, Anne-Cle´mence Veillard, Sue A. Ingles, Jim Zhang, Franc¸oise Mondon, Ge´raldine Gascoin-Lachambre, Ludivine Doridot, The´re`se-Marie Mignot, Re´gis Rebourcet,, Bruno Carbonne, Jean-Paul Concordet, Sandrine Barbaux and Daniel Vaiman. Genetic and epigenetic mechanisms collaborate to control SERPINA3 expression and its association with placental diseases, Human Molecular Genetics, 2012, Vol. 21, No. 9

Christensen BC, Houseman EA, Marsit CJ, Zheng S, Wrensch MR, Wiemels JL, et al. Aging and environmental exposures alter tissue-specific DNA methylation dependent upon CpG island context. PLoS Genet 2009;5(8

Coleman RA, Grob P, King DS, Florens L, Washburn MP, et al. Structural changes in TAF4b-TFIID correlate with promoter selectivity. Mol Cell 2008; 29:81-91

Confidential Enquiry into Maternal and Child Health (CEMACH). *Perinatal Mortality 2007: United Kingdom*. CEMACH: London, 2009 [http://www.cmace.org.uk/getattachment/1d2c0ebc-d2aa-4131-98ed 56bf8269e529/Perinatal-Mortality-2007.aspx]

Constância M, Hemberger M, Hughes J, Dean W, Ferguson-Smith A, Fundele R, Stewart F, Kelsey G, Fowden A, Sibley C, Reik WPlacental-specific IGF-II is a major modulator of placental and fetal growth. Nature. 2002 Jun 27;417(6892):945-8

Crossey PA, Pillai CC, Miell JP. Altered placental development and intrauterine growth restriction in IGF binding protein-1 transgenic mice. J Clin Invest. 2002; 110:411-418

Cruz-Lemini M, Crispi F, van Mieghem T et al. Risk of perinatal death in early-onset intrauterine growth restriction according to gestational age and cardiovascular Doppler indices: a multicenter study. Fetal Diagn Ther 2012; 32: 116-122.

Cui Y, Wang W, Dong N, Lou J et al. Role of corin in trophoblast invasion and uterine spiral artery remodelling in pregnancy. Nature 2012; 484; 246-250

Cutfield WS, Hofman PL, Mitchell M, Morison IM: Could epigenetics play a role in the developmental origins of health and disease? Pediatr Res 2007, 61:68R-75R

Dancause KN, Laplante DP, Oremus C, Fraser S, Brunet A, King S. Disaster related prenatal maternal stress influences birth outcomes: project ice storm. Early Hum Dev 2011;87(12)],

Davey Smith G, Hypponen E, Power C, Lawlor D. Offspring birth weight and parental mortality: Prospective observational study and meta-analysis. Am J Epidemiol 2007; 166: 160-169

Davey Smith G, Sterne JAC, Tynelius, Rasmussen F. Birth characteristics of offspring and parental diabetes: evidence for the fetal insulin hypothesis. J Epidemiol Community Health 2004; 58: 126-128

Dayeh TA, Olsson AH, Volkov P, Almgren P, Rönn T, Ling C. Identification of CpG-SNPs associated with type 2 diabetes and differential DNA methylation in human pancreatic islets. Diabetologia. 2013 May;56(5):1036-46

Divon MY, Haglund B, Nisell H et al. Fetal and neonatal mortality in the post-term pregnancy: the impact of gestational age and fetal growth restriction. Am J Obstet Gynecol 1998; 178: 726-731

Du P, Zhang X, Huang C-C, Jafari N, Kibbe WA, Hou L et al. Comparison of Beta-value and M-value methods for quantifying methylation levels by microarray analysis. Bioinformatics 2010;11:587

Dupuis J, Langenberg C, Prokopenko I, Saxena R, Soranzo N, et al. (2010) New genetic loci implicated in fasting glucose homeostasis and their impact on type 2 diabetes risk. Nat Genet 42: 105–116.

Edghill EL, Bingham C, Slingerland AS, Minton JA, Noordam C, Ellard S,Hattersley AT: Hepatocyte nuclear factor-1 beta mutations cause neonatal diabetes and intrauterine growth retardation: support for a critical role of HNF-1beta in human pancreatic development. *Diabet Med* 23:1301–1306, 2006

Edghill EL, Flanagan SE, Ellard S. Permanent neonatal diabetes due to activating mutations in ABCC8 and KCNJ11. Rev Endocr Metab Disord. 2010 Sep;11(3):193-8

Einstein F, Thompson RF, Bhagat TD, Fazzari MJ, Verma A, Barzilai N, Greally JM. 2010. Cytosine methylation dysregulation in neonates following intrauterine growth restriction. PLoS One 5:e8887

Ek J, Andersen G, Urhammer SA, Hansen L, Carstensen B, Broch-Johnsen K, Drivsholm T, Berglund L, Hansen T, Lithell H, Pedersen O: Studies of the Pro12Ala polymorphism of the peroxisome proliferator-activated receptor-_2 (PPAR-_2) gene in relation to insulin sensitivity among glucose tolerant Caucasians. *Diabetologia* 44:1170–1176, 2001

Ellard S, Bellanne-Chantelot C and Hattersley AT. Best practice guidelines for the molecular genetic diagnosis of maturity-onset diabetes of the young. Diabetologia 2008 51: 546–553

Engel N, West AG, Felsenfeld G, Bartolomei MS. Antagonism between DNA hypermethylation and enhancer-blocking activity at the H19 DMD is uncovered by CpG mutations. Nat Genet. 2004 Aug; 36(8):883-8

Eriksson J, Lindström J, Valle T, Aunola S et al. Prevention of Type II diabetes in subjects with impaired glucose tolerance: the Diabetes Prevention Study (DPS) in Finland. Study design and 1-year interim report on the feasibility of the lifestyle intervention programme. Diabetologia. 1999; 42:793-801

Eriksson JG, Lindi V, Uusitupa M, Forsen TJ, Laakso M, Osmond C, Barker DJ: The effects of the Pro12Ala polymorphism of the peroxisome proliferator-activated receptor-gamma2 gene on insulin sensitivity and insulin metabolism interact with size at birth. *Diabetes* 51:2321–2324, 2002

Eva-Maria D. Nielsen, Lars Hansen, Bendix Carstensen, Søren M. Echwald, Thomas Drivsholm, Charlotte Glu¨mer, Birger Thorsteinsson, Knut Borch-Johnsen, Torben Hansen, and Oluf Pedersen The E23K Variant of Kir6.2 Associates With Impaired Post–OGTT Serum Insulin Response and Increased Risk of Type 2 Diabetes. Diabetes. 2003 Feb;52(2):573-7.

Fajans SS, Bell GI (2011) MODY: history, genetics, pathophysiology, and clinical decision making. Diabetes Care 34: 1878–1884

Fajans SS, Bell GI and Polonsky KS. Molecular mechanisms and clinical pathophysiology of Maturity onset diabetes of the young. New England Journal Medicine (2001); 345: 13,971-984

Feero WG, Guttmacher AE, McCarthy MI (2010) Genomics, Type 2 Diabetes, and Obesity. New England Journal of Medicine 363: 2339–2350.

Feinberg AP. Epigenetics at the Epicenter of Modern Medicine. JAMA 2008; 299: 1345-1350

Florez JC, Jablonski KA, McAteer J, Sandhu MS, Wareham NJ, Barroso I, Franks PW, Altshuler D, Knowler WC.Testing of diabetes-associated WFS1 polymorphisms in the Diabetes Prevention Program. Diabetologia. 2008; 51(3): 451-457

Fowden AL, Sibley C, Reik W, Constancia M. Imprinted genes, placental development and fetal growth. Horm Res 2006; 65:50–8.7

Frankel S, Elwood P, Sweetnam P, Yarnell J, Davey Smith G. Birthweight, body mass index in middle age and incident coronary heart disease. The Lancet 1997; 348: 1478-1480

Frayling TM, Hattersley AT. The role of genetic susceptibility in the association of low birth weight with type 2 diabetes. Br Med Bull. 2001;60:89-101.

Freathy RM, Bennett AJ, Ring SM, Shields B, Groves CJ, Timpson NJ, Weedon MN, Zeggini E, Lindgren CM, Lango H, Perry JR, Pouta A, Ruokonen A, Hyppönen E, Power C, Elliott P, Strachan DP, Järvelin MR, Smith GD, McCarthy MI, Frayling TM, Hattersley AT. Type 2 diabetes risk alleles are associated with reduced size at birth. Diabetes. 2009; 58, 1428–1433

Freathy RM, Mook Kanamori DO, Sovio U, Prokopenko I et al. Variants in ADCY5 and near CCNL1 are associated with fetal growth and birth weight. Nature Genetics May 2010; 42 (5) 430-5

Freathy RM, Weedon MN, Bennett A, Hypponen E, Relton CL, Knight B, Shields B, Parnell KS, Groves CJ, Ring SM, Pembrey ME, Ben-Shlomo Y, Strachan DP, Power C, Jarvelin MR, McCarthy MI, Davey Smith G, Hattersley AT, Frayling TM. Type 2 diabetes TCF7L2 risk genotypes alter birth weight: a study of 24,053 individuals. Am J Hum Genet. 2007 Jun; 80(6):1150-61

Froen JF, Gardosi JO, Thurman A et al. Restricted fetal growth in sudden intrauterine unexplained death. Acta Obstet Gynaecol Scand 2004; 83:801-807

Froguel P, Zouali H, Vionnet N, et al. Familial hyperglycemia due to mutations in glucokinase. Definition of a subtype of diabetes mellitus. N Engl J Med 1993; 328:697–702

Gama-Sosa MA,WangRY,KuoKC, GehrkeCW,EhrlichM.The5-methylcytosine content of highly repeated sequences in human DNA. Nucleic Acids Res 1983; 11(10):3087

Giovannucci E. Alcohol, 1-carbon metabolism and colorectal cancer: recent insights from molecular studies. J Nutr 2004; 134 (9): 2475S-2481S

Godfrey KM, Barker DJ. Fetal nutrition and adult disease. Am J Clin Nutr 2000; 71:1344-52.

Godfrey KM, Barker DJ. Maternal birthweight and diet in pregnancy in relation to the infant's thinness at birth. Br J Obstet Gynaecol 1997;104: 663-7

Grarup N, Sparsø T, Hansen T (2010) Physiologic Characterization of Type 2Diabetes–Related Loci. Current Diabetes Reports 10: 485–497

Guo L, Choufani S, Ferreira J, Smith A, Chitayat D, Shuman C, et al. Altered gene expression and methylation of the human chromosome 11 imprinted region in small for gestational age (SGA) placentae. Developmental Biology 2008; 320:79-91

Gutierrez-Arcelus M, Lappalainen T, Montgomery SB, Buil A, Ongen H, Yurovsky A, Bryois J, Giger T, Romano L, Planchon A, Falconnet E, Bielser D, Gagnebin M, Padioleau I, Borel C, Letourneau A, Makrythanasis P, Guipponi M, Gehrig C, Antonarakis SE, Dermitzakis ET. Passive and active DNA methylation and the interplay with genetic variation in gene regulation. Elife. 2013 Jun 4;2:e00523

Hackett GA, Campbell S, Gamsu H et al. Doppler studies in the growth retarded fetus and predicted necrotising enterocolitis, haemorrhage and neonatal morbidity. Br Med J 1987; 294:13-16

Haffner SM, Kennedy E, Gonzalez C, Stern MP, Meittinen H (1996). Aprospective analysis of the HOMA model: the Mexico City Diabetes Study. Diabetes Care 19: 1138-1141

Hales CN, Barker DJP, Clark PMS et al. Fetal and infant growth and impaired glucose tolerance at age 64 years. BMJ 1991; 303: 1019-1022

Hamad RR, Eriksson MJ, Silveira A, Hamsten A, Bremme K. Decreased flow-mediated dilation is present 1 year after a pre-eclamptic pregnancy. J Hypertens. 2007 Nov;25(11):2301-7

Hampsey M. Molecular genetics of the RNA polymerase II general transcriptional machinery. Microbiol Mol Biol Rev 1998; 62:465-503

Hani EH, Boutin P, Durand E, Inoue H, Permutt MA, Velho G, Froguel P: Missense mutations in the pancreatic islet beta cell inwardly rectifying K_channel gene (KIR6.2/BIR): a meta-analysis suggests a role in the polygenic basis of type II diabetes mellitus in Caucasians. *Diabetologia* 41:1511–1515, 1998

Hanley AJG, Stern MP, Williams K, Haffner SM. Homeostasis model assessment of insulin resistance in relation to the incidence of cardiovascular disease. Diabetes Care 2002; 25: 1177-1184.

Hansen SK, Rose CS, Glumer C, Drivsholm T, Borch-Johnsen K, Jorgensen T, Pedersen O, Hansen T: Variation near the hepatocyte nuclear factor (HNF)-4alpha gene associates with type 2 diabetes in the Danish population.. *Diabetologia* 48:452–458, 2005

Hattersley AT, Tooke JE. The fetal insulin hypothesis: an alternative explanation of the association of low birth weight with diabetes and vascular disease. Lancet 1999; 353: 1789-92

HattersleyAT, Beards F, ballantyne E, Applton M, Harvey R, Ellard S. Mutations in the gluckokinase gene of the fetus result in reduced birth weight. Nat Genetics 1998; 19:268-70

Haycock PC, Ramsay M. Exposure of Mouse Embryos to Ethanol During Preimplantation Development: Effect on DNA Methylation in the H19 Imprinting Control Region. Biol Reprod. 2009

Hegele RA, Cao H, Harris SB, Hanley AJ, Zinman B: The hepatic nuclear factor-1alpha G319S variant is associated with early-onset type 2 diabetes in Canadian Oji-Cree. *J Clin Endocrinol Metab* 84:1077–1082, 1999

Heijmans BT, Tobi EW, Stein AD, Putter H, Blauw GJ, Susser ES, Eline Slagboom P and Lumey LH. PErsistnet epigenetic difference associated with prenatal exposure to famine in humans. PNAS 2008; 105: 17046-17049

Hiby SE, Regan L, Lo W, Farrell L, Carrington M, Moffett A. Association of maternal killer-cell immunoglobulin-like receptors and parental HLA-C genotypes with recurrent miscarriage. Hum Reprod. 2008 Apr;23(4):972-6.

Hiby SE, Walker JJ, O'shaughnessy KM, Redman CW, Carrington M, Trowsdale J, Moffett A. Combinations of maternal KIR and fetal HLA-C genes influence the risk of preeclampsia and reproductive success. J Exp Med. 2004 Oct 18;200(8):957-65

Hirofumi T, Matsumot C, Yamada J, et al. The relationships of cardiovascular disease risk factors to flow mediated dilatation in Japanese subjects free of cardiovascular disease. Hypertension Research 2008; 31: 2019-2025

Holmans P, Green EK, Pahwa JS, Ferreira MA, Purcell SM, Sklar P; Wellcome Trust Case-Control Consortium, Owen MJ, O'Donovan MC, Craddock N. Gene ontology analysis of GWA study data sets provides insights into the biology of bipolar disorder. Am J Hum Genet. 2009 Jul;85 (1):13-24

Holmkvist J, Cervin C, Lyssenko V, Winckler W, Anevski D, Cilio C, Almgren P, Berglund G, Nilsson P, Tuomi T, Lindgren CM, Altshuler D, Groop L: Common variants in *HNF-1a* and risk of type 2 diabetes. *Diabetologia* 49:2882–2891, 2006

Horbar JD, Badger GJ, Ohlsson A, Golan A. Morbidity and mortality among very-low-birth-weight neonates with intrauterine growth restriction. The Vermont Oxford Network. American Journal of Obstetrics and Gynaecology 2000; 182:198-206).

Horikawa Y, Iwasaki N, Hara M, Furuta H, Hinokio Y, et al. (1997) Mutation in hepatocyte nuclear factor-1 beta gene (TCF2) associated with MODY. Nat Genet 17: 384–385

Hung TH, Chen SF, Lo LM, Li MJ, Yeh YL, Hsieh TT. Increased autophagy in placentas of intrauterine growth-restricted pregnancies. PLoS One. 2012;7(7):e40957

Hypponen E, Davey Smith G, Power C. Parental diabetes and birth weight of offspring: intergenerational cohort. BMJ 2003; 326: 19-20

Inoue H, Tanizawa Y, Wasson J, Behn P, Kalidas K, Bernal-Mizrachi E, Mueckler M, Marshall H, Donis-Keller H, Crock P, Rogers D, Mikuni M, Kumashiro H, Higashi K, Sobue G, Oka Y, Permutt MA. A gene encoding a transmembrane protein

is mutated in patients with diabetes mellitus and optic atrophy (Wolfram syndrome). Nat Genet. 1998 Oct;20(2):143-8.

Ishihara N, MatsuoH, Murakushi H et al. Increased apoptosis in the syncytiotrophoblast in human term placentas complicated by either pre-eclampsia or intrauterine growth restriction. Am J OBstet Gynecol 2002; 186: 158-166

Jaenisch R. DNA methylation and imprinting: why bother? Trends Genet 1997;13:323–329

Jaquet D, Tregouet DA, Godefroy T, Nicaud V, Chevenne D, Tiret L, Czernichow P, Levy-Marchal C: Combined effects of genetic and environmental factors on insulin resistance associated with reduced fetal growth. *Diabetes* 51:3473–3478, 2002

Jauniaux E, Poston L, Burton GJ: Placental-related diseases of pregnancy: Involvement of oxidative stress and implications in human evolution. Hum Reprod Update 2006, 12:747-755.

Jirtle RL and Skinner MK. Environmental epigenomics and disease susceptibility. Nature Reviews Genetics. 2007; 8: 253-261

Johansson S, Irgens H, Chudasama KK, Molnes J, Aerts J, Roque FS, Jonassen I, Levy S, Lima K, Knappskog PM, Bell GI, Molven A, Njølstad PR Exome sequencing and genetic testing for MODY. PLoS One. 2012;7(5):e38050

Jones RA, Roberton NR. Small for dates babies: are they really a problem? *Arch Dis Child* 1986;61:877–80

Kafri T, Ariel M, Brandeis M, Shemer R, Urven L, McCarrey J, Cedar H, Razin A. Developmental pattern of gene-specific DNA methylation in the mouse embryo and germ line. Genes Dev 1992;6:705–714.

Kim TG, Kraus JC, Chen J, Lee Y (2004). JUMONJI, a critical factor for cardiac development, functions as a transcriptional repressor. J. Biol. Chem. 278 (43): 42247–55)

Knight B, Shields BM, Hill A et al. Offspring birthweight is not associated with paternal insulin resistance. Diabetologia 2006; 49: 2675-2678

Kobberling J, Tillili H. Empirical risk figures for first degree relatives of non-insuln dependent diabetics. In Kobberling J, Tattersall R (eds) The Genetics of diabetes Mellitus. London

Koon HB, Ippolito GC, Banham AH, Tucker PW. FOXP1: a potential therapeutic target in cancer. Expert Opin Ther Targets. 2007 Jul;11(7):955-65.
Lackman F, Capewell V, Richardson B, daSilva O and Gagnin R. The risks of spontaneous preterm delivery and perinatal mortality in relation to size at birth according to fetal versus neonatal growth standards. Am J Obstet Gynecol 2001; 184:946-953

Laplante DP, Barr RG, Brunet A, Galbaud du Fort G, Meaney ML, Saucier JF, et al. Stress during pregnancy affects general intellectual and language functioning in human toddlers. Pediatr Res 2004;56(3

Lenderman HM. Is maturity onset diabetes at young age (MODY) more common in Europe than previously assumed? Lancet 1995; 345: 648

Leonard S, Murrant C, Tayade C, van den Heuvel M, Watering R, Croy BA. Mechanisms regulating immune cell contributions to spiral artery modification -- facts and hypotheses -- a review. Placenta. 2006 Apr;27 Suppl A:S40-6.

Levy JC, Matthews DR and Herrmans MP (1998). Correct homeostasis model assessment (HOMA) evaluation uses the computer program. Diabetes Care 21 (12): 2191-2

Li H. and Durbin R. (2009). Fast and accurate short read alignment with Burrows-Wheeler Transform. *Bioinformatics*, 25:1754-60.

Lindsay RS, Dabela D, RoumainJ et al. Type 2 diabetes and low birth weight: the role of paternal inheritance in the association of low birth weight and diabetes. Diabetes 2000; 49: 445-9

Ling C, Poulsen P, Simonsson S et al (2007) Genetic and epigenetic factors are associated with expression of respiratory chain component NDUFB6 in human skeletal muscle. J Clin Invest 117:3427–3435

Lithell HO, McKeigue PM, Berglund L, Mohsen R, Lithell UB and Leon DA. Relation of size at birth to non insulin dependent diabetes and insulin concentrations in men aged 50-60 years. BMJ 1996; 312: 406-410

Longo S, Bollani L, Decembrino L, Di Comite A, Angelini M, Stronati M. Short-term and long-term sequelae in intrauterine growth retardation (IUGR). J Matern Fetal Neonatal Med. 2013 Feb;26(3):222-5

Losick, Richard; Watson, James D.; Tania A. Baker; Bell, Stephen; Gann, Alexander; Levine, Michael W. (2008). *Molecular biology of the gene*. San Francisco: Pearson/Benjamin Cummings

Lotfi K, Sund G, Lowe R, Graham J, Landin-Olsson M, Kockum I, Deeb S, Lernmark A: The beta cell glucokinase promoter variant is an unlikely risk factor for diabetes mellitus: Diabetes Incidence Study in Sweden (DISS). *Diabetologia* 40:959 –962, 1997

Love-Gregory LD, Wasson J, Ma J et al (2004) A common polymorphism in the upstream promoter region of the hepatocyte nuclear factor-4α; gene on chromosome 20q is associated with type 2 diabetes and appears to contribute to the evidence for linkage in an Ashkenazi Jewish population. Diabetes 53:1134–1140

Lumey LH. Decreased birth weight in infants after maternal in utero exposure to the Dutch famine of 1944—1945. Paediatr Perinat Epidemiol 1992;6:240– 53

Lunde A,Melve KK, GjessingHK, et al.Genetic and environmental influences on birth weight, birth length, head circumference, and gestational age by use of population-based parent-offspring data. Am J Epidemiol 2007; 165:734–741

Macaulay EC, Weeks RJ, Andrews S, Morison IM. Hypomethylation of functional etrotransposon-derived genes in the human placenta. Mamm Genome 2011;22(11e12):722e35

Maccani MA, Marsit CJ. Epigenetics in the placenta. Am J Reprod Immunol 2009;62(2):78e89.]

Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorff LA, Hunter DJ, McCarthy MI, Ramos EM, Cardon LR, Chakravarti A, Cho JH, Guttmacher AE, Kong A, Kruglyak L, Mardis E, Rotimi CN, Slatkin M, Valle D, Whittemore AS, Boehnke M, Clark AG, Eichler EE, Gibson G, Haines JL, Mackay TF, McCarroll SA, Visscher PM. Finding the missing heritability of complex diseases. Nature. 2009 Oct 8;461(7265):747-53

Mardis ER. Next Generation DNA Sequencing methods.. Annu. Rev. Genomics Hum Genet. 2008; 9:387-402)

Masud S, Ye S: Effect of the peroxisome proliferator activated receptorgamma gene Pro12Ala variant on body mass index: a meta-analysis. *J Med Genet* 40:773–780, 2003

Matthews DR, Hosker JP, Rudenski AS, Naylor BA, Treacher DF, Turner RC: Homeostasis model assessment: insulin resistance and beta-cell function from fasting plasma glucose and insulin concentrations in man. *Diabetologia* 28:412–419, 1985

Maulik D, EvansJF andRagolla L. Fetal growth restrictions: pathogenic mechanisms. Clinical Obstetric and Gynecol 2006;49: 219-227

Maulik D. Fetal Growth Restriction: The Etiology. Clinical Obstetrics and Gynaecology 2006; 49: 228-235

Mayhew TM, Ohadike C, Baker PN, Crocker IP, Mitchell C, Ong SS. Stereological investigation of placental morphology in pregnancies complicated by pre-eclampsia with and without intrauterine growth restriction. Placenta 2003; 24:219–2650

McCarthy M (2010) Genomics, type 2 diabetes, and obesity. N Engl J Med 363:2339–2350

McCowan LM, North RA, Kho EM, Black MA et al. Paternal contribution to small for gestational age babies: a multicenter prospective study. Obesity 2011; 19: 1035-9

McIntire DD, Bloom SL, Casey BM and Leveno K. Birth weight in relation to morbidity among newborn infants. N Engl J Med 1999; 340: 1234-1238

McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, Garimella K, Altshuler D, Gabriel S, Daly M, DePristo MA (2010). The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res*. 20(9):1297-303.

McMaster University Evidence Based Practice Center. Diagnosis, prognosis and treatment of impaired glucose tolerance and impaired fasting glucose. Evidence Report 128

McMinn J, Wei M, Schupf N, Cusmai J, Johnson EB, Smith AC, et al. Unbalanced placental expression of imprinted genes in human intrauterine growth restriction. Placenta 2006;27:540–9

Meglasson MD, Matschinsky FM. Pancreatic islet glucose metabolism and regulation of insulin secretion. Diabetes/ Metabolism Reviews 1986:2: 163-214

Meschia G. Placenta respiratory gas exchange and fetal oxygenation. In: Creasy RK, Resnik R, editors. Maternal Fetal Medicine: Principles and Practice, 1st edition. Philadelphia: WB Saunders, 1987: 274–285

Molteni RA, Stys SJ, Battaglia FC. Relationship of fetal and placental weight in human beings: fetal/placental weight ratios at various gestational ages and birth weight distributions. J Reprod Med 1978;21:327– 334

Mongelli M & Gardosi J. Reduction of false positive diagnosis of fetal growth restriction by application of customized fetal growth standards. Obstet and Gynae 1996; 88:844-8

Monk M, Boubelik M, Lehnert S. Temporal and regional changes in DNA methylation in the embryonic, extraembryonic and germ cell lineages during mouse embryo development. Development 1987;99:371–382.

Morgan AR, Thompson JM, Murphy R, Black PN, Lam WJ, Ferguson LR, Mitchell EA. Obesity and diabetes genes are associated with being born small for gestational age: results from the Auckland Birthweight Collaborative study. BMC Med Genet. 2010 Aug 16;11:125

Morgan, H. D., Sutherland, H. G., Martin, D. I. & Whitelaw, E. Epigenetic inheritance at the agouti locus in the mouse. Nature Genet. 23 , 314–318 1999

Morison IM, Reeve AE. A catalogue of imprinted genes and parent-of-origin effects in humans and animals. *Hum Mol Genet 1998*; *7*: *1599*–609 *Nature Genetics 2009; 42: 105-115*

Mosley AL, Corbett JA, Ozcan S. Glucose regulation of insulin gene expression requires the recruitment of p300 by the beta-cell-specific transcription factor Pdx-1. Mol Endocrinol. 2004 Sep; 18(9):2279-90.

Murrell A, Heeson S, Reik W. Interaction between differentially methylated regions partitions the imprinted genes Igf2 and H19 into parent-specific chromatin loops. Nat Genet 2004; 36:889-93.

Nakamura K, Sakurai M, Miura K et al. Homeostasis model assessment of insulin resistance and the risk of cardiovascular events in middle aged non-diabetic Japanese men. Diabetologia 2010; 53: 1894-1902.

Naylor CD, Sermer M, Chen E, Sykora K. Caesarean section delivery in relation to birth weight and gestational glucose tolerance pathophysiology or practice style? Toronto Trihospital Gestational Diabetes Investigators. JAMA 1996; 275: 1165-70

Naylor R, Philipson LH. Who should have genetic testing for maturity-onset diabetes of the young? Clin Endocrinol (Oxf). 2011 Oct;75(4):422-6.
Neerhof MG. causes of intrauterine growth restriction. Clin Perinatol 1995;22:375-85

Nelissen EC, van Montfoort AP, Dumoulin JC, Evers JL. Epigenetics and the placenta. Hum Reprod Update 2011;17(3):397e417.10,11

Newman B, Selby JV, King MC, Slemenda C, Fabsitz R, Friedman GD. Concordance for type 2 diabetes mellitus in male twins. Diabetologia 1987; 30: 763-8Academic Press, 1982;201-210

Ng PC, Henikoff S. SIFT: Predicting amino acid changes that affect protein function. Nucleic Acids Res. 2003 Jul 1;31(13):3812-4.

Ng SF, LinRC, Laybutt DR, Barres R et al. Chronic high fat diet in fathers programs B cell dysfunction in female rat offspring. Nature 2010; 467(7318): 963-967

Noori M, Donald AE, Angelakopoulou A, Hingorani AD and Williams DJ. A Prospective Study of Placental Angiogenic Factors and Maternal Vascular Function Before and After Preeclampsia and Gestational Hypertension. *Circulation.* 2010;122:478-487

Novakovic B, Saffery R. The ever growing complexity of placental epigenetics - role in adverse pregnancy outcomes and fetal programming. Placenta. 2012 Dec;33(12):959-70

Novakovic B, Wong NC, Sibson M, Ng HK, Morley R, Manuelpillai U, et al. DNA methylation-mediated down-regulation of DNA methyltransferase-1(DNMT1) is coincident with, but not essential for, global hypomethylation in human placenta. J Biol Chem 2010;285(13):9583e93

Novakovic B, Yuen RK, Gordon L, Penaherrera MS et al. Evidence for widespread changes in promoter methylation profile in human placenta in response to increasing gestational age and environmental/stochastic factors. BMC Genomics 2011 Oct 28;12:529

Odom DT, Zizlsperger N, Gordon DB, Bell GW, Rinaldi NJ, Murray HL, Volkert TL, Schreiber J, Rolfe PA, Gifford DK, Fraenkel E, Bell GI, Young RA: Control of pancreas and liver gene expression by HNF transcription factors. *Science* 303:1378 – 1381, 2004

Okorie MI, Bhavsar DD, Ridout D, Charakida M, Deanfield JE, Loukogeorgakis SP, MacAllister RJ. Postconditioning protects against human endothelial ischaemia-reperfusion injury via subtype specific KATP channel activation and is mimicked by inhibition of the mitochondrial permeability transition pore. Eur Heart J. 2011; 32(10): 1266-74

Pallotto EK, Woelnerhanssen B, Putt M et al. Hypoglycaemia of the Newborn. In: Burg FD, Ingelfinger JR, Polin RA, Gersho AA eds. Gellis and Kagan's current paediatric therapy Philadelphia: WB Saunders Co; 2002

Paradisi G, Mattoli MV, Tomei C, Zuppi C, Lulli P, Quagliozzi L, Caruso A. Cardiovascular risk factors in healthy women with previous small for gestational age infants. J Obstet Gynaecol Res. 2011 Oct;37(10):1397-404

Park JH, Stoffers DA, Nicholls RD, Simmons RA. Development of type 2 diabetes following intrauterine growth retardation in rats is associated with progressive epigenetic silencing of *Pdx1*. J Clin Invest. 2008 June 2; 118(6): 2316–2324

Pedersen J. Diabetes and pregnancy: Blood sugar of newborn infants (Ph.D. Thesis) Danish Science Press; Copenhagen: 1952. p.230.

Philips DIW. Insulin resistance is a programmed response to fetal under-nutrition. Diabetologia 1996; 39: 1119-1122

Pijnenborg R, Dixon G, Robertson WB, Brosens I. Trophoblastic invasion of human decidua from 8 to 18 weeks of pregnancy. Placenta. 1980 Jan-Mar;1(1):3-19.

Piper JM, Xenakis EM, McFarland M et al. Do growth retarded premature infants have different rates of perinatal morbidity and mortality than appropriately grown premature infants? Obstet Gynecol 1996; 87:169-174

Pulizzi N, Lysseako V, Jonsson A, et al. Interaction between prenatal growth and high-risk genotypes in the development of type 2 diabetes. Diabetologia. 2009; 52, 825–829

Rakyan VK, Beck S Epigenetic variation and inheritance in mammals. Curr Opin Genet Dev. 2006 Dec; 16(6):573-7

Ramensky V, Bork P, Sunyaev S. Human non-synonymous SNPs: server and survey. Nucleic Acids Res. 2002;30:3894-3900

Razin A, Shemer R. DNA methylation in early development. Hum Mol Genet 1995;4 (Spec No): 1751–1755. [PubMed: 8541875]

RCOG; The Investigation and Management of the Small-for-Gestational-Age Fetus, Royal College of Obstretricians and Gynaecologists (2002)

Reddy KJ, Singh M, Bangit JR, Batsell RR. The role of insulin resistance in the pathogenesis of atherosclerotic cardiovascular disease: an updated review. J Cardiovascular Medicine 2010; 11(9): 633-647

Reeves S and Bernstein I. Effects of maternal tobacco smoke exposure on fetal growth and maternal size. Expert review Obstet Gynecol 2008; 3: 719-730

Roseboom T, de Rooij S, Painter R. The Dutch famine and its long-term consequences for adult health. Early Human Development 2006; 82: 485—491

Ross MG, Beall MH. Adult sequelae of intrauterine growth restriction. Semin Perinatol 2008; 32:213-8.

Russell, Peter. 2002. *iGenetics.* Pearson Education, Inc., San Francisco, pp 187-189

Ryckman KK, Feenstra B, Shaffer JR, et al. Replication of a genome-wide association study of birth weight in preterm neonates. J Pediatr 2012; 160:19e4–24e4

Sandhu MS, Weedon MN, Fawcett KA et al (2007) Common variants in WFS1confer risk of type 2 diabetes. Nat Genet 39:951–953

Sandhu MS, Weedon MN, Fawcett KA, Wasson J, Debenham SL, Daly A, Lango H, Frayling TM, Neumann R, Pharoah PD, McCarthy MI, Walker M, Hitman G, Glaser B, Permutt MA, Hattersley AT, Wareham NJ, Barroso I: *WFS1* is a type 2 diabetes susceptibility gene. *Nat Genet* 39:951–953, 2007

Schalkwyk LC, Meaburn EL, Smith R, Dempster EL, Jeffries AR, Davies MN, Plomin R, Mill J. Allelic skewing of DNA methylation is widespread across the genome. Am J Hum Genet. 2010 Feb 12;86(2):196-212

Schroeder DI, Blair JD, Lott P, Yu HO, Hong D, Crary F, Ashwood P, Walker C, Korf I, Robinson WP, LaSalle JM. The human placenta methylome. Proc Natl Acad Sci U S A. 2013 Apr 9; 110 (15):6037-42

Schroeder JW, Conneely KN, Cubells JC, Kilaru V, Newport DJ, Knight BT, Stowe ZN, Brennan PA, Krushkal J, Tylavsky FA, Taylor RN, Adkins RM, Smith AK. Neonatal DNA methylation patterns associate with gestational age. Epigenetics. 2011 Dec;6(12):1498-504. R

Sermer N, Naylor CD, Gare DJ et al. Impact of increasing carbohydrate intolerance on maternal-fetal outcomes in 3637 women without gestational diabetes: the Toronto Tri-Hospital Gestational Diabetes Project. Am J Obstet Gynaecol 1995; 173: 146-56

Shan J, Wang P, Zhou J, Wu D, Shi H, Huo K. RIOK3 interacts with caspase-10 and negatively regulates the NF-kappaB signaling pathway. Mol Cell Biochem. 2009 Dec; 332(1-2):113-20

Shield JPH. Neonatal diabetes. In: Shield JPH, Baum JD. (eds) Childhood Diabetes, vol 4. London: Balliere Tindall, 1996;681-740

Shields BM, Knight B, Turner M, Wilkins-Wall B et al. Paternal insulin resistance and its association with umbilical cord insulin concentration. Diabetologia 2006; 49: 2668-74

Sibley CP, Coan PM, Ferguson-Smith AC, Dean W, Hughes J, Smith P, et al. Placental-specific insulin-like growth factor 2 (Igf2) regulates the diffusional exchange characteristics of the mouse placenta. Proc Natl Acad Sci U S A 2004;101: 8204–8

Silander K, Mohlke KL, Scott LJ et al (2004) Genetic variation near the hepatocyte nuclear factor-4α gene predicts susceptibility to type 2 diabetes. Diabetes 53:1141–1149

Simchen MJ, Beiner ME, Strauss-Liviathan N et al. Neonatal outcomes in growth-restricted versus appropriately grown preterm infants. Am J Perinatol 2000; 17: 187-192

Simmons RA, Templeton LJ, Gertz SJ. Intrauterine growth retardation leads to the development of type 2 diabetes in the rat. Diabetes. 2001 Oct;50(10):2279-86.
Simmons RA: Developmental origins of beta-cell failure in type 2 diabetes: The role of epigenetic mechanisms. Pediatr Res 2007, 61:64r-67r

Skarfors ET, Selenus KI and Lithell HO. 1991. Risk Factors for developing non-insulin dependent diabetes mellitus: a 10-year follow-up of men in Uppsala.Br.Med.J.303:755-760)

Smith GC, Pell JP, Walsh D (2001) Pregnancy complications and maternal risk of ischaemic heart disease: a retrospective cohort study of 129,290 births. Lancet 357: 2002–2006)

Smyth GK. Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Statistical applications in genetics and molecular biology*. 2004; 3: Article3.

Sparsø T, Andersen G, Albrechtsen A, Jørgensen T, Borch-Johnsen K, Sandbaek A, Lauritzen T, Wasson J, Permutt MA, Glaser B, Madsbad S, Pedersen O, Hansen T. Impact of polymorphisms in WFS1 on prediabetic phenotypes in a population-based sample of middle-aged people with normal and abnormal glucose regulation. Diabetologia. 2008 Sep;51(9):1646-52

Stoffers DA, Desai BM, DeLeon DD, Simmons RA. Neonatal exendin-4 prevents the development of diabetes in the intrauterine growth retarded rat. Diabetes. 2003 Mar;52(3):734-40.

Stride A and Hattersley AT. Differnet, genes, different diabetes: lessons from maturity-onset diabetes of the young. Annals of Medicine 2002; 34:207-216

Stumvoll M, Goldstein BJ, van Haeften TW: Type 2 diabetes: principles of pathogenesis and therapy. *Lancet* 365:1333–1346, 2005

Suter M, Abramovici A, Showalter L, Hu M. In Utero Tobacco Exposure Epigenetically modifies placental CYP1A1 Expression. Metabolism 2010; 59: 1481–1490

Tabano S, Colapietro P, Cetin I, Grati FR, Zanutto S, Mandò C, Antonazzo P, Pileri P, Rossella F, Larizza L, Sirchia SM, Miozzo M. Epigenetic modulation of the IGF2/H19 imprinted domain in human embryonic and extra-embryonic compartments and its possible role in fetal growth restriction. Epigenetics 2010; 5:313–324.

Tallarigo L, Giampetro O, Genno G, Miccoli R, Gergori G, Navalesi R. Relation of glucose tolerance to complications of pregnancy in non-diabetic women. N Engl J Med 1986; 315:989-92

Tamemoto H, Kadowaki T, Tobe K, Yagi T, Sakura H, Hayakawa T, Terauchi Y, Ueki K, Kaburagi Y, Satoh S, Sekihara H, Yoshioka S, Horikoshi H, Furata Y, Ikawa Y, Kasuga M,

Tanner JM, Lejarraga H, Turner G. Within-family standards for birth-weight. Lancet. 1972; 2, 193–197

Taylor S, Accili D, Imai Y. Insulin resistance or insulin deficiency which is the primary casue of NIDDM? Diabetes 1994; 43:735-740

Temple IK, Gardner RJ, Robinson DO et al. Further evidence for an imprinted gene for neonatal diabetes localised to chromosome 6q22-q23. Hum Mol Genet 1996; 5:1117-24

Terauchi Y, Kubota N, Tamemoto H, Sakura H, Nagai R, Akanuma Y, Kimura S and Kadowaki T. Insulin effect during embryogenesis determines fetal growth. A possible molecular link between birth weight and susceptibility to Type 2 diabetes. Diabetes 2000; 49:82-86

The Gene Ontology Consortium. Gene ontology: tool for the unification of biology. Nature. Genetics.. May 2000; 25(1):25-9

Thompson RF, Fazzari MJ, Niu H, Barzilai N, Simmons RA, Greally JM. Experimental intrauterine growth restriction induces alterations in DNA methylation and gene expression in pancreatic islets of rats. J Biol Chem. 2010 May 14; 285(20):15111-8

Tissot van Patot MC,, Murray AJ, Beckey V, Cindrova-Davies T, Johns J, Zwerdlinger L, Jauniaux E, Burton GJ and Serkova NJ. (2010) Human Placental metabolic adaptation to chronic hypoxia, high altitude: hypoxic preconditioning. Am J Physiol Regul Integr Comp Physiol. 298: 166-172

Tora L. A unified nomenclature for TATA box binding protein (TBP)-associated factors (TAFs) involved in RNA polymerase II transcription. Genes Dev 2002; 16:673-5

Toyoda M , Kojima M, Takeuchi T (2000). "Jumonji is a nuclear protein that participates in the negative regulation of cell growth.". Biochem. Biophys. Res. Commun. 274 (2): 332–6

Tsuchiya T, Kominato Y, Ueda M. Human hypoxic signal transduction through a signature motif in hepatocyte nuclear factor 4. J Biochem. 2002 Jul;132(1):37-44

Urhammer SA, Rasmussen SK, Kaisaki PJ, Oda N, Yamagata K, Moller AM, Fridberg M, Hansen L, Hansen T, Bell GI, Pedersen O: Genetic variation in the hepatocyte nuclear factor-1α gene in Danish Caucasians with late onset NIDDM. *Daibetologia* 40:473– 475, 1997

Voight BF, Scott LJ, Steinthorsdottir V, Morris AP, Dina C, et al. (2010) Twelve type 2 diabetes susceptibility loci identified through large-scale association analysis. Nat Genet 42: 579–589.

Wallace TM, Levy JC and Matthews DR (2004). Use and abuse of HOMA modelling. Diabetes Care 27: 1487-1495

Wang HS, Lim J, English J, Irvine L, Chard T. the concentration of insuln like growth factor-1 and insulin-like growth factor binding protein-1 in human umbilical cord serum at delivery: relation to fetal weight. J Endocrinology 1991; 129:459-464

Wannamethee SG, Lawlor DA, Whincup PH, Walker M, Ebrahim S, Davey-Smith G. Birthweight of offspring and paternal insulin resistance and paternal diabetes in late adulthood: cross sectional survey. Diabetologia 2004; 47: 12-18

Weedon MN, Frayling TM, Shields B, Knight B, Turner T, Metcalf BS, Voss L, Wilkin TJ, McCarthy A, Ben-Shlomo Y, Davey Smith G, Ring S, Jones R, Golding J, Byberg L, Mann V, Axelsson T, Syvanen AC, Leon D, Hattersley AT: Genetic regulation of birth weight and fasting glucose by a common polymorphism in the islet cell promoter of the glucokinase gene. *Diabetes* 54:576 –581, 200

Weedon MN, Gloyn AL, Frayling TM, Hattersley AT, Davey Smith G, Ben-Shlomo Y: Quantitative traits associated with the Type 2 diabetes susceptibility allele in Kir6.2. Diabetologia 2003, 46(7):1021-1023

Weedon MN, McCarthy MI, Hitman G, Walker M, Groves CJ, Zeggini E, Rayner NW, Shields B, Owen KR, Hattersley AT, Frayling TM: Combining information from common type 2 diabetes risk polymorphisms improves disease prediction. *PLOS Med* 3:e374, 2006

Weedon MN, Owen KR, Shields B, Hitman G, Walker M, McCarthy MI, Hattersley AT, Frayling TM: A large-scale association analysis of common variation of the HNF1α gene with type 2 diabetes in the U.K. Caucasian population. *Diabetes* 54:2487–2491, 2005

Whincup PH, Kaye SJ, Owen CG, Huxley R, Cook DG, Anazawa S, Barrett-Connor E, Bhargava SK, Birgisdottir BE, Carlsson S, de Rooij SR, Dyck RF, Eriksson JG, Falkner B, Fall C, Forsén T, Grill V, Gudnason V, Hulman S, Hyppönen E, Jeffreys M, Lawlor DA, Leon DA, Minami J, Mishra G, Osmond C, Power C, Rich-Edwards JW, Roseboom TJ, Sachdev HS, Syddall H, Thorsdottir I, Vanhala M, Wadsworth M, Yarbrough DE. Birth weight and risk of type 2 diabetes: a systematic review. JAMA 2008; 300:2886-97

Whitehead CL, Walker SP, Lappas M, Tong S. Circulating RNA coding genes regulating apoptosis in maternal blood in severe early onset fetal growth restriction and pre-eclampsia. J Perinatol. 2013 Aug;33(8):600-4

Williams RC, Long JC, Hanson RJ, Sievers ML, Knowler WC. Individual estimates of European genetic admixture associated with lower body-mass index, plasma glucose, and prevalence of type 2 diabetes in Pima Indians. Am J Hum Genet 2000; 26: 163-75

Winckler W, Burtt NP, Holmkvist J, Cervin C, de Bakker PI, Sun M, Almgren P, Tuomi T, Gaudet D, Hudson TJ, Ardlie KG, Daly MJ, Hirschhorn JN, Altshuler D, Groop L: Association of common variation in the HNF1_ gene region with risk of type 2 diabetes. *Diabetes* 54:2336 –2342, 2005

Winckler W, Weedon MN, Graham RR, McCarroll SA, Purcell S, Almgren P, Tuomi T, Gaudet D, Bostrom KB, Walker M, Hitman G, Hattersley AT, McCarthy MI, Ardlie KG, Hirschhorn JN, Daly MJ, Frayling TM, Groop L, Altshuler D: Evaluation of common variants in the six known maturity onset diabetes of the young (MODY) genes for association with type 2 diabetes. *Diabetes* 56:685– 693, 2007

Xiong Y, Fan C, Kong L, Dong L, Zhu N, Zhang J, Wang L, Qin T, Shen Y, Chen M. Annexin II receptor induces apoptosis independent of Annexin II. Apoptosis. 2013 Aug; 18(8):925-39

Yaghootkar H, Freathy RM. Genetic origins of low birth weight. Curr Opin Clin Nutr Metab Care. 2012 May;15(3):258-64

Zaborowska J, Taylor A, Murphy S.A novel TBP-TAF complex on RNA polymerase II-transcribed snRNA genes. Transcription. 2012 Mar-Apr;3(2):92-104

Zamudio S. (2003). The Placenta at High Altitude. High Altitude Medicine & Biology. 4:171-191

Zeggini E, Mccarthy MI. TCF7L2: the biggest story in diabetes genetics since HLA? Diabetologia 2007: 50:1-4

Zeggini E, Scott LJ, Saxena R, Voight BF, Marchini JL, et al. (2008) Meta-analysis of genome-wide association data and large-scale replication identifies additional susceptibility loci for type 2 diabetes. Nat Genet 40: 638–645

Zhao J, Li M, Bradfield JP, et al. Examination of type 2 diabetes loci implicates CDKAL1 as a birth weight gene. Diabetes. 2009; 58, 2414–2418

Zheng X, Linke S, Dias JM, Zheng X, Gradin K, Wallis TP, Hamilton BR, Gustafsson M, Ruas JL, Wilkins S, Bilton RL, Brismar K, Whitelaw ML, Pereira T, Gorman JJ, Ericson J, Peet DJ, Lendahl U, Poellinger L. Interaction with factor inhibiting HIF-1 defines an additional mode of cross-coupling between the Notch and hypoxia signaling pathways.Proc Natl Acad Sci U S A. 2008 Mar 4;105(9):3368-73

THE END