The Molecular Genetic Investigation of Paediatric Liver Disease

by

Jane Louise Hartley

A thesis submitted to The University of Birmingham
For the degree of Doctor of Philosophy

Section of Medical and Molecular Genetics
Division of Reproductive and Child Health
School of Clinical and Experimental Medicine
The University of Birmingham
March 2010

Abstract

Liver disease in children is rare but often serious, life long, and in many cases leads to death. Advances in diagnosing and treating liver disease (including liver transplant) have improved the outlook for children in many cases however little is known about the molecular pathogenesis of the disease, an understanding of which may identify specific therapeutic options. The aim of this thesis is to investigate the molecular genetics of rare liver disorders as the first step in advancing the understanding of liver disease pathogenesis. As a paediatric hepatologist I have identified cohorts of children in whom there is paucity of knowledge about the disease pathogenesis. I have studied three conditions in detail to encompass different clinical presentations. Chapter 3 summarises the investigation of the multisystem disorder, phenotypic diarrhoea of infancy (PDI), which causes cirrhosis or liver failure. Autozygosity mapping was used to identify the gene *TTC37* in which mutations are associated with the PDI disease phenotype. Further work is now required to characterise *TTC37*, and use knockdown studies to identify whether *TTC37* mutations are causative of the PDI phenotype. Chapter 4 describes the molecular genetic investigation of Jeune asphyxiating thoracic dystrophy (JATD), a chondrodysplasia with extra skeletal manifestations including hepatic ductal plate malformation and renal cyst development. Using autozygosity mapping, *IFT80* was identified in which mutations are associated with the JATD disease phenotype in 4% of cases. The diverse clinical phenotype of JATD limits the utility of autozygosity mapping as it suggests there is genetic heterogeneity. The identification of *IFT80* has led to JATD being classified as a ciliopathy. Chapter 5 is the first description of neonatal liver failure

to be associated with variants in *ABCB11* which previously have only been associated with chronic liver disease and liver disease in pregnancy.

This thesis has described the identification of the molecular genetic basis of rare causes of paediatric liver disease which has provoked many additional research questions. Future work will be to extend our knowledge of molecular genetics to all aspects of paediatric liver physiology so to classify disease according to the molecular pathogenesis such as a ciliopathy or bile salt transport defect.

Acknowledgments

I am indebted to my supervisors Professor Eamonn Maher and Professor Deirdre Kelly for giving me the opportunity to do this PhD and for all the invaluable advice, encouragement and support whilst working on this project and the writing of this thesis. I would also like to thank Dr Paul Gissen for all his advice, support and continuing involvement in my projects.

Many colleagues in the department have given me support and advice during the course of this work. I would particularly like to thank all members of the BWH research lab for their friendship and support. Specifically I would like to thank Dr Neil Morgan, Dr Esther Meyer and Louise Tee for teaching me the practical techniques that I have needed to complete this project and for their patience in explaining molecular genetics to a clinician.

Thanks also to the members of the diagnostic DNA laboratory for use of their equipment and numerous DNA extractions carried out over the years. Thanks also to Dr Alex Knisely who carried out the immunohistochemical investigations which allowed for greater understanding of the molecular genetic findings in chapter 5.

I also thank all the clinicians around the world that have provided samples and details of patients. Particular thanks are due to the various patients and their families for their participation in these studies some of whom I will continue to see as a clinician. For involvement with the families and sample collection I would like to thank Shanaz Pasha and Rachel Gould.

Finally I would like to give special thanks to my husband Martin, for his extreme patience and support during the writing of this thesis, and also to my parents who have always made me believe I can achieve anything I put my mind to.

Contents

3.1 An overview of inherited congenital diarrhoea and hepatic involvement

3.4 The identification of a genetic locus for PDI using autozygosity mapping

Chapter 8 Appendices 326

List of figures

List of tables

Preface

Papers and abstracts which have been published or are under peer review which are derived from work contributing to this thesis.

- Mutations in *TTC37* cause tricohepatoenteric syndrome (phenotypic diarrhoea of infancy). Gastroenterology 2010;138:2388–2398
- Clinical phenotype and autozygosity mapping of phenotypic diarrhoea of infancy. British Society of Human Genetics 2006. Poster presentation
- Clinical phenotype and autozygosity mapping of phenotypic diarrhoea of infancy. American Society of Human Genetics 2006. Poster Presentation
- IFT80, which encodes a conserved intraflagellar transport protein, is mutated in Jeune asphyxiating thoracic dystrophy. Beales PL, Bland E, Tobin JL, Bacchelli C, Tuysuz B, Hill J, Rix S, Pearson CG, Kai M, Hartley J, Johnson C, Irving M, Elcioglu N, Winey M, Tada M, Scambler PJ. Nat Genet. 2007 Jun;39(6):727-9. Epub 2007 Apr 29.
- The identification of renal cysts may implicate primary cilia in the aetiology of biliary atresia. Accepted for publication in JPGN August 2010
- The c.1331C>T / p.V444A *ABCB11* variant in severe intrahepatic cholestasis. American Association for the Study of Liver Disease 2008. Poster presentation.

Abbreviations

A	adenine
A	alanine
A_1AT	alpha 1 antitrypsin
ABC	ATP-binding cassette
AD	autosomal dominant
ADP	Adenosine diphosphate
ADPLD	autosomal dominant polycystic liver disease
ALT	alanine aminotransferase
AR	autosomal recessive
ARC	arthrogryposis-renal dysfunction-cholestasis syndrome
ARPKD	autosomal recessive polycystic kidney disease
ASBT	apical sodium-dependent bile acid transporter
AST	aspartate aminotransferase
ATP	adenosine triphosphate
BA	biliary atresia
BASD	bile acid synthetic defect
BBS	Bardet Beild syndrome
BLAST	Basic local alignment search tool
BLAT	Basic local alignment tool
Bp	base pair
BRIC	benign recurrent intrahepatic cholestasis
BSA	bovine serum albumin

BSEP	bile salt export pump
C	cytosine
cAMP	Cyclic adenosine monophosphate
cDNA	complementary DNA
CEP290	Centrosomal protein of 290 kDa
CFTR	cystic fibrosis transmembrane regulator
cGFR	calculated glomerular filtration rate
CGD	carbohydrate deficient glycoprotein
CI	calcineurin inhibitor
CIS	cryptogenic infantile spasms
cM	Centimorgan
CMV	cytomegalovirus
COACH	cerebellar vermis hypoplasia, oligophrenia, ataxia, coloboma, and hepatic fibrosis
CSD	congenital sodium diarrhoea
CTG	cardiotocography
CT scan	computerised tomography
CTX	cerebrotendinous xanthomatosis
Cu^{2+}	copper
^{o}C	degrees centigrade
del	deletion
DHPLC	denaturing high-performance liquid chromatography
dH_2O	distilled water

DNA	dioxyribonucleic acid
dNTP’s	Deoxyribonucleotide triphosphate’s
DPM	ductal plate malformation
EVC	Allis van Creveld
FACS	Fluorescence-activated cell sorting
FGF	fibroblast growth factor
FHC	familial hypercholesterolaemia
FIC-1	Familial intrahepatic cholestasis 1
FITC	flourescein isothiocyanate
FXR	Farnesoid nuclear receptor
g	grams
G	guanine
GFC	Greenland familial cholestasis
GFR	glomerular filtration rate
γGT	gamma-glutamyl transpeptidase
GWS	genomewide scan
H	hydrogen ion
HCO_3	Bicarbonate
HEK 293	human embryonic kidney 293
Hox	homeobox
HSV	herpes simplex virus
ICP	intrahepatic cholestasis of pregnancy
IFALD	intestinal failure associated liver disease

IFT	intraflagellar transport
IgG	immunoglobulin G
IgM	immunoglobulin M
IMCD3	intramedullary collecting duct 3
IPEX	immune dysregulation-polyendocrinopathy-enteropathy, X-linked
ISBT	ileal sodium-dependent bile salt transporter
IU	international units
JATD	Jeune asphyxiating thoracic dystrophy
K	potassium
Kb	kilobase
Kd	kilodalton
L	litre
LOD	logarithm of odds
LREC	local research ethics committee
M	Moles
MAS	MaCune Albright Syndrome
Mb	Megabases
MDR	Multidrug resistence protein
MgCl	Magnesium chloride
µg	micrograms
ml	millilitres
µl	microlitres
µmol/mmol	micromole

mmol	millimole
MOAT	multiorganic anion transporter
MOI	mode of inheritance
MRI	magnetic resonance imaging
mRNA	messenger RNA
MRP	multidrug resistence associated protein
MRST	mental retardation, spasticity, tapetoretinal degeneration
MVID	microvillus inclusion disease
Na	sodium
NADPH	nicotinamide adenine dinucleotide phosphate-oxidase
NASH	non alcoholic steatohepatitis
NCBI	National Center for Biotechnology Information
NaCl	sodium cholride
NAICC	North American Indian Childhood cirrhosis
NAMR	National autozygosity mapping
NICCD	neonatal intrahepatic cholestasis caused by citrin deficiency
NHE	sodium hydrogen exchanger
NISCH	neonatal sclerosing cholangitis
NTCP	sodium- taurocholate cotransporter polypeptide
OATP	organic anion transporting polypeptide
OCS	open canalicular system
OLT	orthotopic liver transplant
OMIM	Online mendelian inheritance in man

OST	organic solute transporter
PBC	primary biliary cirrhosis
PBS	phosphate buffered saline
PBST	phosphate buffered saline with tween
PCD	primary cilial dyskinesia
PCR	polymerase chain reaction
PDGF	platelet derived growth factor
PDI	phenotypic diarrhoea of infancy
PFIC	progressive familial intrahepatic cholestasis
pH	power of hydrogen
PN	parenteral nutrition
PRP	plateley rich plasma
PS	phosphatidylserine
PSC	primary sclerosing cholangitis
PSIC	position specific independent count
PT	prothrombin
Rev	revolutions
RNA	Ribonucleic acid
RT	Rothmond Thompson syndrome
RT PCR	reverse transcriptase polymerase chain reaction
RTS	Rothmund Thompson syndrome
RXR	retinoid nuclear receptor
SGPT	Serum glutamic pyruvic transaminase

SHP	short heterodimer partner
SiRNA	short interfering RNA
SLE	systemic lupus erythematosis
SNARE	Soluble NSF Attachment
SNP	single nucleotide polymorphism
T	thymine
TAP	Tandem Affinity Purification
TBE	Tris/Borate/EDTA
TBS	Tris-Buffered Saline
TEM	transmission electron microscopy
TEMED	Tetramethylethylenediamine
THE	tricohepatoenteric syndrome
TN	trichorrhexis nodosa
TPR	tetratrichopeptide repeat
TTC37	tetratricopeptide repeat domain 37
TTD	trichothiodystrophy
TTV	transfusion transmissible virus
TXA	Tranexamic acid
UCSC	University California, Santa Cruz
V	valine
VGEF	vascular endothelial growth factor
vWF	Von Willebrand factor
2-D	two dimensional

Wnt	wingless
WT	wild type

Chapter 1

Introduction

Contents

“nor is there any better way to advance the proper practice of medicine than to give our minds to the discovery of the usual law of nature by the careful investigation of cases of rarer forms of disease”

Dr William Harvey, 1657

This chapter provides an introduction to molecular genetics and to liver embryology and physiology.

The introduction to molecular genetics commences with a history of the past 144 years from the first description of genetic traits in peas by Gregor Mendel, to the Human Genome Project and now the recognition that in diseases previously thought to be environmental there may also be a genetic susceptibility.

I have studied rare paediatric liver diseases which are inherited in an autosomal recessive manor. Studying this form of inheritance in those from consanguineous union provides a large amount of genetic information which can be utilised to identify mutations in genes which are associated with disease. A description of the technique of autozygosity mapping and its limitations is provided in this introduction as it is a technique I have used throughout this thesis. Other molecular genetic techniques which are used on single occasions are described within the relevant chapters.

The liver is a complex organ with many functions many of which are not fully understood. This introduction provides firstly a description of the embryology concentrating on bile duct development which is pertinent to the understanding of ductal plate malformations described in chapter 4. Thereafter I have concentrated on the physiology of the liver with emphasis on bile transport which is relevant to all chapters, in particular to chapter 5.

The introduction provided here is a general overview applicable to all chapters. As I have studied three diverse disorders affecting the liver a more detailed introduction is provided at the beginning of each chapter. Hence, chapter 3 introduction provides detail on inherited causes of intractable diarrhoea, chapter 4 provides an introduction to primary cilial structure and ciliopathies, and chapter 5 introduces the diseases associated with bile salt transport defects.

At the end of this chapter is the philosophy underpinning the thesis is discussed and poses the question ‘why study the molecular genetic investigate of paediatric liver disease?’

1.1 Introduction to molecular genetics

1.1.1 History of molecular genetics

In 1866 Gregor Mendel published his work on pea plants describing specific traits which had been inherited from the previous generation of peas so providing the first description of the inheritance of specific discrete genetic material which he termed alleles.

The earliest description of the autosomal recessive inheritance of human disease was in alkaptonuria in 1905 by Garrod (Garrod & Hele, 1905) where the term inborn error of metabolism was coined.

DNA was first identified in 1869 by Miescher who described nuclei acid in pus from open wounds. In the 1920’s DNA was found to be a major component of chromosomes.

Each chromosome was thought to be composed of one polymer unit until the 1950's when the composition of chromosomes was found to differ depending on source. In 1953 Watson and Crick determined the detailed double stranded DNA structure of chromosomes and described the replication of DNA through the formation of complementary strands and the manufacture of proteins from specific genes (Watson and Crick, 1953).

In the 1970's regulator genes were identified and in 1984 Hox genes which regulate limb development were described (Hogrefe, 1984).

In 2001 the Human Genome Project presented preliminary results of an accurate chemical map of the genome and showed that only about 1% of nucleotides within the genome are transcribed and they reside in approximately 30,000 genes (Venter *et al*, 2001).

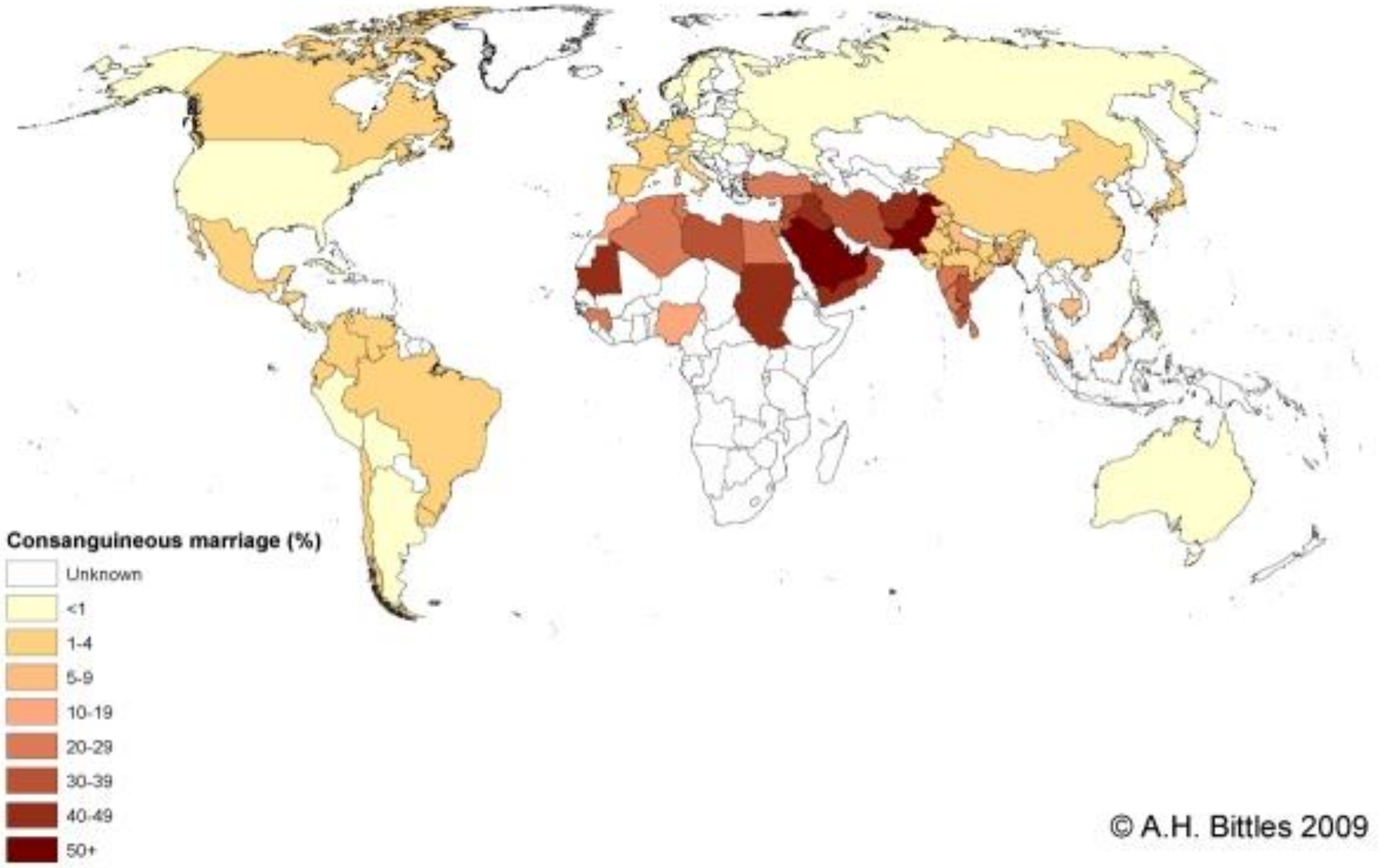
The transcript of the sequence of the human genome is a major factor in facilitating the identification of disease genes by allowing specific targeting of genetic regions.

The concept of genetic inheritance has expanded from classical inheritance patterns within families to population genetics in which features previously thought to be environmental are found to be associated with specific genetic variants within the genome. These variants may directly cause the trait or confer susceptibility. For example high blood pressure or the development of type 2 diabetes (Smushkin and Vella, 2010).

1.1.2 Autosomal recessive inheritance

Autosomal recessive disorders represent an important cause of morbidity and mortality, particularly within the paediatric age range. As many of the conditions are either lethal in utero e.g. Meckel Gruber syndrome (Salonen and Norio, 1984); lethal in early childhood e.g. JATD (Turkel *et al*, 1985); cause developmental delay e.g. Niemann Pick C disease (Garver *et al*, 2007); or become clinically apparent during childhood e.g. cystic fibrosis (Sing *et al*, 1982). Most people have a number of autosomal recessive mutations which affect a single allele and therefore are not disease causing (the disease occurs when both alleles are mutated). An offspring of the carrier parent will only be affected if the other parent is also a carrier and both carrier alleles are inherited i.e. there is a 1 in 4 chance of a child being affected when both parents are carriers. Most mutations are extremely rare making the chances of both parents being carriers unlikely. However in specific populations and ethnic groups a number of autosomal recessive conditions reach significant frequency due to consanguinity. This is more common in those from the Indian sub-continent (Bundey and Alam, 1993) or from isolated communities e.g. Wilsons disease is common in Sardinia (1 in 7000) (Figus *et al*, 1985; Zappu *et al*, 2008) whilst in America the incidence is 1 in 30,000; and Tay-Sachs disease is more common in the Ashkenazi Jewish population than in others (the carrier rate in Ashkenazi Jews is 1 in 27 whilst in the general population it is estimated to be 1 in 250) (Kaback e*t al*, 1977). Where a disease occurs in all ethnicities, specific mutations may be more common in certain populations due to a founder effect e.g. cystic fibrosis is much more common in people of Northern European descent, for which a single amino acid deletion named ΔF508 accounts for up 70 % of the mutant alleles, compared with only one-third of Turkish or Arab CF patients (Lemna *et al*, 1990).

Inbreeding or consanguineous union is common is certain areas of the world primarily due to social and religious beliefs where it is believed family ties would be stronger and health and financial uncertainty can be avoided by a close kin marriage. This may be first cousin or uncle-niece union (Bittles, 2008). The rate of consanguineous marriage in each country throughout the world is shown in figure 1.1.1.



© A.H. Bittles 2009

Figure 1.1.1 shows the percentage number of consanguineous marriages throughout the world

Bundey and Alam in 1993, found among British Pakistanis, 69% of couples were related and 57% were first cousins. Amongst the British Pakistani children the birth prevalence of all congenital and genetic disorders was 7.9%; almost double that of Northern European children (4.3%) where 0.4% of couples are related.

1.1.3 Gene identification using autozygosity mapping

The genetic implications of a consanguineous marriage are related to the proportion of the children's gene pairs that are identical because they are inherited from a common ancestor. For example, in a first cousin union, their children inherit one-eighth of their genes from their common grandparents, and so one-sixteenth of their genes (6.25%) are identical-by-descent. Similarly if the affected offspring has second cousin parents then they would share 1/64 of their genome.

The technique of autozygosity mapping to identify rare disease genes utilises the genetic information that can be gained from studying consanguineous union families and the identical by descent genetic material which is inherited form a common single ancestor. This technique was described in 1987 by Lander and Botstein using microsatellite markers. The technique has now been adapted to incorporate single nucleotide polymorphism (SNP) genome wide scans.

Autozygosity mapping is divided into steps:-

- Identify a cohort with a homogeneous phenotype ideally from the same extended family with multiple affected probands
- Use microsatellite DNA markers to exclude potential known disease loci

- Genomewide linkage scan (GWS) – in this thesis scans of 10 or 250 thousand SNP's were used for linkage
- Identify regions of homozygosity that are shared by all affected probands
- Fine map the candidate region(s) using polymorphic microsatellite markers
- Candidate gene analysis from the region of interest by direct sequencing

A homogeneous phenotype

Ascertaining families to identify disease genes using autozygosity mapping can be difficult due to the rarity of the disease. The genetic information derived from a consanguineous union, especially if there are multiple affected children, is far greater than from families of non-consanguineous unions. It is essential to phenotype the probands accurately to ensure only one disease state is being studied. If more than one disease is studied at the same time the technique of autozygosity mapping will not identify the disease gene as a common locus of homozygosity will not be recognised. This can be difficult for diseases in which the clinical spectrum can be variable such as with JATD.

Microsatellites to exclude potential known disease loci

Known disease genes or loci that might cause the disease phenotype should be excluded initially using microsatellite markers flanking the locus or gene.

GWS and locus identification

SNPs are the most frequent DNA sequence variations in the human genome, found on average once every 1,200 base pairs (bp) of DNA. SNPs can be used to map regions for gene identification in autozygosity mapping and can also identify susceptibility haplotypes for complex disorders (Johnson *et al,* 2001). SNP's are bi allelic therefore the results from a GWS identify either SNP A or SNP B at each position. In areas of homozygosity both alleles are identical i.e. AA or BB, areas of heterozygosity, and therefore not identical by descent, are AB or BA. The pathogenic gene is likely to be in a region where all the probands (from consanguineous union) are homozygous. In those from the same family the homozygous haplotype should be identical i.e. both AA or BB. When analysing probands from different families the homozygous haplotype may not be identical (although they may be identical if there is a founder effect). The larger the region of homozygosity the more likely it is to harbour the gene of interest. A homozygous region of 3 centiMorgan (cM) or greater with an average size of 27cM is most likely to harbour recessive disease gene (Woods *et al,* 2006).

Fine mapping using microsatellites

Microsatellite markers are short tandem repeats usually of less than 10bp which are polymorphic and stably inherited (although they are occasionally prone to slippage). The segregation of highly polymorphic genetic markers in family pedigrees enables linkage to be either (a) excluded or (b) the families are said to be compatible with linkage. For autozygosity mapping it is usual to look for heterozygosity of particular markers which flank a gene of interest so to exclude the gene.

Selection of candidate genes

When the candidate region is narrowed as much as possible and there is evidence for a conserved haplotype either in one family or between families candidate genes can be identified from genome databases with selection determined by position of the gene, the predicted gene function (compared to what is known about the disease pathogenesis) and the areas of expression of the gene in the body. The candidate genes are screened for pathogenic mutations by direct sequencing using the ABI 3730 capillary sequencer as described in the methods section. A change in the DNA sequence from the wild type is either a:-

Known SNP – will be identified in public databases of SNP's

Novel SNP – the change is likely to be found in control samples from the same ethnic background or is predicted to be a benign change by bioinformatic tools such as Polyphen

Pathogenic mutation – these changes are likely to be significant such as a change at a splice site or the formation of a stop codon. Computer prediction tools such as Polyphen will identify the change as potentially or probably damaging. The sequence change will segregate within families so that the parents will be carriers and the unaffected siblings will be either carriers or homozygous wild type. The sequence change will not be seen (or only very rarely) in controls of the same ethnic origin. Other mutations within the same gene lead to an identical disease phenotype or trait including compound heterozygotes from non consanguineous unions.

1.1.4 Limitations of autozygosity mapping

Although autozygosity mapping is a powerful tool for gene identification it has limitations and pit falls.

Due to the rarity of many of the conditions it may be difficult to ascertain a large enough cohort of affected probands. The ideal family for study would be a multiple consanguineous with multiple affected probands from a common ancestor. In many parts of the world however small family units are more common which limits autozygosity mapping.

Autozygosity mapping will only be successful if mutations in only one gene are responsible for the disease phenotype. In a number of autosomal recessive conditions there is more than one gene (locus heterogeneity) which can lead to the same phenotype (e.g. Bardet Biedl syndrome in which many different genes have been identified and lead to the same phenotype) and indeed mutations in more than one region can lead to a disease state again as in Bardet Beidl syndrome in which triallelic inheritance may also occur (Katsanis *et al*, 2001). These pitfalls can be avoided if a single family with multiple affected probands are studied.

1.2 The liver

Liver disease is a significant cause of morbidity and mortality worldwide. In childhood it is rare but is potentially a life long and life limiting disorder, the treatment of which may include liver transplantation and therefore is a significant cause of morbidity.

The aetiology of liver disease varies according to age and figure 1.2.1 demonstrates the different aetiology of liver diseases in adults as compare to children.

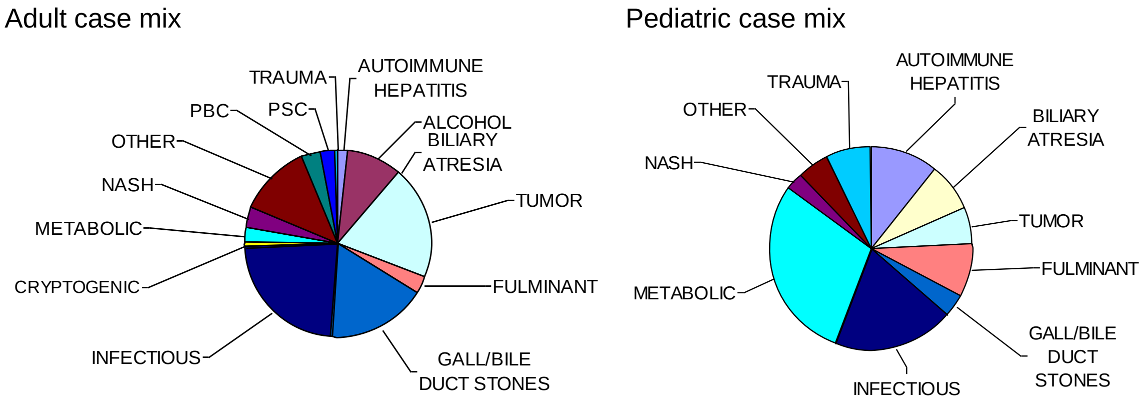


Figure 1.2.1 pie charts demonstrating the different aetiology of liver disease in adults as compared to children.

In childhood it can be seen that metabolic conditions, fulminant hepatic failure and biliary atresia, which potentially have a genetic component to the aetiology, are more common than in adults.

1.2.1 Embryology of the liver

During the third week of gestation the liver bud (hepatic diverticulum) develops as an outgrowth from the endodermal epithelium of the ventral foregut. The bud consists of rapidly developing cell strands which grow through the septum transversum (a mesodermal plate between the pericardial cavity and the yolk sac from which connective tissue, haemopoietic cells and Kupffer cells derive). The strands continue to grow through the septum forming thick plates of hepatocytes whilst the connection with the foregut narrows and becomes the common bile duct. The blood supply to the liver influences the growth of the liver. The two vitelline veins (which ultimately join to become the portal vein) initially supply the developing liver. Later in embryogenesis initially both umbilical veins provide blood to the liver and then the right disappears leaving the left to perfuse the whole liver. The left umbilical vein supplies the left lobe sinusoids, mixes with left branch portal blood flow to provide retrograde flow to the right lobe sinusoids (50% of the blood to the right lobe is supplied from the portal vein which is nutrient poor) and also flows into the inferior vena cava via the ductus venosus (Sadler, 1990).

The arterial supply develops later in embryogenesis and is closely related to the development of the bile ducts. The external biliary system (including the gallbladder) is derived from the hepatic diverticulum. The intrahepatic bile ducts develop independently of the extrahepatic and only towards the end of foetal development do the two structures join. The hepatoblasts which surround the mesenchyme of the portal vein tracts become smaller and form a single cell sheath around the portal vein known as the ductal plate. A second layer of cells forms and under the influence of signals from the portal mesenchyme and the genes NOTCH2, JAG1, HES1, bile ducts develop (Lamaigre, 2008).

Transcription factors regulate the development and differentiation of the liver.
The factors which are known are:-
Homeobox gene (Hhex) which is involved at multiple stages of hepatobiliary development (Hunter *et al*, 2007).

Homeodomain-containing proteins (HNF-1 alpha, HNF-1 beta) (Nagaki and Moriwaki, 2008).

Winged helix family proteins HNF-3 alpha, HNF-3 beta, and HNF-3 gamma (also called FoxA1, 2, and 3) (Lee *et al*, 2005).

Nuclear hormone receptor family (HNF-4, COUP-TFII, LRH-1, FXR alpha, and PXR) (Kamiya *et al*, 2003).

Basic leucine zipper-containing factor C/EBP alpha (Hsu *et al*, 1991)

Homeodomain protein HNF-6 (Clotman *et al*, 2002).

The hepatic artery development is influenced by vascular endothelial growth factor (VEGF) from developing bile duct cholangiocytes and angiopoietin-1 from hepatoblasts regulates remodelling of the hepatic artery (Fabris *et al*, 2008).

1.2.2 Physiology of the liver

The liver is a complex organ with many diverse functions. When the liver fails the complexity of the organ means there is no dialysis system which can replace it and therefore despite being able to support the liver, for many, liver transplantation is the only option. To develop specific therapies and so reduce the number of transplantations and deaths due to liver failure, it is essential to extend our knowledge of liver physiology in health as well as in disease. The investigation of rare causes of liver disease amenable to molecular genetic investigation may extend knowledge of hepatic intracellular pathways and interactions.

Unlike many other organs in the body the liver has the capacity to self repair and regenerate.

The physiology of the liver can be arbitrarily divided into:-

- Formation and excretion of bile
- Synthesis of clotting factors and proteins
- Immune function – formation of immune products and defence
- Excretion of hormones, drugs and waste products
- Storage of vitamins, minerals and nutrients
- The processing of nutrients from digestion

Bile acid physiology

Bile salts are formed by or circulated through hepatocytes. They are excreted into bile canaliculi where they mix with water and ions to form bile. The bile canaliculi join in ever increasing sizes to form bile ducts. Bile is stored in the gallbladder which in the presence of cholecystokinin (secreted due to a fat contain meal) empties bile into the small intestine to aid digestion and absorption of dietary fats including fat soluble vitamins. Bile salts are then recycled through the intestine back to the liver. A bile salt can circulate through the enterohepatic circulation up to 18 times. To replenish those lost in faeces cholesterol is converted to bile acids (Guyton, 1991).

The formation and secretion of bile is a unique function of the liver. The purpose of bile is (Lefebvre *et al*, 2009):-

- Bile salts in the intestine are essential to emulsify fats and form micelles with the lipid particles in the intestine to make them more soluble and creating a larger surface area. The larger surface area promotes the effective hydrolysis by lipases. The micelles direct the lipid to the intestinal mucosa for absorption including the fat soluble vitamins A, D, E and K
- The secretion of bile is an excretory pathway for cholesterol, bilirubin, porphyrins, drugs and infectious agents
- The production of bile from cholesterol is an important route for elimination of cholesterol
- Bile in the intestine regulates pancreatic secretions and gastrointestinal polypeptides

- They are implicated in pathways regulating apoptosis, mucin secretion and biliary ductular secretion
- Signalling molecules

Bile is synthesised in the hepatocyte from either dietary cholesterol or cholesterol synthesised from fat metabolism in the liver.

The basic structure of bile acids is a cyclopentanoperhydrophenanthrene (ABCD-ring) nucleus with a five carbon atom side chain and a terminal carboxylic acid.

The two primary bile acids are cholic acid (3α, 7α, 12α-trihydroxy-5β-cholan-24-oic) and chenodeoxycholic acid (3α, 7α-dihydroxy-5β-cholan-24-oic). Bile acid synthesis from cholesterol is catalyzed by individual enzymes and any deficiencies in these will prevent bile acid production and lead to the build up of other metabolities which may be hepatotoxic (Russell, 2003).

There are two pathways by which bile acids are formed

- Neutral (classical) pathway which is initiated by 7 α-hydroxylation of cholesterol and is the rate limiting step. The gene for this enzyme reaction is *CYP7A1*. This reaction is specific to the liver. Further hydroxylation results in the formation of chenodeoxycholic acid and cholic acid
- An alternative pathway is the side chain oxidation of cholesterol by 27-hydroxylase (*CYP27*). *CYP27* is expressed throughout the body including vascular endothelium where it may be important in the removal of cholesterol and therefore reducing vascular atherosclerosis. The oxidised cholesterol is then

converted to chenodeoxycholic acid by the liver. Mutations in CYP27 results in cerebrotendinous xanthomatosis which is characterised by the accumulation of cholesterol in tissue and atherosclerosis in vascular endothelium (Cali *et al*, 1991).

The bile acids are then conjugated with an amino acid either glycine or taurine (a smaller amount with sulfates and glucuronides) to form conjugated bile acids which are water soluble and excreted in bile. Any unconjugated bile acids are excreted in the urine. The pH within the intestine ionises conjugated bile acids usually with sodium so forming primary conjugated bile salts. Bacteria in the intestine dehydroxylate the primary conjugated bile salts to form secondary conjugated bile salts lithocholic acid from chenodeoxycholic acid and deoxycholic acid from cholic acid.

Regulation of bile acid synthesis

CYP7A1has numerous transcription factor binding sites within the promoter region (Chiang *et al*, 1998). It is down regulated by the Farnesoid nuclear receptor (FXR) through a series of reactions shown in figure 1.2.2 (Rizzo *et al*, 2005).

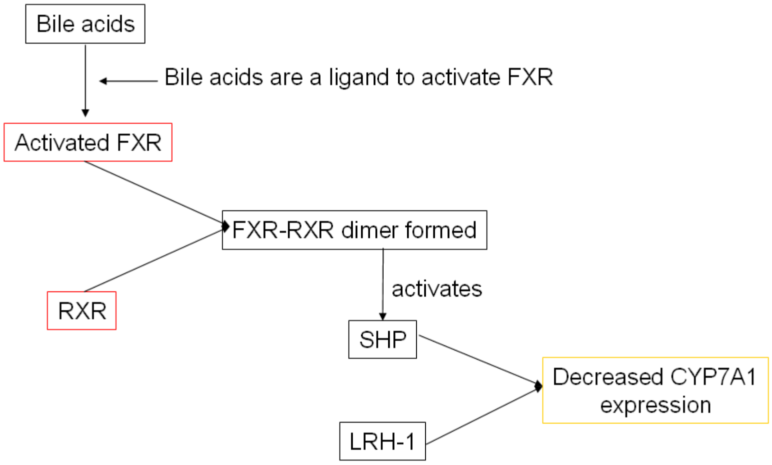


Figure 1.2.2 A diagram of the regulation of CYP7A1. Bile acids are a natural ligand for activation of FXR (Frankenberg *et al*, 2008). When activated the FXR forms a dimer with another nuclear receptor, retinoid nuclear receptor (RXR). This dimer activates short heterodimer partner (SHP). In association with liver receptor homolog-1 (LHR-1) CYP7A1 is repressed (Lu *et al*, 2000).

An increase in bile salts in the ileum stimulates the transcription of fibroblast growth factor (FGF) via FXR (Holt *et al*, 2003). FGF circulates and binds to FGF receptor 4 on the hepatocyte plasma membrane leading to a down regulation of CYP7A1.

FXR also stimulates bile salt export pump (BSEP) and organic solute transporter (OST) α-β whilst down regulating sodium-taurocholate cotransporter polypeptide (NTCP) and organic anion transporting polypeptides (OATPs) via SHP (Sanyal *et al*, 2007).

Bile acid transport across the apical membrane

The conjugated bile acids are transported into the canaliculus against a strong gradient. This is achieved by ATP binding cassette (ABC) superfamily of proteins which have a conserved intracellular domain which binds ATP and couples unidirectional movement of compounds with ATP hydrolysis (Weinman *et al*, 1998). This is shown schematically in figure 1.2.3.

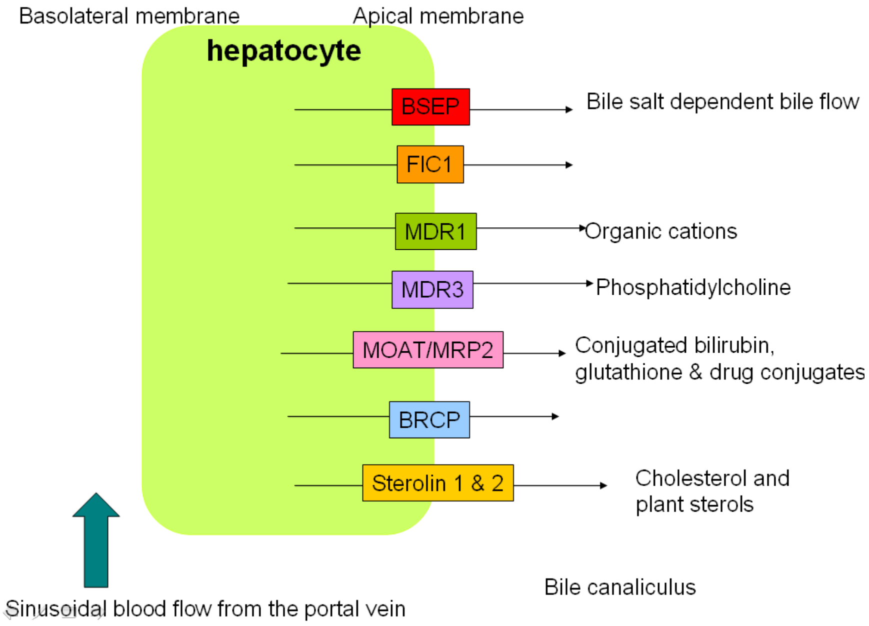


Figure 1.2.3 A schematic drawing of the apical membrane transport proteins of bile salts. Except for familial intrahepatic cholestasis 1 (FIC1) all the other transporters on the apical canalicular membrane of the hepatocyte are members of the ABC super family.

BSEP encoded by *ABCB11* determines bile salt dependent bile flow (Trauner and Boyer, 2003).

The multidrug resistant protein (MDR) family, transport the secondary bile acid lithocholic acid. MDR1 is a P-glycoprotein which is encoded by *ABCB1* and transports organic cations.

MDR3 is a phospholipid export pump which excretes phosphatidylcholine and is encoded by *ABCB4* mutations in which results in PFIC3.

Multi organic anion transporter (also known as multidrug resistence associated protein 2) (MOAT / MRP2) excretes conjugated bilirubin as well as glutathione and drug conjugates. This transporter is independent of bile flow. The gene encoding this protein is *ABCC2*. Mutations in MRP2 causes Dubin-Johnson syndrome. In these patients there is no hepatotoxicity suggesting alternative mechanisms must exist (Elferink and Groen, 2002).

Breast cancer resistance protein (BRCP) encoded by *ABCG2* transports similar compounds to MOAT (Choudhuri and Klaassen, 2006).

Sterolin 1 and 2 (*ABCG5* & *ABCG8*) form a heteromeric transporter for cholesterol and plant sterols (Yu *et al*, 2002).

FIC1 (an ATPase) is a phosphatidylserine flippase encoded by ATP8B1.

Efflux of bile salts across the basolateral membrane into the sinusoids

There is also a flow of bile salts through the basolateral membrane back into the sinusoids. This pathway increases in the presence of cholestasis. The schematic drawing in figure 1.2.4 demonstrates the transport proteins involved and also that FXR can also increase the expression of some of the proteins.

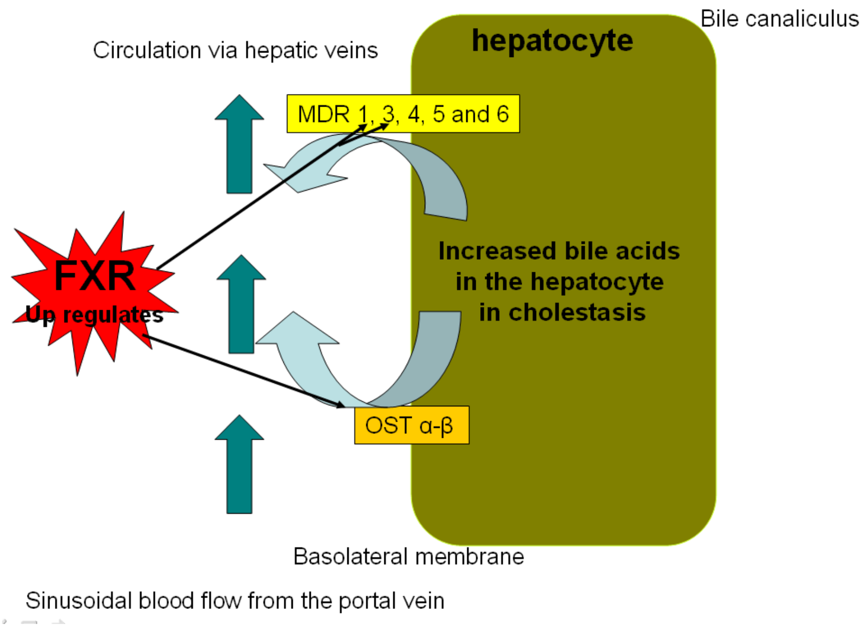


Figure 1.2.4 A schematic diagram demonstrating the transport proteins which efflux bile from the hepatocyte across the basolateral membrane into the sinusoids.

There is efflux of bile from the hepatocyte into the liver sinusoid and therefore entry into the general circulation via the portal veins. MDR 1, 3, 4, 5 and 6 on the basolateral membrane are important in this reaction. MRP3 (*ABCC3)* and MRP4 (*ABCC4*) are upregulated in cholestasis so removing bile acids from the hepatocyte but increasing the bile acid concentration in the circulation. OST α-β is also upregulated in liver disease via FXR (Kullak-Ublick *et al*, 2004).

Cholehepatic circulation

Within the large bile ducts there is passive as well as active transport of bile salts across the apical cholangiocyte border to the basolateral membrane and back into the sinusoids. The reason for this circulation is unknown but it may play a role in the conservation of bile salts or the bile salts may act as signalling molecules to regulate proliferation of bile ducts within the biliary tree (Trauner and Boyer, 2003).

Enterohepatic circulation

The intestinal bile acids are taken up by enterocytes into the portal venous circulation to the liver sinusoids for uptake through the hepatocyte basolateral membrane and is shown schematically in figure 1.2.5. If the absorption into the hepatocytes is not complete bile salts stay in the circulation and jaundice develops. Bile salts can be recirculated approximately 18 times before being lost in the faeces.

Different parts of the intestine have varying capacity to uptake bile salts.

The distal ileum expresses Na-dependent carrier uptake of taurine and glycine conjugate bile salts through the apical sodium dependent bile acid transporter (ASBT, also known as ileal sodium dependent bile acid transporter (ISBT) (Dawson *et al*, 2005). Mutations in *SLC10A2* which encodes ASBT, result in a syndrome of primary bile salt malabsorption with severe diarrhoea and failure to thrive through poor fat absorption (Oelkers *et al*, 1997).

The transport of bile salts from the enterocyte to the splanchnic circulation (and so to the liver sinusoids via the portal vein) is by OST α-β transport proteins.

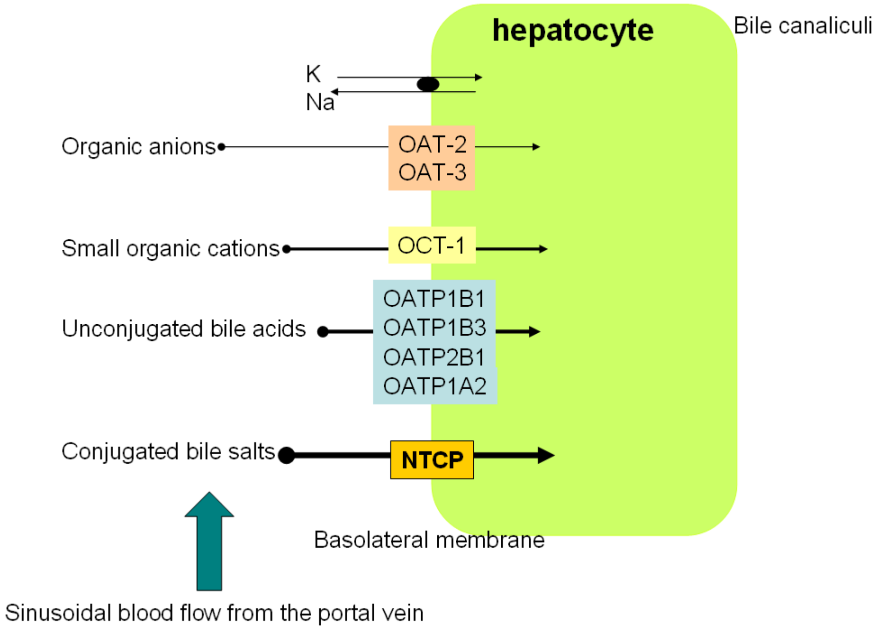


Figure 1.2.5 A schematic drawing of transport proteins at the basolateral membrane of the hepatocyte

The major route for conjugated bile salts to be transported into the hepatocyte is by NTCP which is encoded by *SLC10A1* and has a high affinity for taurocholate (Splinter *et al*, 2006). This is sodium dependent and therefore relies on a Na gradient created by Na/K ATPase. The transcription of NTCP can be increased such as by prolactin in the postpartum state or down regulated as in the presence of cholestasis, oestrogens or endotoxins (Donner *et al*, 2007).

The unconjugated bile acids are transported by OATP's, consisting of OATP1B1, OATP1B3, OATP2B1 and OATP1A2 which are encoded by *SLCO1B1, SLCO1B3,* *SLCO2B1* and *SLCO1A2* respectively. The OATP1B1 is the major sodium independent mechanism for the uptake of bile acids.

Small organic cations are taken up by organic cation 1 transporter OCT-1 *(SLC22A1).* Some organic anions are taken up by the organic anion transporters OAT-2 and 3 (*SLC22A6 and SLC22A8)* (Zhou and You, 2006).

Renal regulation of bile salts

Bile acids are filtered through the glomerulus and reabsorption occurs at the apical surface of the proximal tubule cells. This is a sodium dependent process by the transport protein ASBT. Unconjugated as well as conjugated bile acids are reabsorbed although sulphate conjugates are not excreted in the urine. At the basolateral surface transport into the circulation is predominantly by OST α-β and a small amount by MRP3 (Ballatori *et al,* 2005).

Inherited diseases of intrahepatic cholestasis

Intrahepatic cholestasis use to be an umbrella term for children who had neonatal hepatitis but with the advances in molecular medicine many of these cases now have a specific diagnosis and a large amount knowledge about the liver and metabolic pathways has been gained from studying cholestatic genes and protein expression in the normal as well as the pathological state. Examples of such genes are shown in Table 1.2.1.

Mechanism of disease	Disorder	Gene	Protein, function, substrate
Disorders of canalicular transport	PFIC-2, BRIC-2	*ABCB11* (Strautnieks *et al*, 1998)	Bile salt export pump to transport bile acids through the canalicular membrane
	PFIC-3, ICP, cholelithiasis	*ABCB4* (Deleuze *et al*, 1996)	Multi drug resistance protein 3, phospholipid flippase
	Dubin-Johnson syndrome	*ABCC2* (Tsujii *et al*, 1999)	Multi drug resistance associated protein 2. Regulates canalicular transport of GSH conjugates
Multi organ disorders	PFIC-1, BRIC-1, RFCF-1, GFC	*ATP8B1* (Bull *et al*, 1998)	Familial intrahepatic cholestasis-1. P-type ATPase
	NISCH	*CLDN1* (Hadj-Rabia *et al*, 2004)	Claudin-1; tight junction protein
	ARC syndrome	*VPS33B* (Gissen *et al*, 2004)	Protein trafficking
Altered ion transport	Cystic fibrosis	*CFTR* (Riordan *et al*, 1989)	cystic fibrosis transmembrane conductance regulator; chloride channel with ATP binding cassette ;regulates chloride transport
Disorders of embryogenesis	Alagille syndrome	*JAG1, NOTCH2* (Oda *et al*, 1997; McDaniell *et al*, 2007)	Transmembrane, cell-surface protein that interacts with NOTCH receptors to regulate hepatocyte to bile duct cell fate during embryogenesis
	ARPKD	*PKHD1* (Ward *et al*, 2002)	Fibrocystin; protein involved in ciliary function
	ADPLD	*PRKCSH* (Li *et al*, 2003)	Hepatocystin; protein assembles with glucosidase II apha subunit in endoplasmic reticulum

Metabolic diseases	A_1AT deficiency	*SERPINA1*	A_1AT; accumulation of mutant PiZZ in hepatocytes; decreased anti-proteolytic activity
	BASD: neonatal cholestasis with giant cell hepatitis	*AKR1D1* (Lemonde *et al*, 2003)	3-oxoΔ-4-steroid 5b-reductase; enzyme that regulates bile acid synthesis
	BASD: chronic intrahepatic cholestasis	*HSD3B7* (Cheng *et al*, 2003)	3β-hydroxy-5-C27-steroid oxido-reductase (C27-3β-HSD): enzyme that regulates bile acid synthesis
	BASD: neonatal cholestasis with giant cell hepatitis	*CYP7B1* (Setchell *et al*, 1998)	Oxysterol 7α-hydroxylase; enzyme that regulates the acidic pathway of bile acid synthesis
	FHC	*TJP2* (Carlton *et al*, 2003)	Tight junction protein-2; a family of guanylate kinase homologous that are involved in the organisation of epithelial and endothelial intracellular junction; regulates paracellular permeability
	FHC	*BAAT* (Carlton *et al*, 2003)	Bile acid CoA:amino acid N-acyltransferase; enzyme that transfers the bile acid moiety from the acyl-CoA thioester to either glycine or taurine
	FHC	*EPHX1* (Zhu *et al*, 2003)	Epoxide hydrolase-1; microsomal epoxide hydrolase regulates the activation and detoxification of endogenous chemicals
	Wilsons	*ATP7B* (Bull *et al*, 1993)	ATPase Cu^{2+}-transporting-beta; P-type ATPase; function as copper export pump
	NICCD	*SLC25A13* (Ohura *et al*, 2001)	Citrin; mitochondrial aspartate glutamate carrier involved in the malate-aspartate NADH shuttle
	Niemann Pick type C	*NPC1* (Carstea *et al*, 1997)	Abnormal cholesterol sterification and storage

Unclassified	NAICC	*CIRH1A* (Chagnon *et al*, 2002)	Cirhin; protein involved in cell signalling
	Villin deficiency	*VIL1* (Phillips *et al*, 2003)	Villin; protein involved in structural integrity of canalicular microvilli
	MAS	*GNAS1* (Weinstein *et al*, 1991)	Postzygotic activating mutations of arginine 201 leading to the constitutive activation of the guanine-nucleotide-binding protein (G protein) α subunit

Table 1.2.1 Provides a list of clinical conditions in which intrahepatic cholestasis presenting at any age is a dominant feature and a genetic defect has been identified. PFIC, progressive familial intrahepatic cholestasis; BRIC, benign recurrent intrahepatic cholestasis; ICP, intrahepatic cholestasis of pregnancy; RFCF, recurrent familial cholestasis of the Faeroe Islands; GFC, Greenland familial cholestasis; NISCH, neonatal ichthyosis-sclerosing cholangitis; ARC, arthrogryposis-renal dysfunction-cholestasis; ARPKD, autosomal recessive polycystic kidney disease; ADPLD, autosomal dominant polycystic liver disease; A_1AT, alpha-1-antitrypsin deficiency; BASD, bile acid synthetic defect; FHC, familial hypercholanemia; NICCD, North American Indian childhood cirrhosis; MAS, McCune-Albright syndrome
(adapted from Balistreri and Bezerra, 2006)

Hepatic synthesis of clotting factors

A table of the clotting factors which are made in the liver is shown in table 1.2.2. Those factors with an asterix also require vitamin K which may be deficient in liver disease due to poor absorption with a lack of bile salts and long chain fat absorption. Vitamin K is a cofactor for gamma-glutamyl carboxylase which adds a carboxyl group to glutamic acid residues of clotting factors II, VII, IX, X, protein S, protein C and protein Z (Blanchard *et al*, 1981).

FACTOR	NAME	SOURCE	PATHWAY
I	Fibrinogen	Liver	Common
II	Prothrombin (enzyme)	Liver *	Common
V	Proaccererin (heat labile cofactor)	Liver and Platelets	Extrinsic and Intrinsic
VII	Proconvertin (enzyme)	Liver *	Extrinsic
IX	Christmas factor(plasma thromboplastin component)	Liver *	Intrinsic
X	Stuart Prower factor (enzyme)	Liver *	Extrinsic and Intrinsic
XI	Plasma thromboplastin antecedent (enzyme)	Liver	Intrinsic
XII	Hageman factor	Liver	Intrinsic; also activates plasmin
XIII	Fibrin stabilizing factor	Liver	Retards fibrinolysis

Table 1.2.2 A table of coagulation factors made by the liver

A complication of liver disease is hypersplenism which may compound abnormal clotting by causing low platelets.

Hepatic synthesis of albumin

Albumin synthesis begins in the nucleus, where genes are transcribed into messenger ribonucleic acid (mRNA). The mRNA is secreted into the cytoplasm, where it is bound to ribosomes, forming polysomes that synthesize preproalbumin. Preproalbumin is an albumin molecule with a 24 amino acid extension at the N terminus. The amino acid extension signals insertion of preproalbumin into the membrane of the endoplasmic reticulum (Rhodes *et al*, 1989). Once inside the lumen of the endoplasmic reticulum, the leading 18 amino acids of this extension are cleaved, leaving proalbumin (albumin with the remaining extension of 6 amino acids). Proalbumin is the principal intracellular form of albumin. Proalbumin is exported to the Golgi apparatus, where the extension of 6 amino acids is removed prior to secretion of albumin by the hepatocyte. Once synthesized, albumin is secreted immediately; it is not stored in the liver.

Albumin transports various substances, including bilirubin, fatty acids, metals, ions, hormones, and exogenous drugs. One consequence of hypoalbuminemia is that drugs that are usually protein bound are free in the plasma, allowing for higher drug levels, more rapid hepatic metabolism, or both (Klammt *et al* 2007).

Liver immunity

The liver is constantly exposed to large varieties of antigens that are derived from the gastrointestinal tract, including dietary antigens, pathogens, and toxins. Its function as a major immune organ is now being appreciated. The liver lymphocyte population is enriched in macrophages (*ie,* Kupffer cells), natural killer and natural killer T cells, which constitute the innate immune system (Bilzer *et al*, 2006).

Hepatic excretion of hormones and drugs

The conjugation process in the liver also plays a major role in excreting cholesterol, hormones, and drugs from the body.

Hepatic processing and storage of vitamins, minerals and nutrients

The liver plays an important role in metabolizing nutrients such as carbohydrates, proteins, and fats. The liver helps metabolise carbohydrates in three ways:

- Through the process of glycogenesis, glucose, fructose, and galactose are converted to glycogen and stored in the liver.
- Through the process of glycogenolysis, the liver breaks down stored glycogen to maintain blood glucose levels when there is a decrease in carbohydrate intake.
- Through the process of gluconeogenesis, the liver synthesizes glucose from proteins or fats to maintain blood glucose levels.

Liver cells also chemically convert amino acids to produce ketoacids and ammonia, from which urea is formed and excreted in the urine. Digested fat is converted in the intestine to triglycerides, cholesterol, phospholipids, and lipoproteins. These substances are converted in the liver into glycerol and fatty acids, through a process known as ketogenesis (Guyton, 1991).

1.3 Thesis philosophy

Having worked in paediatric hepatology I have had first hand experience of the devastation liver disease can have on children and their families. Paediatric liver disease is uncommon and therefore relatively little is known in the general population. From the initial realisation a child has liver dysfunction to making a definitive diagnosis can be an extremely difficult time.

Why study the molecular genetic investigation of paediatric liver disease?

It is hypothesised that by investigating the molecular genetics of paediatric liver disease and thereby identifying mutations which are associated with the disease phenotype an accurate diagnosis can be made.

For the families and those affected the detection of a mutation within an identified gene can potentially:

- Provide accurate information for parents about their child’s condition
- Prevent unnecessary investigations
- Allow specific medical management and counselling as to the likely progression of the disease
- Enable counselling of the genetic risk for other family members and the risk to future children. Preimplantation or prenatal diagnostic testing may also be available
- Provide a specific diagnosis that is often required for parents to be able to access support services such as educational, social and financial as well as support groups

- Enable a greater understanding of the condition and therefore permit the identification of complications
- Prevent iatrogenic disease by pharmacogenetics

From a scientific point of view the identification of disease genes provides further information of intracellular pathways and cellular interactions. This allows greater understanding of normal and disease states and may also allow identification of potential therapeutic targets.

The title of this thesis is the molecular genetic investigation of paediatric liver disease and covers a number of different classifications of liver disease. Chapter 3 is the investigation of the multisystem disease phenotypic diarrhoea of infancy (PDI) which involves the liver in numerous ways, presenting with cirrhosis in infancy or liver failure from iron overload. Significant liver disease secondary to the need for parenteral nutrition may develop. Personally in my first paediatric post 12 years ago I looked after a child with PDI who sadly died from the complications of parenteral nutrition. The investigation of the molecular genetics of PDI has enabled me to identify the gene which is associated with PDI which is the first step in elucidating the molecular pathogenesis of the condition. Chapter 4 describes the role of cilia in liver disease. The knowledge regarding primary cilia is fast growing following the identification of primary cilia in the pathogenic process of polycystic kidney disease. It is likely to be a field which continues to grow and the list of conditions in which primary cilia are affected may increase. I have hypothesised that biliary atresia, a more common paediatric liver disease may also be a ciliopathy as the phenotype is also that of a ductal plate malformation, hepatic fibrosis, renal cyst development, and in the embryological form of biliary atresia, situs inversus. Chapter 5 investigates the role of bile salt transport proteins and their defects and describes a novel phenotype of acute neonatal liver failure associated with abnormality in bile salt transport which has not previously been described.

The identification of a mutation associated with the disease phenotype however is the first step in elucidating the molecular pathogenesis of the condition, the role of the identified gene in the normal and diseased state and whether the pathways involved are common to other forms of liver disease. As such this thesis has provided the basis for exciting future work in paediatric liver disease.

Chapter 2

Materials and methods

Contents

This chapter commences initially with a description of how the patients were ascertained for each project and the ethical approval information for each study.

The methods and materials in this chapter are those I have personally used. Techniques and materials used by collaborators are embedded in the relevant chapter and are highlighted in italics.

The materials, methods, equipment, computer software and mathematical prediction tools are discussed for each molecular genetic technique employed. The chapter in which the technique is used is also stated.

The methods have been written in list format rather than prose to enable the techniques to be replicated easily.

2.1 Subjects

2.1.1 Phenotypic diarrhoea of infancy (PDI)

I have personally been involved in the care of 9 patients with PDI in my career as a paediatric hepatologist and is the reason for my interested in studying this rare condition. The methods I used to identify further patients are:-

1. Those patient I personally know through clinical work were the majority of the research cohort (9 patients)
2. Identified as having PDI during the period of study by collaborating gastroenterologists (5 patients)

3 The presenting of clinical features and outline of this study at the European Society of Paediatric Gastroenterology, Hepatology and Nutrition, Hepatology Summer school in Hungary 2006 (0 patients)

4 The presenting of clinical features and outline of this study at the European Society of Paediatric Gastroenterology, Hepatology and Nutrition, Young Investigators meeting, Bavaria 2006 (1 patient)

5 A poster presentation of clinical features and outline of study at British Society of Human Genetics meeting, York 2006 (1 patient – who was already known to me)

6 A poster presentation of clinical features and outline of study at American Society of Human Genetic meeting, New Orleans 2006 (0 patients)

7 Writing to all leads clinical geneticists in the UK (1 patient – already known to me)

8 Writing to all the lead consultants in paediatric parenteral nutrition in the UK (0 patients)

9 Writing to the correspondence of clinical cases published in the literature (1 patient)

To ensure all the patients had the same clinical features a proforma of clinical details was completed by the referring clinician and a sample of hair was sent with the DNA sample. The proforma can be found in appendix 8.1. When screening genes for mutations, DNA from 4 probands was used, all of whom I had personally phenotyped to have the classical features of PDI.

2.1.2 Jeune asphyxiating thoracic dystrophy (JATD)

This large cohort of patients with JATD was ascertained prior to my commencing the project. All the patients had x-rays which had been examined by an expert in chondrodysplasias with the majority seen by Dr C Hall, Consultant Paediatric Radiologist, Great Ormond Street Hospital.

On commencing the study I reviewed the patient clinical details and all the patients had features which were compatible with JATD although the extra skeletal manifestations were very varied.

During the course of my investigation of JATD further samples were sent from Clinical Geneticists who were aware of the project.

One patient was known to me personally.

2.1.3 Biliary atresia

All children with biliary atresia are looked after in one of three centres in the country with Birmingham Children's Hospital being one. All children with biliary atresia are followed up long term for complications.

This cohort of patients therefore I have personally been involved in their clinical care, reviewed the histology of the liver and the liver biochemistry.

2.1.4 Neonatal liver failure

As with biliary atresia, due to Birmingham Children's Hospital being a designated paediatric hepatology centre these patients I was personally involved in their care when they have been referred for specialist hepatology management.

2.1.5 Control samples

Control DNA samples were provided in anonymised 96 well plates by the West Midlands Regional Genetics Laboratory. The plates are a by product of screening for cystic fibrosis. Each plate was composed of DNA from a single ethnicity (i.e. 96 well plate of Caucasian or 96 wells plate of Asian controls) to ensure the correct controls were used.

2.2 Ethical approval

PDI and JATD: National Autozygosity Mapping Research (NAMR) - molecular pathology of human genetic disease. South Birmingham Research Ethics approval (LREC number: CA/5175)

Biliary atresia and neonatal liver failure: Molecular genetics of multifactorial liver disease. South Birmingham Research Ethics approval (LREC number 2002/036)

Biliary atresia: South Birmingham ethics approval: Control of human liver stem cell growth and differentiation: An alternative source of hepatocytes (LREC 2002/158) and Liver Disease Registry and Tissue Bank (LREC 2003/028).

2.3 Materials and methods

A stepwise format has been used in this chapter to describe the methods so as to provide a practical approach to each experiment.

2.3.1 Extraction of DNA

DNA was extracted from either whole blood (by me or by the West Midlands Regional Genetics Service). Chapter 3 proband cousin of 8C had DNA extracted from cultured fibroblasts by myself (the fibroblasts had been stored and then cultured at Birmingham Children's Hospital).

These techniques chosen were those commonly used in the lab and therefore all reagents were available. DNA extracted using this method was used in all chapters of this thesis.

Materials used for DNA extraction from whole blood:-

Lysis buffer	Sigma
Protinase K	Boehringer Mannheim
Phenol	Sigma
Chloroform	Sigma
0.2M NaCl	Fischer Scientific
Ethanol	Fischer Scientific

Method to extract DNA from whole blood:-

DNA was extracted from blood using the Puregene Genomic DNA Purification Kit (Gentra systems) according to the manufacturer's instructions.

Materials used for DNA extraction from cultured cell fibroblasts:-

Trypsin	Invitrogen
Cell lysis solution	Sigma
RNase A solution	Sigma
Protein precipitation solution	Promega
Isopropanolol	Fischer Scientific
Ethanol	Fischer Scientific

Method to extract DNA from cultured fibroblasts:-

1. add 400µl trypsin and swill around the flask then aspirate out
2. Add 1µl trypsin and incubate at 37 ^{o}C for 5 minutes
3. remove all the fluid and the cells from the flask and put into eppendorphs
4. Spin each epindorf at 12,000 rev for 5 minutes to form a pellet of cells

Step 1: cell lysis

1. remove supernatant from eppendorph leaving 20µl in base
2. vortex vigorously to resuspend
3. add 300µl cell lysis solution and pipette up and down to lyse cells

Step 2: RNase treatment

4. add 1.5µl RNase A solution
5. invert tube 25 times and incubate at 37 ^{o}C for 5 minutes

Step 3: protein precipitation

6. place in ice for 1 minute

7. add 100µl Protein Precipitation solution
8. Vortex vigorously for 20 seconds
9. centrifuge 13,000 rev for 1 minute to form a tight pellet

Step 4: DNA precipitation

10. pour off supernatant and put protein pellet into clean 1.5ml ependdorf with 300µl 100% isopropranolol
11. invert 50 times
12. centrifuge at 13,000 rev for 1 minute
13. Pour off supernatant and drain on paper. Add 300µl 70% ethanol and invert several times
14. centrifuge at 13,000 rev for 1 minute and pour off supernatant
15. drain for 5 seconds on paper

Step 5: Hydration

16. add 50µl DNA hydration solution
17. vortex (medium) for 5 seconds
18. Incubate at 65 ^{o}C for 5 minutes
19. vortex for 5 seconds at medium speed
20. pulse spin down
21. store at -20 ^{o}C long term

The standard stock concentration of DNA extracted is 500µg/µl. The amount of DNA used for each PCR reaction was 40µg. To form a working solution:-

$$\frac{\text{What is wanted}}{\text{Concentration}} \times \text{volume} = \frac{20\mu g}{500\mu g} \times 100 = 4\mu l \text{ into } 100\mu l = 20\mu g/\mu l$$

2.3.2 Extraction of RNA

RNA was extracted from whole blood using Qiagen RNease mini kit as per manufacturer's instructions.

RNA was used for the gene expression microarray in section 3.5.

2.3.3 Conversion of RNA to cDNA

RNA was converted to cDNA for reverse transcriptase PCR:-

All reagents come in a single kit - Reverse transcription System - Promega

25 mM MgCl	4µl
RT buffer	2µl
dNTP's	2µl
RNAsin	0.5µl
AMVRTase	0.75µl
Random primer	1µl
dH_2O	7.75µl

Added all reagents to 2µl RNA

Incubate at 42 ^{o}C for 20 minutes and then terminate the reaction by increasing the temperature to 94 ^{o}C for 2 minutes.

Reverse transcriptase PCR was used to investigate splice site mutations in section 3.6.

2.3.4 Whole genome amplification

DNA from some of the patients with JATD had been used by other researchers and little DNA was left for this study. To increase the DNA quantity I amplified the DNA (whole genome amplification) using Genomphil kit from Amersham Biosciences. I followed the manufacturer's recommendations however I found the poor results from the advised precipitation therefore precipitated again using chloroform and phenol.

1. To the DNA add equal quantities of phenol and chloroform and vortex to make an emulsion
2. Spin for 5 minutes at top speed
3. Take off the aqueous part to use ~450mls
4. Add 4µl sodium acetate 0.2µM (Abbey Chemicals)
5. Add 2.5 x volume of 100% ethanol
6. Spin for 10 minutes then remove supernatant
7. Spin again and remove supernatant
8. Repeat process by adding sodium acetate and then 100% ethanol – this was then left in the -80^{o}C for 24 hours
9. Spin and take off supernatant
10. Add 70% Ethanol and spin for 10 minutes take of supernatant and allow drying in air
11. Resuspend in water

All DNA. RNA and cDNA was stored at -80^{0}C for long term storage whilst aliquots for daily use were stored at -20^{0}C.

2.3.5 Polymerase chain reaction (PCR)

The PCR reaction is the most important molecular genetic technique that I used in this project. It forms the basis of microsatellite markers, single nucleotide polymorphism genome wide scan and the sequencing of genes.

During the time period of this Ph.D. I used two different standard techniques. The initial method used separate reagents. This was prone to poor PCR product especially when the DNA was not optimal (when the DNA had been used on multiple occasions or was many years old). Due to the difficulties a trial of a premixed solution was used (Biomix red) which was found to be superior even with sub optimal DNA and I adopted this as my preferred method for PCR. The annealing temperatures required re optimising after switching to using Biomix red as in general a low temperature was required. The PCR for all PDI investigations was made using Biomix red. Both methods were used in the investigations of JATD and neonatal liver failure.

Primers were optimised using control DNA prior to using patient DNA.

Primers were designed using Exon Primer (http://ihg2.helmholtz-muenchen.de/ihg/ExonPrimer.html) which derives suggested primers from analysing cDNA and genomic DNA alignment. When choosing the primer recommendations from Exon primer I used the following criteria:-

- Ensure 60 bp from the start of the exon
- Ensure the length is 19-22 bp
- Ensure there is a C or G at the start and finish
- Ensure there is at least 50% C or G

Exon3 is another online site to identify primers however I did not find this site as easy to use and I found the primers from Exon Primer generally worked well. The primers were ordered on line through Sigma Genosys (I had an excellent quote for each base and all other companies were more expensive).

Initial PCR reagents using ABGene Taq

DNA 20µg/µl	1µl	
10x buffer	1µl	ABGene
MgCl 25mM	0.6µl	ABGene
dNTP's 10mM	1µl	Bioline
dH_2O	5.9µl	Sigma
Forward and reverse primer 10mM	0.4µl (0.2µl of each)	Sigma
ABGene Taq	0.1µl	ABGene

Biomix red PCR reaction

DNA, 20µg/µl	1µl	
Biomix red	5µl	Bioline
Forward and reverse primer 10mM	1µl	Sigma
dH_2O	3µl	Sigma

PCR reaction

1. denature at 95^{o}C -5 minutes
2. 95 ^{o}C -45 seconds
3. annealing 56 ^{o}C -45 seconds
4. extension 72 ^{o}C -45 seconds
5. repeat steps 2 to 4 for 27 reactions (total 28 reactions)
6. final extension 72 ^{o}C -5 minutes

To ensure the PCR product is optimal the product is run on an agarose gel electrophoresis. A clear single product should be seen.

The PCR product is not optimal if:-

There is more than one band

There is no band or only a faint band seen (microsatellites markers only require a small amount of PCR product and therefore even a faint band PCR product can be used. For sequencing a clear band is required).

If the product is not optimal then techniques which can be used to improve the product:-

Annealing temperature decrease to improve sensitivity

Annealing temperature increased to improve specificity

Other techniques employed to optimise PCR product

- increasing the magnesium concentration to improve sensitivity

- added acetamide (25%, Fluka) to improve specificity (I did not find this a useful technique although I tried it on a number of occasions)
- use different *Taq*
    - Hot Star - more specific
    - Bioexact – makes the reaction more specific

<u>PCR using Bioexact *Taq*</u>: PCR reaction as above

DNA 20µg/µl	1µl	
10x buffer	1.5µl	
MgCl 25mM	1.2µl	
dNTP's 10mM	1.5µl	
dH_2O	5.6µl	
Enhancer	3µl	Qaigen
Forward and reverse primer 10mM	0.5µl (each)	
Bioexact Taq	0.2µl	Qaigen

<u>PCR using Hot Star *Taq*</u>

DNA	1µl	
10x buffer	2.5µl	Bioline
5 x Q solution	5µl	Bioline
dNTP's	2µl	
f primer	1µl	
r primer	1µl	
dH_2O	12.85µl	
Hot star Taq	0.15µl	Bioline

Hot Star Taq PCR reaction

1. denature for 15 minutes at 95°C
2. 95°C for 45 seconds
3. annealing temp 57°C for 45 seconds
4. extension for 1 minute at 72°C
5. repeat steps 2-4 34 times
6. final extension at 72°C for 10 minutes

Hot Star and Bioexact were used for some of the genes when investigating JATD.

Nested PCR

The technique of nested PCR was used to make the PCR product specific to the region of interest. The initial PCR creates a large PCR product size repeating the PCR 28 times. More specific primers then are used with the initial PCR product as the DNA template.

RT-PCR

In sequencing large genes cDNA was used in chapter 4. cDNA was also sequenced .

2.3.6 Gel electrophoresis

Reagents for gel electrophoresis:-

Agarose	Invitrogen
Ethidium bromide	Sigma
10x TBE electrophoresis buffer	Invitrogen
100 bp ladder	Norgen
Bromophenol blue loading buffer	Promega
Hyperladder	Bioline

Method for gel electrophoresis:-

1g agarose is diluted in 100ml 1xTBE buffer. 3µl ethidium bromide is added and the gel set in a horizontal electrophoresis plate with combs.

When using ABgene *Taq* PCR method 5µl of PCR product is mixed with 5µl bromophenol blue loading buffer.

When using Biomix red PCR method the 5µl of the PCR product was directly inserted into the wells. A 100 base ladder was loaded for product size reference.

The electrophoresis was run at 170 volts for 15-20 minutes (until visually the PCR product could be seen to have moved an adequate distance in the gel).

The DNA was visualised using a UV light transilluminator.

If there is a single band then the PCR can be used for microsatellite markers or sequencing.

2.3.7 Microsatellite markers

An explanation of microsatellite markers is provided in section 1.2.3.4. Microsatellites were identified using UniSTS (http://www.ncbi.nlm.nih.gov.ezproxye.bham.ac.uk/sites/entrez?db=unists) and ordered from Sigma Genosys using FAM (blue) as the fluorescent marker attached to the primer (I found the blue marker to be the clearest). The initial PCR was carried out as described in 2.3.5. For JATD this was carried out using ABGene Taq technique. The PCR of microsatellites for PDI were carried out using biomix red technique.

The PCR product was diluted 140µl dH_2O (i.e. 1:15 dilution)

1µl of diluted PCR reaction was added to

10µl HiDi (Applied Biosystems)

and

0.05µl Genescan500 LIZ size standard (Applied Biosystems).

Denature at 95^o for 5 minutes and cooled on ice.

The microsatellites were run on ABI 3730 DNA Analyser (Applied Biosystems) and the PCR product sizes were analysed using Genemapper v3.0 (http://products.appliedbiosystems.com) as described in chapter 1.

Microsatellites were used for mapping in chapter 3 to refute linkage to the ARC syndrome locus. Microsatellites were then used to further explore regions of shared homozygosity which had been identified using SNP genome wide scans and in this way regions could be either excluded or deemed to be a region of interest.

In chapter 4 microsatellites were again used to explore the regions of shared homozygosity identified by the SNP genomewide scans. Microsatellites were also used to look for linkage to other complex B genes in children from consanguineous families.

Appendix 8.2 provides the sequences of the microsatellites used.

2.3.8 LOD score

When mapping DNA this tool predicts linkage. The score is obtained by comparing the likelihood of the results occurring due to linkage compared to occurring by chance. To infer statistic significance therefore a likelihood of more than 1 in 1000 which equates to a LOD score of 3 suggests there is linkage. The larger the LOD score the more significant the linkage and the less likely it has occurred by chance.

The LOD score can be calculated using the formula:-

$$\begin{aligned} LOD = Z &= \log_{10} \frac{\text{probability of birth sequence with a given linkage value}}{\text{probability of birth sequence with no linkage}} \\ &= \log_{10} \frac{(1-\theta)^{NR} \times \theta^{R}}{0.5^{(NR+R)}} \end{aligned}$$

When calculating LOD scores for this thesis I used the computer program Superlink to ascertain the LOD score.

(http://bioinfo.cs.technion.ac.il/superlink-online)

When using a SNP genome wide scan in conjunction with microsatellites I did not find using a calculation of LOD scores useful except to confirm there was linkage to a region.

2.3.9 Single nucleotide polymorphism genomewide scan

The SNP genome wide scans were carried out by Louise Tee laboratory technician. I inserted the results into an Excel spreadsheet for analysis. I manually analyzed the results to identify regions of homozygosity which were shared between affected probands. The Affymetrix analyser was used and all recommended consumables.

The basic principles of SNP array are the convergence of DNA hybridization, fluorescence microscopy, and solid surface DNA capture. The three mandatory components of the SNP arrays are:

1. The array that contains immobilized nucleic acid sequences or target
2. One or more labeled allele-specific oligonucleotide probes
3. A detection system that records and interprets the hybridization signal.

Each SNP on the array is interrogated with different probes. Affymetrix extended their SNP array range during the course of this research such that the initial SNP arrays were of 10K SNP's (initial used in chapter 3 and 4) whilst recent SNP arrays contained 250K SNP's.

2.3.10 Sequencing of genes

All genes were directly sequenced by me. Sequencing was carried out in all chapters. The initial step is a PCR reaction as described above.

The product is then run on an electrophoresis gel to identify the PCR reaction has successfully created a single product.

The PCR product is then cleaned. There are two techniques which were used during the course of this project. The initial method was using Exosap. This was timely and expensive so I trialled Microclean which I found to be much quicker and cheaper.
In some cases a single product was not obtained by PCR optimisation and therefore by identifying the desired product from the size ladder on the gel electrophoresis the band was cut out and DNA extracted from the band.

Exosap enzyme USB

7.5 µl PCR product was added to 3µl Exosap

The enzyme was incubated at 37^{o}C for 30 minutes and then deactivated at 80 ^{o}C for 15 minutes.

Purification by gel extraction Qiagen

1. Mix 7.5µl PCR reactant with 5µl loading buffer
2. Run DNA on gel 1% gel
3. Cut out band and place in a clear tube
4. Add 450µl GQ buffer to tube
5. Incubate at 55˚ bath for 10 mins
6. Add 150µl isopropanol (propan-2ol)
7. Put sample in the QIAquick columns and clear tube
8. Centrifuged 13,000r for 1 minute, discard supernatant fluid
9. Add 500µl QG buffer
10. Centrifuged 13,000r for 1 minutes, fluid discarded

11. Add 750µl PE buffer
12. Stand for 2 minutes
13. Centrifuge at 13,000r for 1 minute, discard fluid
14. Centrifuge again 13,000r for 1 minute
15. Put tube into a clean vial
16. 50µl water through membrane
17. Centrifuge for 1 minute with elute used for sequencing reaction

microCLEAN enzyme Webscientific

Add 5µl PCR product to 5µl microCLEAN. Stand for 10 minutes.

Centrifuge 4000rev for 40 minutes.

Inverse spin to remove supernatant at 500rev for 30 seconds.

Resuspend in 10µl water and allow to stand for 10 minutes prior to use.

The ‘cleaned’ PCR product was then sequenced using the Big Dye method in all cases.

<u>Reagents for sequencing</u>

2µl Big dye buffer	Applied biosystems
1µl dH_2O	
0.5µl Big Dye sequence terminator	Applied biosystems
2µl 10mM oligonucleotide primer either forward or reverse (2mM concentration was initially used but better results were obtained using a higher concentration)	
4.5µl purified PCR product	

1. 96°C for 30 seconds
2. 96°C for 30 seconds
3. 50 °C for 15 seconds
4. 60 °C for 4 minutes
5. repeat from step 2 29 times
6. end

The sequenced product was then precipitated using a standard technique throughout the research.

Reagents for precipitation of DNA

0.5M EDTA diluted to 0.125M	Sigma
Ethanol	Fischer Scientific

2µl 0.125M EDTA added to sequence reaction.

Centrifuged at 2000 rev for 30 seconds.

30µl 100% ethanol added and left for 10 minutes prior to centrifuging at 2000rev for 20 minutes. Supernatant removed by inversing the plate and spinning up to 250 rev only. 90µl 70% ethanol added and centrifuged at 2000 rev for 10 minutes. Repeat inversed spin. Sample allowed to dry in air for 5 minutes. Add 10µl HiDi formamide and centrifuge at 2000rev for 30 seconds. Denature at 95 °C for 5 minutes and then cooling on ice.

The ABI 3730 sequencer was used to analyse all sequencing of this project. The data was then manually analysed by me using Chromas (www.technelysium.com.au/**chromas**) against a reference sequencing provided by either a control DNA sample or the reference sequence from NCBI. Chromas is a program available to illustrate the electropherogram produced by the electrophoresis of DNA in an automated DNA sequencing machine. The electropherogram reading is a series of discrete peaks corresponding to each base. By convention

Adenine = green

Cytosine = blue

Guanine = black

Thymine = red

When I identified a variant I initially looked in NCBI and UCSC databases to see if it was a known SNP. I then used the tools Polyphen, BLAST and BLAT:-

NCBI – National Center for biotechnology information (www.ncbi.nlm.nih.gov)

UCSC – University of California Santa Cruz (http://genome.cse.ucsc.edu)

Ensembl – www.ensembl.org

Polyphen – Polyphen is a computerised program to predict the effect a change in amino acid would have on the structure and function of the protein. It takes into account the amino acid site, whether it forms part of a domain, what the structure will change to and whether it will change any of the interactions with ions. Using BLAST it looks at the likelihood of the amino acid being in the position of the base change on the likelihood of the changed amino acid being in this position. This then forms a PSIC score (position specific independent count). Less than 0.5 is likely to be benign, 0.5-1.5 possibly damaging and more than 1.5 probably damaging.

BLAST - Finds regions of local similarity between sequences. The program compares nucleotide or protein sequences to sequence databases and calculates the statistical significance of matches. BLAST can be used to infer functional and evolutionary relationships between sequences as well as help identify members of gene families. (http://blast.ncbi.nlm.nih.gov/Blast.cgi)

BLAT - This is a similar tool to BLAST but is quicker as it uses a stored nucleotide sequence rather than searching databases. This however means it is only good for primates. (www.genome.ucsc.edu)

Splice site prediction tool – Berkeley Drosophila Genome project (www.fruitfly.org). This was used in chapter 3 to ascertain the effect of sequence variants near the intron/ exon boundary to ascertain the effect on the splice site.

In all cases control samples were analysed to identify if the sequence variant was a previously unknown SNP.

2.3.11 Western blotting

Western blotting was used in chapter 4. As this was not a technique I was using regularly I found that it was time consuming, technically challenging and the results were not very satisfactory.

Reagents used

ProtoFLOW gel (30% acrylamide, 0.3% *bis)*	Flowgen
2.5M Tris pH 8.8	Sigma

10% SDS (sodium Dodecyl Sulphate)	Invitrogen

add just before use:-

10% APS	USB
TEMED	Fischer Biotech
Isopropranolol	Fischer Scientific
CAPS solution (3-[cyclohexylamino]-1-propanesulfonic acid)	Sigma Aldrich
Rainbow coloured protein molecular weight markers	Amersham
Methanol	Fischer Scientific
10x PBS (phosphate buffered saline)	Invitrogen
Tween	Sigma
BSA (bovine serum albumin)	Sigma
ECL Western blotting analysis system	Amersham
Developer and fixer	Kodak
X-ray film	Kodak

Antibodies

Cenexin was supplied by the Gull lab at Oxford University.

Alpha tubulin	
Rabbit anti-goat (horseradish peroxidase conjugate)	Sigma Aldrich
TRIS pH8.0	Sigma
1% IGE-PAL CA 630	Sigma
Glycerol	Sigma
Protease inhibitor	Sigma

PMSF (phenylmethylsulfryl fluoride)	Sigma
Coomasie Plus Protein Standard reagent	Pierce

Method

1. clean glass plates with ethanol
2. put plates together and seal bottom with 2% gel (0.2g in 10mls – put a line on the glass plate and then stand the two plates joined by bulldog clips on top)
3. make up SDS gel in universal containers (6% is good for most proteins but increase percentage for smaller proteins)

6% resolving gel (10ml):

30% acrylamide, 0.3% *bis*	2 ml
2.5M Tris pH 8.8	1.5 ml
10% SDS	0.1 ml
dH_2O	6.4 ml
Add just before use:-	
10% APS	0.067 ml
TEMED	0.01 ml

6 % stacking gel (5ml):

30% acrylamide, 0.3% *bis*	1 ml
0.5M Tris pH6.8	1.25 ml
10% SDS	0.05 ml

dH_2O	2.65 ml
add just before use:-	
10% APS	0.05 ml
TEMED	0.005 ml

4. 10mls of the resolving gel for 1 sheet and 5mls of the loading gel
5. before pouring on the resolving gel put in combs an mark 0.5cm below this to show how far to fill it.
6. Pour in resolving gel the pour on isopropanolol to a depth of 2mm to make the gel flat
7. When set (can tell this by the part left in the universal container) pour off isopropanolol and then add a small amount of water to clean it out. Use blotting paper to get rid of any moisture
8. Pour on stacking buffer and put in combs, allow to set
9. Take off clips and sealing gel from bottom, put into tank with windows inwards. Tighten so 10x buffer will stay in the central chamber.
10. denature proteins at 96 degrees for 3 mins then put on ice
11. load gels and a ladder (protein)
12. fill tank to cover electrodes with 1% SDS buffer solution, 100v until moved through stacker then 120v to see a good band

To transfer

13. Put gel in CAPS solution for 15-20 mins to equilibrate
    a. 100ml 10x CAPS
    b. 800 ml water
    c. 100ml methanol
14. equilibrate membrane
    a. 30 seconds methanol
    b. 2 minutes in shaker
    c. 1x CAPS for 10 minutes
15. soak sponges in CAPS
16. make up transfer order
    1. pad
    2. filter paper
    3. gel
    4. membrane
    5. filter paper
    6. pad
17. transfer in CAPS solution for 45mins at 50volts
18. place membrane on filter paper and dry

Immunodetection

19. Incubate filter paper with primary antibody in a blocking buffer (5% dried milk, 0.01% tween-20 in PBS) for 1 hour – rocking

20. rinse with PBS quickly then 3 x 5minute washes with PBS on rocker

21. incubate with secondary antibody (peroxidase conjugate) diluted 1:10000 in blocking buffer for 30 minutes on rocker

22. quickly rinse with PBS and then 3x 5 minute washes with PBS on rocker

23. detect antibody using ECL
    a. make 1:40 dilution of reagent B: reagent A
    b. pipette directly onto membrane
    c. leave for 5 minutes and then dab off excess
    d. cover with cling film

24. in a dark room expose x-ray film to membrane varying length from 30 seconds to 10 minutes

Chapter 3

Identification of the causative gene for Phenotypic Diarrhoea of Infancy

Contents

3.1 An overview of inherited congenital diarrhoea and hepatic involvement

An estimated 5.6 million children die per year from malnutrition with diarrhoea being a major contributory factor (Moszynski, 2006). The development of oral rehydration solution in 1970's revolutionised the management of outbreaks of infective diarrhoea (Fonatine *et al*, 2007) and the charity Medicins Sans Frontieres are now lobbying for the use of energy dense ready to use foods (such as peanut paste) to provide nutrition for those in famine stricken areas without the requirement of adding potentially contaminated water (Moszynski, 2006). These interventions manage the devastating symptoms and consequences of diarrhoea but do not prevent the occurrence or specifically treat the cause.

The genetic investigation of intestinal failure has resulted in a greater understanding of the pathways involved in gut function and pathogenesis. A genetic diagnosis aids patient management and enables prognostication. The genetic investigation of other rare causes of intestinal failure will provide further insight into the complexity of intestinal function and may further facilitate identification of potential therapeutic targets.

3.1.1 Inherited diarrhoea presenting in infancy

Phenotypic diarrhoea of infancy (PDI)

Diarrhoea presents within the first two months of life. It is associated with dysmorphic facial and hair features and a variable immunodeficiency. The phenotyping and genetic investigation leading to the identification of the causative gene *TTC37* is the subject of this chapter.

Congenital tufting enteropathy

This is a rare (1 in 50,000-100,000 live births in Western Europe (Goulet *et al*, 2007)) autosomal recessive condition initially described in 1994 (Reifen *et al*, 1994). It presents in the neonatal period with diarrhoea, failure to thrive and electrolyte disturbance. Prolonging life and nutrition can only occur through the use of parenteral nutrition which may be required life long. The histology of the intestine is abnormal with epithelial cell dysplasia, villous atrophy without any inflammatory component. Using the technique of autozygosity mapping in one Mexican-American consanguineous family in which two second cousins are affected, the gene for congenital tufting enteropathy was identified (Sivagnanam *et al*, 2008). The epithelial cell adhesion molecule (*EpCAM*) is at 2p21 and is comprised of two different transcripts with the largest consisting of 10 exons. *EpCAM* is expressed throughout the intestine and also in intestinal carcinomas where it acts as a calcium independent adhesion molecule and is being investigated as a potential immunomodulatory target for cancer treatment (Varga *et al*, 2004).

Congenital chloride diarrhoea

The diarrhoea in this condition has a high concentration of chloride. The gene was identified to be the Down-regulated adenoma gene (*DRA*) also known as *SLC26A3* (Hoglund *et al*, 1996). The protein encoded by this gene is a transmembrane glycoprotein which exchanges bicarbonate with chloride mainly in the lower intestinal tract. The reduction in bicarbonate exchange results in systemic alkalosis. The gene consists of 21 coding exons and is located at 7q31 and is adjacent to *CFTR* which is mutated in cystic fibrosis.

Microvillus inclusion disease (MVID)

There is a bimodal presentation of MVID with one group presenting in the neonatal period and another group presenting at 3-4 months of age. The diarrhoea is severe requiring parenteral nutrition to maintain life. Interestingly this congenital diarrhoea may improve allowing reduction and even discontinuation of nutritional support in older children. The histological features are diagnostic with shortened or absent apical villi and inclusions seen on electron microscopy. There is accumulation of PAS positive granules. The gene was identified using autozygosity mapping in a single consanguineous family from Turkey with two affected children who were first cousins (Muller *et al*, 2008). The region of interest contained 79 genes and *MYO5B* was considered a candidate gene as previous phenotyping of the condition showed a deficiency of myosin in the brush border in children with MVID (Carruthers *et al*, 1985).

Congenital sodium diarrhoea (CSD)

This abnormality of sodium-proton exchange of the brush border of the intestine, results in diarrhoea with excessive sodium loss and serum acidosis. It may present in the antenatal period with polyhydramnios and oedematous intestine (Koh *et al*, 1998). The severity varies with some reports of resolution of symptoms (Fell *et al*, 1992). The known sodium-proton exchange genes have been investigated and no mutations have been identified (Muller *et al*, 2000). *SPINT2* has recently been identified using autozygosity mapping techniques in a syndromic form of CSD (associated with hypertelorism, corneal erosions and anal or choanal atresia (Heinz-Erian *et al*, 2009)).

IPEX syndrome

Immundysregulation, polyendocrinopathy, enteropathy X-linked syndrome is caused by mutations in foxhead box P3 (*FOXP3*) (Bennett *et al*, 2001; Wildin *et al*, 2001). The gene was first identified using an animal model approach. The Scurfy mouse was found to have similar symptoms to that found in children with IPEX syndrome. The gene was identified in the mouse to be *Foxp3* (Brunkow *et al*, 2001) and the human ortholog was identified with mutations identified in children with IPEX syndrome. The syndrome develops due to a lack of regulatory T-cells.

Congenital enteropeptidase deficiency

Enteropeptidase is found on the brush border of the proximal small intestine. The serine protease activity is essential to cleave trypsinogen to an active form. *PPRS7* is at 21q21.1 and consists of 27 coding exons (Holzinger *et al*, 2002).

Congenital glucose-galactose malabsorption

Sodium-glucose cotransporter gene (*SGLT1*) has been identified as being mutated in this severe diarrhoeal disease which presents in the neonatal period when fed glucose, lactose and galactose. The mutations are predicted to cause truncation of the protein and loss of function (Martin *et al*, 1996). The histology of the intestine is normal.

Congenital lactase deficiency

This severe diarrhoea is rare worldwide but with an increased incidence in Finland of 1 in 60,000 (Norio, 2003). The diarrhoea presents when the infant takes lactase containing milk and is due to mutations in *LCT* (Kuokkanen *et al*, 2006). Hypercalcaemia and nephrocalcinosis can also occur in this condition (Saarela *et al*, 1995).

Junctional epidermolysis bullosa and pyloric atresia

This congenital diarrhoea is due to the desquamation of the epithelium of the small intestine. The skin is usually severely affected and is often the life limiting symptom however cases have been reported where there is no skin involvement but diarrhoea and pyloric atresia only (Salvestrini *et al*, 2008). In this case the association of pyloric atresia led to the investigation of the epidermolysis genes and mutations were identified in *ITGB4*. The histology showed abnormal IgG at the intestinal basement membrane. *ITGB4* is an integrin which mediate cell-cell or cell-matrix adhesion and *ITGB4* is specifically a receptor for laminins. It is an important gene in the development of intestinal carcinoma (Sashiyama *et al*, 2002).

Mitochondrial respiratory chain complex deficiency

Single case reports have described chronic diarrhoea with villous atrophy in infancy due to mutations resulting in mitochondrial complex I or complex III deficiencies (Cormier-Daire *et al*, 1994).

Cerebrotendinous xanthomatosis (CTX)

CTX is characterised initially by presenile cataracts, chronic diarrhoea and xanthoma formation especially over tendons (the Achilles is the most common) and progresses to involve the nervous system with cerebellar ataxia, spinal cord paresis and dementia. The condition is easily treated in the early stages by supplementing with chenodeoxycholic acid and therefore it is important to diagnose early. The gene for this condition was identified following the investigation of the cholesterol synthesis pathway showing that in these patients there is reduction in 27 sterol which is encoded by *CYP27A* (Skrede *et al*, 1986; Cali *et al*, 1991).

Carbohydrate deficient glycoprotein (CGD)

Diarrhoea can present at any age in CDG although typically it is infancy. In type 1 is associated with neurological features and facial dysmorphism however in infants this may not be prominent and the child may present with gastrointestinal or hepatic involvement. Up to 66% have hepatic fibrosis in childhood, others have steatosis and bile duct hamartomas have also been reported (Damen *et al*, 2004). There is a deficiency of glycosylation of serum and other glycoproteins. The diarrhoea is due to a protein losing enteropathy. The gene for CDG Ib is *MPI* and consists of 7 coding exons.

Immunodeficiency

All forms of immunodeficiency may present with diarrhoea and is commonly one of the presenting features of an immunological disease. For some this will be due to susceptibility to infection whilst in others diarrhoea is part of the pathogenesis.

Motility and multifactorial disorders resulting in infantile diarrhoea

The above is not a comprehensive list of causes of intestinal failure but those disorders in which a genetic aetiology has been characterised or is very clearly the case. Many other disorders result in intestinal failure in which the genetic involvement is complex and the aetiology multifactorial. For example in Hirschsprungs Disease, an abnormality in the innervation of the bowel resulting in varying lengths of aganglionosis, more than 16 genes have been associated with the disease, which may indicate the complexity of the embryological development of the intestine. Pseudo obstruction in which the muscle component of the intestine is affected is also likely to have a genetic component which has not yet been delineated. In others environmental factors play a role e.g. gastroschisis and necrotising enterocolitis but a genetic component is likely.

3.1.2 Hepatic involvement in congenital diarrhoea

The liver may or may not be involved in the primary pathological process. In PDI excessive iron accumulation can result in liver failure and in CGD Ib hepatic fibrosis may be present at diagnosis leading to portal hypertension and the development of varices. The liver may be involved in immunodeficiency secondary to hepatic infection such as with cytomegalovirus or cryptosporidium, or through the drug therapies used to treat the conditions. In CTX the enzyme deficiency in the cholesterol synthesis pathway is in the hepatocytes however the liver remains histologically and functionally normal.

Congenital diarrhoeas result in intestinal failure with huge water loss, electrolyte imbalance and malnutrition. Without parenteral nutrition (PN) these infants will die.

Long term PN can have deleterious effects on the liver with infants developing intestinal failure associated liver disease (IFALD) (Carter and Karpen, 2007). Cholestasis develops and the liver becomes fibrotic with portal hypertension developing. The onset of IFALD is variable but is influenced by the amount of enteral feeding which is being concurrently given, infections of central venous catheters and also the formulation of the PN itself (Diamanti, 2008). For those whose liver disease progresses and the intestinal failure remains, a small bowel and liver transplant is required (Mazariegos *et al*, 2008).

When a small bowel and liver transplant are considered it is essential to have a precise diagnosis as to the cause of the intestinal failure so to understand the long term prognosis of intestinal function. Mutation identification results in a specific diagnosis and aids specific and tailored clinical management of the child.

3.2 A literature review of phenotypic diarrhoea of infancy and the reported clinical features

3.2.1 Introduction

PDI is a rare cause of congenital diarrhoea with an estimate incidence of 1 in 300,000-400,000 in Western Europe although this may be higher in areas in which consanguinity is frequent.

Within the literature the syndrome has been labelled by three different names which demonstrate the multisystem involvement of the syndrome. The names used have been

- Phenotypic diarrhoea of infancy (PDI)
- Syndromic diarrhoea
- Tricho-hepatic-enteric syndrome (THE syndrome) (OMIM %222470)

For this thesis the name used throughout is phenotypic diarrhoea of infancy (PDI).

The syndrome was first reported by Stankler et al. in 1982, in which two siblings from non consanguineous union with unusual facies and abnormal scalp hair who developed unexplained diarrhoea causing dehydration, malnutrition and death. Post mortem findings included normal histology of the intestine and a cirrhotic liver.

Over a 27 year period a total of 25-27 PDI cases have been reported in 9 papers and the clinical findings described.

This chapter is an overview of the published literature on PDI and collates the reported clinical findings.

3.2.2 Stankler *et al*, 1982

This was the first published case report of siblings who had features of what is now known as PDI. The reported children were the second and third pregnancies of healthy non consanguineous parents, with the first pregnancy having resulted in an anencephalic still birth with severe spinal rachischisis.

The first affected infant developed diarrhoea aged 15 days and died aged 33 days from malnutrition. The child had facial features of low-set ears, prominent eyes, broad flat nose and large mouth. The post mortem showed massive intestinal haemorrhage. Other findings were thymic atrophy, abnormal lobulation of the lungs and an increased number of Islets of Langerhans. The intestinal histology was normal. The liver was found to be coarsely lobulated with bile staining and haemorrhage. Histologically the liver was extensively fibrotic with bile duct proliferation, occasional giant cells and regenerative parenchymal nodules. There was haemosiderosis of the liver, exocrine pancreas and thymus. The second child had similar facial features and developed diarrhoea aged 15 days. Death from malnutrition occurred aged 87 days and the post mortem findings were similar to that of the sibling. Differences were normal lung formation and the development of microcortical renal cysts in the second child.

The abnormal woolly hair was extensively investigated. On light microscopy many abnormal features were identified and included pili torti, partial and complete breaks in the hair and trichorrhexis nodosa. On electron microscopy the hair showed abnormal bud like appearances and termed trichorrhexis blastysis. The amino acid composition of the hair demonstrated low cystine and proline, and a raised aspartate and leucine whilst that of the parents was normal.

The authors concluded this was the first presentation of a previously unrecognised syndrome.

3.2.3 Girault *et al*, 1994

In 1994 the first case series of PDI was published and consisted of 8 patients. They were all born small for gestational age with the typical facial features and trichorrhexis nodosa. The onset of diarrhoea was extremely variable between 6 days and 6 months although the majority were in the first 2 months of life. All the children required parenteral nutrition (PN) to maintain life of which two were able to stop with improved enteral feed tolerance. All the children had moderate to severe villous atrophy on initial biopsies and two also had inflammation of the lamina propria. None of the children responded to immunisations and in contrast to subsequent publications none of this series had low serum immunoglobulins. Liver fibrosis was detected in two patients prior to the onset of PN and two other children died of cirrhosis after receiving PN. Three children had mental retardation.

3.2.4 Verloes *et al*, 1997

This paper described two siblings born to non consanguineous parents who had the typical facial and hair appearance of PDI. The first had progressive liver disease thought to be secondary to PN whilst the second child had liver dysfunction from birth. Both children died, the first from liver failure and sepsis and the second from CMV infection following liver transplant. At post mortem of the first child and explant examination of the second child, the liver was found to be severely iron over loaded. It was most severe in the hepatocytes but was also found in the Kupffer cells, biliary epithelium and portal macrophages. Extrahepatic iron was prominent in the thyroid, adrenal cortex, pancreas and pituitary glands in the first child but the thyroid was spared in the second. There was no iron in the reticuloendothelial system including the spleen. The distribution of iron is similar to that found in neonatal haemochromatosis.

The authors concluded that this syndrome is a form of neonatal haemochromatosis due to the iron overload of the liver and the distribution of extrahepatic iron sparing the reticuloendothelial system.

3.2.5 Goulet *et al*, 1998

This paper is a review of all infants with intractable diarrhoea and reports 6 of the 8 children presented by Girault as well as 2 new children with PDI which may be included in the paper by Martinez-Vinson. No new clinical findings were identified.

3.2.6 de Vries *et al*, 2000

In this single case report a child developed diarrhoea aged 5 weeks. He had intrauterine growth retardation, weighing 1345g at 35 weeks gestation. He had tetrology of Fallot. He required parenteral nutrition to maintain nutrition. He also had mild mental retardation. He underwent extensive immunological investigation. Cell mediated immunity and T lymphocytes were normal. Humoral immunity was found to be abnormal with low IgG and poor antibody specific response to tetanus and pneumococcal immunisations. Immunoglobulin electrophoresis showed oligoclonal gammopathy in IgM, IgG1K, IgA1K and IgA2K. The gammopathy improved after the administration of IgG and also with improvement in the child's general clinical condition.

The patient in this case report consented to be included in this genetic research to identify the gene for PDI.

3.2.7 Landers and Schroeder, 2003

This single case report was published in *Pediatric Dermatology* to emphasise the abnormal hair findings. This child had the typical facial appearance of PDI. The hair had trichorrhexis nodosa on every shaft. Severe diarrhoea started at 2 weeks of age and the child required parenteral nutrition. The histology of the bowel at two months of age showed subtotal villous atrophy and focal inflammation however on subsequent biopsies was normal. She developed elevated hepatic transaminases and hepatomegaly. A liver biopsy revealed cirrhosis and mild periportal inflammation. She had difficulties with fine motor movements and had mental retardation. The child had recurrent infections however the immunoglobulin levels were consistently normal although the total T and B cell levels were low.

The authors concluded that dermatologists should be aware of this condition as they may be referred patients with PDI due to the hair findings.

3.2.8 Teitelbaum *et al*, 2004

A poster at the American Society of Human Genetics conference 2004, presented two siblings with facial features in keeping with those described by Stankler et al. In these siblings the iron content of the organs was investigated using MRI. This identified increased iron deposition within the liver and thyroid but not elsewhere in the body. In both children the platelets were enlarged and further investigations showed canalicular dilatation, vacuolation, abnormal granule content and in the second sibling a complete absence of microtubules.

The family consented to the use of DNA from both siblings for this genetic research.

3.2.9 Barabino *et al*, 2004

This report follows up on one patient reported by Girault and also presents a new case. The follow up patient in the initial series remained dependent on PN aged 5 years however full enteral feeding was established at 6 years although intestinal absorption still remained difficult. The child had short stature and delayed puberty. The new child had a milder form of PDI. She presented with the same facial dysmorphism and hair findings and had severe diarrhoea from age 15 days. The intestine improved and PN was discontinued aged 1 year. She continued to have episodes of diarrhoea but otherwise was well. She had elevated IgA but no other immune dysfunction identified. She had short stature but normal pubertal development

3.2.10 Martinez-Vinson *et al*, 2005

A poster at the European Society of Pediatric Gastroenterology, Hepatology and Nutrition conference in 2005 described 8 children with syndromic diarrhoea. Despite coming from the same institution as Girault, the series was said to be novel. In this series all the typical clinical features of PDI were identified. Encouragingly 7 of the 8 patients were alive at the time of publication which may indicate better medical care of children on long term parenteral nutrition and the treatment of infections.

3.2.11 Dweikat *et al*, 2007

This case reports a child born to consanguineous parents who had all the features previously described and suggested it may be an inherited condition.

3.2.12 Fabre *et al*, 2007

This reports two cases of PDI. The clinical features are those already published but highlights THE syndrome and syndromic diarrhoea are the same disease entity and there is some heterogeneity between the liver findings.

The authors suggest that the clinical findings which are diagnostic are

- Characteristic facial dysmorphism
- Woolly hair
- Severe diarrhoea requiring parenteral nutrition
- Intra-uterine growth retardation
- Immunodepression

3.2.13 Egritas *et al*, 2009

This single case report is of a mildly affected patient who only presented aged 4 years with diarrhoea since birth and failure to thrive. She had the typical facial and hair features of PDI. Histology of the colon showed mild colitis with cryptitis and cryptic abscesses which has not been a typical feature of this condition previously.

3.2.14 Fabre *et al*, 2009

This is the first paper looking at the potential genetic cause of PDI. The cohort of patients studied consisted of 8 affected children from 7 families in which 4 were consanguineous. 7 candidate genes were selected by their known function. *EGFR* which encodes the epidermal growth factor receptor, *HRAS* which encodes a Rho-GTPase involved in the MAP-kinase pathway and is mutated in Costello syndrome (mental retardation with skin and hair anomalies), *JUP* and *DSP* which encode proteins within the desmosome (an intracellular adhesive junction involved in many different tissues). *CTNNB1*, *EPPK1* and *PLEC1* were also investigated. All genes were sequenced and no germline mutations identified.

3.2.15 Summary of published clinical features of PDI

Table 3.2.1 a,b,c provides a summary of the published clinical findings. From the published material a maximum of 29 patients with PDI (also called THE syndrome and syndromic diarrhoea) have been reported since 1982. The number of reported affected children is likely to be inaccurate due to the reporting of numerous series from the same institution and therefore likely to represent the same patient cohort.

The facial and hair dysmorphism are a constant finding. The onset of diarrhoea varies hugely from day 1 to 168 days. In all children there was failure to thrive and it may be the later diagnosis reflects less severe diarrhoea and therefore was not initially identified. Villous atrophy was seen on small bowel biopsies but no other diagnostic features were identified on histopathology examination. In 24 patients development had been assessed. Half had mild learning difficulties (12 patients), 1 child had features of autism and one child had difficulties with fine motor movements and learning difficulties. The other children were reported to have normal development. Mild developmental delay may be secondary to prolonged hospitalisation, being physically attached to continuous infusions and recurrent infections and not necessarily part of the clinical phenotype of PDI.

Three different heart defected were identified (atrial septal defect, ventricular septal defect and tetrology of Fallot).

In 18 children the liver was commented on. In all of these cases there was fibrosis or cirrhosis. In 4 of these cases excessive iron was noted. Excess iron was also identified in the pancreas, thymus, spleen, thyroid, adrenal glands, kidneys, Langerhans cells and pituitary gland.

The immune system was commented on in 24 cases. Of these, 18 had poor response to vaccinations. 8 were also reported to have low immunoglobulins however this appeared to improve with time and led to the development of monoclonal gammopathies. One child was reported to have normal immunity and one had an isolated increased IgA. Of the 29 children reported 13 were reported to have died. This includes the initially reported 2 siblings who did not receive PN. The others died between 6 months and 5 years. 5 died secondary to liver complications including one child following a liver transplant. The others died secondary to infection.

Table 3.2.1 part a: the demographics, birth details, facial and hair findings and intestinal histology of published cases

Paper	Year	Consanguineous parents	Birth weight	Gestation	Day diarrhoea diagnosed	Small bowel histology	PN required	Broad flat nose	Hypertel -orism	Low set ears	Woolly hair
Stankler	1982	no	1680	38	17	normal		Yes	Yes	Yes	Yes
		no	1620	39	15	normal		Yes	Yes	Yes	Yes
de Vries	2000	no	1345	35	35	moderate villous atrophy. Variable mononuclear cellularity of lamina propria	Yes	Yes		Yes	No
Landers	2003	yes		40	14	subtotal villous atrophy with focal inflammation	Yes	Yes	Yes	Yes	Yes
Girault	1994	no	2100	34	168	moderate to severe villous atrophy	Yes	Yes	Yes	Yes	
		no	1520	37	21	moderate to severe villous atrophy	Yes	Yes	Yes	Yes	
		no	1480	37	6	moderate to severe villous atrophy	Yes	Yes	Yes	Yes	
		no	1940	40	14	moderate to severe villous atrophy	Yes	Yes	Yes	Yes	
		no	2670	40	21	moderate to severe villous atrophy	Yes	Yes	Yes	Yes	
		no	1180	35	56	moderate to severe villous atrophy. Increased mononuclear cells in the lamina propria	Yes	Yes	Yes	Yes	
		yes	2000	40	42	moderate to severe villous atrophy	Yes	Yes	Yes	Yes	
		yes	1950	39	10	moderate to severe villous atrophy. Increased mononuclear cells in the lamina propria	Yes	Yes	Yes	Yes	

Paper	Year	Consanguineous parents	Birth weight	Gestation	Day diarrhoea diagnosed	Small bowel histology	PN required	Broad flat nose	Hypertel -orism	Low set ears	Woolly hair
Teitelbaum	2004	no		40		not diagnostic	Yes	Yes	Yes	Yes	Yes
		no		40		not diagnostic	Yes	Yes	Yes	Yes	Yes
Martinez-Vinson	2005	yes	small		within 1 month	severe villous atrophy	Yes	Yes	Yes	Yes	Yes
		yes	small		within 1 month	severe villous atrophy	Yes	Yes	Yes	Yes	Yes
		yes	small		within 1 month	partial villous atrophy	Yes	Yes	Yes	Yes	Yes
		no	small		within 1 month	partial villous atrophy	Yes	Yes	Yes	Yes	Yes
		no	small		within 1 month	partial villous atrophy	Yes	Yes	Yes	Yes	Yes
		no	small		within 1 month	partial villous atrophy	Yes	Yes	Yes	Yes	Yes
		no	small		within 1 month	partial villous atrophy	Yes	Yes	Yes	Yes	Yes
		no	small		within 1 month		Yes	Yes	Yes	Yes	Yes
Verloes	1997	no	1410	34	7		Yes	Yes	Yes	Yes	Yes
		no	1860	37	7	villous atrophy	Yes	Yes	Yes	Yes	Yes
Dweikat	2007	yes	3250	40	50	flattened villi	Yes	Yes	Yes	Yes	Yes
Goulet	1997	no	1400	37	6		Yes				
		no	1760	39	10						
Fabre	2007	no	1240	32		total villous atrophy		Yes	Yes		Yes
		yes	2200	40	1	moderate non specific inflammation	Yes	Yes	Yes		Yes
Egritas	2008	no	1600	32	84	normal duodenum	No	Yes	Yes		Yes
Barabino	2004	no	1800	40	15	normal	Yes	Yes	Yes		Yes

Table 3.2.1 part b: the hair histology, development, cardiac abnormalities, pancreas and liver histology of published cases

Paper	Hair histology	Development	Heart	Pancreas	Liver
Stankler	Trichorrhexis nodosa		normal	Islet cell hyperplasia	Extensive fibrosis, bile duct proliferation, occasional giant cells and regenerative parenchymal nodules.
	Trichorrhexis nodosa		normal	Islet cell hyperplasia	Extensive fibrosis, bile duct proliferation, occasional giant cells and regenerative parenchymal nodules.
de Vries	Normal	mild mental retardation	tetrology of fallot		
Landers	Trichorrhexis nodosa	difficulty with fine motor movements and learning difficulties			Cirrhosis with mild periportal inflammation
Girault	Trichorrhexis nodosa	normal	normal		Cirrhosis
	Trichorrhexis nodosa	normal	normal		Cirrhosis
		normal	normal		
		mental retardation	normal		
		normal	normal		
	Trichorrhexis nodosa	mental retardation	normal		Cirrhosis
	Trichorrhexis nodosa	mental retardation	normal		Cirrhosis
	Trichorrhexis nodosa	normal	normal		
Teitelbaum	Trichorrhexis nodosa		VSD	normal	Fibrosis
	Trichorrhexis nodosa		normal		Fibrosis

Paper	Hair histology	Development	Heart	Pancreas	Liver
Martinez-Vinson		mental retardation	normal		Cirrhosis
		mental retardation	normal		Cirrhosis
		mental retardation	normal		Cirrhosis
		mental retardation	normal		
		mental retardation	normal		
		mental retardation	normal		
		mental retardation	normal		
		normal	normal		
Verloes	Trichorrhexis nodosa	normal	ASD	Islet cell hyperplasia	Cirrhosis, severe cholestasis with rosetting, ductular proliferation, giant cells and extramedullary hematopoiesis. Severe iron deposition
	Trichorrhexis nodosa	normal			Hepatic dysfunction from birth. Cirrhosis, severe cholestasis with rosetting, ductular proliferation, giant cells and extramedullary hematopoiesis. Severe iron deposition
Dweikat	Trichorrhexis nodosa	normal	normal		Portal oedema and fibrosis. Marked deposition of iron
Goulet					
Fabre	Trichorrhexis nodosa		normal		Cholestasis and hepatomegaly, micronodular cirrhosis with paucity of intrahepatic ducts. No iron overload
	Trichorrhexis nodosa	mental retardation and developmental delay with mild autism	normal		Cirrhosis at 1 month of age with some iron overload (prior to PN)
Egritas		normal			Hepatomegaly with histology macrovesicular steatosis, mild to moderate mononuclear infiltrate in portal areas. No iron
Barabino	Trichorrhexis nodosa	mild impairment	normal		

Table 3.2.1 part c: the abnormal iron distribution, the immunodeficiency and the age and cause of death in published cases

Paper	Iron	Immunity	Age at death	Cause of death
Stankler	Liver, pancrease, thymus		33 days	Malnutrition
	Liver, pancrease, thymus, spleen		87 days	Malnutrition
de Vries		Low specific antibody responses vaccines. Oligoclonal gammopathy of IgM, IgG and IgA. Improvement with time		
Landers		Mildly decreased T and B cells. Immunoglobulins normal		
Girault		No response to immunisations and immunoglobulin levels were normal or raised with a monoclonal gammopathy of IgA		
		No response to immunisations and immunoglobulin levels were normal or raised with a monoclonal gammopathy of IgA		
		No response to immunisations	38 months	Pneumonia
		No response to immunisations	26 months	Sepsis
		No response to immunisations	20 months	Pneumonia following bone marrow transplant
		No response to immunisations and immunoglobulin levels were normal or raised with a monoclonal gammopathy of IgA	26 months	Cirrhosis
		No response to immunisations	58 months	Cirrhosis
		No response to immunisations		
Teitelbaum	Liver and thyroid	Mild immunodeficiency	6 months	Sepsis and liver dysfunction
	Liver and thyroid	Mild immunodeficiency	died	

Paper	Iron	Immunity	Age at death	Cause of death
Martinez-Vinson		Absence of response to vaccines and hypogamaglobulinaemia which resolved	died	Infection
		Absence of response to vaccines and hypogamaglobulinaemia which resolved		
		Absence of response to vaccines and hypogamaglobulinaemia		
		Absence of response to vaccines and hypogamaglobulinaemia		
		Absence of response to vaccines and hypogamaglobulinaemia		
		Absence of response to vaccines and hypogamaglobulinaemia		
		Absence of response to vaccines		
		Absence of response to vaccines		
Verloes	Liver, thyroid, adrenal cortex, pancreas and pituitary glands		6 months	Liver failure and sepsis
	Liver, adrenal glands, kidneys, langerhans cells - not thyroid		6 months	CMV hepatitis post liver transplant
Dweikat	Liver		10 months	Diarrhoea
Goulet				
Fabre	No	Poor response to immunisations		
	No	Decreased IgA and IgG initially then an increase in IgG - monoclonal gammopathy		
Egritas		Normal		
Barabino		Elevated IgA		

3.3 Clinical features of a novel PDI cohort

3.3.1 Introduction

The identification of patients with PDI and ascertaining consent and DNA samples was the initial focus of work towards this thesis. Patients were identified by:-

- Known personally through clinical work to the investigator (JH)
- Identified as having PDI during the period of study by collaborating gastroenterologists
- The presenting of clinical features and outline of this study at the European Society of Paediatric Gastroenterology, Hepatology and Nutrition, Hepatology Summer school in Hungary 2006, by JH
- The presenting of clinical features and outline of this study at the European Society of Paediatric Gastroenterology, Hepatology and Nutrition, Young Investigators meeting, Bavaria 2006, by JH
- A poster presentation of clinical features and outline of study at British Society of Human Genetics meeting, York 2006, by JH
- A poster presentation of clinical features and outline of study at American Society of Human Genetic meeting, New Orleans 2006, by JH
- Writing to all leads clinical geneticists in the UK
- Writing to all the lead consultants in paediatric parenteral nutrition in the UK
- Writing to the correspondence of clinical cases published in the literature

Using these methods 16 patients with PDI were identified and consented for the study. Due to writing to correspondence of published case reports, three patients within this cohort had previously been reported in the literature.

In all families a unique identifying number was given and a letter to designate the person within the family.

A=father

B=mother

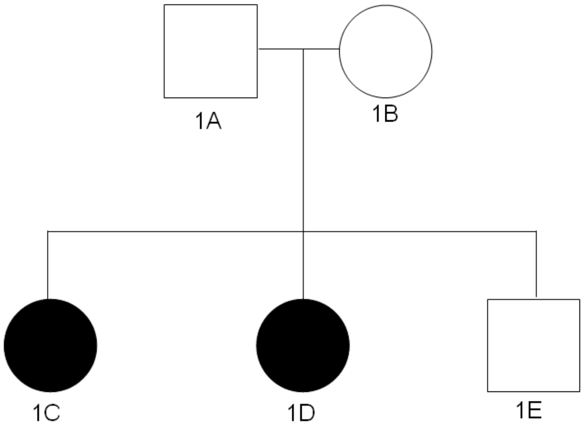
C=affected child

D and onwards = siblings

Table 3.3.1 provides a summary of the clinical findings of this study cohort of PDI patients.

3.3.2 Family pedigrees and clinical features

Family 1

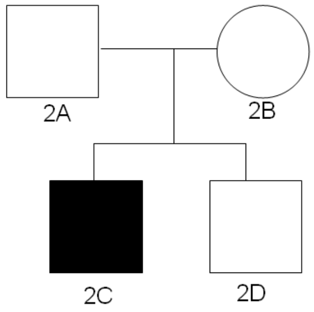


The two affected children in this non consanguineous family were initially presented as a poster at the American Society of Human Genetics conference 2004 by Teitelbaum *et al*.

The parents were originally from a small village in India and although consanguinity was denied it is possible they may be distantly related. The two affected children were typical of PDI. Unusual features were the severity of the iron within the Kuppfer cells of the liver which contributed to death and iron was also noted in the thyroid gland.

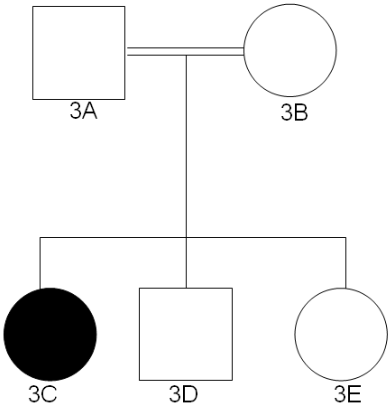
Both children were noted to have enlarged platelets and electron microscopy of the platelets was abnormal.

Family 2



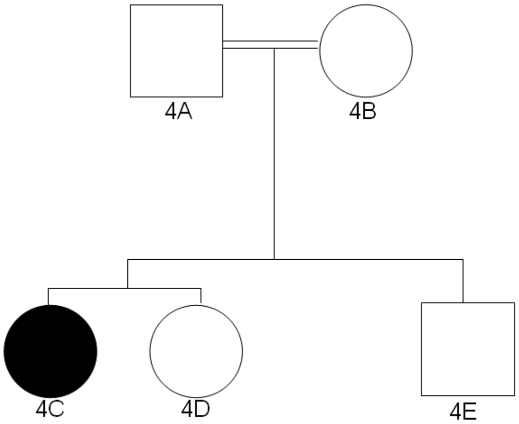
This Dutch Caucasian non consanguineous family were reported in the literature by de Vries *et al* in 2000. This child had all the typical features of PDI but also had Tetrology of Fallot.

Family 3



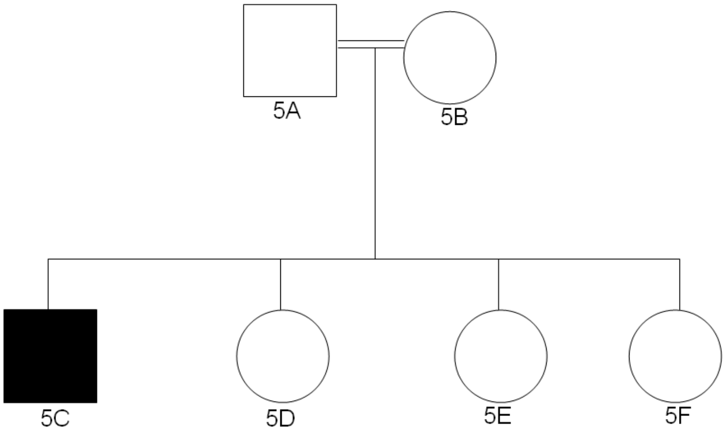
This family originated from the Mirpur region of Pakistan. The affected child has the typical features of PDI. In addition she has mild aortic insufficiency and mild ureteric reflux. She also has von Willibrands factor deficiency which has not been reported in any other affected child.

Family 4



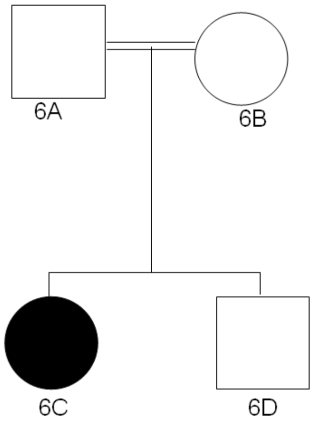
This family are from Iraq and are Kurdish in origin. The twins are dizygotic with only one being affected. A cousin was also affected by diarrhoea and liver disease but as the family lived in the rural hills of Iraq they did not have access to investigations or medical care and the child died. Child 4C had all the clinical features typical of PDI at the time of this study however her initial presentation in the neonatal period was with an enlarged and nodular liver. She was initially investigated for infective causes of liver disease but on biopsy was found to have cirrhosis. During these investigations she had normal stools and developed diarrhoea following discharge from hospital. Despite cirrhosis the liver function has remained normal.

Family 5



This consanguineous family originate from Mirpur region of Pakistan. The father has marked vitiligo. This child has all the typical features of PDI. He also has mild pulmonary stenosis which has not required intervention.

Family 6

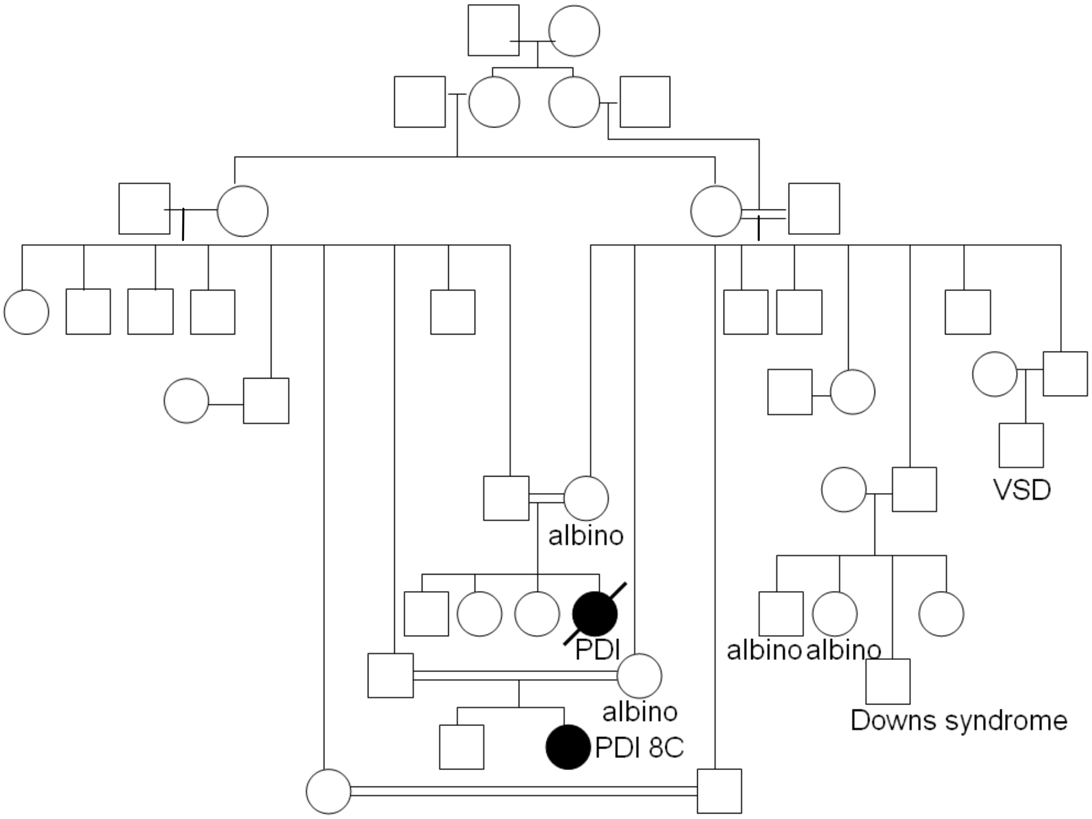


This family originate from the Mirpur region of Pakistan. The single affected child has the typical features of PDI.

Family 7

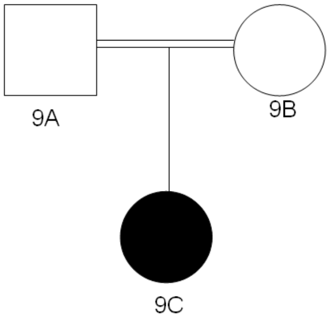
The single affected child of this consanguineous family died aged 8 years from complications of central venous catheters. This study commenced 4 years after his death and although DNA was available and the family consented for it’s use in this research project none of the other family members were available to provide DNA samples.
The family originated from Pakistan and this child was the only affected in the family. He had typical features of PDI.

Family 8



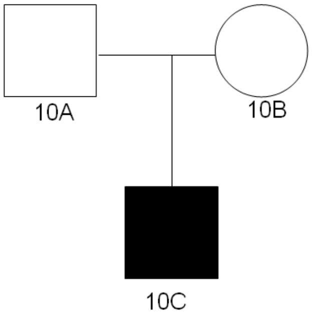
This multiple consanguineous family originate from Pakistan. The initial child with PDI investigated is labelled 8C and the other is known as the cousin of 8C. This family is unusual in the degree of albinism. PDI patients have been noted to have reduced pigmentation as compared to others in their families however true albinism has not been a feature. This may reflect a separate pathology in a multiple consanguineous family however the albinism seems to have occurred in both mothers who have had children with PDI and therefore may be important in the pathogenesis. Two children with albinism do not have PDI. Both affected children in this group had severe liver disease with excessive iron.

Family 9



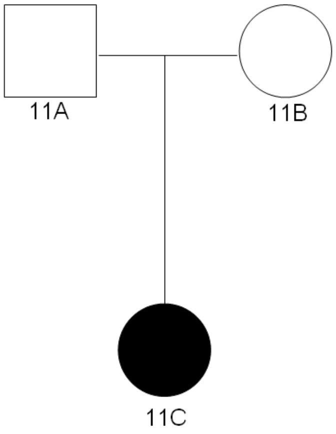
This consanguineous family originate from Italy. The parents are second cousins. This child had mild liver disease with some steatohepatitis on biopsy and no fibrosis. The immunodeficiency is similar to that found in other PDI patients with a monoclonal band of IgM and then IgG identified. This affected case is the oldest surviving in this cohort and any previously reported in the literature and is currently aged 21 years.

Family 10



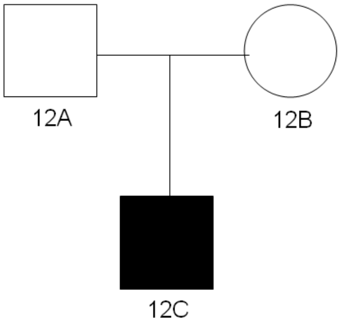
This non consanguineous family are Caucasian English. The affected child was mildly affected by PDI with PN only required for 4 months. The immunodeficiency of low IgG resolved with time. Other clinical features are as those seen in other patients with PDI.

Family 11



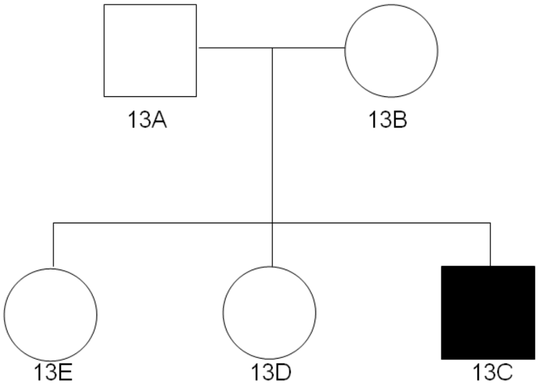
This family are non consanguineous Caucasian and originate in Italy. This child has diffuse hypopigmentation as compared to her family. She also has aortic insufficiency. She has hepatomegaly but has not had a liver biopsy. She has thrombocytopenia the cause of which is unknown.

Family 12



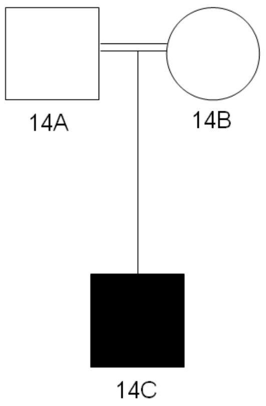
This family are Caucasian Flemish. The child has severe diarrhoea diagnosed aged 2 weeks and continues to require PN every night at the age of 3 years. He has mild developmental delay. He is diffusely hypopigmented. He has been noted to have thrombocytosis.

Family 13



This English Caucasian family are non consanguineous. The affected child is unusual in that he has upper motor neuron signs as well as developmental delay. There is a history of birth asphyxia which may account for these findings. This may be a novel feature of PDI or these signs may indicate that this child is not the same phenotype as PDI. Again unlike the rest of the cohort this child has intermittent neutropenia but immunoglobulins are normal. There appears to be dysmotility of the gut which again has not been seen in other patients. The diarrhoea is severe and the child has required a small bowel transplant. Although this child has been included in this genetic study it is with the caveat that there are subtle differences.

Family 14



This first born child of consanguineous family presented during the course of this study and had not been fully investigated for intractable diarrhoea. The child had abnormal hair with trichorrhexis nodosa and typical facies. The child had failure to thrive and was being treated with a trial of enteral feed manipulation but it was thought that he may require PN. At the time of investigation there were no immunodeficiencies identified. The liver was found to be heterogenous on ultrasound scan which may indicate there is a fibrosis

Table 3.3.1 is a tabulated summary of the clinical features in this cohort of PDI patients

Identification number	1C	1D	2C	3C	4C	5C	6C	7C	8C	9C	10C	11C	12C	13C	14C
Ethnicity	Indian	Indian	Dutch	Pakistani	Kurdish	Pakistani	Pakistani	Pakistani	Pakistani	Italian	English	Italian	Flemish	English	Pakistani
Consanguinity	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	No	No	No	Yes
Family history	Sibling	Sibling	No	No	Cousin	No	No	No	Cousin	No	No	No	No	No	No
Sex	Female	Female	Male	Female	Female	Male	Female	Male	Female	Female	Male	Female	Male	Male	Male
Current age	Died 6/12	died 6/12	12 years	11 years	2 1/2 years	3 1/2 years	14 months	died 8 years	9 months	21 years	3 years	13 years	2 years	4 years	1 year
Gestation	40/40	40/40	35/40	34/40	34/40	30/40				39/40	37/40	33/40	34/40	29/40	35/40
Birth weight	IUGR	IUGR	1345g	1410g	1220g	980g				1960g	3580g	780g	1700g		1375g
Dysmorphology															
Wide forehead	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Hypertelorism	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Café au lait spots				Yes				Yes		Yes		Diffuse hypopigmentation	fair skin	Yes and general pale skin	
TN	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Nutrition															
Age when PN started			7 months	1 month	2 months	3 months	1 month	5 months	1 month	1 month	4 months	1 month	2 weeks		
Current PN regime			off	5 nights	3 months	5 nights	7 nights	NA	7 nights	off	off	5 nights	7 night	7 nights	
PN stopped			3 years					did Not stop		4 years	5 months				
GI biopsies															
Villous atrophy	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Haematology															
Low immunoglobulins	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Intermittent neutropenia	No
Thrombocytosis	Yes	Yes	No	No	No	Yes	No			No	No	No	Yes	No	No
Large platelets	Yes	Yes		Yes	Yes									No	No
Liver and spleen															
Fibrosis	Yes	Yes	No	Yes	Yes	Yes				fatty changes	No	No	No	Yes	No
Haemosiderosis	Yes	Yes	No	No	No	No					No		No	Yes	
Splenomegaly	No	No	Yes	No	No	Yes				No	No		No	Yes	No
Other systems															
Cardiac anomalies	VSD	Nil	Fallots tetrology	Mild aortic insufficency	Nil	pulmonary stenosis	Nil	Nil	Nil	Nil	Nil	Aortic insufficiency	Nil	Mild pulmonary artery stenosis	Nil
Skeletal anomalies	Nil	Nil	Perthes disease	Nil	Nil	Nil	Nil	Nil	Nil	Nil	Nil	Nil	Nil	Nil	Nil
Renal anomalies	Nil	Nil	Small right kidney	Mild reflux	Nil	Nil	Nil	Nil	Nil	Nil	Nil	Nil	Nil	Nil	Bilateral inguinal hernia.
Development	Unknown	Unknown	Severe delay	Delayed	Nil	Mild delay	Normal	Delayed	Nil	Nil	Mild delay	Delayed	Mild delay	Neurological impairment	Unknown

3.3.3 Hair in PDI

3.3.3.1 Normal hair

Hair is unique to the mammal and is likely therefore to be a recent development in evolution. The skin is composed of two layers, the epidermis which is derived from the embryonic ectodermal sheets, and the dermis which is formed from the mesoderm. Hair follicles have components in both layers. The first follicles begin to appear at the end of the second gestational month and hair tends to develop cephalocaudally.

Hair cells are undifferentiated in the hair bulb but as they migrate in the direction of the skin surface they form the different specialised layers of the hair (Muller *et al*, 1991).

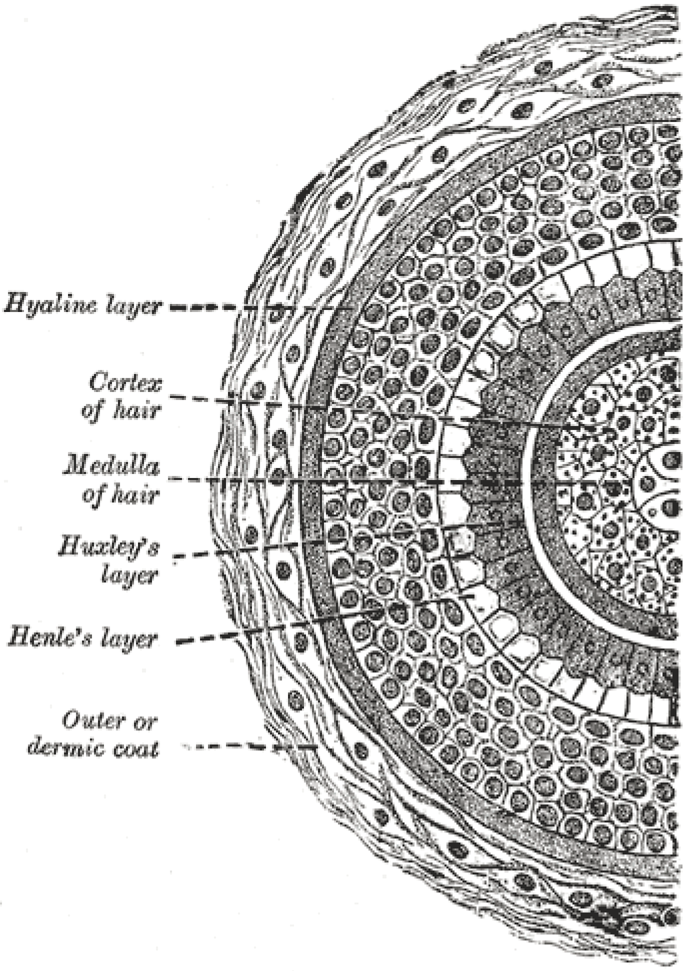
The inner root sheath cells contain an amorphous mass called trichohyalin granules (rich in arginine). During development these increase in numbers and filaments develop (rich in citrulline). The function of the sheath is unknown but may contribute to the formation of the hair shape. The inner root sheath is degraded at the level of the sebaceous duct probably due to proteolytic enzymes.

The outer layer of hair is made of flattened cuticle cells and is resistant to proteolytic digestion but can be solubilised after oxidation or reduction indicating a large proportion of cystine.

The cortex forms the largest proportion of the hair. Fibrous proteins aggregate to form fibrils and the cell elongates with the filaments in parallel. During this process of keratinisation there is a deposition of electron rich granules which disintegrates along with all of the cell organelles except for the nuclear membrane which is detectable in the centre of the keratinised cells (this is unlike keratinisation of the epidermis in which the nucleus is lost).

The medulla is the central core of hair. The amount of medulla varies and determines the diameter of the hair shaft. It is composed of random directional filaments and coalescent vesicles. Projections of cortical cells run between the medulla cells to separate them. The medulla may be a vestigial remnant as in humans it is none functioning whilst in other mammals it contains air to provide insulation (Messenger and Dawber, 1997). The structure of normal hair is illustrated in figure 3.3.1.

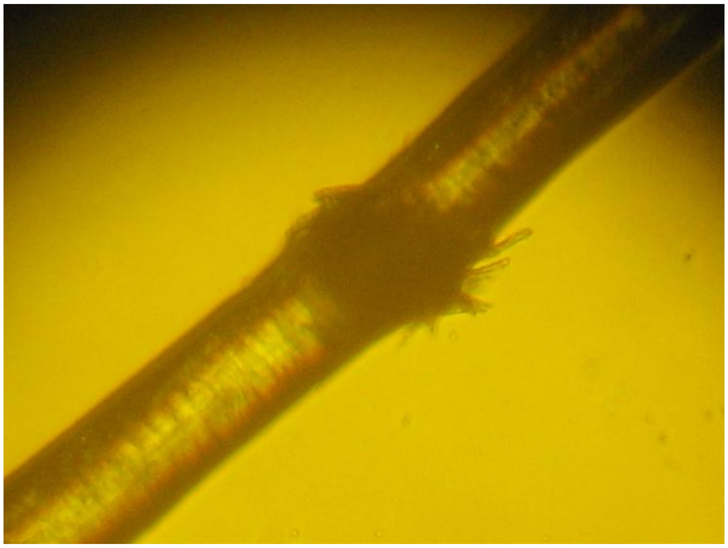
Figure 3.3.1 A schematic diagram showing a normal hair shaft.



3.3.3.2 Trichorrhexis nodosa (TN)

TN is a consistent finding in all probands with PDI. The head hair is sparse, woolly and coarse to touch. This may affect all head hairs or may be patchy with normal surrounding hair. Microscopy of the abnormal hair patches shows trichorrhexis nodosa, an abnormal break in the cuticle with cortical hair protruding as if two paint brushes are being pushed together which is illustrated in figure 3.3.2.

Figure 3.3.2 shows a hair shaft from patient 13C with typical findings of trichorrhexis nodosa which looks as if two paint brush ends are being pushed together as the cortex extrudes through the broken cuticle.



Although TN is a consistent finding in all cases of PDI it is not specific to this syndrome. An understanding of the other causes of TN may provide indications as to the underlying pathogenesis of PDI and a list of other conditions is provided in table 3.3.2

Table 3.3.2 Conditions in which trichorrhexis nodosa is a feature

	Clinical features	**Pathology**	**Inheritance pattern**	**Gene**
Trichothiodystropthy	Ichthyosis, intellectual impairment, decreased fertility and short stature	Nucleotide excision repair and transcription	AR	*XPD*
				XPB
				TTD-A
				TTDN1
Argininosuccinic aciduria	Hyperammonaemia resulting in encephalopathy and neurological sequelae	Urea cycle defect	AR	*ASL*
Conradi-Hunermann-Happle syndrome	Icthyosis, atrophoderma, splitting of nails and patchy alopecia, rhizomelic limb shortening, unilateral facial hypoplasia and scoliosis.	Abnormal sterol formation	X-linked dominant	*EBP*
Nethertons syndrome	Congenital ichthyosiform erythroderma, and atopic diathesis, hypogammaglobulinemia, failure to thrive, and enteropathy	Reduced serine protease inhibitor	AR	*SPINK5*
Menke's disease	Focal cerebral and cerebellar degeneration and seizures and kinky hair	Deficiency of copper transport	X-linked recessive	*ATP7A*
Bazex syndrome	Hypotrichosis, follicular atrophoderma and basal cell neoplasm development		X-linked dominant	
Laron syndrome	Short stature, delayed bone age, occasional blue sclera and hypoglycaemia	Abnormality in growth hormone receptor	AR	*IGF1*
Giant axonal degeneration	Polyneuropathy and kinky hair	Abnormal intermediate filament organisation	AR	*GAN*
Kabuki syndrome	Intellectual impairment, short stature, eversion of the lateral 1/3rd of the lower eye lid, long palpebral fissures, cleft or high arched palate, persistence of finger pads, short 5th finger, congenital heart defects, occasionally dysplastic kidneys, liver fibrosis and biliary atresia	.	AD	
Oculo-dento-digital dysplasia	Microphthalmia, abnormal small nose, hypotrichosis, dental anomalies, fifth finger camptodactyly, syndactyly of the fourth and fifth fingers (type III syndactyly), and missing toe phalanges	Misassembly of channels or altered channel conduction properties		*GJA1*
Carbohydrate deficient glycoprotein	Severe encephalopathy with axial hypotonia, abnormal eye movement, pronounced psychomotor retardation, peripheral neuropathy, cerebellar hypoplasia, retinitis pigmentosa, peculiar distribution of subcutaneous fat, nipple retraction, hypogonadism, severe infections, liver insufficiency, and cardiomyopathy.	Enzymatic defects in the synthesis and processing of asparagine (N)-linked glycans or oligosaccharides on glycoproteins	AR	*PMM2*

TN in normal hair: excessive weathering and use of cosmetics results in trauma to the hair and the appearance of TN. This is more pronounced in African hair. When trauma is the cause the TN it is sparse and mainly found on the extremities of the hair which is in contrast to the syndromic causes of TN when it is found throughout the hair shaft.

Trichothiodystrophy (TTD)

This is the term used to describe brittle hair with low sulphur content. On polarized light microscopy the hair shows a pattern of light and dark bands (known as tiger tail banding) as well as trichorrhexis nodosa, trichoschisis, ribboning and a reduced cysteine content. The associated anomalies from the ectodermal and neuroectodermal origins vary from fragile hair only to ichthyosis, intellectual impairment, decreased fertility and short stature (Price *et al*, 1980). Half are affected by photosensitivity but not a predisposition to cancer. The molecular abnormality is a defect in nucleotide excision repair and transcription with mutations identified in the genes *XPD, XPB and TTD-A* in photosensitive TTD and *TTDN1*in the non photosensitive form. Amish brittle hair-brain syndrome, Sabinas brittle hair syndrome and Pollitt syndrome (which includes mental retardation) are all forms of non photosensitive TTD.

Argininosuccinic aciduria

This urea cycle defect is due to a deficiency in the enzyme argininosuccinate lyase resulting in hyperammonaemia. All children with this condition have TN and no other hair anomalies have been described. *ASL* on chromosome 7 is the causative gene (Walker *et al*, 1990).

Conradi-Hunermann-Happle syndrome

This syndrome is also known as X-linked chondroplasia punctata type 2. Icthyosis, atrophoderma, splitting of nails and patchy alopecia are the clinical features. As well as the chondroplasia punctata the other skeletal findings are rhizomelic limb shortening, unilateral facial hypoplasia and scoliosis. The *EBP* gene at Xp11.22-11.23, encodes emopamil-binding protein, mutations in which result in deficiency of 3β-hydroxysterol Δ8, Δ7-isomerase activity and consequently abnormal sterol synthesis (Derry *et al*, 1999).

Nethertons syndrome

Both TN and trichorrhexis invaginata (bamboo hair) are seen in this condition of congenital ichthyosiform erythroderma, and atopic diathesis. Some have hypogammaglobulinemia, hypernatremic dehydration (due to transcutaneous fluid loss), failure to thrive, and enteropathy. Mutations in *SPINK5* result in a reduction of the serine protease inhibitor, LEKT1 (Chavanas *et al*, 2000).

Mitochondrial disorders

The defect of oxidative phosphorylation has a wide range of clinical presentations with the most common being neurological and muscular signs. Silengo *et al* (2003) described sparse, thin and fragile hair in 8 of 25 children, one of whom had chronic diarrhoea with recurrent infections and failure to thrive.

Menke’s disease

This X-linked recessive disorder of copper deficiency is due to mutations in the gene encoding Cu(2+)-transporting ATPase, alpha polypeptide, *ATP7A.* Clinical features are due to a lack of copper required for metabolic reactions. Clinical features are predominantly neurological with focal cerebral and cerebellar degeneration and seizures. The condition is also known as ‘kinky hair syndrome’ with TN as well as pill torti and monilethrix. The disease is progressive and death occurs in early childhood (Chelly *et al*, 1993; Mercer *et al*, 1993).

Biotin deficiency

This may result in the formation of TN which is reversible with biotin supplementation. In those with other causes of TN (Menkes, argininosuccinic aciduria and Pollit syndrome) biotin supplementation did not correct the hair changes.

Hypothyroidism

On microscopic examination the coarse wiry hair of hypothyroidism has TN. This is an acquired form of TN.

Other

TN has also been described in sporadic occasional cases of Bazex syndrome (Vabres *et al*, 1995), Laron syndrome (Berg *et al*, 1993), giant axonal degeneration (Bomont *et al*, 2000), Kabuki syndrome (Maas *et al*, 2007), Oculo-dento-digital dysplasia (Paznekas *et al*, 2003) and carbohydrate deficient glycoprotein (Imbach et al, 1999).

3.3.3.3 Amino acid composition of the hair from PDI patients

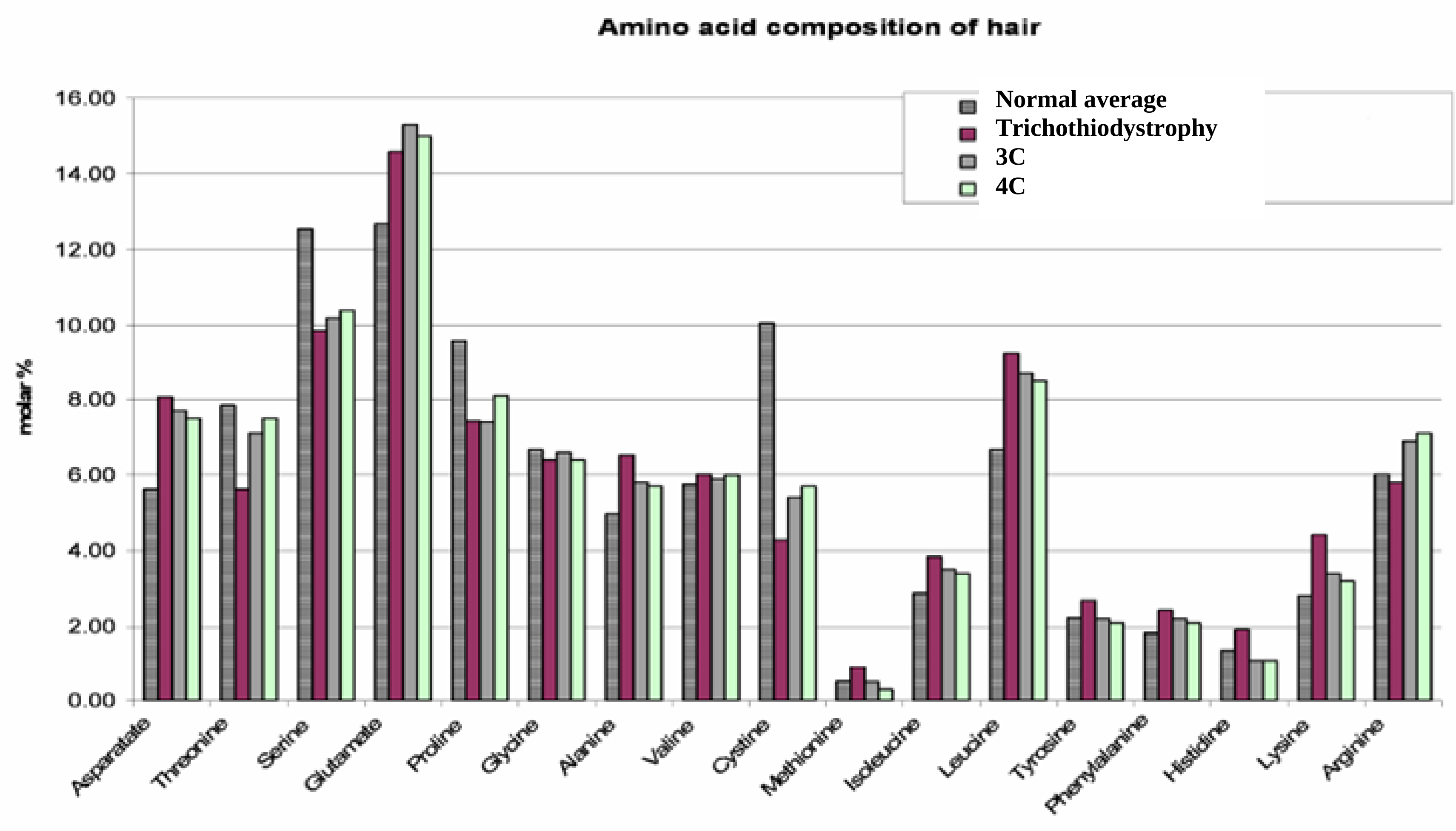
Due to the finding of TN in all patients with PDI I hypothesised there may be an abnormality of the hair which is either specific for PDI or the amino acid composition would provide an indication as to the underlying pathogenesis of PDI.

I contacted Dr R Pollitt, Sheffield Children's Hospital, who had previously carried out the hair analysis in the original PDI paper by Stankler *et al*, who carried out the analysis.

A specimen of hair was hydrolysed in a sealed tube with hydrochloric acid to denature the proteins and the amino acids then examined by electrophoresis prior to being measured quantatively on an amino acid analyzer.

Amino acid analysis of hair from two of the children showed reduced cystine. This is a constant finding in hair that is morphologically abnormal. The sulphur bonds between the cystine aids structural stability, so when reduced leads to the formation of trichorrhexis nodosa. Other sulphur rich amino acids are also reduced – serine and proline. Aspartate, glutamate and leucine are increased. These findings are comparable to those of trichothiodystrophy. Although trichorrhexis nodosa is seen in trichothiodystrophy the constant finding of tiger tail banding is not present in PDI patients. A bar chart of the results is shown in table 3.3.3.

Table 3.3.3 The amino acid composition of PDI hair. For each amino acid there a four bars shown. Normal hair composition is bar 1, trichothiodystrophy is shown in bar 2 and the amino acid composition of patients 3C and 4C are shown in bars 3 and 4.



3.3.4 Investigation of platelet structure and function

During initial investigation into the cause of the diarrhoea and immunodeficiency in siblings 1C and 1D the platelets were examined. The initial blood film showed enlarged platelets. This is not a common feature to all affected with PDI but has subsequently been identified in three other children, 3C, 4C and 5C at the time of central venous line infection. There is no clinical indication that there is either a bleeding diathesis or a hypercoagulable state.

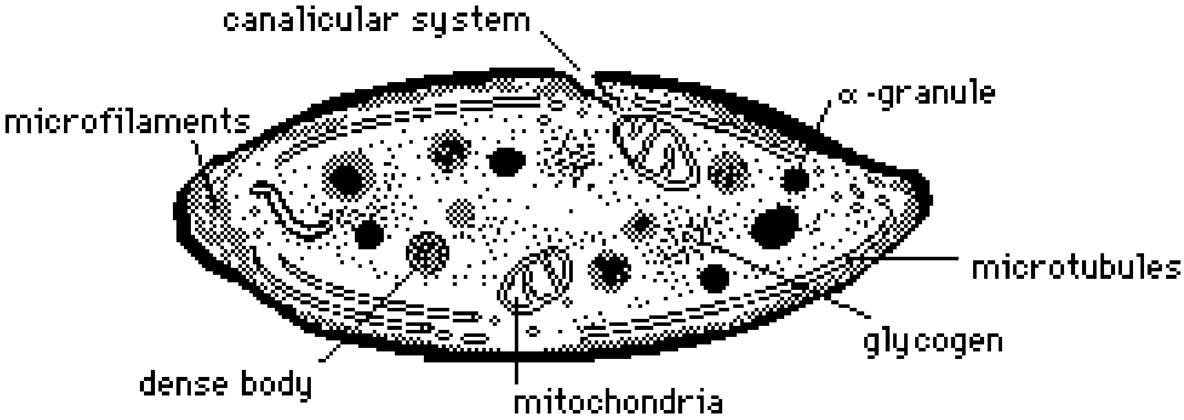


Figure 3.3.3 A schematic drawing of the normal morphology of a platelet.

Alpha granules contain platelet derived growth factor, platelet factor 4, factor V and XIII, and fibrinogen and vWF (Holmsen, 1994). Alpha granules within the platelet are shown in figure 3.3.3.

During platelets activation the alpha granule contents are centralized and then the contents are discharged into the open canalicular system for release into the exterior. ADP, epinephrine and TXA 2 stimulation can cause release of alpha granules as well as collagen (Ciferri *et al*, 2000).
P-selectin is an alpha granule specific protein to which there is a monoclonal antibody.

Storage pool disorders may be restricted to platelets and cause mild haemostasis defects or be part of a systemic syndrome of defective granule assembly and packaging.
Delta storage pool disorders appear to be autosomal recessive. They can also be an acquired disorder – SLE (causes premature release of granules induced by circulating immune complexes), myeloproliferative disorders, myelodysplasia and acute leukaemia.

Grey platelet syndrome is a possible inherited defect in the secretion of alpha granules. In this condition there are few alpha granules seen and in the place in the cytoplasm are vacuoles (Smith *et al*, 1997).

Actin and myosin are the major platelet proteins which for a 3D network throughout the cytoskeleton. Shorter actin fibers in 2D retain the discoid shape of the platelets. Surrounding the organelle zone is a membrane system which invaginates into the platelets membrane therefore known as the open canalicular system (OCS) and creates increased amount of membrane during activation (Berry *et al*, 1989).

3.3.4.1 Morphology of platelets from PDI patients

Five children (1C, 1D, 3C, 4C and 5C) of 10 patients evaluated were noted intermittently to have enlarged platelets on light microscopy of blood films. Platelets from patients 1C and 1D at the time of platelet enlargement and patient 4C when the platelets were of a normal size were examined by thin section transmission electron microscopy (TEM) (work carried out by Prof W Kahr, Children's Hospital, Toronto). Blood was available for the investigation of the structure of platelets by electron microscopy as previous described[17]*, in three patients (1C, 1D and 4C). 3.2% citrate anticoagulant blood was centrifuged (150g for 20 minutes) to obtain platelet rich plasma (PRP). PRP was fixed with 2.5% glutaraldehyde (Electron Microscopy Sciences, Hatfield, PA) in PBS pH 7.4 at* 4^oC *for 1 hour. Platelets were washed with 0.1M phosphate buffer (pH 7.4), followed by* dH_2O*. Platelets were then postfixed with 2% osmium tetroxide, dehydrated in graded acetones and embedded in Epon (Electron Microscopy Sciences). Thin sections were examined with JEOL JEM-1011 electron microscope with uranyl acetate and lead citrate staining (Electron Microscopy Sciences) Digital images were captured with a side mounted Advanced Microscopy Techniques (AMT) Advantage HR CCD cameras (Advanced Microscopy Techniques Corp., Danvers, MA).*

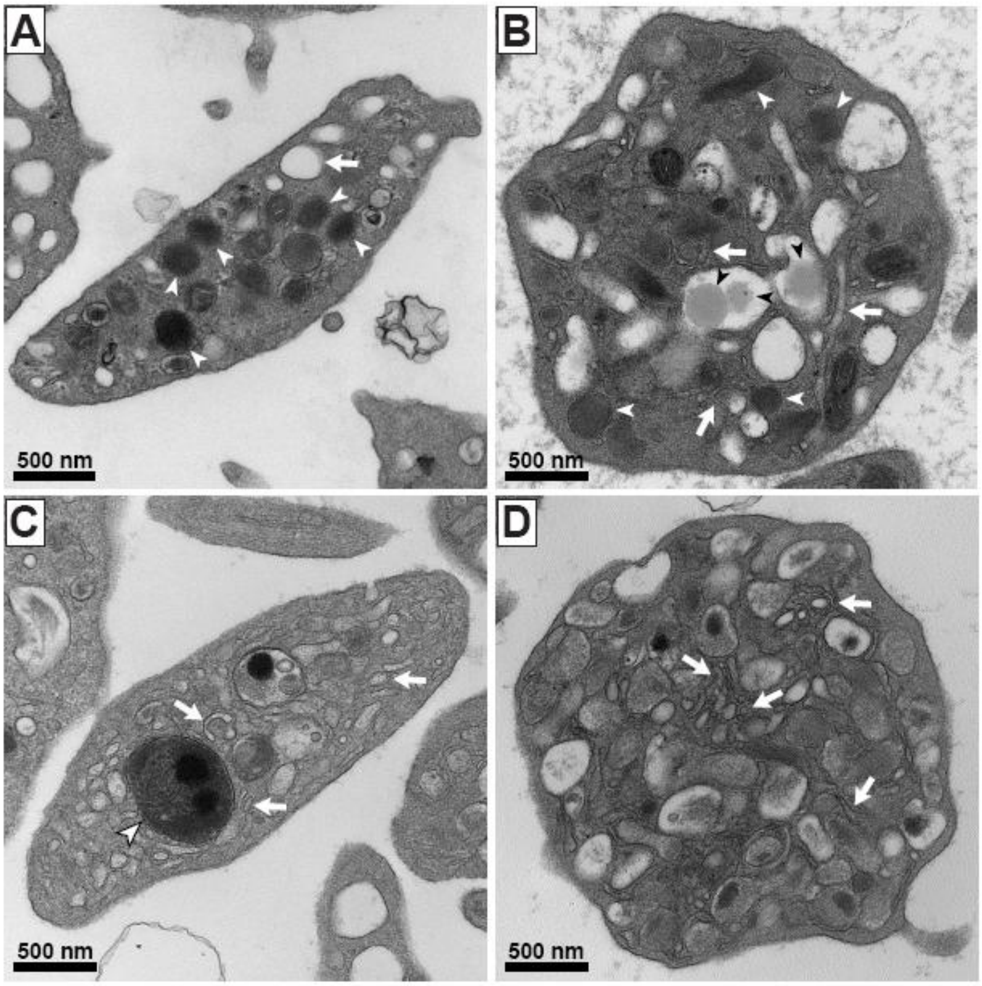


Figure 3.3.4 shows the transmission electron micrographs of a representative normal platelet (A) and of platelets from 3 different PDI patients (B-D). The black bar represents 500nm. Normal α-granules (white arrowheads) are observed in the control platelet (A) and occasionally in a PDI patient (B) but are frequently absent in PDI platelets (C, D). Whereas the membrane surface–connected canalicular system appears normal in control platelets (A, arrow) it was disrupted, with prominent tubules and small membranous vesicles, in PDI platelets (B-D, arrows). Lipid inclusions were frequently observed in PDI platelets (B, black arrowheads). Electron-dense lysosomal bodies fused with an α-granule were often seen (C, arrowhead).

3.3.4.2 Investigation of platelet alpha granules

This work was carried out by JH under the direction of Dr B Danwood and Prof S Watson. The protocol and methods had been developed by Dr B Danwood.

Platelet aggregation was measured in response to PAR1-specific peptide (SFLLRN, Alta Bioscience, Birmingham, UK), ADP (Sigma-Aldrich, Poole, UK) and collagen (Nycomed Austria, Linz, Austria) in two patients (4C and 5C). Secretion from dense granules was measured in a dual channel lumi-aggregometer (460VS, Chronolog) (Chrono-log Corporation, Havertown, PA) using a luciferase assay that detects released ATP (Chrono-log Corporation, Havertown, PA) [18]*. The level of expression of CD62P (P-selectin; α granule secretion indicator) was measured by flow cytometry using a specific antibody (Fluorescein isothiocyanate -conjugated anti mouse P-selectin antibody, Emfret Analytics, Wuorzburg, Germany) following stimulation by a collagen related peptide (CRP) (Dr Richard Farndale Cambridge University, UK).*

<u>Aggregation Studies results</u>

The platelets of patients 4c and 5c showed aggregation responses to high concentrations of a peptide specific to the PAR 1 thrombin receptor (100 μM), ADP (100 μM) and an intermediate collagen concentration (3 μg/ml) were similar to that of the control. The results are shown in figure 3.3.5.

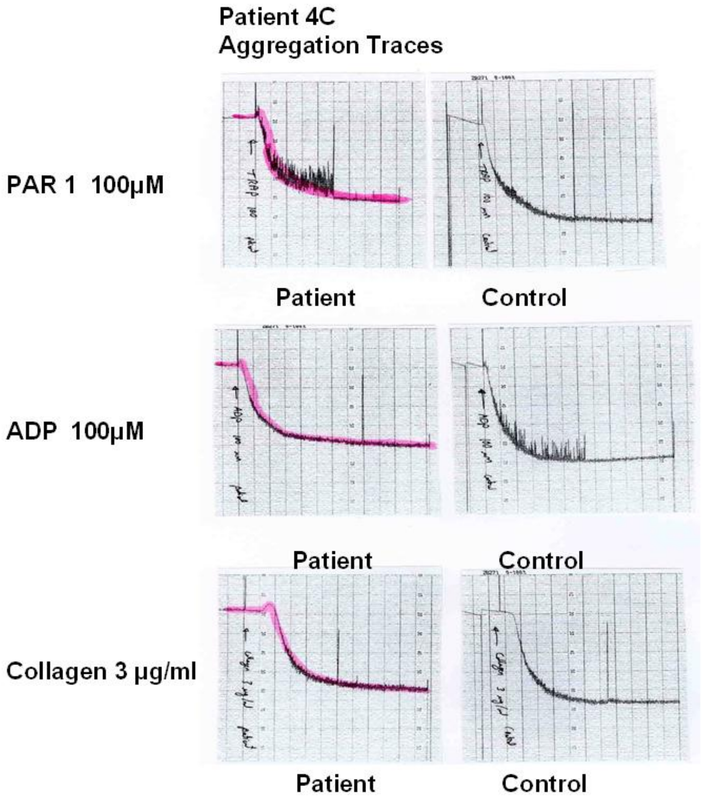
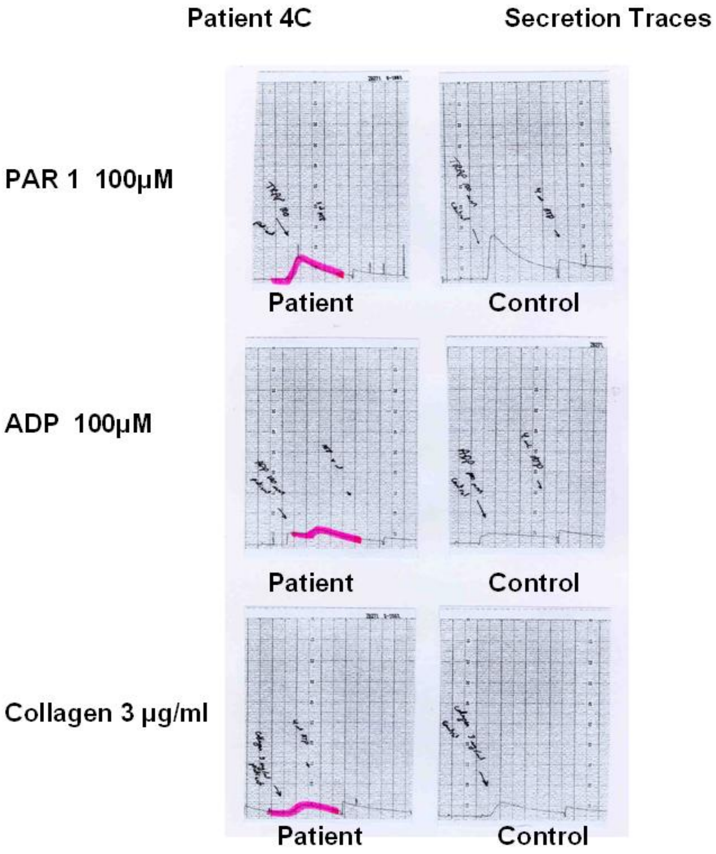


Figure 3.3.5 The aggregation of platelets in the control sample and patient 4C in response to PAR, ADP and collagen.

<ins>ATP secretion from dense granules results</ins>

Secretion from dense granules was measured in the lumi-aggregometer using a luciferase assay (ATP). No secretion defect from dense granules was detected in response to any of the three agonists as measured by release of ATP. The greater secretion seen in ATP secretion for the TRAP samples can be explained by the higher platelet count – all results to these high concentrations of agonists lie within the normal range. This is shown in figure 3.3.6.



3.3.6. shows the normal secretion from dense granules from the control and patient 4C

α-granule secretion (analysed by FACS)

The level of expression of CD62P during stimulation showed 2 populations of platelets, of which one expressed P-selectin and the other did not. The level of expression in a non-aged matched control measured on the same day was within the normal range. This is shown in figure 3.3.7.

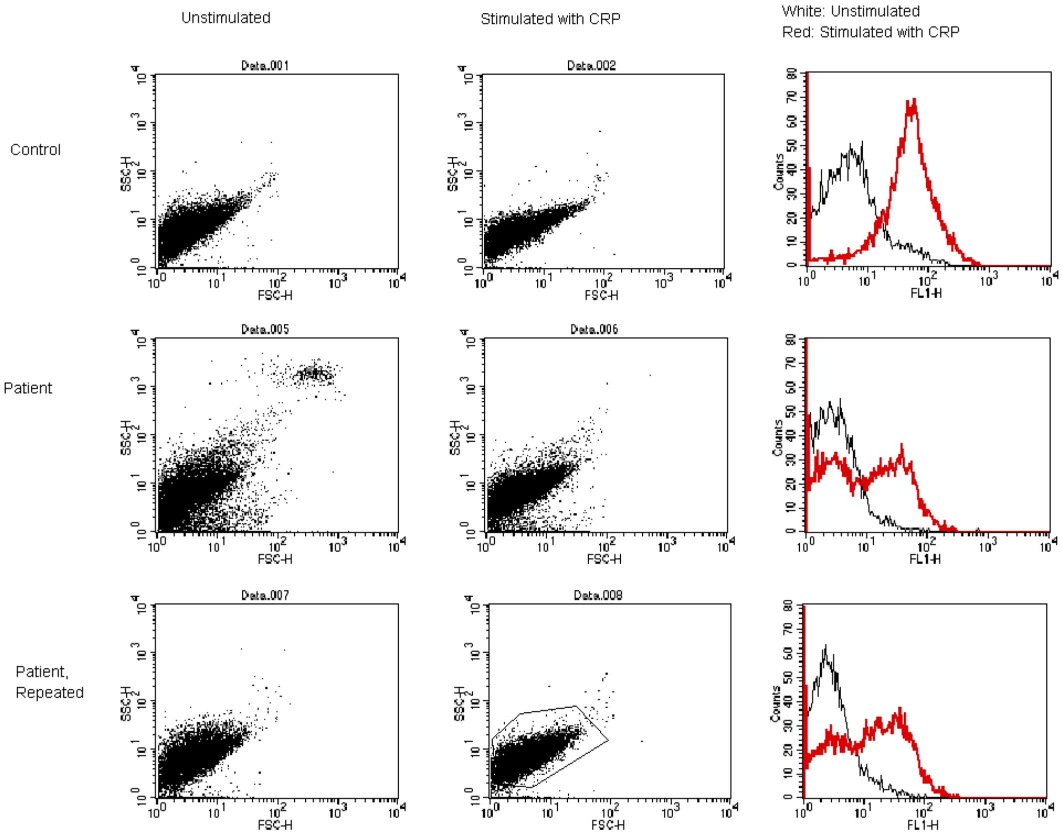


Figure 3.3.7 The abnormal release of CD62P from the alpha granules of patient 4C as compared to control. The patient showed two populations of platelets one which expresses P selectin and the other did not.

The ultrastructural analysis suggests an abnormality in platelet α-granule formation and secretion of α-granule contents. An abnormal α-granule secretion pattern was also observed with agonist-stimulated PDI platelets whereby only a subpopulation of platelets expressed P-selectin on their surface. It is likely that only those platelets containing some α-granules released their contents whereas platelets without α-granules could not. Clinically children with PDI do not have a bleeding diathesis which can be explained by the identification of two platelet populations, one of which functions normally.

3.3.5 Summary

This chapter describes 14 families of children with PDI. Of these families 10 are of consanguineous union. The clinical features of the cohort are homogeneous. Two families have additional phenotypes – 8C has maternal albinism and although hypopigmentation is seen in the other patient's albinism has not previously been seen. 13C has motor neurological features. These features have been attributed to hypoxia at delivery secondary to maternal severe acute asthma. Both of these families have been considered to have clinical PDI.

Platelet α-granule deficiencies are very rare and have only previously been described in Grey platelet syndrome (GPS) in which the molecular basis has not been elucidated (Nurden and Nurden, 2007), Quebec platelet disorder when there is abnormal break down of α-granule proteins (Haywood *et al*, 1997) and as a feature of Arthrogryposis-Renal dysfunction-Cholestasis (ARC) syndrome (MIM #208085) (Gissen et al, 2004). ARC syndrome is caused by mutations in *VPS33B* which is associated with abnormal vesicular trafficking and mislocalisation of polarised membrane proteins. As well as α-granule abnormalities, children with ARC syndrome also have diarrhoea and liver disease which overlaps with clinical features of PDI. The hypothesised role of TPR proteins in protein-protein interactions and the shared characteristics with ARC syndrome raises the possibility of an abnormal protein localisation in PDI.

3.4 The identification of a genetic locus for PDI using autozygosity mapping

3.4.1 Chapter overview

Autozygosity mapping utilising the genetic information gained from those families of consanguineous union was used to ascertain a genetic region for investigation.

With some PDI clinical features and the α-granule abnormality similar to ARC syndrome, linkage to the ARC syndrome locus was initially sought and refuted.

A 10K SNP GWS was initially of two of the consanguineous probands identified potential regions of interest. To increase accuracy the original probands and additional affected patients then underwent a 250K SNP GWS. Regions of shared homozygosity were ranked according to length.

250K SNP GWS from other conditions were compared to the PDI GWS to identify regions of the genome which are homozygous in all and therefore unlikely to be solely pertinent to PDI.

Microsatellite markers for the largest regions of shared homozygosity were used to ascertain or refute linkage. The largest region on chromosome 19 was excluded whilst microsatellites for a region on chromosome 5 identified a critical region in all affected children from D5S1462 to D5S433 (96,406,286 – 103,990,534). With further analysis limiting it to those probands of Pakistani origin, extended the proximal region to D5S1725 (89,202,358).

3.4.2 Linkage to ARC syndrome locus

The investigation of clinical features in children with PDI showed some similarity to ARC syndrome. This led to the investigation of linkage to the ARC locus using microsatellite markers from D15S996 to D15S963. Table 3.4.1 shows the microsatellite marker results in patients 1C, 1D, 3C, 4C and 5C. In siblings 1C and 1D the microsatellites were of different sizes suggesting that the inherited alleles are different and therefore they are unlikely to have inherited identical copies of genes in this region (1C and 1D are reported to be non-consanguineous but the parents come from the same small village in India suggesting that they may be distantly related). None of the other probands had homozygous alleles throughout the region and in those alleles which were homozygous the marker was not fully informative.

Linkage to ARC syndrome was therefore not detected and no further investigation of this region was carried out.

		1A parent		1B parent		**1C affected**		**1D affected**		1E sibling		3A parent		3B parent		**3C affected**		4A parent		4B parent		**4C affected**		4D sibling		5A parent		5B parent		**5C affected**	
D15S996		196	201	194	196	196	196	194	201	188	201			197	197	197	201	192	199			192	194	192	195	192	201	192	196	192	192
ARC 2		203	207	201	201	203	201	201	207	201	207	201	201	197	197	197	201					197	203			201	197	201	201	201	201
D15s127		142	123	123	123	142	123	123	123	123	123	131	131	131	131	131	131	130	130			130	140	130	130	124	144	124	124	124	124
VPS33B	89342907 to 89366817																														
ARC 4		311	330	328	330	311	330	328	330	311	330	313	313	324	328	313	328	313	328			328	328	324	328	328	330	311	311	311	330
D15S158		98	98	98	100	98	100			98	98	98	98	98	98	98	98	82	86			82	86	82	86	98	98	98	98	98	98
D15S963		275	269	275	269	275	269	275	269	275	269	270	270	270	270	270	270									269	270	269	277	269	269

Table 3.4.1 A table showing the results of microsatellite markers at the ARC syndrome locus in 5 consanguineous probands with PDI. *VPS33B* is the gene for ARC syndrome and is shown in red. Those microsatellites which are homozygous in the affected children are shaded green.

3.4.3 Single nucleotide polymorphism (SNP) genomewide scans (GWS)

An initial 10K SNP GWS of two probands 3C and 5C both of whom are from Pakistan, identified two regions of extensive homozygosity in chromosome 19 and chromosome 5.

All the SNP GWS (10K and 250K) were carried out by L. Tee, research technician.
The principle of the technique is

- E*ach DNA sample is quantitated using a dual beam spectrophotometer and a negligible amount is used to perform QC check*
- *After adaptor ligation a PCR reaction is set up to amplify 250 – 2000bp fragments. The resulting amplified DNA is then fragmented, denatured, and tagged using a terminal deoxynucleotidyl transferase end labelling reaction.*
- *Use a DNA chip with 10 or 250K different DNA sequences immobilised at different positions on the surface*
- *Patient DNA is hybridised to the chip*
- *A signal corresponding to the specific base is detected i.e. different signals for either base*

I interpreted the results of all the scans.

10K SNP GWS

The interpretation of the results:-

In chromosome 19 the region spanned 23MB (16,317,260 - 39,351,986) and there were two haploidentical regions within this:-

20,152,263 – 32,869,643	12.7MB
33,176,984 – 34,722,418	1.8MB

The region on chromosome 5 was 19.2MB (86,830,251 – 106,117,493) with two haploidentical regions:-

91,111,908 – 94,499,325	3.3MB
95,851,505 – 98,085,824	2.2MB

250K SNP GWS

With further identification of PDI cases and an increased number of SNP’s available on a GWS, the DNA of all consanguineous probands (including 3C and 5C) underwent a 250K SNP GWS.

Chromosome 19

In chromosome 19 the region reduced to 9.5MB (23,551,431 – 33,078,226) with proband 4C being mostly heterozygous in the region whilst the other consanguineous probands had homozygous alleles.

Chromosome 5

In chromosome 5 the region changed to 15.1MB (87,858,428 – 102,981,136). In this region all those of Pakistani origin except 8C and cousin of 8C were completely homozygous. Proband 9C, of Italian origin was also extensively homozygous. Probands 1C, 1D and 4C all had regions within this of homozygosity whilst 8C had smaller regions of homozygosity which did not correspond with the cousin. This may be due to either both 8C and the cousin having a different locus for PDI or just the cousin not having the same clinical phenotype (phenotyping in this patient was limited).

Forming this region were overlapping regions of extensive homozygosity from individuals which is shown in table 3.4.2

Proband	**Ethnicity**	**Length of homozygosity MB**
7C	Pakistani	44
6C	Pakistani	34.4
5C	Pakistani	24.9
14C	Pakistani	21.5
9C	Italian	18.8
3C	Pakistani	15.1
4C	Kurdish	13.4

Table 3.4.2 shows the extent of allele homozygosity in individual consanguineous probands, which overlap so forming a common region of shared homozygosity in chromosome 5.

The haplotype of the alleles was not identical which suggests there is no common ancestor for these alleles even in those of the same ethnicity.

Chromosome 16

A new region was identified from the 250K SNP GWS on chromosome 16 which spanned 15.1MB (31,629,915 – 46,785,363) and the allele haplotypes were identical in all probands. This had not been identified on the 10K SNP GWS due to a paucity of SNP’s in this region.

3.4.4 Microsatellite Markers

Microsatellite markers chromosome 19

Microsatellite markers D19S714, D19S898, D19S560, D19S568, D19S433, D19S414, D19S245 showed homozygosity in 3C and 5C so confirming the original 10K SNP GWS but the other affected consanguineous probands had only occasional allele homozygosity for these markers. The region was excluded by the identification of an identical allele haplotype in an unaffected sibling in family 4C in two fully informative markers D19s433 and D19S245. The results are shown in table 3.4.3.

	1A parent		1B parent		1C PDI		1D PDI		1E sib		3A parent		3B parent		3C PDI		3D sib	
D19S714		251	251	249	249	249	249	249	251	251	246	250	238	250	238	250		
D19s898	171	179	150	179	179	179	179	179	171	175	175	178	150	179	175	179		
D19s560	170	202	192	202	170	202	170	202	170	202	182	193	202	196	182	196		
D19S568	248	248	248	248	248	248	248	248	248	248	248	240	248	248	248	248		
D19S433	199	199	207	209	199	207	199	207	199	209	205	209			205	205	205	209
D19S414	168	187	183	187	168	187	168	187	187	187	187	188	183	187	187	187		
D19S245	201	201	193	201	193	201	193	201	201	201	201	201			201	201		

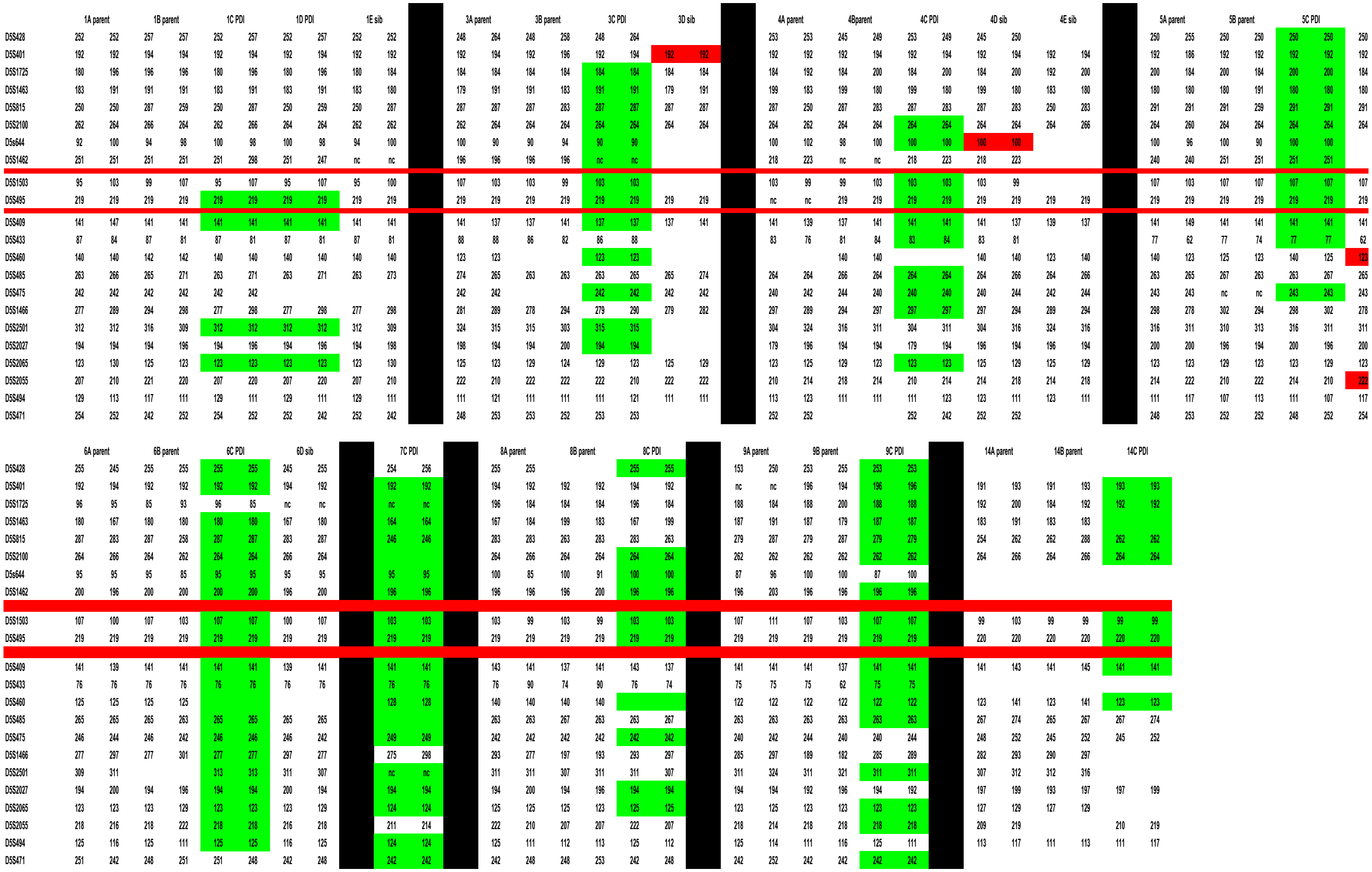
	4A parent		4B parent		4C PDI		4D sib		4E sib		5A parent		5B parent		5C PDI		5D sib		5E sib	
D19S714	246	254			257	257	258	258			250	250	253	253	253	253				
D19s898	150	178	178	178	175	178	150	178			176	176	176	176	176	176				
D19s560	171	171	171	200	?181	?188	171	171			202	202	172	202	202	202				
D19S568	257	268	245	245	257	257	257	257			250	269	248	248	248	249				
D19S433	199	207	195	199	199	199	199	207	199	199	201	211	201	213	199	199	199	207	199	203
D19S414	187	187	168	187	168	187	169	187	168	187	189	189	167	189	189	189				
D19S245	197	201	197	201	197	197	197	201	197	197					205	205	201	205	193	205

	6A parent		6B parent		6C PDI		6D sib		7C PDI		8A parent		8B parent		8C PDI		9A parent		9B parent		9C PDI	
D19S714																						
D19s898																						
D19s560																						
D19S568																						
D19S433	211	213	199	203	203	213	199	213	199	199	199	203	196	200	195	203	199	207	199	203	199	199
D19S414	168	168	168	189	168	189	168	168	183	187	171	183					187	187	187	168	187	187
D19S245	201	201	193	201	201	201	193	201	197	201	201	201					201	197	201	197	201	201

Table 3.4.3 shows the microsatellite marker allele sizes for the region on chromosome 19 in all consanguineous PDI families. Highlighted green are the alleles in affected probands which are homozygous. The boxes shaded red are in an unaffected sibling in family 4 who has inherited identical alleles to the affected proband 4C in markers which are fully informative. Sib=sibling

Microsatellite markers chromosome 5

Microsatellite markers for the region on chromosome 5 confirmed linkage in all those from consanguineous union. The proximal limit was determined by 4C with heterozygous alleles at D5S1462. The distal limit was determined by 8C with heterozygous alleles at D5S409. This resulted in a critical region between 96,406,286 and 102,757,533 a region of 6.3MB. The results are shown in table 3.4.4.



	1A parent		1B parent		1C PDI		1D PDI		1E sib		3A parent		3B parent		3C PDI		3D sib		4A parent		4Bparent		4C PDI		4D sib		4E sib		5A parent		5B parent		5C PDI		
D5S428	252	252	257	257	252	257	252	257	252	252	248	264	248	258	248	264			253	253	245	249	253	249	245	250			250	255	250	250	250	250	250
D5S401	192	192	194	194	192	194	192	194	192	192	192	194	192	196	192	194	192	192	192	192	192	194	192	194	192	194	192	194	192	186	192	192	192	192	192
D5S1725	180	196	196	196	180	196	180	196	180	184	184	184	184	184	184	184	184	184	184	192	184	200	184	200	184	200	192	200	200	184	200	184	200	200	184
D5S1463	183	191	191	191	183	191	183	191	183	180	179	191	191	183	191	191	179	191	199	183	199	180	199	180	199	180	183	180	180	180	180	191	180	180	180
D5S815	250	250	287	259	250	287	250	259	250	287	287	287	287	283	287	287	287	287	287	250	287	283	287	283	287	283	250	283	291	291	291	259	291	291	291
D5S2100	262	264	266	264	262	266	264	264	262	262	262	264	264	264	264	264	264	264	264	262	264	264	264	264	264	264	264	266	264	260	264	264	264	264	264
D5s644	92	100	94	98	100	98	100	98	94	100	100	90	90	94	90	90			100	102	98	100	100	100	100	100			100	96	100	90	100	100	
D5S1462	251	251	251	251	251	298	251	247	nc	nc	196	196	196	196	nc	nc			218	223	nc	nc	218	223	218	223			240	240	251	251	251	251	
D5S1503	95	103	99	107	95	107	95	107	95	100	107	103	103	99	103	103			103	99	99	103	103	103	103	99			107	103	107	107	107	107	107
D5S495	219	219	219	219	219	219	219	219	219	219	219	219	219	219	219	219	219	219	nc	nc	219	219	219	219	219	219	219	219	219	219	219	219	219	219	219
D5S409	141	147	141	141	141	141	141	141	141	141	141	137	137	141	137	137	137	141	141	139	137	141	141	141	141	137	139	137	141	149	141	141	141	141	141
D5S433	87	84	87	81	87	81	87	81	87	81	88	88	86	82	86	88			83	76	81	84	83	84	83	81			77	62	77	74	77	77	62
D5S460	140	140	142	142	140	140	140	140	140	140	123	123			123	123					140	140			140	140	123	140	140	123	125	123	140	125	123
D5S485	263	266	265	271	263	271	263	271	263	273	274	265	263	263	263	265	265	274	264	264	266	264	264	264	264	266	264	266	263	265	267	263	263	267	265
D5S475	242	242	242	242	242	242					242	242			242	242	242	242	240	242	244	240	240	240	240	244	242	244	243	243	nc	nc	243	243	243
D5S1466	277	289	294	298	277	298	277	298	277	298	281	289	278	294	279	290	279	282	297	289	294	297	297	297	297	294	289	294	298	278	302	294	298	302	278
D5S2501	312	312	316	309	312	312	312	312	312	309	324	315	315	303	315	315			304	324	316	311	304	311	304	316	324	316	316	311	310	313	316	311	311
D5S2027	194	194	194	196	194	196	194	196	194	198	198	194	194	200	194	194			179	196	194	194	179	194	196	194	196	194	200	200	196	194	200	196	200
D5S2065	123	130	125	123	123	123	123	123	123	130	125	123	129	124	129	123	125	129	123	125	129	123	123	123	125	129	125	129	123	123	129	123	123	129	123
D5S2055	207	210	221	220	207	220	207	220	207	210	222	210	222	222	222	210	222	222	210	214	218	214	210	214	214	218	214	218	214	222	210	222	214	210	222
D5S494	129	113	117	111	129	111	129	111	129	111	111	121	111	111	111	121	111	111	113	123	111	111	111	123	123	111	123	111	111	117	107	113	111	107	117
D5S471	254	252	242	252	254	252	252	242	252	242	248	253	253	252	253	253			252	252			252	242	252	252			248	253	252	252	248	252	254

	6A parent		6B parent		6C PDI		6D sib		7C PDI		8A parent		8B parent		8C PDI		9A parent		9B parent		9C PDI		14A parent		14B parent		14C PDI	
D5S428	255	245	255	255	255	255	245	255	254	256	255	255			255	255	153	250	253	255	253	253						
D5S401	192	194	192	192	192	192	194	192	192	192	194	192	192	192	194	192	nc	nc	196	194	196	196	191	193	191	193	193	193
D5S1725	96	95	85	93	96	85	nc	nc	nc	nc	196	184	184	184	196	184	188	184	188	200	188	188	192	200	184	192	192	192
D5S1463	180	167	180	180	180	180	167	180	164	164	167	184	199	183	167	199	187	191	187	179	187	187	183	191	183	183		
D5S815	287	283	287	258	287	287	283	287	246	246	283	283	263	283	283	263	279	287	279	287	279	279	254	262	262	288	262	262
D5S2100	264	266	264	262	264	264	266	264			264	266	264	264	264	264	262	262	262	262	262	262	264	266	264	266	264	264
D5s644	95	95	95	85	95	95	95	95	95	95	100	85	100	91	100	100	87	96	100	100	87	100						
D5S1462	200	196	200	200	200	200	196	200	196	196	196	196	196	200	196	196	196	203	196	196	196	196						
D5S1503	107	100	107	103	107	107	100	107	103	103	103	99	103	99	103	103	107	111	107	103	107	107	99	103	99	99	99	99
D5S495	219	219	219	219	219	219	219	219	219	219	219	219	219	219	219	219	219	219	219	219	219	219	220	220	220	220	220	220
D5S409	141	139	141	141	141	141	139	141	141	141	143	141	137	141	143	137	141	141	141	137	141	141	141	143	141	145	141	141
D5S433	76	76	76	76	76	76	76	76	76	76	76	90	74	90	76	74	75	75	75	62	75	75						
D5S460	125	125	125	125					128	128	140	140	140	140			122	122	122	122	122	122	123	141	123	141	123	123
D5S485	265	265	265	263	265	265	265	265			263	263	267	263	263	267	263	263	263	263	263	263	267	274	265	267	267	274
D5S475	246	244	246	242	246	246	246	242	249	249	242	242	242	242	242	242	240	242	244	240	240	244	248	252	245	252	245	252
D5S1466	277	297	277	301	277	277	297	277	275	298	293	277	197	193	293	297	285	297	189	182	285	289	282	293	290	297		
D5S2501	309	311			313	313	311	307	nc	nc	311	311	307	311	311	307	311	324	311	321	311	311	307	312	312	316		
D5S2027	194	200	194	196	194	194	200	194	194	194	194	200	194	196	194	194	194	194	192	196	194	192	197	199	193	197	197	199
D5S2065	123	123	123	129	123	123	123	129	124	124	125	125	125	123	125	125	123	125	123	123	123	123	127	129	127	129		
D5S2055	218	216	218	222	218	218	216	218	211	214	222	210	207	207	222	207	218	214	218	218	218	218	209	219			210	219
D5S494	125	116	125	111	125	125	116	125	124	124	125	111	112	113	125	112	125	114	111	116	125	111	113	117	111	113	111	117
D5S471	251	242	248	251	251	248	242	248	242	242	242	248	248	253	242	248	242	252	242	242	242	242						

Table 3.4.4 shows the microsatellite marker allele size in all consanguineous families for the chromosome 5 region. The alleles which are homozygous in affected probands are shaded green. The unaffected siblings who are homozygous for informative markers are shaded red. The two horizontal lines show the limits of the critical region for investigation. The region is limited proximally by 4C and distally by 8C. Sib=sibling

The 250K SNP GWS of the chromosome 5 region showed extensive homozygosity in those of Pakistani origin and smaller areas in those of other ethnicities. This led the investigation of the microsatellite markers in only those from Pakistan. This increased the critical region to lie between D5S815 to D5S409, with both the proximal and distal limits set by 8C. This is shown in table 3.4.5.

The region of interest on chromosome 5 when limited to those of Pakistani origin lay between 91,026,202 and 102,757,753 which is 11.7MB.

	3A parent		3B parent		**3C PDI**		3D sib		5A parent		5B parent		5C PDI		5D sib		5E sib		5F sib		6A parent		6B parent		6C PDI		6D sib		7C PDI		8A parent		8B parent		8C PDI		14A parent		14B parent		14C PDI	
D5S428	248	264	248	258	248	264			250	255	250	250	**250**	**250**	**250**	**250**	**250**	**250**	**250**	**250**	255	245	255	255	255	255	245	255	254	256	255	255			**255**	**255**						
D5S401	192	194	192	196	192	194	**192**	**192**	192	186	192	192	**192**	**192**	**192**	**192**	186	192	186	192	192	194	192	192	192	192	194	192	**192**	**192**	194	192	192	192	194	192	191	193	191	193	**193**	**193**
D5S1725	184	184	184	184	**184**	**184**	**184**	**184**	200	184	200	184	**200**	**200**	184	200	184	200	184	200	96	95	85	93	96	85	nc	nc	**nc**	**nc**	196	184	184	184	196	184	192	200	184	192	**192**	**192**
D5S1463	179	191	191	183	**191**	**191**	179	191	180	180	180	191	**180**	**180**	180	191	180	180	**180**	**180**	180	167	180	180	**180**	**180**	167	180	**164**	**164**	167	184	199	183	167	199	183	191	183	183		
D5S815	287	287	287	283	**287**	**287**	**287**	**287**	291	291	291	259	**291**	**291**	291	259	**291**	**291**	**291**	**291**	287	283	287	258	**287**	**287**	283	287	**246**	**246**	283	283	263	283	283	263	254	262	262	288	**262**	**262**
D5S2100	262	264	264	264	**264**	**264**	**264**	**264**	264	260	264	264	**264**	**264**	**264**	**264**	260	264	260	264	264	266	264	262	**264**	**264**	266	264			264	266	264	264	**264**	**264**	264	266	264	266	**264**	**264**
D5s644	100	90	90	94	**90**	**90**			100	96	100	90	**100**	**100**							95	95	95	85	**95**	**95**	**95**	**95**	**95**	**95**	100	85	100	91	**100**	**100**						
D5S1462	196	196	196	196	**nc**	**nc**			240	240	251	251	**251**	**251**							200	196	200	200	**200**	**200**	196	200	**196**	**196**	196	196	196	200	**196**	**196**						
D5S1503	107	103	103	99	**103**	**103**			107	103	107	107	**107**	**107**	**107**	**107**	103	107	103	107	107	100	107	103	**107**	**107**	100	107	**103**	**103**	103	99	103	99	**103**	**103**	99	103	99	99	**99**	**99**
D5S495	219	219	219	219	**219**	**219**	**219**	**219**	219	219	219	219	**219**	**219**	**219**	**219**	**219**	**219**	**219**	**219**	219	219	219	219	**219**	**219**	**219**	**219**	**219**	**219**	219	219	219	219	**219**	**219**	220	220	220	220	**220**	**220**
D5S409	141	137	137	141	**137**	**137**	137	141	141	149	141	141	**141**	**141**	141	149	149	141	149	141	141	139	141	141	**141**	**141**	139	141	**141**	**141**	143	141	137	141	143	137	141	143	141	145	**141**	**141**
D5S433	88	88	86	82	86	88			77	62	77	74	**77**	**77**	62	74	62	74	62	77	76	76	76	76	**76**	**76**	**76**	**76**	**76**	**76**	76	90	74	90	76	74						
D5S460	123	123			123	123			140	123	125	123	140	125	**123**	**123**			**123**	**123**	125	125	125	125					**128**	**128**	140	140	140	140			123	141	123	141	**123**	**123**
D5S485	274	265	263	263	263	265	265	274	263	265	267	263	263	267	265	263	263	267	263	267	265	265	265	263	**265**	**265**	**265**	**265**			263	263	267	263	263	267	267	274	265	267	267	274
D5S475	242	242			**242**	**242**	**242**	**242**	243	243	nc	nc	**243**	**243**	**243**	**243**	**243**	**243**	**243**	**243**	246	244	246	242	**246**	**246**	246	242	**249**	**249**	242	242	242	242	**242**	**242**	248	252	245	252	245	252
D5S1466	281	289	278	294	279	290	279	282	298	278	302	294	298	302	278	294	278	294	278	302	277	297	277	301	**277**	**277**	297	277	275	298	293	277	197	193	293	297	282	293	290	297		
D5S2501	324	315	315	303	**315**	**315**			316	311	310	313	316	311	311	316	311	316	311	316	309	311			**313**	**313**	311	307	nc	nc	311	311	307	311	311	307	307	312	312	316		
D5S2027	198	194	194	200	**194**	**194**			200	200	196	194	200	196	200	194	200	194	200	196	194	200	194	196	**194**	**194**	200	194	**194**	**194**	194	200	194	196	**194**	**194**	197	199	193	197	197	199
D5S2065	125	123	129	124	129	123	125	129	123	123	129	123	123	129	123	129	123	125	123	129	123	123	123	129	**123**	**123**	123	129	**124**	**124**	125	125	125	123	**125**	**125**	127	129	127	129		
D5S2055	222	210	222	222	222	210	222	222	214	222	210	222	214	210	**222**	**222**	**222**	**222**	222	210	218	216	218	222	**218**	**218**	216	218	211	214	222	210	207	207	222	207	209	219			210	219
D5S494	111	121	111	111	111	121	111	111	111	117	107	113	111	107	117	113	117	113	117	107	125	116	125	111	**125**	**125**	116	125	**124**	**124**	125	111	112	113	125	112	113	117	111	113	111	117
D5S471	248	253	253	252	253	253			248	253	252	252	248	252	254	251	254	251	254	252	251	242	248	251	251	248	242	248	**242**	**242**	242	248	248	253	242	248						

Table 3.4.5 is a table of the microsatellite marker sizes of chromosome 5 in those families who originate from Pakistan. Markers which are homozygous in affected probands are shaded in green. Siblings who are homozygous in informative markers are shaded red. The proximal limit of the region of interest is a horizontal blue line and the distal limit is shown by a horizontal red line. Sib=sibling

LOD scores for chromosome 5 region

A LOD score using Superlink was calculated using pedigrees of Pakistani origin. The following criteria were used:-

Mode of inheritance (MOI) = Recessive 0.99

Number of markers 5

Distances between markers = 3.74,0.53,2.78,0

Disease mutant gene frequency = 0.001

Marker Names = D5S2100, D5S644, D5S1462, D5S1503, D5S495

Points to calculate LOD score = -n 2 1 5 -o 10.0000

	Trait position from marker #1 (in cM)	Ln(Likelihood)	LOD-SCORE
	-10.0000	**-158.3699**	**4.5652**
	-5.0000	**-157.0412**	**5.1422**
D5S2100	**0.0000**	**-155.5882**	**5.7733**
	1.8700	**-156.1735**	**5.5191**
D5S644	**3.7399**	**-159.9254**	**3.8897**
	4.0049	**-156.0066**	**5.5915**
D5S1462	**4.2699**	**-155.3254**	**5.8874**
	5.6599	**-155.3702**	**5.8679**
D5S1462 & D5S495	**7.0499**	**-155.3939**	**5.8577**
	12.0499	**-156.8532**	**5.2239**
	17.0499	**-158.1898**	**4.6434**

Table 3.4.6 shows the LOD scores for markers and between markers for the cohort originating from Pakistani.

The maximum LOD score was at D5S1462 with similar scores at D5S2100, D5S1462 and D5S495. The LOD scores were reassuring as to the linkage to this region. The scores were high throughout the region and therefore did not give an indication of which locus to focus on.

Microsatellite markers chromosome 16

Chromosome 16 had a region of 15.1MB from 31,629,915 to 46,785,363 of haploidentical alleles. All markers in this region were uninformative. Comparing this region of the genome to SNP GWS investigated for other diseases (JATD, biliary atresia, PFIC) showed that this region of the genome is homozygous in all conditions and not specifically homozygous in children with PDI. No further investigation of this region was carried out. The results are shown in table 3.4.7.

	1A parent		1B parent		1C PDI		1D PDI		1E sib		3A parent		3B parent		3C PDI		3D sib	
D16S753	260	260	249	265	260	265	260	265	260	265	256	260	256	260	260	260		
D16S3183	197	199	197	199	197	197	197	197	197	199	197	199	197	199	197	199	197	197
D16S3232	202	220	208	220	220	280	220	280	202	208	216	220	210	208	220	208		
D16S3321	188	191	194	196	191	194	191	194	188	196	195	183	183	189	183	189	195	183
D16S2964	326	326	326	326	326	326	326	326	326	326	326	326	326	326	326	326	326	326
D16S3409	219	219	213	219	219	219	219	219	219	219	219	219	219	219	219	219	219	219
D16S746	278	280	280	280	278	280	278	280	280	280	280	280	280	280	280	280	280	280
D16S3105	183	185	183	185	183	185	183	185	183	183	183	185	187	183	185	183	183	187
D16S3044	192	194	186	192	194	192	194	192	192	194	192	190	192	183	190	183	190	183

	4A parent		4B parent		4C PDI		4D sib		4E sin		5A parent		5B parent		5C PDI		5D sib		5E sib		5F sib	
D16S753	264	257	257	257	257	264	257	257			253	257	261	261	257	261	257	261	257	261	253	261
D16S3183	197	208	197	197	197	197	197	208	197	208	197	197	197	197	197	197	197	197	197	197	197	197
D16S3232	212	214	202	214	202	212	202	214	212	214	216	218	208	219	209	219	217	219	209	219	216	218
D16S3321	190	198	195	199	190	199	198	198	195	198	190	194	187	189	187	189	187	189	187	189	187	190
D16S2964	326	326	326	326	326	326	326	326	326	326	326	326	326	326	326	326	326	326	326	326	326	326
D16S3409	219	219	219	219	219	219	219	219	219	219	216	219	216	222	219	222	219	222	219	222	216	222
D16S746	280	280	282	280	280	280	280	280	280	282	280	280	280	280	280	280	280	280	280	280	280	280
D16S3105	183	180	183	183	183	183	180	183	180	183	183	187	183	183	187	183	187	183	187	183	183	183
D16S3044	190	194	194	190	190	194	190	190	190	194	192	194	192	194	192	194	192	194	192	194	192	192

	6A parent		6B parent		6C PDI		6D sib		7C PDI		8A parent		8B parent		8C PDI		9A parent		9B parent		9C PDI	
D16S753	265	265	257	265	257	265	265	265									253	268	253	257	253	268
D16S3183	197	199	197	197	197	197	197	197			202	197	197	199	199	202	197	197	196	200	197	197
D16S3232	214	214	216	214	216	214	214	216			218	220	212	204	204	218	226	210	210	210	210	210
D16S3321	190	189	200	193	193	190	190	200			184	192	181	187	187	184	188	192	192	190	192	192
D16S2964	326	326	326	326	326	326	326	326	326	326	326	326	326	326	326	326	326	326	326	326	326	326
D16S3409	219	219	219	219	219	219	219	219			219	219	219	219	219	219	219	219	219	219	219	219
D16S746	280	280	280	280	280	280	280	280			280	280	280	280			280	290	280	290	280	290
D16S3105	183	183	183	183	183	183	183	183	183	183	183	187	183	187	183	187	183	183	183	183	183	183
D16S3044	190	196	192	187	192	190	190	187	192	192	194	190	190	192	192	194	186	188	192	192	192	188

Table 3.4.7 shows the microsatellite marker sizes for chromosome 16. The homozygous alleles for the affected probands are shown in green. The markers are non informative. Sib= sibling

3.4.5 Summary of chapter

The 10K and 250K SNP GWS both detected extensive areas of shared homozygosity in chromosome 5 and chromosome 19. In addition the 250K SNP GWS also identified an area on chromosome 16. Microsatellite markers excluded linkage to chromosome 19 region. Chromosome 16 could not be excluded by microsatellite markers as none were informative. When the area was compared to GWS carried out for other research projects the area was found to be a common area of homozygosity to all. This region was therefore not further explored.

Linkage was confirmed to chromosome 5 therefore identifying a locus for PDI. The extent of the shared homozygosity was narrowed to the smallest region by microsatellite markers. The markers were analysed for all consanguineous families (region 1) and then also by limiting to just those who originate from Pakistan which extended the locus (region 2).

The probands were not haploidentical suggesting the locus has not descended from a single common ancestor.

3.5 Gene expression microarray analysis of PDI patients

Whole genome expression profiles provide insight into the pattern of gene expression in specific tissue. By comparing RNA expression in disease states to that of normal controls differences in specific gene expression can be identified as being either upregulated or downregulated.

I hypothesised that a gene expression microarray would facilitate the selection of a candidate gene from the region of interest.

I sourced Cambridge Genomic Services to run the microarray.

RNA was transferred to Cambridge Genomic Services on dry ice. Despite careful handling the RNA suffered some degradation which required multiple samples to be analysed in the first stages.

The technique used involved:-

The technology uses a bead chip which contains 1.8 million beads (HumanRef-6 v2). Each bead has several thousand copies of a specific oligo probe.

There are 5 steps in the process

1. *RNA isolation*
2. *cDNA synthesis*
3. *in vitro transcription amplification*
4. *hybridization to the Illumina RNA array HumanRef-6 v2*
5. *the bead chip is then imaged using the beadchip reader*

A schematic diagram of the probe and cDNA attaches to it is shown in figure 3.6.1

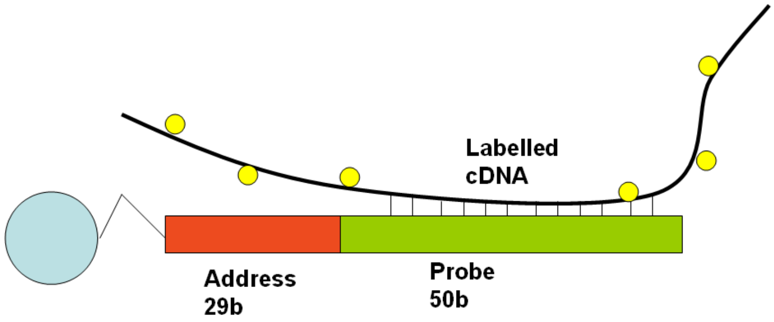


Figure 3.6.1 A schematic drawing of the relationship of the RNA expression analysis oligo probe to the applied cDNA

The level of expression of RNA is compared to the expression of control samples. No normal non-diseased controls were available and therefore PDI was compared to three disease groups with clinical features distinct to PDI – Rothman Thompson syndrome (RT), cryptogenic infantile spasms (CIS) and mental retardation, spasticty and taptoretinal degeneration (MRST) syndrome.

3.5.1 Data analysis

I analysed the results which were provided as raw expression data. PDI expression data was then compared to the expression for the control conditions and differences in expression compared in order to identify genes which are either over expressed or under expressed compared to other conditions.

Comparison (shown in table 3.5.1):-

1. ranked in order of degree of over expression compared to CIS
2. ranked in order of degree of over expression compared to RT + CIS
3. ranked in order of degree of over expression compared to MRST + CIS
4. from the top 100 over expressed and under expressed genes
   a. compare to the chromosome 5 region of interest
   b. using the data from the GWS, identify the position of the top ranked genes to recognise if a ranked gene fell in a region of homozygosity shared by all the affected probands in which the RNA was analysed

GENE	CHROMOSOME	START	MRST + CIS	RT + CIS	CIS	3C	5C	6C
C7ORF54	7	52821138	1	11	26	0.3 Mb	5.5 Mb	heterozygous
DDB1	11	60823495	2	19	48	0.1 Mb	0.4 Mb	0.2 Mb
LOC728153	5	60706424	3	12	22	heterozygous	19.7 Mb	0.2 Mb
RAB6B	3	135025769	6	1	1	15.6 Mb	0.04 Mb	0.04 Mb
CDC14A	1	100590611	20	44	11	7.5 Mb	0.1 Mb	0.4 Mb
LOC440456	17	40869050	22	22	12	1.4 Mb	heterozygous	1.8 Mb
MLL3	7	151462947	28	3	2	0.3 Mb	3.5 Mb	0.5 Mb
C7ORF28A	7	5904867	32	16	30	0.2 Mb	0.1 Mb	heterozygous
STX1A	7	72751472	33	42	72	0.8 Mb	1.1 Mb	heterozygous
HECTD2	15	93160081	41	13	33	heterozygous	0.1 Mb	heterozygous
C15ORF28	15	25686427	42	26	82	0.04 Mb	heterozygous	0.1 Mb
RAPGEF1	9	13344197	8	6	18	0.25 Mb	0.3 Mb	heterozygous
C9ORF38	9	6460369	58	7	16	heterozygous	0.4 Mb	heterozygous
DPRXP4	17	301860	78	24	58	3.4 Mb	heterozygous	0.6 Mb
POLR2J4	7	37175317	82	15	42	0.03 Mb	0.6 Mb	0.18 Mb
PML	15	72074067	85	41	40	heterozygous	heterozygous	heterozygous

Table 3.5.1. A table showing the ranking of the first 16 PDI gene expression against differing combinations of unrelated conditions. Each gene has been located on the GWS and the region of homozygosity surrounding the gene has been recorded for each proband (3C, 5C and 6C)

3.5.2 Summary of results

These results show that none of the genes with altered expression are found within the PDI locus on chromosome 5. The results also show that there were no large areas of shared homozygosity in the three probands at the position of the ranked genes. This suggests that the protein expression of the affected gene is not altered when using RNA. It can therefore be hypothesised that the molecular defect resulting in PDI may be due to an abnormality in post translational modification.

The results of the RNA expression analysis are not conclusive and do not aid the identification of the causative gene for PDI assuming the hypothesis that PDI is caused by a single gene which in consanguineous families both copies of the gene in this autosomal recessive inherited condition are identical by descent.

3.6 Direct sequencing of genes within the PDI locus

3.6.1 Chapter overview

Genes within the initial region of linkage between 96,406,286 and 103,990,534 (region 1) were further investigated by direct sequencing as in this region all probands were homozygous. The NCBI and UCSC databases were used to identify the genes in region 1. 30 genes lay within this region. There were no known genes for intestinal diseases. Prioritisation of gene sequencing was determined mainly by putative gene function, expression and size.

Predicted function which suggested high priority for sequencing were

DNA repair
Ion transport
Immunoregulatory

Further priority was then given to the size of the genes with smaller genes sequenced first.

The annotation of the gene within the public databases also influenced prioritising with the sequence being

Known genes with a documented function
Known genes with the function unknown
Hypothetical genes

Region 1 did not contain any genes with mutations in the affected probands.

The linkage data was re-analysed using only those probands who originated from Pakistan. This extended the region of shared homozygosity to include an additional 40 genes (region 2). No genes within region 2 were known to cause intestinal disease. Genes within the region were prioritised as in region 1.

Mutations were identified in the hypothetical gene *KIAA0372*. During the course of the investigation of the region the gene name was changed to *TTC37*.

This chapter describes the sequencing of genes within regions 1 and 2. A detailed description of the mutations identified in *TTC37* is found in chapter 3.7.

3.6.2 Investigation of genes in region 1

In region 1 (96,406,286 – 103,990,534) all affected probands were homozygous. The haplotype was not identical even for common ethnicity.

Using NCBI public database 30 genes were annotated in this region. Table 3.6.1 summarises the gene types within the chromosome 5 region.

Gene type	**Number in region 1**
Pseudogene	13
Hypothetical gene	4
Known function	11
Function unknown	2

Table 3.6.1 A summary of the type of genes within region 1

Genes were initially prioritised in the following order:

1. Known function
2. Function unknown
3. Hypothetical genes. Two major properties of DNA which distinguishes between coding and non coding are

- High evolutionary conservation
- Expression to give RNA transcripts

In vertebrate DNA there may also be CpG islands associated with a gene. Using these methods areas of the human genome which could potentially be coding DNA can be identified. In some cases the sequence is similar to known genes and a putative function can therefore be ascribed. In exploring the PDI region, hypothetical genes have been the last to be sequenced as they are not fully annotated.

4. Pseudogenes were excluded and not sequenced. A pseudogene is a sequence which has characteristics of one or more paralogous genes. The pseudogene sequence differs from the paralogous gene at crucial points and is therefore non functional due to non transcription, non translation or the production of a non functional protein (Mighell *et al* 2000). The exact number of pseudogenes throughout the genome is unknown. Gene families are likely to arise from a single ancestor gene via gene duplication of which some will be non functional, this is known as a nonprocessed pseudogene. 22.6% of nonprocessed pseudogenes are

within 500kb of the functional gene (Bischof *et al* 2006). Another method of pseudogene formation is by retrotransposition (processed pseudogene) in which a double stranded sequence is inserted randomly into the genome formed from single-stranded RNA and therefore has no introns.

Pseudogenes are nonfunctioning genes and therefore have not been sequenced to look for somatic mutations resulting in PDI.

It has to be remembered however that in a small number of cases a pseudogene has been reported to be involved in causing disease. This can occur by two methods:-

- Regulation of gene expression
- Gene conversion – the transformation of the sequence of one gene to that of another arising during genetic recombination i.e. a functional gene acquiring some of the sequence of a pseudogene during recombination due to the close locality (or co-localisation) of the two genes and their genetic similarity. Conditions in which this has occurred include chronic granulomatous disease (OMIM 608512), Schwachman-Bodian-Diamond (OMIM 607444) syndrome and autosomal dominant polycystic kidney disease (OMIM 601313).

Within each category genes were then prioritised according to a putative function which could result in the phenotype of PDI and the tissue expression of the gene. There were no genes with known function within the region which were good candidates for PDI. The expression of the gene for PDI is likely to be ubiquitous due to the many systems affected.

The initial four genes sequenced due to the putative function of the gene (size of the gene did not influence the selection of these genes as the proposed function made the gene a candidate for PDI) were:-

CHD1: The CHD family of proteins is characterized by the presence of chromo (chromatin organization modifier) domains and SNF2-related helicase/ATPase domains. CHD genes alter gene expression possibly by modification of chromatin structure thus altering access of the transcriptional apparatus to its chromosomal DNA template. *CHD1* is an ATP dependent chromatin-remodelling factor that also functions as a chromatin assembly factor which is thought to function during elongation as it displays physical and genetic interactions with numerous elongation factors.

CHD1 contained 35 exons and due to the large size it was initially sequenced using RT PCR with 8 RT PCR primer pairs to cover all the gene cDNA. All fragments were identified on gel electrophoresis in the probands investigated – 3C, 4C, 5C and 6C. The PCR products were then directly sequenced and no mutations were identified.

CAST: this is a calpastatin which is an endogenous inhibitor of calpains in the presence of calcium. Calpains are non-lysosomal calcium-dependent cysteine proteinases that selectively cleave proteins in response to calcium signals and thereby control cellular functions such as cytoskeletal remodelling, cell cycle progression, gene expression and apoptotic cell death. The ubiquitous expression and variety of functions of calpains and the inhibitor calpastatin make *CAST* a potential candidate gene for PDI.

CAST is composed of 31 exons which were directly sequenced using genomic DNA. Two SNPs were detected which were homozygous.

***SLC**O4C1* and *SLCO6A1*: solute transporters are associated with intractable diarrhoea of infancy and therefore mutations may potentially result in the phenotype of PDI. *SLCO4C1* was sequence using cDNA and RT PCR whilst *SLCO6A1* was sequenced using genomic DNA. Homozygous SNP's were identified.

Following exclusion of the above genes the other known genes within region 1 were sequenced using genomic DNA in order of number of exons commencing with the smallest genes.

The DNA from four probands (3C, 4C, 5C, 6C) was used to screen each of the genes. This cohort had been personally phenotyped by the investigator (JH) and DNA volume and quality was good.

Table 3.6.2 summarises the findings in each of the genes sequenced.

Gene	Proband	Change
LIX1		no changes
RIOK2	6C	rs2544773
	5C, 6C	rs160632
	4C, 5C	rs12188395
	5C, 6C	rs8654
RGMB	3C, 4C, 5C, 6C	rs2547973
	6C	?rare SNP
FLJ35946	3C, 4C, 5C	start codon
CHD1		no changes
LOC441066		no changes
LOC728104	4C, 6C	rs2460669
TMEM157		no changes
ST8SIA4	6C	?rare SNP
SLCO4C1	5C	rs10479190
SLCO6A1	3C	rs6884141
	3C, 6C	rs11746217
	6C	rs17150488
	6C	rs10073333
PAM		no changes
LOC134505		no changes
FLJ20125		no changes
HISPPD1		no changes
LOC90355		no changes
NUDT12	4C	rs7734923
	4C	rs7723689
	4C, 6C	rs10045774

Table 3.6.2 A summary of the sequence changes found within the genes of region 1.

Fourteen SNP’s which are annotated with an rs number were identified in 6 different genes. Only one SNP (rs2547973) in *RGMB* was seen in all 4 screening probands. The identification of a SNP in one proband and not the others made this gene an unlikely candidate for PDI which is hypothesised to be caused by a single gene.

In two genes *RGMB* and *ST8SIA4*, proband 6C had point mutations which were not seen in the other probands and also had not previously been described. This suggests that these mutations are likely to be newly identified SNP’s.

RGMB

The RGMB gene (Repulsive guidance molecule B) is expressed in the development of the central and peripheral nervous system and may also be involved in the differentiation of intestinal epithelium and therefore a potential gene candidate for PDI.

A sequence variant ATC-ATT in exon 5 of proband 6 results in a synonymous amino acid change leucine to leucine. The nucleotide change segregated appropriately within the family with both parents being heterozygous as well as the unaffected sibling. 160 Asian controls were directly sequenced and no heterozygous or homozygous identical nucleotide changes were found.

No identical sequence variants were found in the other affected patients and no other changes were identified in the other exons. This synonymous change is therefore unlikely to be a significant disease causing mutations for PDI and is likely to be a rare SNP which has not previously been described.

ST8SIA4

The family of sialyltransferases regulate the linkage between neural cell adhesion molecules and polysialic acid which modulates the adhesive properties. Polysialic acid has been implicated in numerous normal and pathologic processes, including development, neuronal plasticity, and tumour metastasis. It has no known pathology within the intestine and is therefore not an obvious candidate gene but was sequenced as part of sequencing of region1.

Proband 6C had a single base change resulting in a non synonymous amino acid change from leucine to methionine. The change segregated within the family with both parents and sibling being heterozygous. The sequence variant was not seen in any other proband. 104 Asian control chromosomes also did not identify this sequence change. This suggests the change did not result in the clinical phenotype of PDI and is likely to be a rare SNP which has not previously been described.

In summary, the direct sequencing of all genes in region 1 did not identify any sequence variants which could plausibly result in the clinical phenotype of PDI.

3.6.3 Investigation of genes in the region 2

Region 2 was investigated following the reanalysis of the microsatellite data using exclusively those probands which had originated from Pakistan.

This extended the region to include a further 40 genes. None of the genes within the region were excellent candidate genes and therefore the same approach as to region 1 was applied. The 6 pseudogenes within the region were not sequenced. The region also contained 2 microRNA genes.

MicroRNAs (miRNAs) are small RNA molecules of about 22 nucleotides which act as regulators of other genes. Human MiR583 has homology with Drosophila MiR278 which is important in adipocyte regulation and insulin resistance. The MIRN583 gene was sequenced in the exploration of the PDI region.

A summary of the type of genes in region 2 is shown in table 3.6.3

Gene type	**Number of genes**
Genes with known function	14
Genes with no known function	11
MicroRNA genes	2
Hypothetical genes	7
Pseudogenes	6

Table 3.6.3 A summary of the gene types in region 2.

Table 3.6.4 Summary of sequence changes identified in genes in region 2

Gene	Proband	SNP identification number
CETN3	4C	rs11554603
	5C	rs4873
LOC153364	3C, 5C,6C	rs2162986
POLR3G		no changes
LYSMD3	5C	rs10069050
ARRDC3		no changes
C5orf21		no changes
C5orf36		no changes
ANKRD32	4C	novel SNP
FAM81B	5C, 6C	rs17853328
TTC37		mutations
ARSK	3C, 5C	rs17084927
RFESD		no changes
SPATA9		no changes
RHOBTB3	4C, 5C	rs34896
	3C, 4C, 5C	rs34898
	3C	rs41276257
GLRX		no changes
FIS	6C	Novel splice site
	4C	Novel splice site
ELL2	3C	rs17685249
	3C, 4C	rs3777204
MIRN583		no changes
PCSK	6C	rs6233
	3C	rs6235
CAST	3C, 4C	rs11558594
	3C	rs2290678
LNPEP	5C	novel SNP
	5C	rs11311774
	6C	rs1174632
LRAP	3C	rs73150323
	4C, 6C	rs2549782
	4C, 6C	rs2548438
	4C, 6C	rs41276277
	4C, 6C	rs61731306
	6C	rs2549796
	6C	rs1056893
	6C	rs73152140

Twenty-five SNP were identified in region 2. The SNP in *CETN3* was heterozygous in 4C and suggested therefore that this region was no longer homozygous and was the upper limit of the region. The genes prior to *CETN3* were not sequenced (*LOC645323, MIRN9-2*, *MEF2C*, *LOC729011*).

Not all genes had complete sequencing analysed before the discovery of mutations in TTC37.

Genes in which sequencing was not completely analysed:-

NR2F1

LOC729040

GPR150

MCTP1

LOC441097

FLJ25680

Those genes which had no analysis:-

ARTS

GPR98

Further investigations of FIS

This gene is poorly annotated in the databases with no exons delineated.

The gene was extensively investigated.

1. Four exons were identified by comparing the cDNA sequence with genomic DNA sequence.
2. The identified exons were directly sequenced and many changes were identified including the formation of premature stop condons.

3. BLAST of the genome with individual exons and protein sequence was used to identify alternative transcripts– no alternative transcripts within the genome but different assemblies of the genome (HuFA and Celera) have different transcripts which were identified by BLAST. BLAT did not reveal any alternative sequence within the genome
4. cDNA fragments to cover alternative start and stop codons and to also cover the identified sequence variants. Control cDNA was also sequenced.
5. FIS was sequenced in all affected probands.
6. FIS was sequenced in control DNA samples.

<u>Results</u>

Exon 2: insertion of 4 bases – identified in a different transcript of the gene

Exon 2: GCC->GAC – seen in control samples

Exon 4: GAA->TAA (stop) – also seen in control samples therefore likely to be an alternative transcript

3.6.4 Summary of chapter

All the genes in the region linked by all probands did not identify any mutations which could be ascribed to causing PDI. When the region of linkage was examined to include just those patients from Pakistan the region extended to include a further 40 genes. In total 39 genes were fully sequenced. Known SNP’s and novel SNP’s were identified. Mutations were identified in *TTC*37 and the further investigation of *TTC*37 is described in chapter 3.7.

3.7 Mutation identification in *TTC37*

3.7.1 Chapter overview

All sequencing was analysed using Chromas and compared to a reference sequence from Ensembl release 52 – Dec 2008.

The DNA from 4 affected children with classical features (3C, 4C, 5C and 6C) was initially sequenced and when germline mutations were identified the other probands with PDI were sequenced. Where available, the affected exon of the proband was analysed in the parents and siblings to look for segregation within families.

To ensure the sequence changes were not due to rare SNP's more than 350 Asian control samples were sequenced for each mutation.

3.7.2 *TTC37* sequencing results

Family 1

This family are reported to be non consanguineous of Indian origin but originate from the same village and therefore there may be potential distant consanguinity. The probands died from liver disease with excessive hepatic iron overload. In affected siblings 1C and 1D a homozygous base change from c.2808G>A changes the amino acid codon from tryptophan (TGG: W) to a stop codon (TGA: X) at amino acid position 936 (W936X). The unaffected parents and unaffected sibling were heterozygous as shown in figure 3.7.1.

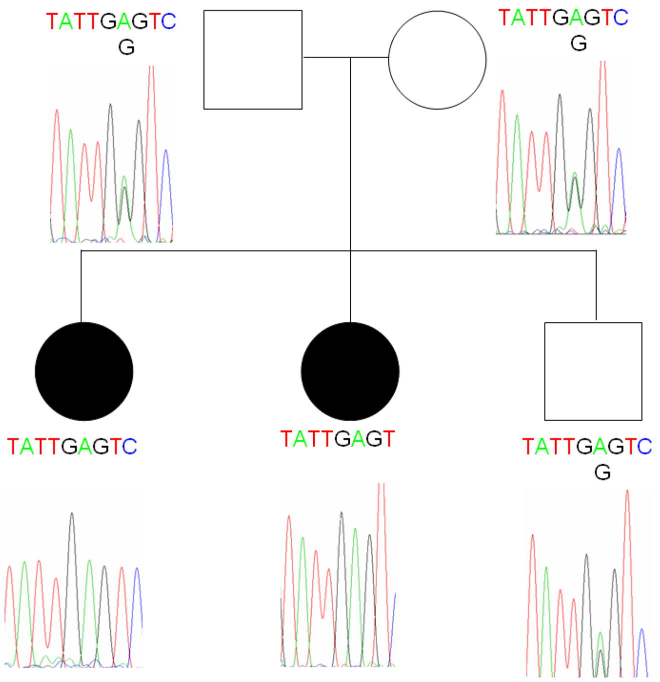


Figure 3.7.1. A chromatogram showing the segregation within family 1 of the base change TGG->TGA (stop)

A premature stop codon at position 936 will truncate the protein.

The amino acid at this position is conserved in mammals and flies but not in zebra fish (Danio rario).

Homo sapian	928	EGALGYAYWVCTTLQDKSNRETELYQYNILQMNAIPAAQVILNKYVERIQ	977
Canis lupus	926	EGAIGYAYWVCTTLQDKSNRDTELYRYNIVQMNAIPAAQVVLSKYIERIQ	975
Bos taurus	928	EGAIGYAYWVCTTLQDKSNRDTELYRYNILQMNAIPAAQVVLSKYVERIQ	977
Mus musculus	928	EGAIGYAYWVCTTLQDKSNRETELYQYNILEMNAIPAAQGVLCKYVERIQ	977
Rattus norvegicus	928	EGAIGYAYWVCTTLQDKSNRETELYQYNILQMNAVPAAQGVLCKYVERIQ	977
Gallus gallus	926	EGAKGYAHWVCSTLQDKSNRDTEQYLYNIVEMNAIPAAQVVMSKYTERNP	975
Danio rario	904	EGVKGYAYCVCSTLLDRSNRDSELYLYNIVQMNAVSAAQVALSKYTERIQ	953
Fly	872	EAALGFAHWVCEMLSTPGSFDKPRIKHAIEHMYADVLALDAINWYVQNEE	921
Anopheles	873	EAALGYAHWVCSIVNEDNYHENERYRFAIDAMAALPVAHDAIGWHCADLA	922

Figure 3.7.2 The conservation of amino acid 936 of *TTC37* in other species

Family 2

This non-consanguineous proband of Dutch origin has a heterozygous sequence change p.Asp1283Asn as also found in a Flemish proband. This may be a founder sequence variant. It is predicted to be a benign change in Polyphen. The child is also heterozygous for the SNP rs17084873 which is also predicted to be a benign change by Polyphen.

Family 3

Family 3 originates from the Mirpur region of Pakistan and are of consanguineous union. Proband 3C has a homozygous intronic change c.2779-2G>A. The parents are heterozygous for this splice site change as shown in figure 3.7.3.

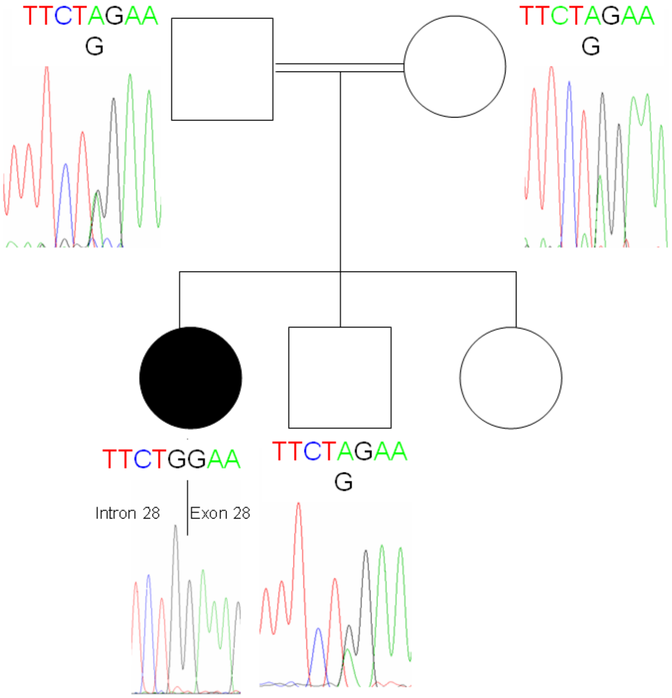


Figure 3.7.3 A chromatogram showing the splice site mutation segregating within family 3

Splice site prediction for this acceptor site predicts the wild type splice site with a score of 0.96 but the splice site can not be identified with the base change from A->G.

Reverse transcriptase PCR of cDNA with a forward primer in exon 26 and reverse primer in exon 31 showed a loss of the normal size band on gel electrophoresis. A second transcript was also seen in 3 affected children and faintly in control 2. This is shown in figure 3.7.4.

Both 3C and 5C have the same mutation c.2779-2G>A. This suggests that this mutation results in skipping of exon 29 creating a smaller RT PCR product.

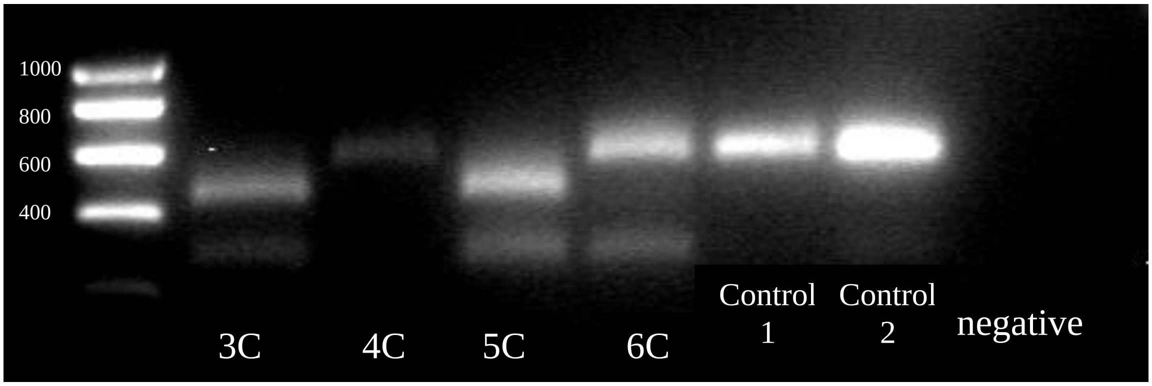


Figure 3.7.4 The RT-PCR products for patients 3C and 5C are smaller suggesting truncation of the protein secondary to the mutations.

The RT PCR product was sequenced to identify the missing sequence and is shown in figure 3.7.5. There is background sequencing due to the second transcript but despite this it can be seen that exon 28 goes straight to exon 30 and exon 29 has not been expressed.

The mutations identified in 4C and 6C do not affect this splice site and the exon 29 is normally expressed.

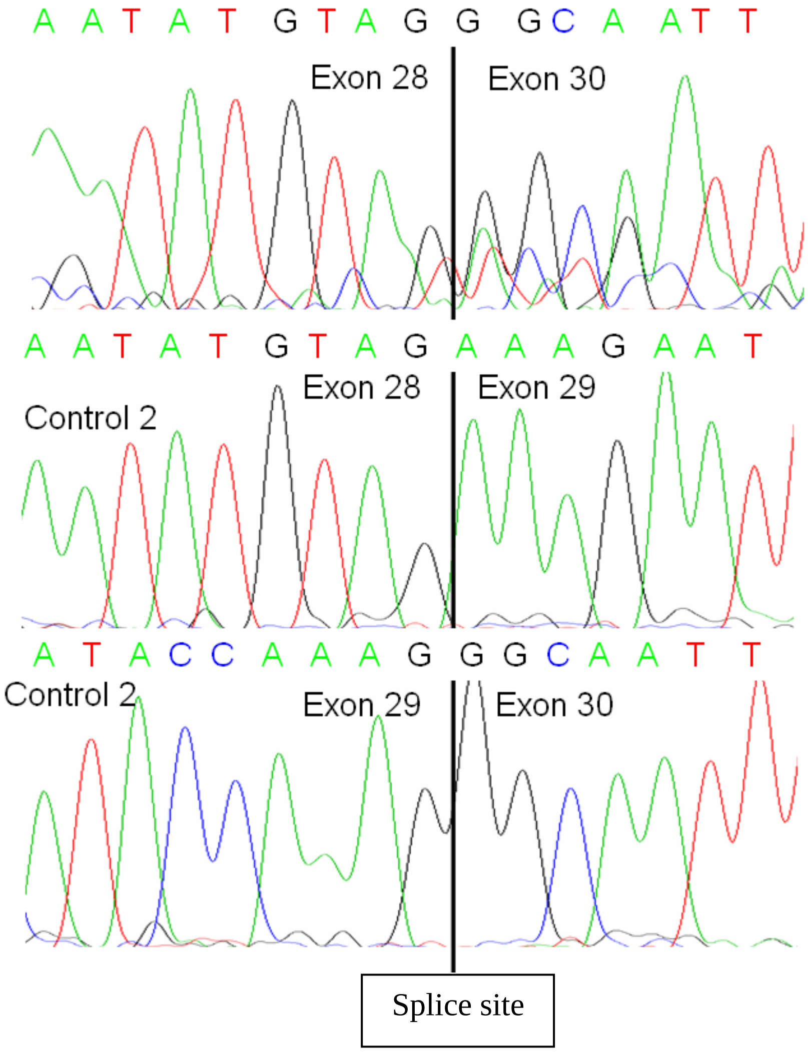


Figure 3.7.5 The chromatograms of the cDNA PCR product of 3C. The splice site mutation in 3C causes skipping of exon 29

The splice site mutation also results in a frame shift which is likely to create an abnormal protein. The frameshift also results in the formation of a premature stop codon at amino acid 19 in exon 30 so truncating the abnormal protein.

Figure 3.7.6 Shows wild type cDNA transcript with codons shown in alternating colours followed by the cDNA transcript for 3C highlighting that skipping of exon 29 causes the formation of a premature stop codon on exon 30 and therefore truncation of the protein:-

Exon 28

CTGCACAACATTGCAAGATAAAAGCAACAGAGAAACAGAGCTGTACCAGTACAACATCCT

CCAGATGAAT

GCTATTCCAGCAGCACAAGTTATTTTGAATAAATATGTAG

Exon 29

AAAGAATTCAGAATTATGCCCCAGCTTTCACAATGTTGGGTTACTTAAACGAACATCTAC

AACTGAAAAAGGAAGCAGCAAATGCATACCAAAG

Exon 30

GGCAATTTTGTTGTTACAGACTGCAGAAGACCAAGATACTTACAATGTTGCAATAAGAAA

TTACGGCAGATTGTTATGTTCCACTGGTGAATATGATAAAGCTATCCAGGCTTTTAAGTC

AACACCCCTTGAAGTG

3C cDNA transcript with codons shown in alternating colours and the formation of a stop codon in exon 30 shown in red

Exon 28

GCTATTCCAGCAGCACAAGTTATTTTGAATAAATATGTAG

Exon 30

GGCAATTTTGTTGTTACAGACTGCAGAAGACCAAGATACTTACAATGTTGCAATAAGAAA

TTACGGCAGATTGTTATGTTCCACTGGTGAATATGATAAAGC

Family 4

Family 4 are of Kurdish origin and are of consanguineous union. 4C has a homozygous deletion of a single base G at the splice site of exon17 / intron17, c.1632+1delG as shown in figure 3.7.7. The parents and unaffected dizygotic twin are heterozygous and the unaffected sibling is wild type.

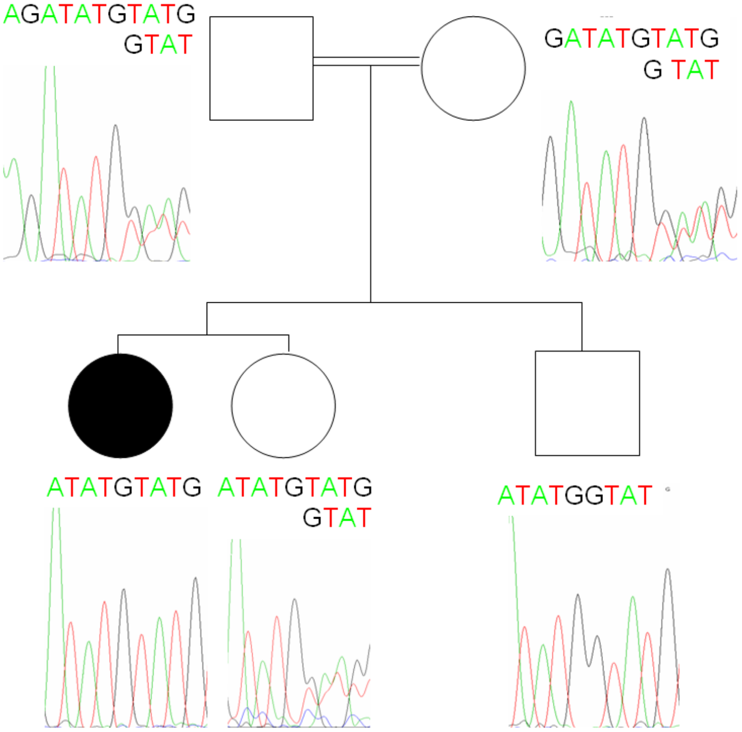
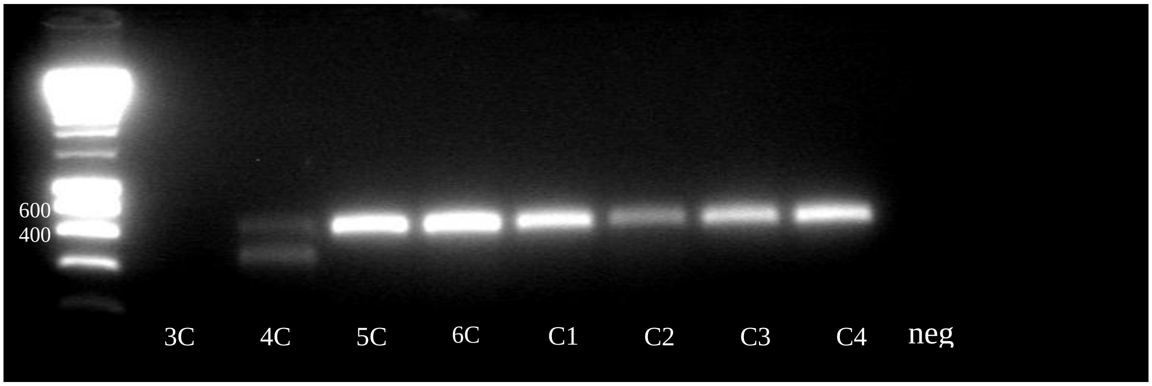


Figure 3.7.7 Shows the chromatogram for family 4 with segregation of the splice site mutation

The deleted base is either the last in exon17 or the first in the intron. In either case the splice site is affected with splice site predictor show a score of 0.98 for the wild type splice site and a reduction to 0.75 with the base deletion.

If the base deletion is in the exon there is a frame shift resulting in the formation of abnormal protein in exon 18 and the formation of a stop codon, p.Glu545Phefs*40

Figure 3.7.8 RT PCR using primers in exons 15 and 19 on gel electrophoresis showed a smaller band for 4C as compared with the wild type and other affecteds with different mutations. A faint band was seen the same size as the wild type and this is thought to be due to leaky splicing.



To investigate for contamination in 4C, 3 different samples of RNA from patient 4C were converted to cDNA (chapter 2) and the same RT PCR primers run. This showed that all three samples had the two bands and again this was not seen in the controls. The predominant band is the smaller band. In this RT PCR a sample from 3C was re run as in the initial sample no product was identified. On repeat 3C had a single wild type band as shown in figure 3.7.9.

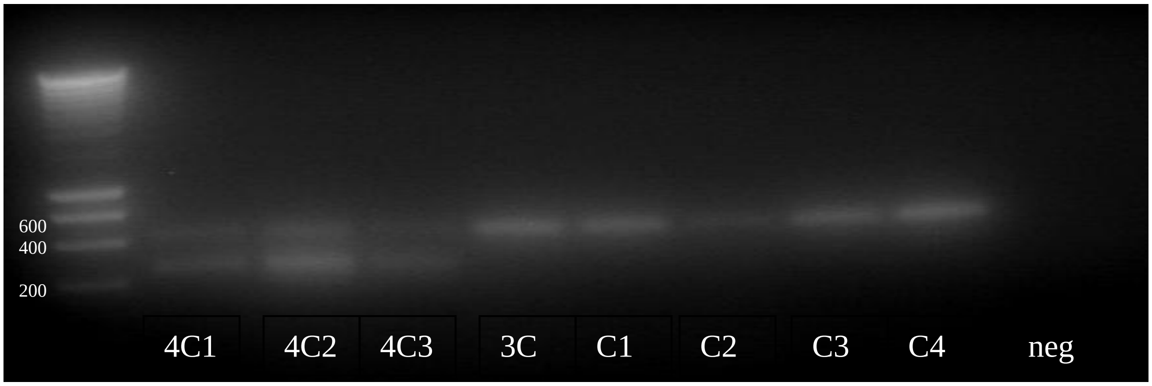


Figure 3.7.9 RT PCR of 3 different RNA samples for 4C to ensure the findings were not due to contamination.

The RT PCR product was sequenced which identified a loss of expression of exon 17. Although a small amount of exon 17 sequence can also be identified as shown in figure 3.7.10.

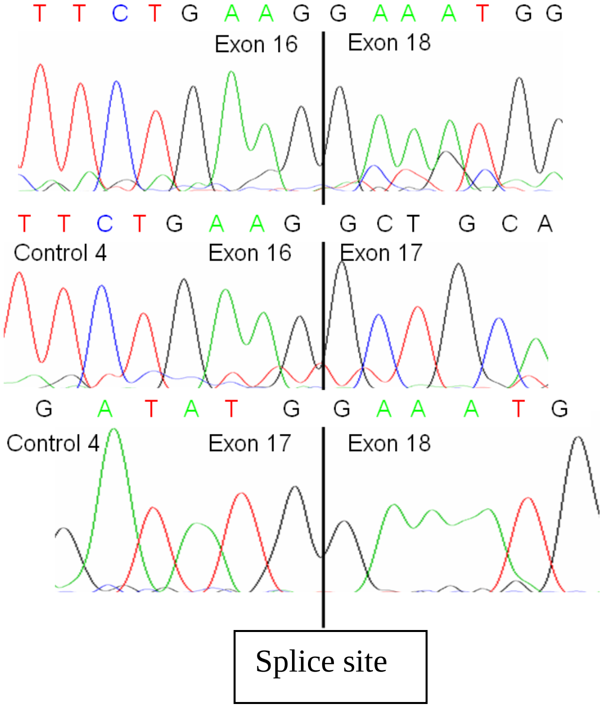


Figure 3.7.10 Sequencing of the RT PCR product identified skipping of exon 17 due to the splice site mutation.

Family 5

Family 5 are consanguineous and originate from the Mirpur region of Pakistan. The affected child 5C is homozygous for the intronic mutation c.2779-2G>A. This is the same mutation as identified in child 3C. The parents are heterozygous and the three unaffected siblings are all heterozygous as shown in figure 3.7.11.

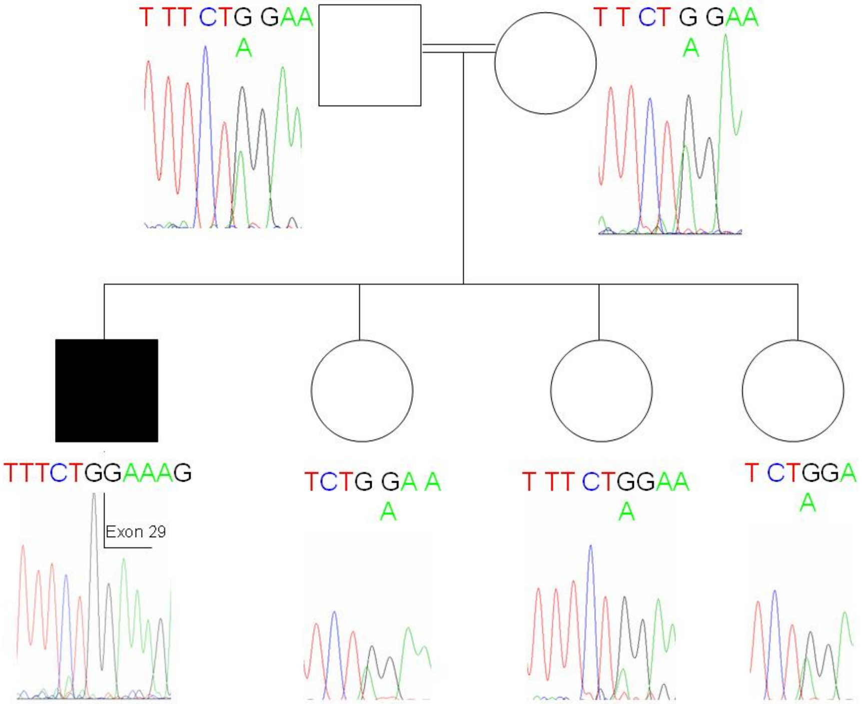


Figure 3.7.11 The chromatogram of family 5 showing segregation of the splice site mutation

This base change in intron 28 predicts a splice site change which results in skipping of exon 29 as shown in figure 3.7.12. The RT PCR of 5C has been shown in 3.7.4 where the wild type electrophoresis band is lost and a smaller band has been formed. This is identical to those of 3C who has the same mutation.

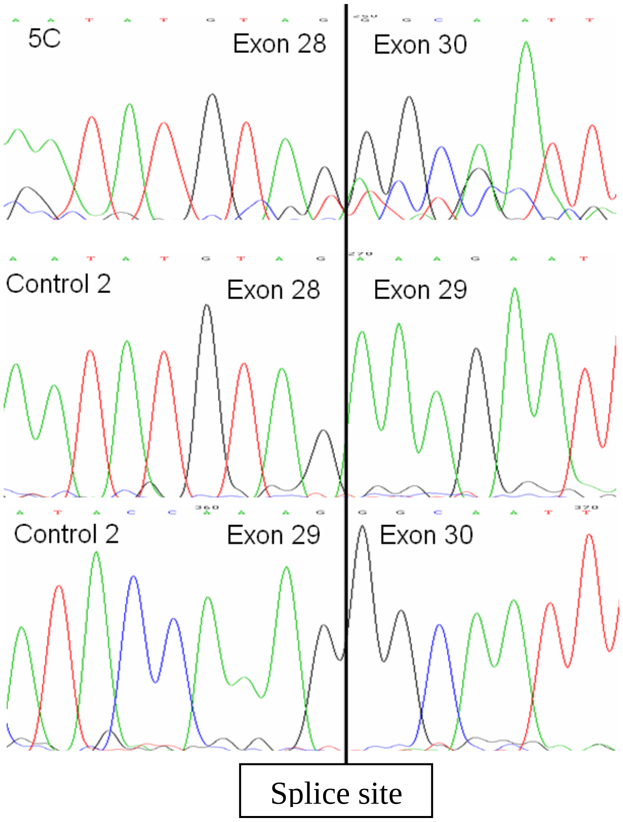


Figure 3.7.12 Sequencing of the RT PCR for 5C confirms the skipping of exon 29.

Family 6

This family are of Pakistani origin and of consanguineous union. A homozygous base change in exon 10 at the exon / intron boundary was identified, c.751G>A shown in figure 3.7.13. This was confirmed by sequencing of the RT PCR product which showed skipping of exon 10 (figure 3.7.15). This is predicted to cause a splice site alteration and also an amino acid change G251R which is predicted using Polyphen to be probably damaging with a PSIC score difference of 2.168.

Figure 3.7.13 A chromatogram of family 6 showing segregation of the splice site mutation

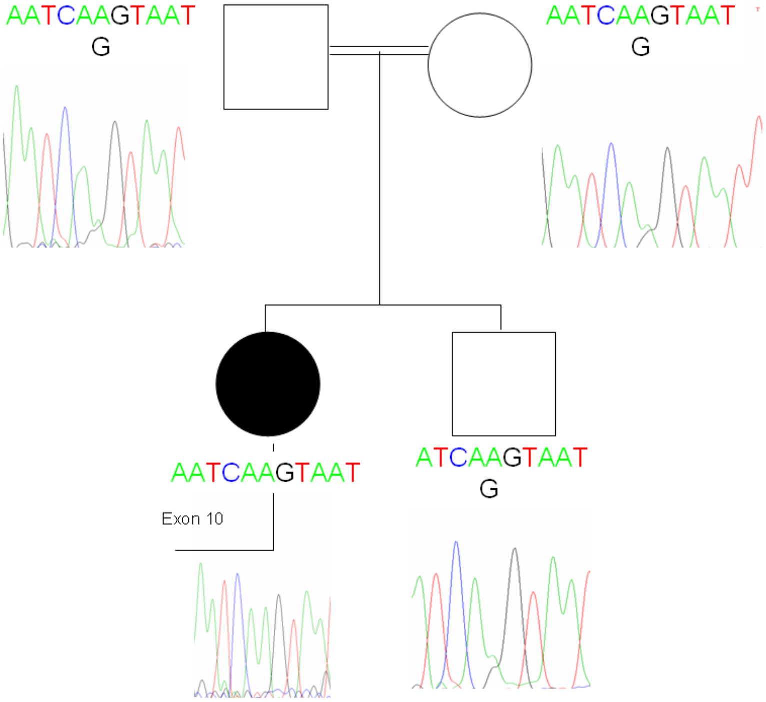


Figure 3.7.14 The splice site prediction was investigated using RT PCR and showed a smaller PCR product for 6C as compared to other affecteds with different mutations and control samples.

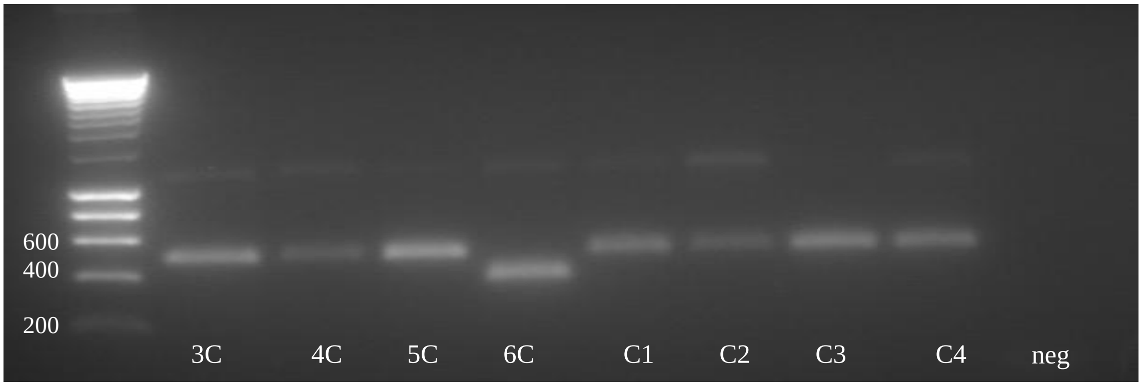
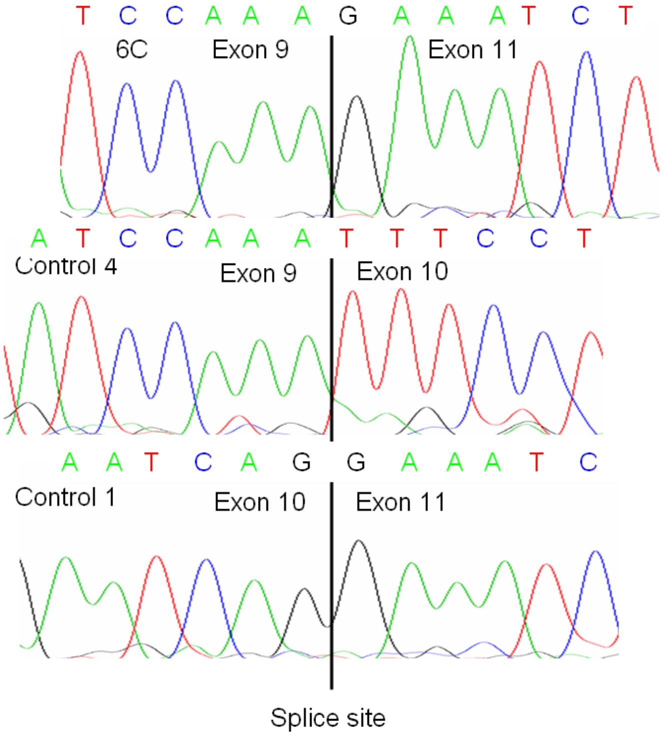


Figure 3.7.15 Sequencing of the RT PCR fragment shows that exon 10 is not expressed. This then causes a frame shift and the formation of a stop at codon 14 in exon 11



Family 7

This family originate from Pakistan and are of consanguineous union. A homozygous base change TGG->TGA (c.2808G>A) causes the formation of a stop codon W936X in exon 28. This is seen in figure 3.7.16. This is the same mutation as seen in family 1 and also family 14.

Only DNA from the affected child is available and therefore segregation within the family can not be ascertained.

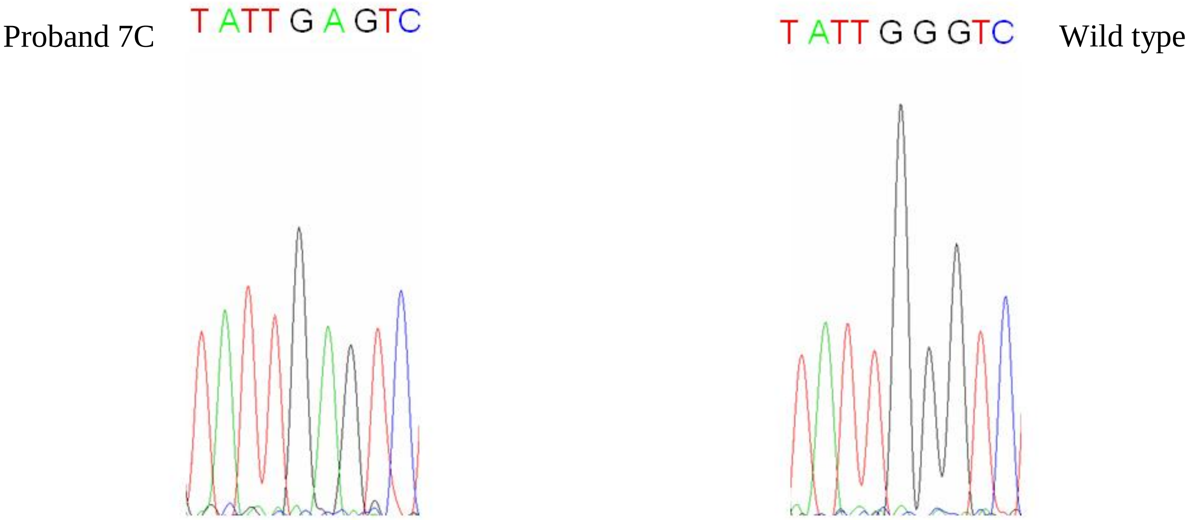


Figure 3.7.16 Show the chromatogram for patient 7C and a control as extended family DNA was not available

Family 8

This family originate from Pakistan and are of consanguineous union. The original genome wide scan and microsatellite studies showed linkage to the region. A cousin died from liver disease and had a similar phenotype. A genomewide scan of the cousin did not show homozygosity to the region of interest. No mutations have been identified in this family. Two SNP's were identified in 8C both of which were heterozygous, rs17084873 in exon 17 and rs2303650 in exon 37.

This suggests that there may be locus heterogeneity and another gene also causes the PDI phenotype. An alternative explanation is the two SNP's result in a reduction in protein expression causing the disease phenotype. Polyphen predicts R1296S to be possibly damaging with a PSIC score difference of 1.895 whilst L437V is predicted to be benign with a score difference of 0.092.

Family 9

Initially a single base deletion was identified but not confirmed on resequencing and therefore no mutation was identified in this family.

Family 10

This family are Caucasian English and not of consanguineous union. Only DNA from the child is available.

This initially showed 4 sequence variants:-

Exon 6: heterozygous L106H. This variant was identified in many controls suggesting it is an unannotated SNP. This is predicted to be potentially damaging by Polyphen.

Exon 15: a 2 base deletion which causes a frame shift and the formation of a stop codon c.1300_1301delAA, p.Lys434Lysfs*14 The amino acid which is deleted is conserved in mammals but not in chicken, flies or zebra fish.

Intron 21-4: heterozygous change in intron 21, 4 bases before exon 22. 290 controls are all wild type. The prediction is that this causes no effect on the splice site indeed the score for the wild type is 0.95 and the sequence variant 0.97. This may also be a rare SNP.

Exon 42: heterozygous change c .4514T>C, p.Leu1505Ser which is predicted to be probably damaging by Polyphen. 140 control chromosomes are all wild type. The amino acid is conserved in mammals and in zebra fish but the sequence is not found in flies.

```
Homo sapian      RKMGARETRRLLERIVYQ--TGYPSSIVSAARWYLLRHLYAKDDPELIDV 1540
Canis lupus      ---DIIETRRLLERVVYQ--TGYPNSIVSTARWYLLRHLHAKDDHELIDV 1448
Bos taurus       RKMGARETRRLLERVVYQ--PGYPKSIVSTARWYLLRHLHAKNDYELIDV 1539
Mus musculus     RKMGARETRRLLERVVYQ--PGYPKSIVSTARWYLLRHLHAKNDYELIDV 1537
Rattus norvegicusRKMGARETRRLLERVVYQ--PGYPKSIASTARWYLLRHLYAKDDYELIDV 1541
Gallus gallus    RKMGARETRRMLERVVYQ--PGNPETIVSVARWYLLQHLYAKDDYELIDV 1538
Danio rario      VKMGARETRRLLERIVYASALGGSETIASVARWYLLRHLHAKDDLELIDT 1502
Fly              --------------------------------------------------
Anopheles        --------------------------------------------------
```

Figure 3.7.17 Shows conservation between species of the amino acid 1505

Family 11

This family are non consanguineous and are from Italy. This patient is a compound heterozygote with two base changes both resulting in the formation of a stop codon.

From the heterozygous father a CAA->TAA in exon 21, c.2251C>T, G751X and from the heterozygous mother CAG->TAG in exon 8, c.439C>T, G147X. These are shown in figure 3.7.18.

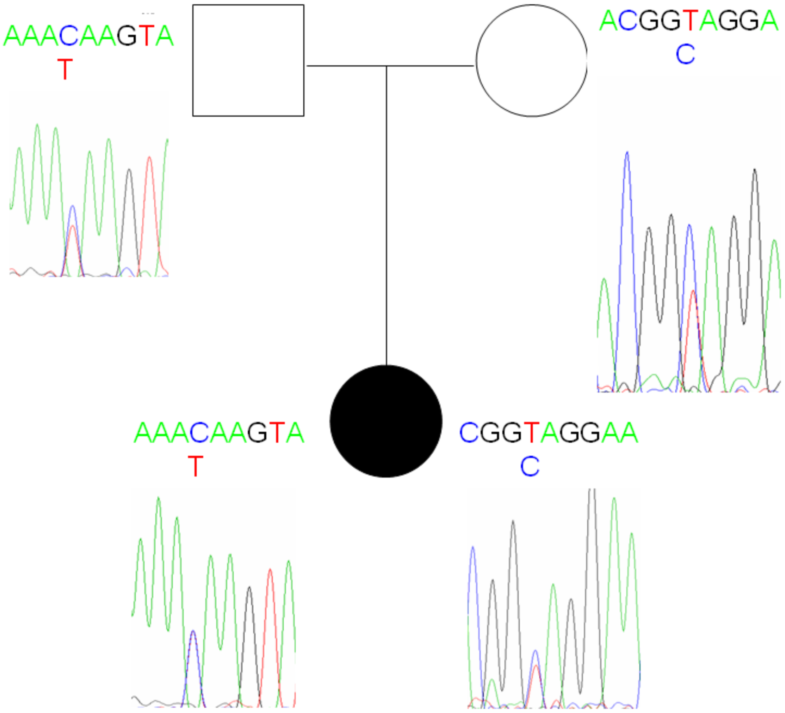


Figure 3.7.18 Chromatograms showing the segregation of the two mutations CAA->TAA and CAG->TAG within family 11

Family 12

In this Flemish non consanguineous family only one mutation has been identified, D1283N. This mutation was also identified in patient 2C. Polyphen predicts this to be a benign change. The amino acid is conserved in mammals but not in chicken, zebra fish or flies which is shown in figure 3.7.19. As in child 2C the other mutation in this child has not been identified. Child 2C is Dutch and there may be a rare SNP in these two families.

Homo sapian	TAEDKSNTALKTIQKAAFLSPDDPAVWAGLMAACHADDKLALLNNTQPKR	1295
Canis lupus	TAEDKSSTALKTIQKAAFLSPDDPAVWAGLMAACHADDKLALVSNTQPKR	1295
Bos taurus	SAETEKNLALKTIQKAALLSPGDPAVWAGLMAACHADDKLALVNNTQPKR	1294
Mus musculus	SAEDEKNTALKTIQKAALLSPGDPAVWAGLMAACHADDILALVSSTQPKR	1292
Rattus norvegicus	SAEDEKNTALKTIQKAALLSPGDPAIWAGLMAACHADDKLALVNNTQPKR	1296
Gallus gallus	MLEDERNNPLKNIQKAIHICPDNPAAWAVLMAACHAENTVVCLNNTQPKR	1293
Danio rario	SGEDRRHNALKTIQRAVLLCPDDPAGWAGLMAACHTENTACFLTGSTPHR	1303
Fly	VSAVDKTCSMKLLQRAILLSPTDQRARQLLSAIIANS-------------	1233
Anopheles	--	

Figure 3.7.19 shows conservation of the amino acid D1283N in other species

In the parents with sequence change was not identified and this change may represent a de novo mutation.

Family 13

This family are English, non consanguineous. No sequence variant has been identified in this child. This may indicate a second locus for the PDI phenotype. This child has a neurological signs which may be part of the clinical component or may represent a separate insult at the time of birth. If the neurology is part of the syndrome then this is distinct to the other affected children and may represent a novel phenotype.

Family 14

This consanguineous family are from the Mirpur region of Pakistan. The affected child 14C is homozygous for TGG->TGA, c.2808G>A, p.Trp936X in exon 28. The parents are heterozygous as seen in figure 3.7.20. This mutation has also been identified in the two Indian siblings 1C and 1D and also 7C who also originated from Pakistan.

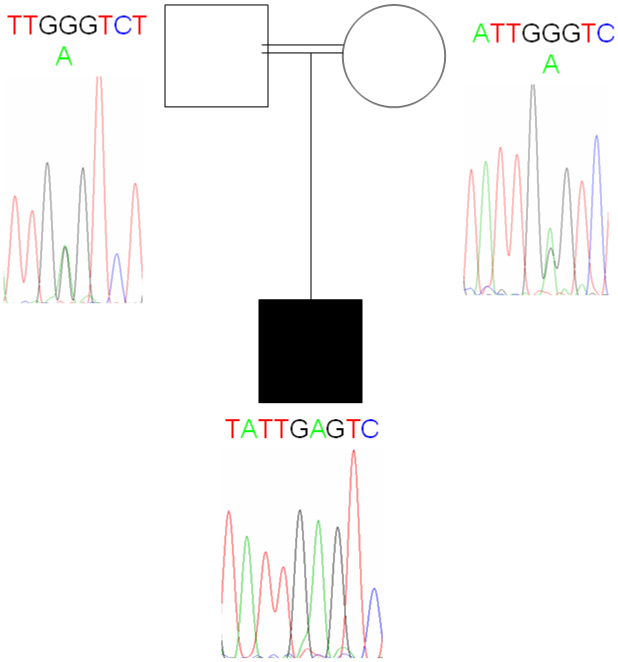


Figure 3.7.20 The chromatograms for family 14 showing appropriate segregation of p.Trp936X

3.7.3 Summary of results

Mutations in *TTC37* are associated with the clinical condition PDI.

A nonsense mutation in exon 28 (c.2808G>A: p.W936X) was detected in three apparently unrelated families (1, 7 and 14) of South Asian origin. The SNP haplotypes in the affected individuals were identical over a 974 kb region (from rs255375 to rs34897) containing *TTC37*, consistent with a founder mutation. An intron 28 splice site mutation was detected in two apparently unrelated families from Pakistan (3 and 5) with identical

SNP haplotypes over a 457 kb interval encompassing *TTC37* (from rs116286 to rs7736948).

RNA studies were undertaken to investigate putative splice site mutations in 4 affected individuals. In families 3 and 5 an intronic c.2779-2G>A sequence change resulted in skipping of exon 29 and a predicted truncated protein (p.Glu974Glyfs*19). In family 4, a c.1632+1delG variation at the first nucleotide of intron 17 resulted in skipping of exon 18 producing a frameshift and a premature stop codon, p.Glu545Phefs*40. In family 6 the c.751G>A substitution, predicted to cause a missense substitution (p.Gly251Arg) involved the final nucleotide of exon 10 and resulted in skipping of exon 10 (figure 5c) resulting in a frameshift and predicting a truncated protein (p.Phe215Glufs*14). A summary of results is shown in table 3.7.1.

Table 3.7.1 provides a summary of mutations detected in *TTC37*

Family Identifier	**Number of affected individuals**	**Ethnicity**	**Consanguinity**	**Mutation 1**	**Mutation 2**	**Changes which may be rare SNP's**
1	2	Indian	No	c.2808G>A p.Trp936X	c.2808G>A p.Trp936X	
2	1	Dutch	No			c.3847G>A p.Asp1283Asn
3	1	Pakistani	Yes	c.2779-2G>A p.Glu974Glyfs*19	c.2779-2G>A p.Glu974Glyfs*19	
4	1	Kurdish	Yes	c.1632+1delG p.Glu545Phefs*40	c.1632+1delG p.Glu545Phefs*40	
5	1	Pakistani	Yes	c.2779-2G>A p.Glu974Glyfs*19	c.2779-2G>A p.Glu974Glyfs*19	
6	1	Pakistani	Yes	c.751G>A p.Phe215Glufs*14	c.751G>A p.Phe215Glufs*14	
7	1	Pakistani	Yes	c.2808G>A p.Trp936X	c.2808G>A p.Trp936X	
10	1	English	No	c.1300_1301delAA p.Lys434Lysfs*14	c.4514T>C p.Leu1505Ser	c. p.Leu106His
11	1	Italian	No	c.439C>T p.Gln147X	c.2251C>T p.Gln751X	
12	1	Flemish	No			c.3847G>A p.Asp1283Asn
14	1	Pakistani	Yes	c.2808G>A p.Trp936X	c.2808G>A p.Trp936X	

3.8 *in silico* analysis of *TTC*37

3.8.1 Gene position and structure

The gene *TTC37* is found on chromosome 5q15 at position 94,799,599-94,890,709

TTC37 was previously known as *KIAA0372*. Analysis of the domains of *KIAA0372* led to the finding that the gene contained tetratricopeptide repeat domains leading to the renaming of the annotated gene in 2008.

There is a single transcript of the gene which is found on the reverse strand.

TTC37 consists of 43 exons of which exons 1, 2 and 3 are non coding. *TTC37* spans 5,704 bases and encodes 1564 amino acids.

The nucleotide and amino acid sequence of TTC37 is provided in appendix 8.4.

The relationship with other genes in the area shows that *TTC37* does not incorporate any other genes. *TTC37* is also not in the region of a superfamily domain.

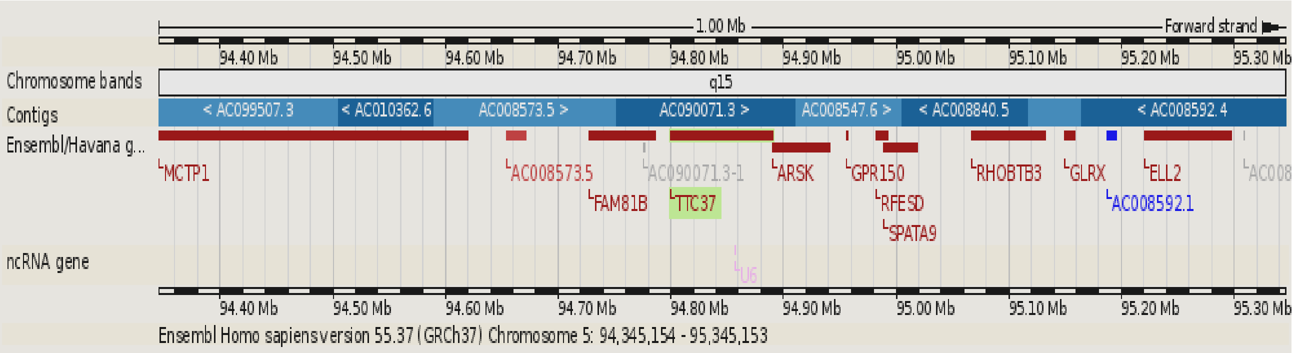


Figure 3.8.1 Shows the relationship of *TTC37* to other genes in the region as seen in Ensembl. There is no superfamily domain or bases which overlap with other genes. The position of *TTC37* therefore does not indicate the gene product function.

3.8.2 Known polymorphisms within the coding region

5 polymorphisms have been annotated in the exons of *TTC37*.

Exon 1non coding region in the start codon amino acid position 1

Exon 9 synonymous change mRNA position 855 rs73147944 heterozygosity rate unknown G->A Lys->Lys codon position 1 amino acid 195

Exon 15 missense change mRNA position 1579 rs17684873 heterozygosity rate 0.342 C->G leu->val codon position 1 amino acid position 437

Exon 37 Missense change mRNA position 4158 rs2303650 heterozygosity rate 0.333 G->T arg->ser codon position 3 amino acid position 1296

Exon 38 missense change mRNA position 4288 rs1062020 heterozygosity rate unknown T->X codon position 1 amino acid 1340

3.8.3 Conservation through species

There are no paralogues. There are 50 orthologues documented in Ensembl including:-

Anopheles	27%
Bos Taurus	91%
Caenorhabditis elegans	18%
Canis familiaris	91%
Danio rerio	55%
Drosophila melanogester	25%
Felis catus	64%
Gallus gallus	66%
Macaca mulatta	90%

Mus musculus	84%
Pan troglodytes	98%
Rattus norvegicus	83%
Xenopus tropicalis	60%

3.8.4 Tetratricorepeat domains (TPR)

Within the literature tetratricopeptide repeat domains are abbreviated to TTC, TCC or TPR domains. Throughout this thesis I have used the abbreviation TPR.

TPR typically consists of 34 amino acids. The consensus of the sequence is:-

[WLF]-X(2)-[LIM]-[GAS]-X(2)-[YLF]-X(8)-[ASE]-X(3)-[FYL]-X(2)-[ASL]-X(4)-[PKE]

The sequence is found in many organisms including yeast, bacteria, cyanobacteria, fungi, plants and in humans in various subcellular locations.

TPR’s are involved in protein: protein interactions the exact method of which has not been identified. The domains may be involved in chaperone, cell cycle transcription and protein transport complexes.

5-6 repeats of the TPR sequence generate a right hand helical structure with an amphipathic channel that accommodates an alpha-helix of a target protein. The target protein may be WD-40 repeat proteins or other tetratricopeptide repeats to form multiprotein complexes.

It can be repeated many times and the repeats may be separated by non motif sequence.

There are 8 conserved residues in each TPR protein.

-W-L-G-Y-A-F-A-P-

The rest is very diverse and corresponds to the diverse function of TPR containing proteins.

The structure is likely to form 2 α helical domains with –W-L-G-Y- forming a hydrophobic gap into which a phenylalanine side chain fits -A-F-A-. This provides a mechanism for protein to protein interactions (TPR-TPR interactions).

In individual TPR’s there is a greater degree of homology.

Examples of TPR containing proteins include:-

Cdc16p, Cdc23p and Cdc27p which are responsible for cell cycle progression through mitosis in Saccharomyces cerevisiae.

Cyclosome/ APC responsible for targeting mitotic regulators

Pex5p/Pas10p involved in targeting sequences of peroxisomes

Tom70p involved in mitochondrial biogenesis

3.8.5 Disease involving tetratricopeptide repeat (TPR) domain containing genes

Fanconi’s anaemia

This disorder is characterised by developmental anomalies, progressive pancytopenia and a predisposition to tumour formation. At least 8 genes are involved in DNA protection in this condition with mutations in any resulting in the phenotype. *FANCG* consists of 7 TPR domains and is crucial in binding to *FANCA* and is hypothesised to stabilise the FA protein core complex. Mutations at the specific position 8 in 4 of the TRP domains prevent the binding to *FANCA* (Blom *et al*, 2004).

Zellwegers syndrome

This peroxisomal biogenesis disorder is inherited in an autosomal recessive fashion and is caused by mutations in one of at least 12 known *PEX* genes encoding peroxins. Peroxisomal matrix proteins have a carboxyl terminal peroxisomal targeting signal which has a conserved amino acid sequence which bind to the *PEX5* receptor via a TPR domain to import protein into the peroxisome (McCollum et al, 1993).

Bardet Beidl syndrome (BBS)

Many genes have been attributed to cause the clinical phenotype of BBS. *BBS4* has a TPR domain. BBS is a syndrome caused by abnormalities in cell cilia and is therefore termed a ciliopathy (Ansley *et al*, 2003).

Chronic granulomatous disease

$p67^{phox}$ is an essential component of the NADPH oxidase multiprotein complex involved in producing superoxide ions in response to microbes. Abnormalities in the formation of these complexes results in severe bacterial and fungal infections due to the inability of phagocytes to respond to microbes. The complex formation is mediated by insertion between two TPR domains (Grizot *et al*, 2001).

Polycystic ovary syndrome

Unbound androgen receptors are composed of heat shock proteins, co chaperones and TPR containing proteins and assists in ensuring the androgen receptors are functioning normally (Goodarzi *et al,* 2008).

Nephronophthisis

This cause of renal failure is an oligogenic inherited condition affecting cilia. To date 6 different genes have been identified with *NPHP3* having a TPR domain and is being expressed in the central node of the developing mouse embryo (Fliegauf and Omran, 2006).

Williams Syndrome

This is a developmental abnormality resulting from haploinsufficiency of 7p11.23. Different genes within this region are responsible for the clinical features if William's syndrome. The gene *FKBP6* consists of 3 TPR domains and is deficient in all cases of William's syndrome and therefore may be responsible for the common features of short stature and hypercalcaemia. The specific role of the TPR domains is unknown but is predicted to be involved in hormone binding (Meng *et al*, 1998).

3.8.6 Chapter summary

*TTC*37 has recently been described to contain TPR domains. The finding of TPR suggests the protein may be involved in protein-protein interaction or in the formation of multiprotein complexes.

TPR domains are found throughout the genome and have a wide range of roles with no specific pathways involved.

The role of *TTC*37 in the pathogenesis of PDI can not be predicted from the *in silico* analysis of this specific gene or from the domains within the gene.

3.9 Further investigation of *TTC*37

The tetratricopeptide repeat domains of the TTC37 protein are predicted to be involved in protein-protein interactions and in the formation of multiprotein complexes. There is no further knowledge known about the interactions of *TTC37* within intracellular pathways to enable a hypothesis to be formed as to the molecular pathogenesis of PDI.

With no indication as to the localisation of *TTC37* the main tissues affected in PDI, intestine and liver, were investigated by immunohistochemistry to identify any potential disruption to protein localisation as a result of *TTC37* mutations.

This hypothesis was investigated by collaboration with experts within the field and the work undertaken in their laboratory by their staff.

Dr Alex Knisely, Kings College Hospital, London, UK, looked at immunohistochemistry of the liver and Dr Mark Donowitz, Johns Hopkins University, Baltimore, Maryland, USA, investigated protein localisation within the small intestine.

This is a summary of the methods used and the preliminary findings.

3.9.1 Immunohistochemistry method for jejunal specimens

4µm sections of archival material originally obtained for clinical diagnosis were deparaffinized and heat fixed. Slides were microwaved for antigen recovery in 10mM sodium citrate buffer, pH6 (Sigma Chemical Company, St Louise, MO) at power level setting 9 (Panasonic Model NN-C980B Conventional Microwave Oven, Secaucus, NJ), for 2-5minutes. After cooling for 30 minutes, sections were washed in phosphate buffered saline (PBS) and blocked with 5% normal goat serum in phosphate buffered saline –

tween (PBS-T). Sections were incubated with rabbit polyclonal antibodies against NHE2 (Ab597) and NHE3 (Ab1381) as well as with a monoclonal antibody against villin, as previously descibed. After 2 washes with PBS, sections were incubated with AlexaFluor 488 anti-rabbit and AlexaFluor 568 anti-mouse secondary antibodies (Invitrogen Ltd, Paisley, UK). Sections were washed, counterstained with Hoechst 33342 (Invitrogen Ltd, Paisley, UK) for nuclei, and mounted with glass cover slips mounting medium. Immunofluorescence images were obtained using a 63x water objective and a LSM 510 Meta confocal microscope. Similar settings for laser power, gain and resolution were used for each sample analysed. Jejunal samples from four patients and 20 aged matched controls (archival material initially taken for the clinical consideration of coeliac disease and found to be normal on histological examination) were analysed by a single investigator blinded to the clinical diagnosis.

3.9.2 Immunohistochemistry method for liver specimens

Samples of liver, sectioned at 4-5µm and picked up on glass slides, were immunostained and counterstained as described previously for the canalicular transport proteins bile salt export pump (BSEP) and multidrug resistance-associated protein 2 (MRP2).

The hepatic proteins BSEP, MDR2, MDR3 and GGT were examined. All immunoflourescence showed normal expression of these proteins and normal localisation within the liver. BSEP, MDR3 and GGT are apical proteins whilst MDR2 is a basolateral protein.

3.9.3 Protein expression results

In patients 3C, 4C, 6C and 13C, the apical cytoskeleton marker, villin was expressed normally at the brush border of enterocytes. However, expression of the apical transporter proteins NHE2 and NHE3 was reduced and was not localised to the enterocyte brush border but remained in the subapical region in patients 3C and 6C (Figure 3.9.1). No abnormality of NHE2 or NHE3 staining was detected in 20 age matched controls, p=0.022).

There was no abnormality detected in expression and localisation of the bile salt transporters BSEP and MRP2 in liver biopsies.

Patient 13C was found to have the greatest reduction in NHE2 and NHE3 expression and yet mutations in TTC37 were not identified despite a similar although not identical phenotype. It could be hypothesised that this patient may have

- *mutations in a different gene within the same pathway as that involving TTC37*
- *the mutations are within the unsequenced non coding intron*
- *There is a sequence change within a promoter region for TTC37*

Figure 3.9.1 Immunofluorescence of villin, NHE2 and NHE3 expression in jejunal samples of patient 3C compared with age matched control samples

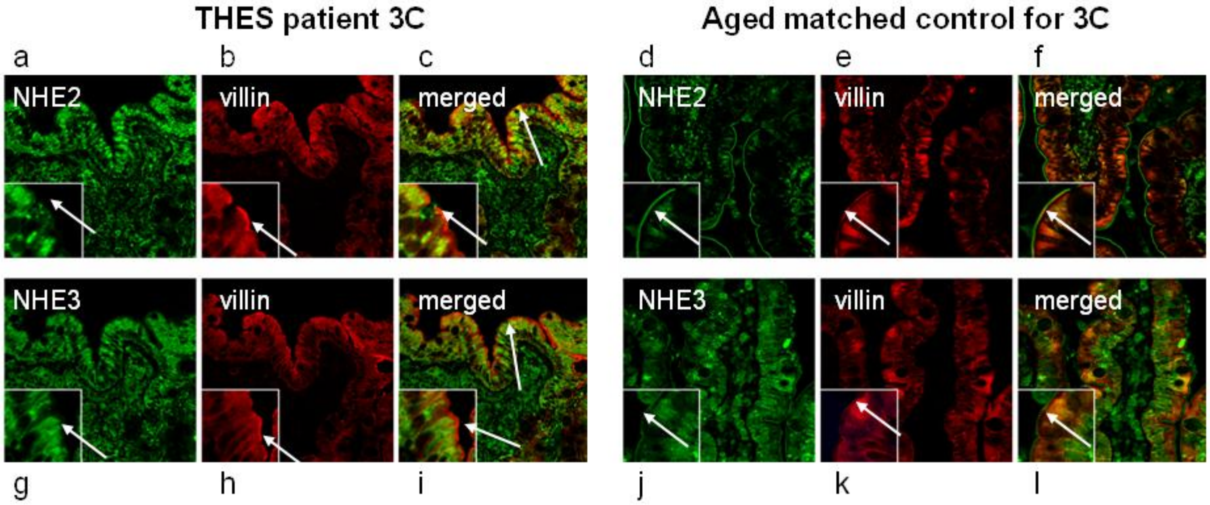


Figure 3.9.1

Immunofluorescence photomicrographs (confocal microscopy); villin, NHE2 and NHE3 expression in control material and THES patient 3C. The arrows point to the brush borders. 6a and 6g show NHE2 and NHE3 expression (in green) respectively at the apical border of the enterocytes is absent as compared to the control 6d and 6j in which the brush border is clearly seen to express NHE2 and NHE3. In both the THES patient and control samples villin which fluoresces red is seen indicating in both samples the brush border is present. With merging the images, the green dye of NHE2 and NHE3 is not seen at the brush border but remains in the subapical region in the THES patient (6c and 6i) whilst villin in red continues to be seen. In the control sample 6f and 6l, with image merging, villin and NHE2 and NHE3 fluoresce together indicating co localisation at the brush border.

3.9.4 NHE2 and NHE3

The sodium bicarbonate transporter 2. The greatest expression is within the gastrointestinal tract mainly the jejunum, ileum and colon (Dudeja *et al*, 1996). NHE2 is also expressed in the skeletal muscle, kidneys, brain, uterus, testis and heart (Bookstein *et al*, 1997). The location of NHE2 is within the plasma membrane or the apical membrane and it is not trafficked within the cell. NHE2 is involved in Na^+:H^+ transport in a 1 to 1 ratio (exchanging one intracellular proton with one extracellular sodium) and therefore maintaining a neutral ionic intracellular balance (Gawenis *et al*, 2001). NHE2 interacts with CHP1 and CHP2.The NHE2 knockout mouse has very few symptoms in early life but develops increasing gastritis.

NHE3 is found in the kidneys and the intestinal tract as well as other epithelia. It localises to the apical boarder of epithelial cells and it may be involved in trafficking in early recycling of endosomes (D'souza *et al*, 1998). NHE3 exchanges one extracellular Na^+ with one intracellular proton to maintain a neutral balance. It is also involved in HCO^{3-} transport, the acidification of early endosomes and the renal absorption of albumin.

In the duodenum NHE3 can be blocked by amiloride which increases the secretion of HCO^- through the cystic fibrosis transmembrane regulator (CFTR). NHE3 is also found in the cholangiocytes and hepatocytes (Kiela *at al*, 2009).

The NHE3 knockout mouse has the phenotype of congenital sodium diarrhoea.

<u>Further investigations</u>

The identification of NHE2/3 disruption in *TTC37* mutations requires further investigation. Both proteins are found on the apical membrane. The effect on the basolateral membrane by looking at the Na/HCO^{3-} transporter proteins will be important in identifying if the effect on NHE2/3 is a generalised trafficking disorder.

3.9.5 Summary of investigations

Mutations in *TTC37* result in an abnormal expression of intestinal protein expression in some but not all affected patients. The expression of hepatic proteins was not affected. Although the abnormality in intestinal proteins could explain the diarrhoeal symptoms in some cases the abnormality was not seen in all probands suggesting the affected pathway may not be directly influencing NHE3 / NHE2 expression.

It is also unclear how the multisystem clinical features could be explained by NHE2 / NHE3 and it will be important to look at the effect of other congenital diarrhoeas on NHE2 / NHE3 expression to ascertain if this is primarily due to the effects of *TTC37* mutations or secondary to intestinal epithelial damage.

It may be that *TTC37* is involved in a ubiquitous pathway which influences the expression of many different proteins.

Further investigation of the *TTC37* protein will aid the elucidation of the pathways involved.

3.9.6 Future direction

The identification of *TTC37* as having mutations associated with the disease phenotype PDI is the initial step in understanding the molecular pathogenesis of this multisystem disease.

The first stage however is to confirm that mutations in *TTC37* are the cause of the observed phenotype in humans.

This can be achieved by using siRNA to knockdown the gene in animal models. The best model to use for this is probably mice. With this model the effects on the hair can be observed. Zebrafish are unlikely to be a good model due to the lack of body hair and a number of the pathological mutations identified in this cohort of humans were not in a conserved in the zebrafish.

It has to be remembered however that there may be species specificity meaning that mutations in humans may not have the same effect in other animals.

If a disease phenotype is developed from siRNA it will then be important to ensure that the phenotype can be rescued by then providing *TTC37* RNA.

The development of an antibody to the protein product of *TTC37* will enable the protein expression to be evaluated in animal models as well as in the intestinal and liver specimens from the human cohort.

The observed abnormality in alpha granules of platelets requires more investigation as this is a rare finding and little is known about alpha granule pathology.

As mutations in *TTC37* were not responsible for all the PDI phenotypes in this cohort of patients studied it can be hypothesised that the *TTC37* protein is likely to interact with other proteins within the cell. It may be that mutations in one of the interacting proteins will also harbour pathogenic mutations leading to the same phenotype.

Understanding the interactions of the *TTC37* protein with other proteins will be essential in unravelling the intracellular pathway which has led to this multisystem disease.

Chapter 4

Ciliopathies and the liver

Contents

4.1 Introduction to primary cilia

Cilia are antenna like projects which extend out from many differentiated cells and are evolutionary conserved organelles.

Cilia were first reported in 1898 by Zimmerman using a light microscope, but little more could be delineated until the advent of the electron microscope in the 1950's and the cutting of ultra thin sections. Using these technologies Barnes in 1961 first described in detail the ultrastructure of the cilia of the adenohypophyseal cell. In the past decade there has been increasing interest in primary cilia with the discovery that pathology of the components of primary cilia lead to the development of a wide spectrum of human disease which are termed ciliopathies. Cilia are classified into whether they are motile or are immotile/ primary cilia.

The development of cilia from ancestral eukaryotic cells can be explained by two hypotheses

1. symbiotic theory: cilia were originally formed from the fusion of a motile Spirochete with a host cell
2. autogenous theory: cilia have developed from pre-existing components of the cell

Homology between eukaryocytes and the lack of homology with prokaryocytes suggests that the autogenous theory is more likely (Hartman & Smith, 2009).

4.1.1 Structure of primary cilia

Cilia are microtubular structures which extend from the cell membrane of non proliferating cells which stem from a basal body within the cell. During cell division under tight control cilia are reabsorbed and therefore they can influence cell division.

A schematic drawing of the structure of primary cilia are shown in figure 4.1.1. The microtubules which originate from the mother centriole of centrosomes are made of nine microtubule doublets. In general motile cilia also have a central pair of microtubules which is lacking in primary cilia however this is not exclusive and some primary cilia have two central microtubules and some motile cilia do not have any.

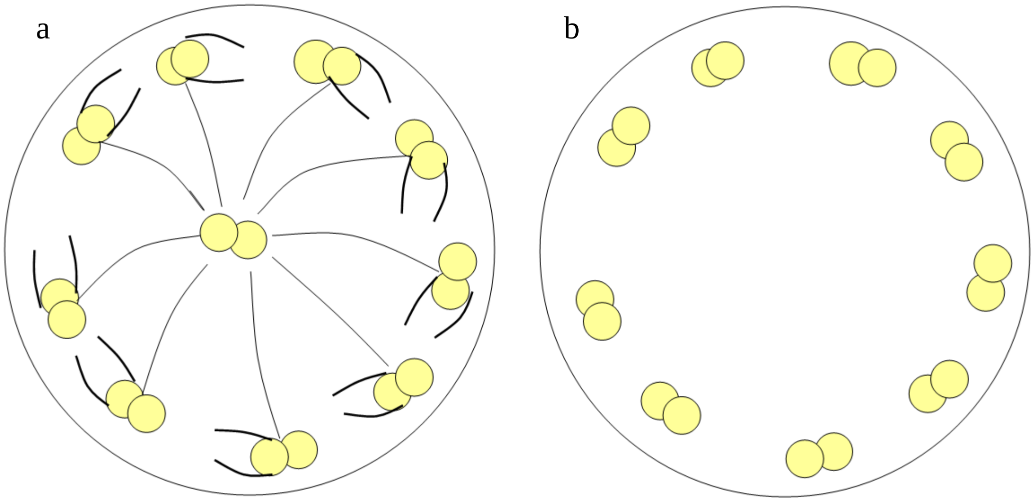


Figure 4.1.1 is a schematic diagram of a cross sectional view of a cilium.

a) cross section of a motile cilium with two central microtubules. This cilium also contains spokes of dynein and dynein project from the 9 surrounding microtubules.

b) Is a primary cilium with nine peripheral microtubules and no central microtubules (Cardenas-Rodriguez and Badano, 2009).

Protein synthesis is unable to take place in the cilia and therefore proteins which are made in the cell body are transported into (anterograde) and out (retrograde) of the cilia by an intraflagellar transport system (IFT) to provide building blocks for the continuously assembled and disassembled of the cilial tip. Figure 4.1.2 shows a schematic diagram of the cilium and the IFT transport systems. IFT is essential for cilia function and disruption of any of the components can lead to the development of a disease phenotype. The anterograde transport motor is kinesin-II (composed of 2 subunits KIF3A and KIF3B) and the retrograde motor is dynein 2 (Baldari and Rosenbaum, 2009).

.

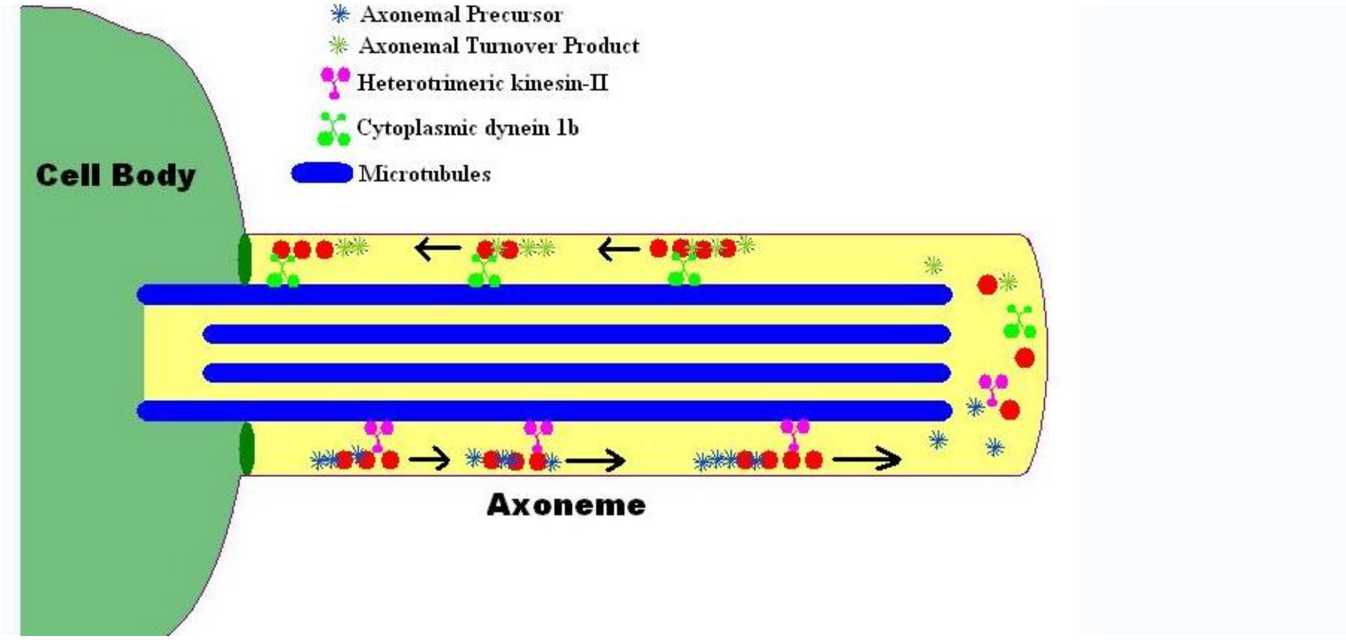


Figure 4.1.2 shows a schematic drawing of a cilium projecting from the cell body. The antegrade flow of particles (or intraflagellar transport) is modulated by kinesin II

and carries the precursors from which to make the tip of the cilia. The retrograde flow (back to the cell body) is mediated by Dynein 1b and carries the waste products away from the tip.

4.1.2 Intraflagellar transport (IFT)

IFT are protein complexes which move bidirectionally along the axoneme coordinated by IFT motors dynein 1b and kinesin II. Complex A is involved in retrograde transport and IFT complex B is involved in antegrade flow.

The rate of movement along the axoneme is approximately 1 micron per second.

Complex A consists of 6 different proteins or subunits and deficiency of complex A in *Chlamydomonas reinhardtii* shortened flagellar filled with IFT particles.

Complex B consists of 12 subunits with deficiency leading to shortened or absent cilia.

IFT particles are predicted to bind proteins, as sequence analysis has revealed specific motifs which are predicted to form protein-protein binding domains.

These include:-

tetratricopeptide repeats – involved in transient protein-protein interactions
WD-40 repeats – involved in transient protein-protein interactions
Coiled-coiled domains – involved in stable protein-protein interactions
WAA repeats – which are specific to IFT

The stable interactions may form the permanent complex structure whilst the tetratricopeptide repeats and WD-40 repeats may be involved in the trafficking of protein particles to and from the flagellar tip.

Complex B is formed of subunits (shown schematically in figure 4.1.3), of which 7 subunits form the core of complex B. The core is composed of IFT81 which complexes with IFT72 and 74 and is stable, whilst IFT46, IFT52, IFT88 and IFT27 weakly interact.

The other components of complex B are IFT57, 20, 80 and 172 which can dissociate easily with increases in ionic strength which may facilitate the dynamic movement of IFT particles.

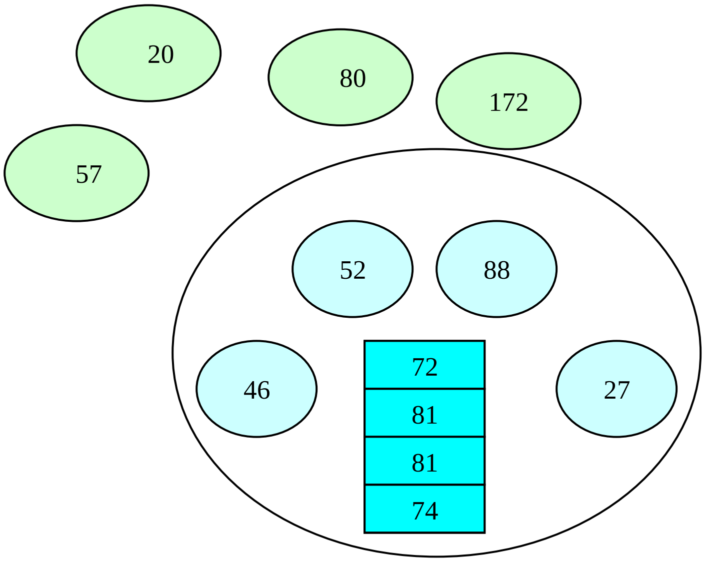


Figure 4.1.3. Diagram of the components of Complex B

IFT 81, 72 and 74 form a stable core. IFT 46, 52, 88 and 27 also form part of the core but can dissociate with increasing ionic strength. IFT 57, 20, 80 and 172 are part of complex B but dissociate easily. (Pedersen and Rosenbaum, 2008).

4.1.3 Function of cilia

Primary cilia are involved in sensing chemical, mechanical such as luminal flow and osmotic stimuli. They are involved in signal transduction pathways including hedgehog signalling, Wnt, PDGFER-α and integrin signalling.

Primary cilia regulate important cellular functions such as proliferation and maintenance of planar cell polarity and mitotic spindles orientation so ensuring normal epithelial function and normal diameter of tubular structures.

(Nigg and Raff, 2009).

4.1.4 Diseases of motile cilia

Disease of motile cilia affect the respiratory system where the lack of mucus clearance causes bronchiectasis in a condition named primary cilial dyskinesia (PCD, OMIM 244400) which is also associated with hydrocephalus, sinusitis and immotile sperm. When it is associated with situs inversus it is known as Kartageners syndrome. Left-right axis is determined by motile cilia at the embryonic node which influences the flow of the extra embryonic fluid (Pennarun *et al*, 1999).

4.1.5 Diseases of primary cilia

Diseases which are secondary to defects in primary cilia are known as ciliopathies. The first diseases to be ascribed to primary cilia dysfunction were the autosomal dominant and recessive polycystic kidney diseases. Since then other conditions have come to light to be due to abnormalities in cilia. They share common clinical features and therefore a

disease in which the pathogenesis is not know can be predicted to be a ciliopathy if there are renal cysts, liver fibrosis, polydactyly and abnormal brain development.
JATD has many of these findings but the key clinical feature of abnormal skeletal development was not thought to be a component of cilial disease. Section 4.2 describes the identification of the first gene for JATD and led to it being classified as a ciliopathy.

4.2 Jeune's asphyxiating thoracic dystrophy clinical features

JATD was first described in 1954 by Jeune *et al* and is a multisystem autosomal recessive disorder (ATD, MIM 208500). It has a variable presentation and is classified in lethal, severe, mild and latent forms. Most patients die in the neonatal period from the asphyxia resulting from hypoplastic lung development. In those who survive the neonatal period (approximately 1 in 5) then the extraskeletal manifestations become more prominent.

The abnormal skeletal development (which is illustrated by the X-rays in figure 4.2.1) is characterised by:-

- long, narrow "bell shaped" thorax
- short abnormal ribs
- abnormalities of the pelvis are diagnostic and include
    - small ilia
    - trident acetabulum
    - medial and lateral bony projections from the acetabulum
- metaphyseal irregularities
- short long bones (involving predominately ulnae, radii, fibulae and tibiae)

- handlebar shaped clavicles
- cone shaped epiphyses of the hands
- polydactyly of hands and feet and short phalanges

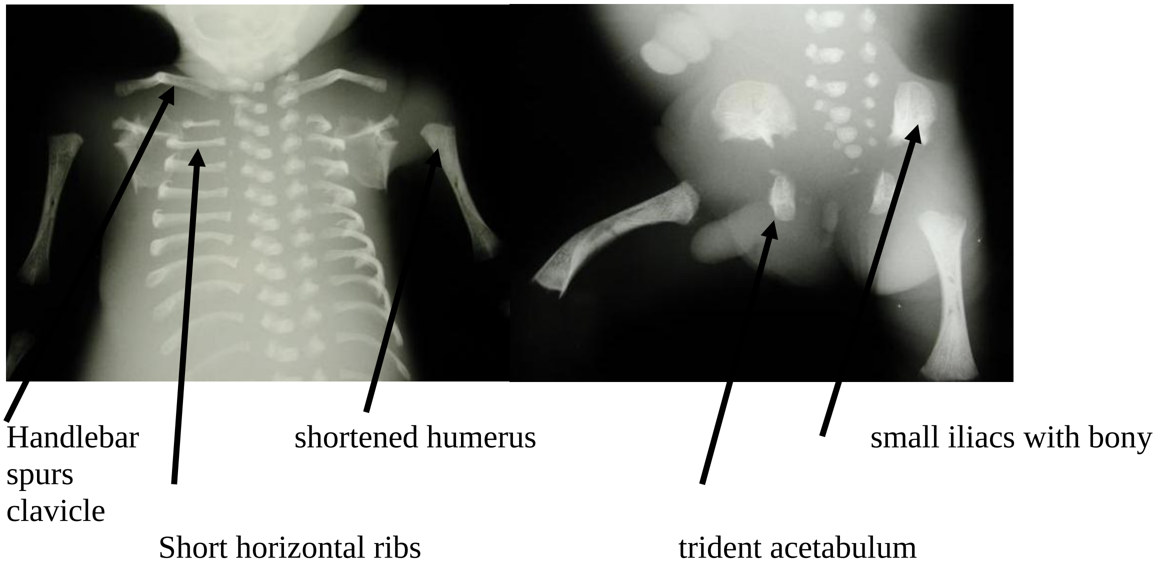


Figure 4.2.1 The skeletal features of JATD. The x-rays are of proband C1 with the skeletal manifestations marked by arrows. The X-ray was taken in the initial clinical investigation of the child and reviewed by Dr C Hall, radiologist, to confirm the skeletal features of JATD. Parents of C1 consented for the use of the x-ray for research.

The extraskeletal manifestations are common but more varied and include

- Renal cyst formation, periglomerular fibrosis and interstitial nephritis may lead to chronic renal failure
- Prolonged neonatal cholestasis can be a presenting feature if the skeletal manifestations are mild. Liver fibrosis and cirrhosis secondary to ductal plate malformation may be severe requiring liver transplantation
- Pancreatic cysts development
- Retinal dystrophy (which is usually not present at birth but develops later in life)

4.2.1 JATD study cohort

During the course of this study 32 probands with a clinical diagnosis of JATD were studied. The x-rays were assessed by an independent radiologist with expertise in skeletal dysplasias. All post mortems were carried out by paediatric and neonatal pathologists. Of these 16 were from consanguineous union and the others were sporadic cases in non consanguineous families. Table 4.2.1 provide a comprehensive list of the clinical features of those probands from consanguineous union. Table 4.2.2 provides a list of clinical features of sporadic cases of JATD. All phenotyping of the patients had taken place prior to my commencing working on JATD

	C17	C18	C19	C20	C21	C22	C23	C24	C25	C26	C27	C28	C29	C30	C31	C32
gestational age (weeks)				19/40	19/40									23/40		
birth weight (grams)																
consanguinity	no	no	no	no	no	no	no	no	no	no	No	no	no	no	no	no
mechanical ventilation		no	no	na	na						No					
skeletal																
short thorax		+	+	+	+						+			+		
bell shaped thorax																
horizontal ribs		+	+								+					
iliac wings																
trident acetabulum		+									+					
iliac spurs		+														
short limbs				+	+						+			+		
short phalanges		+		+	+						+					
polydactyly																
increased bone maturation																
Liver																
neonatal cholestasis											+					
portal hypertension			+													
bile duct paucity			+								+					
renal																
cystic dysplasia											+					
nephronopthisis		+														
interstitial nephritis																
Pancreas																
cysts																
Cranial																
dilated ventricles																
dochocephaly																

Table 4.2.2 The clinical features of JATD who are from non consanguineous families

The most common clinical features are skeletal and only 6 of the 32 (18%) have renal involvement and may indicate those patients who have survived long enough to develop renal disease. Hepatic manifestations are as common as renal disease with 7 of the cohort having abnormalities of the liver.

4.2.2 Summary of the clinical features of the JATD study cohort

The clinical features identified in the study cohort are in keeping with those previously described in the literature except for the dilated ventricles and dochocephaly in two patients. The extraskeletal manifestations are rare and highly variable which may indicate a heterogenous group with variable aetiologies as to the underlying pathogenesis of JATD.

4.3 Identification of IFT80 as a causative gene for JATD

4.3.1 Background molecular genetic investigation of JATD

Prior to my undertaking of investigation into JATD a region of interest had been identified by Dr Colin Johnson. Using the technique of autozygosity mapping a 10K SNP GWS was undertaken in 9 probands from consanguineous union which identified a region of interest on chromosome 9. The region was confirmed using microsatellite markers which also identified the boundaries of the region to 126,409,079-128,607,022. Other centres were also working on JATD and collaborations had been formed to share information on different regions of interest and share DNA from affected probands.

Our lab focused on chromosome 9 and 15 (chromosome 15 region had been excluded following sequencing of genes within the region by Dr Neil Morgan)

Dr P Beale, University College London (UCL) Institute of Child Health focused on chromosome 3.

Dr T Attié-Bitach, Hôpital Necker-Enfants Malades, Université Paris Descartes focused on chromosome 12.

4.3.2 Candidate gene selection from region on Chromosome 9

I commenced investigating JATD by interrogating the chromosome 9 region for candidate genes. Within the previously delineated region using the NCBI public database there are 42 genes. Two genes had homeoboxes within the sequence which are associated with skeletal development and therefore were good candidate genes. Another 15 genes had no known function attributed to them and therefore required exclusion. The other genes within the region were either pseudogenes and therefore not sequenced or the known gene function did not correspond to the clinical phenotypic findings of JATD.

To summaries genes were sequenced in the priority

- Putative function
  - Known function likely candidate
  - Known function unlikely candidate
  - Function unknown and therefore can not be excluded
- Pseudogenes (not sequenced)

Two genes were good candidate genes

LMX1B

LIM homeobox transcription factor 1 beta is implicated in renal disease through transcription of podocytes and skeletal anomalies including nail patella syndrome. The sequence of *LMX1B* contains a homeobox which is also a basis for it being a candidate gene.

Homeoboxes are sequences of approximately 180 bases which encode a protein which binds to DNA and regulates patterns of development especially limbs and therefore may be important in the formation of polydactyly seen in JATD.

DNM1

DNM1 also contains a homeobox and therefore has a potential to disrupt limb formation.

Genes with unknown function

Another 6 genes were also investigated due to the function being unknown and therefore could not be excluded.

WDR34, C9orf74, C9orf16, Nihan like protein, *ZNF79* and *ZNF297*

4.3.3 Results from sequencing candidate genes

DNA from two probands was used to sequence candidate genes to identify mutations.

DNA from C3 and C5 was used.

Sequencing was carried out as specified in chapter 2

No mutations were identified in the candidate genes LMX1B, DNM1 or the genes with no known function.

4.3.4 IFT80 mutations cause JATD

Collaborators (P. Beales and P Scambler) identified mutations in *IFT80* in 3 affected families including a proband from Birmingham. The proband C1 was from a multiple consanguineous family from Pakistan and found to have a deletion of a single amino acid at position 549.

In work for this Ph.D. I confirmed the mutation in the Birmingham family (shown in figure 4.3.1) and demonstrated appropriate segregation within the family. An affected sibling also had the mutation. I sequenced 200 chromosomes from Pakistani Asian controls which were all wild type.

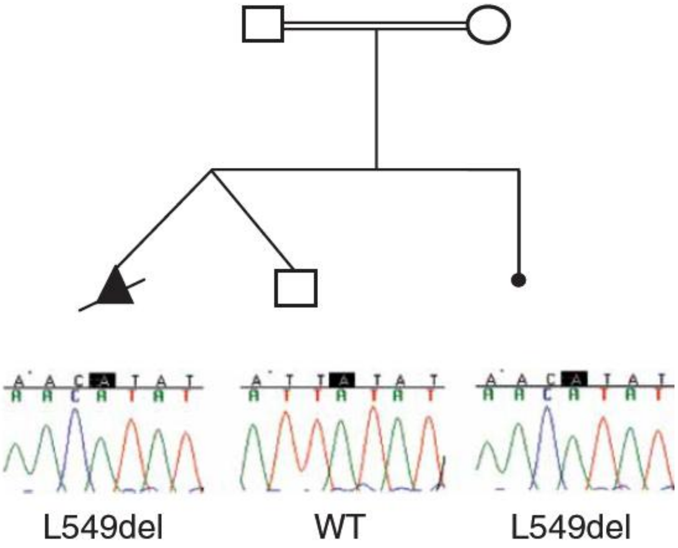


Figure 4.3.1 Electropherograms of mutations in *IFT80* in proband C1 and family.

Following identification of *IFT80* as a causative gene for JATD all the cohort of non consanguineous probands were sequenced and no other affected patient had any mutation identified. The consanguineous families underwent linkage studies and only those who showed linkage underwent sequencing.

No other affected patients had mutations in *IFT80* identified.

IFT80 encodes a 777 amino acid seven WD40 domain protein component of intraflagellar transport (IFT) complex B (Uniprot Q9P2H3). IFT is essential for the development and maintenance of motile and sensory cilia. Interestingly, the *C. Elegans* orthologue of *IFT80* shows a weak genetic interaction with bbs-8, a protein involved in IFT. Human mutations in *BBS8* cause the Bardet-Biedl syndrome, another ciliopathy which shares features of retinal degeneration, cystic kidney disease and polydactyly.

The gene was shown by collaborators to affect the structure and function of primary cilia. The identification of *IFT80* as the first causative gene for JATD also characterised JATD as a ciliopathy.

4.3.5 Discussion of results

IFT80 is the first identified gene in which mutations are associated with the phenotype of JATD. It was identified using autozygosity mapping techniques however the clinical features are variable and the cohort not homogenous therefore only one of this study cohort and 4% of all affected children have mutations in *IFT80*. This illustrates the limitations of autozygosity mapping as a tool to identify rare genes.

The small number of JATD patients affected by *IFT80* mutations suggests there are many other genes which may be implicated in the pathogenesis of JATD. The discovery of *IFT80* has been critical in understanding the underlying pathogenesis of JATD and classifying it as a ciliopathy. This will allow the search for further genes to target those involved in cilial formation and function.

4.4 Investigating for other genes resulting in the JATD phenotype

With only family C1 with JATD have mutations in *IFT80* from a cohort of 32 affected children equating to only approximately 3% of cases have mutations in *IFT80*. JATD therefore is likely to be secondary to many different gene mutations which also reflects the heterogeneity of clinical features and is in keeping with other ciliopathies such as Bardet Biedl Syndrome in which at least 12 genes have been described.

A number of processes have been used to investigate for other genes which may lead to the clinical phenotype of JATD in this study cohort

- 250K SNP GWS in affected probands from consanguineous families to identify regions of shared homozygosity which were not identifiable from the 10K SNP GWS
- Group probands into ethnicities to ascertain shared regions of homozygosity
    - From these regions to then identify from public databases any genes which have a known or putative function within the cilium

- To focus on genes which are involved in retrograde IFT transport (complex B) and identify if any of the probands are homozygous at the position of these genes in the GWS

4.4.1 Homozygous regions identified by 250K SNP GWS

All consanguineous probands underwent a 250K SNP GWS. From this any area of homozygosity of 3cM or greater was taken to be significant. All individual probands had multiple extensive regions of homozygosity examples of which are given in table 4.4.1.

For example:-

Table 4.4.1 shows the regions of homozygosity for proband C6 which are greater than 1 Mb in length.

Chromosome	Start	End	Mb
3	210748	2808324	2.5
3	15506753	116758491	1.25
3	117248975	118286357	1.03
5	66239371	133948902	67.7
5	160382763	161469708	1.08
7	12456780	130236163	5.6
9	36431	3998954	3.9
11	72787283	73815199	1.02
14	57091406	73649222	16.5
14	79682157	88340677	8.6
16	31624102	46841950	15.2
20	4161663	57915771	1.6
20	51108814	58696025	7.6
22	29863796	31205871	1.3

There were no areas of extensive homozygosity which was shared between all the affected probands even when analysed for specific ethnic grouping.

For example:-

Two probands are Lebanese siblings (C13 and C14) and are from a consanguineous family. During the course of this project another two siblings originating from Lebanon had DNA available (labelled CM). 250K SNP GWS were carried out on these siblings to identify regions of shared homozygosity. The areas of homozygosity are shown in table 4.4.2. Lebanon is a very mixed ethnic country and it may be there is no link between these families therefore each sibling pair was analysed independently.

Table 4.4.2 This table shows the regions of homozygosity shared by each Lebanese sibling pair that are greater in size than 1Mb.

None of the regions were shared by both sibling pairs which suggest the families are not related.

Sibling pair	**Chromosome**	**Start**	**Finish**	**Length**
CM	3	131838928	133512910	1.67
C13/C14	4	163938136	167600952	3.66
C13/C14	7	155726693	158819393	3.0
C13/C14	10	54060307	58993694	4.3
CM	12	6216045	7854479	1.6
C13/C14	12	100045651	108761297	8.7
C13/C14	13	38049139	49974381	11.9
C13/C14	18	61364137	64047182	2.68
C13/C14	21	39395769	46351620	6.9
CM	22	26370974	27789634	1.4

None of the regions were shared by both sibling pairs which suggest the families are not related.

4.4.2 Cilial genes identified within the regions of homozygosity

Interrogation of the NCBI database identified genes within the regions of homozygosity which were identified for each individual proband which are associated with cilial formation or function:-

TMEM2C – within a region of 1.78Mb of homozygosity for C7 and C8 who are cousins originating from Bangladesh

NEK2 – in a homozygous region for C7 (1.6Mb) and also the siblings CM (1.67Mb)

ODF2 – within a homozygous region of 16.5Mb for C6. The protein product of *ODF2* is cenexin.

KIAA0586 – this was a good candidate gene in a region of homozygosity for C6. The chicken ortholog has polydactyly and it is a known cilial gene.

None of the genes formed part of the retrograde transport system complex B.

4.4.3 Genes which form intraflagellar transport complex B

Ten genes are known to make up intraflagellar transport complex B which is involved in retrograde flow of particles. Table 4.4.3 shows the IFT complex B gene positions and the corresponding areas of homozygosity in probands with JATD.

Gene	**Chromosome location**	**Start**	**Finish**	**Regions of homozygosity**
IFT172	2	27667240	27712571	No
IFT88	13	21141208	21265576	No
IFT81	12	108530506	114544839	CM 8.7Mb region
IFT80	3	161459482	161600014	C1 mutations
IFT74/72	9	26947037	27062931	C8 16.5 Mb but no homozygosity in cousin C7
IFT57/55	3	107879659	107941417	C7 34Mb but no homozygosity in C8
IFT46	11	118415258	118436750	C7 24.4 but no homozygosity in C8
IFT20	17	26655353	26662495	No
IFT52	20	41652993	41709276	No areas greater than 1Mb but 6 probands had small areas of homozygosity C13, C8, C3, C11, C4, C15

Table 4.4.3 This table shows the complex B genes and the corresponding areas of homozygosity identified from the 250K SNPs.

Microsatellite markers

Fine mapping of five of these genes using microsatellite markers was used to corroborate or refute the GWS linkage. The markers used were placed as close to the known gene position as possible with a marker above and below the gene. Using this technique the gene could be excluded if no linkage was ascertained or individuals who were not homozygous at this region could be excluded. Table 4.4.4 and 4.4.5 shows the results of the microsatellite markers for the genes of interest.

Table 4.4.4 Shows the results of microsatellite markers in the regions of the complex B genes in those probands who were homozygous in these regions. The marker name is in the first column and the physical position. The identification of the probands are given in blue and in the top row F, M and S denote the mother, father and siblings of the proband. Those boxes in yellow highlight the homozygous markers.

Marker	physical location	M	F	C10	C3	M	F	S	C11	M	F	C2	M	F	C13	M	F	C9	M	C6	M	F	C5	M	C4	M	F	C13	C14
TMEM2L	**78858767-79031054**																												
D15S1510	78607698	243 252	249 252	249 252		258 252	258 252	258 252	258 252		258 261	263 266	244 252	248 248	248 252	242 248	248 252	248 252	252 252	244 252	269 269	267 270	266 270	251 251	251 251	244 248	244 248	248 248	244 244
D15S206	79991342	267 269	265 267	267 267		269 270	268 269	269 270	269 270				263 266	266 270	266 269	257 268	271 271	257 273	263 277	263 266				267 267	267 267	263 272	260 268	264 270	264 270
IFT52	**41652993-41709276**																												
D20S1121	41612806	126 126	126 126	126 126	127 127	126 126	126 126	126 126	126 126	125 125		125 125	125 125	125 125	125 125	126 126	126 126	126 126	125 125	125 125	125 125	125 125	125 125	125 125	125 125	125 125	125 125	125 125	125 125
D20S150	41929128	246 253	249 253	253 253		245 249	245 245	245 249	245 249	245 249		245 253	245 253	249 249	249 253	245 253	249 253	245 249	245 253	245 253	245 253	245 249	245 245	245 245	245 245	245 245	245 245	245 245	245 245
NEK11	**132231147 - 132551993**																												
D3S1587	132218515	218 223	217 220	217 233	216	218 223	216 223	216 221	216 218		217 217	217 217	217 223	214 223		216 217	216 217	216 216	216 216	216 216	217 217	216 216	218 218	214 223	214 218	218 221	218 221	218 218	218 221
D3S621	132281515		204 206	206 206		204 204	203 205	203 203	204 204		208 208	210 210	204 204	204 404		208 208	204 208	208 208		206 206		204 210	210 210	206 206					
D3S3548	132670878	216 216	217 219	220 218	216	216 218	216 218	216 218	216 218		219 217		219 219	218 218		216 216	219 219	218 218	216 219	217 219	216 216	217 219	216 217	216 218	216 218			217 219	217 217
KIAA0586	**58894710-59015549**																												
D14s66	56120368			183 185	183 185					183 183			183 187	185 185			183 183	183 183	183 183	183 183	185 187	185 191	185 187	187 190	183 190	183 187		183 187	183 183
D14s980	56222231			165 167		157 157		157 157		165 171			157 173	165 165			157 165	165 167	151 157	157 157		155 169	169 173	155 167	155 167	165 167			
D14s274	56729090	143 146		143 146	143 143	141 150	143 152			143 150		143 143	143 143	143 143			143 150	143 150	145 147	145 145	143 143	143 143	143 143	143 143		137 145	143 147		145 147
D14s696					214 225																								
D14s1038	58692967	225 225		221 225	221 221	223 236		220 236	221 223	236 236		225 236	226 226	217 219		219 221	219 221	219 221	225 225	225 225	219 221	221 236	221 236	213 219	220 224	219 221			221 221
D14s994	59754388	213 218	212218	212 218	212 218	218 218	215 218	215 218		212 214		214 214	212 218	212220		218 218	218 218	218 218	212 212	212 212	220 220	218 218	218 220			216 218	216 218	216 218	218 218
IFT81	**108530506-114544839**																												
D12S1300	97003411	115 128	115 128	115 128	116							112 124													114 126				
PAH	101784625	240 244	236 244	240 244	244 244							236 240								240 244					227 239				
D12S353	106534007	91 91	91 99	91 99	95 105	93 103		93 105	93 105	89 93		93 93												93 103	93 101				
D12S84	107524647	199 213	215 215	199 215	215 215	202 224?			201 205			207 213	204 212	200 218	200 212	198 216	198 204	198 204				212 218		216 218	199 216		198 200	198 198	198 198
D12S1583	108287449	239 243	237 239	239 239	234 236	222 224	224 236	224 236	222 236	234 240		234 240	221 239	223 233	221 223	223 237	219 237	219 237	233 241		223 245	223 233	223 223	240 240	238 240	223 233	223 244		223 223
D12S1339	108530506			268 268	272 272	268 268	270 276	268 276	268 276	268 276		268 268												270 276	268 270				
D12s79	114544839	160 166	164 166	166 166	162 167	167 167	173 179	167 173	167 179	169 181	163 169	163 169	160 165	167 171		163 167	163 167	163 167	163 167	167 167	159 163	167 173	159 167	162 162	162 179	165 169	167 173	165 173	169 173
D12S2070	114567092	86 86	8292	86 92	90 98							90 98								96 96					87 93				

Marker	physical location												
TMEM2L	**78858767-79031054**	M	F	S	S	C7		M	F	S	S	S	C8
D15S1510	78607698												
D15S206	79991342	251 260	244 248	248 256	244 260	244 260		244 252	244 248	248 255	244 248	244 244	no call
IFT52	**41652993-41709276**	265 270	266 266	266 266	266 270	262 266		265 270	265 265	265 265	270 270	257 262	264 265
D20S1121	41612806												
D20S150	41929128	125 125	125 125	125 125	125 125	125 125		125 125	125 125	125 125	125 125	125 125	125 125
NEK11	**132231147 - 132551993**	245 249	245 245	245 245	245 245	245 249		245 249	245 245	245 245	245 249	245 249	245 249
D3S1587	132218515												
D3S621	132281515	217 217	217 217	217 217	217 217	217 217		217 217	217 217	217 217	217 217	217 217	216 217
D3S3548	132670878	204 208	204 208	204 206	204 208	204 208		208 210	204 208	206 208	204 210	208 210	208 210
ODF2	**130258253-130303060**	218 218	216 217	217 218	216 219	217 219		218 218	216 217	216 216	216 218	218 218	218 218
D14s66	56120368												
D14s980	56222231	185 187	185 187					183 187	183 189		183189	183 189	183 183
D14s274	56729090	151 165	151 165	151 151	151 151				161 169	151 161	151 169	151 169	151 161
D14s696		143 143	143 145	143 143	143 143	143 145					143 143	143 143	143 143
D14s1038	58692967												
D14s994	59754388	219 221	219 221	219 219	220 221	219 220		219 223	219 219		221 223	219 219	219 223
IFT81	**108530506-114544839**	218 218	210 217	210 217		218 218		219 219		218 218	212 218	218 218	218 218
D12S1300	97003411												
PAH	101784625												
D12S353	106534007												
D12S84	107524647												
D12S1583	108287449												
D12S1339	108530506												
D12s79	114544839												
D12S2070	114567092												

Table 4.4.5 Shows the results of microsatellite markers for two related Bangladeshi families in the regions of the complex B genes of interest. The marker name is in the first column and the physical position. The identification of the probands are shown in blue and in the top row F, M and S denote the mother, father and siblings of the proband. Those boxes in yellow highlight the homozygous markers.

Summary of results of microsatellite markers

There are three possible outcomes from the microsatellite marker results – linked, cannot be excluded and does not link, which are shown in table 4.4.6.

Gene	Proband	Linked or cannot be excluded
TMEM2L	Linked	C4
	Can not be excluded	C10
IFT52	Linked	C10, C3, C5, C4, C13, C14
	Can not be excluded	C11, C13, C9, C6, C7, C8
NEK2	Linked	C2, C9, C6, C5
	Can not be excluded	C10, C11, C13, C14, C7, C8
KIAA0586	Linked	C6
	Can not be excluded	C3, C2, C5, C13, C14, C7, C8
IFT81	Linked	C10
	Can not be excluded	C3, C2, C6, C5, C13, C14, C7, C8

Table 4.4.6 gives a summary of those probands that link or can not be excluded

In those that could not be excluded or there was positive linkage to the region further investigations was carried out by direct sequencing of the gene or by protein expression.

Sequencing of candidate genes

The initial genes sequenced were those involved in complex B IFT transport – *IFT81* and *IFT52* and also *KIAA0586* as C6 strongly linked to the region.

IFT81

Probands C10, C3, C2, C6, C5, C13 and C14 underwent direct sequencing as described in Chapter 2. No mutations were identified in any of the probands.

KIAA0586

C6 underwent sequencing of *KIAA0586.* Three homozygous known SNP's were identified. No mutations were detected.

IFT52

Sequencing of IFT52 showed no mutations in any of the linked probands.

Protein expression using Western blot

To investigate the protein expression of *ODF2*, an antibody to cenexin, the protein product of *ODF2* was used to identify protein expression in cells from JATD patients. The Western blot is shown in figure 4.4.1. The method used has been described in chapter 2.

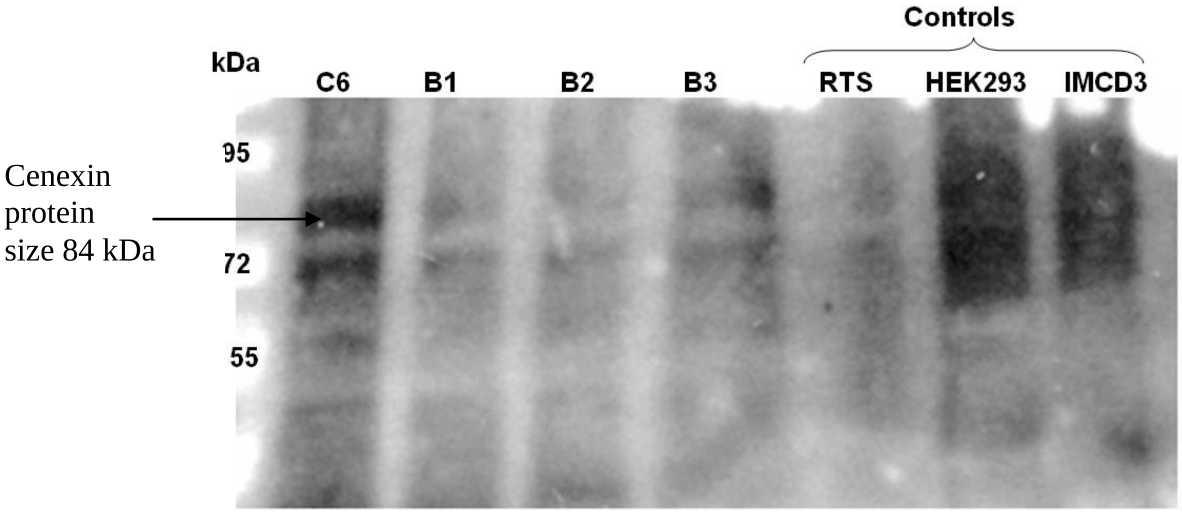


Figure 4.4.1 is a Western blot of 4 JATD patient cells and 3 control cells labelled using an antibody against cenexin which is the protein product of *ODF2.* B1, B2 and B3 are cell samples only and no DNA is available for investigation. RT = Rothmund Thompson syndrome, HEK293 human embryonic kidney cell line, IMCD3= inner medullary collecting duct.

The Western shows that all the bands were seen in all JATD patients which corresponded to the bands seen in the controls suggesting there is no loss of protein expression in these patients.

As the Western blot did not show any loss of expression i.e. was normal, the gene was then sequenced to ensure no heterozygous changes were identified (in which case the function may not have been lost but only decreased) although both mutations should be identical if they have come from a common ancestor in these consanguineous families. The other limitation of the Western blot is that cells were only available for 4 patients.

C3, C2, C6, C5, C13, C14 and C7 therefore underwent sequencing of *IFT52*. No mutations were identified therefore confirming the Western Blot.

4.4.4 JATD research in progress

The strategy used here to detect novel genes for JATD has not yielded any further mutations. It may be that there will be many genes in which mutations result in the phenotype of JATD. The other possibility is that mutations in genes known to cause other ciliopathies may also result in the phenotype of JATD as there are many similar clinical features.

This work is now in progress to look at all known cilial genes for all conditions and sequence all patients with ciliopathies in a multicentred resequencing study.

4.5 Liver involvement in ciliopathies

Liver involvement in hepatorenal fibrocystic diseases are secondary to abnormalities in the development of the ductal plate, so called ductal plate malformation (DPM) and table 4.5.1 provides a list of ciliopathies which have DPM's. This causes the clinical manifestations of congenital hepatic fibrosis, Caroli disease and polycystic liver disease. Only the cholangiocytes of the liver have cilia.

Clinically DPM do not usually cause liver insufficiency or abnormal liver biochemistry. However they tend to lead to the development of portal hypertension and the resulting manifestations of oesophageal varices and hypersplenism. Caroli's disease also has a tendency to develop cholangitis which can be a life threatening infection. Polycystic liver disease may also cause morbidity due to the mass effect from the huge hepatomegaly.

The biliary epithelium has cilia which are essential to maintain ductular function and diameter. Cilia detect luminal fluid flow which regulates cell proliferation and maintenance of planar cell polarity and mitotic spindle orientation.

Congenital hepatic fibrosis: This is a histopathological diagnosis which consists of DPM, abnormal portal veins and progressive fibrosis of the portal tracts. This affects the microscopic bile ducts.

Caroli's disease: This is diagnosed on imaging of the biliary system where macroscopic saccular or fusiform dilatations of the medium and large sized intrahepatic bile ducts can be identified. These are in continuity with the biliary tract and therefore non obstructed.

Caroli's syndrome: is Caroli's disease in association with congenital hepatic fibrosis and may be a continuum of the same disease. This affects the medium and large bile ducts.

Polycystic liver disease: this is also diagnosed on imaging in which the cysts are not in continuity and develop from hamartomas from the biliary tract. This affects the large bile ducts. Polycystic liver disease is a component of autosomal dominant polycystic kidney disease.

DPM are most commonly found as part of a multisystem ciliopathy. It may not be apparent in infancy or early childhood as the progressive fibrosis is time dependent. Figure 4.5.1.provdies a schematic diagram of a ductal plate malformation (adapted from Lamaigre, 2008).

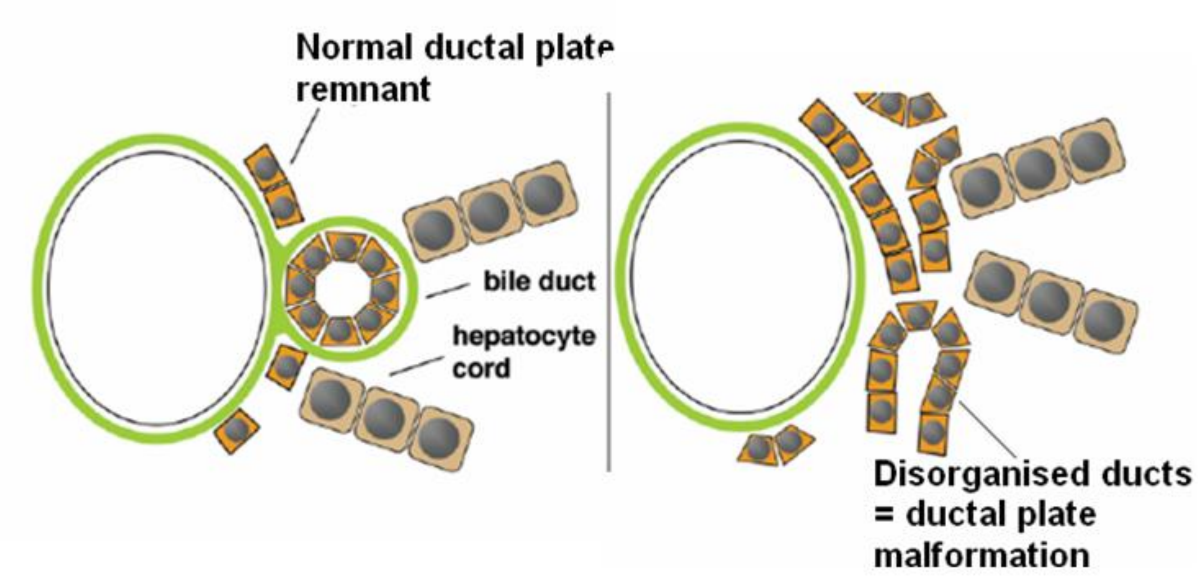


Figure 4.5.1 is a diagram of the formation of the ductal plate in the normal and when there is a ductal plate malformation.

Table 4.5.1 Gives a list of the ciliopathies which are associated with DPM's.

Disease	**Gene**
Autosomal recessive polycystic kidney disease	*PKHD1*
JATD	*IFT80*
Autosomal dominant polycystic kidney disease	*PKD1, PKD2*
Meckel-Gruber syndrome	*MKS1, TMEM67, CEP290, RPGRIP1L, CC2D2A*
Jouberts syndrome and related conditions eg COACH syndrome	*AHI1, NPHP1, CEP290, TMEM67, RPGRIP1L, ARL13B, CC2D2A*
Bardet-Biedl syndrome	*BBS1, BBS2, ARL6, BBS4, BBS5, MKKS, BBS7, TTC8, BBS9, BBS10, TRIM32, BBS12, MKS1, CEP290*
Oral-facial –digital syndrome	*OFD1*
Renal-hepatic-pancreatic dysplasia	*NPHP3*
Ellis-Van Creveld syndrome	*EVC, EVC2*
Nephronophthisis	*NPHP1, INVS, NPHP3, NPHP4, IQCB1, CEP290, GLIS2, RPGRIP1L, NEK8*
Glomerulocystic kidney disease	*HNF-1β*

4.5.1 Cilia of cholangiocytes

It is thought that the major role of cilia in cholangiocytes is as a sensory organelle.

Mechanosensors: In intrahepatic bile ducts (as in renal tubules) flow across cilia cause the cilia to bend and results in PC-1 and PC2 to form a complex which leads to an increase in intracellular calcium and suppression of cAMP. This is shown schematically in figure 4.5.2.

Osmosensors: Transient receptor potential vaniloid 4 receptor is sensitive to changes in osmolality and is stimulated bile a decrease in tonicity causing a change in intracellular calcium affecting ATP release.

Chemosensors: ATP and AFP affect the cilial receptor $P2Y_{12}$ which is activated by biliary nucleotides and causes changes in intracellular calcium affecting the cAMP signalling cascade.

Defects in cilia structure and function result in decreased intracellular calcium and increased cAMP leading to hyperproliferation and may be the mechanism by how cysts are formed.

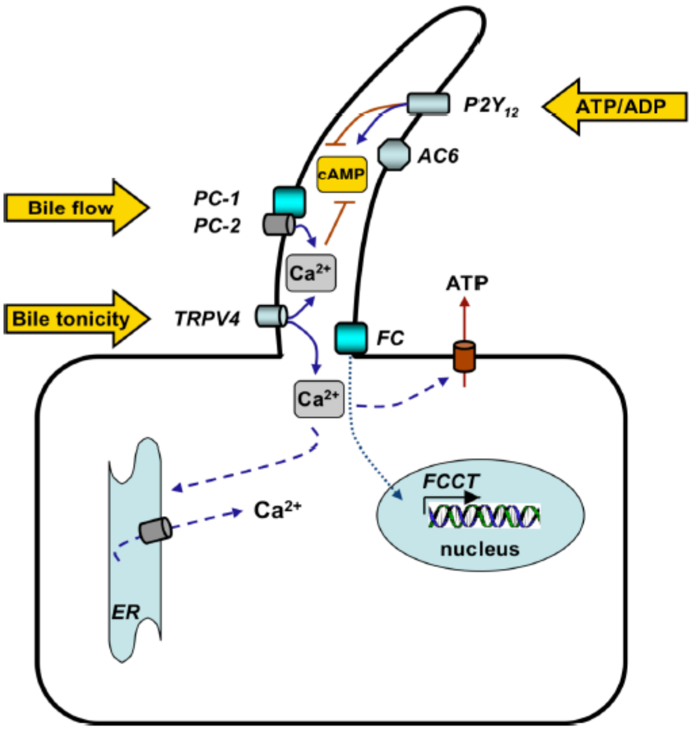


Figure 4.5.2 The cartoon shows a working model for the sensory functions of primary cilia (adapted from Masyuk *et al*, 2008). FCCT is the C-terminal tail of cleaved fibrocystin which is translocated into the nucleus under certain conditions where it may regulate expression of genes involved in cholangiocyte function.

4.5.2 Predicting ciliopathies

A constellation of clinical phenotypes is common to ciliopathies. These include

Cystic kidneys

Left right asymmetry defects

Hepatic developmental anomalies

Heart disease

Diabetes

Gonadal malformation

Developmental delay

Obesity

Polydactyly

Retinal regeneration

Skeletal defects

Central nervous system malformation

(Badano *et al*, 2006)

4.6 The investigation of cilia in the pathogenesis of biliary atresia

Biliary atresia shares common features with ciliopathies in that there is ductal plate malformation, cardiac development anomalies, situs inversus and described as part of this Ph.D. the development of renal cysts.

From the constellation of clinical features of biliary atresia the hypothesis could be drawn that primary cilia may play a role in the pathogenesis of biliary atresia.

4.6.1 The description of renal cysts in children with biliary atresia

Renal cysts were identified by abdominal imaging (either ultrasound scan or CT scan) used routinely during patient care. To ascertain the effect of transplantation on the development of renal cysts all children post transplant with differing primary aetiologies were also investigated for comparison.

Renal function was calculated using the Schwartz formula (height (cm) x 40/ creatinine (µmol/l)) as an approximation for glomerular filtration rate GFR (cGFR) with renal dysfunction being less than 60ml/min/1.73m^2.

355 children had biliary atresia of which 206 have subsequently undergone liver transplant (58%). No child had renal cysts identified at the time of diagnosis of biliary atresia.

9/355 children were identified as having developed renal cysts either by abdominal ultrasound scan (8 children) or abdominal CT scan (1 child). All 9 children had the perinatal (non syndromic) form of BA. There was no reported family history of liver or renal disease and no children developed hepatic cysts. 8/9 children have undergone liver transplantation. Child 1 had an atrial septal defect repaired in the neonatal period.

Child 1 has had long term success from the Kasai portoenterostomy and has not required liver transplantation and has not developed renal dysfunction and child 7 developed renal cysts prior to transplantation.

The median age at detection of renal cysts was 9 years (range 0.25 to 14 years). The natural history of the cysts was variable ranging from a single cyst which did not change with time, to a rapid increase in size and number of cysts.

8/9 children had received a liver transplant at a median age of 1.08 years (range 0.66 – 5 years). No child prior to liver transplant had detectable renal dysfunction. Of the 8 transplanted children, 4 developed renal dysfunction following transplantation whilst on calcineurin inhibitors (CI) (cyclosporine or tacrolimus). Two children required renal support for renal failure. On stopping the CI (and commencing mycophenolate mofetil (MMF) for immunosuppression) in 3 children the renal function stabilised including child 9 who was able to discontinue renal dialysis whilst one has continued to require renal support until renal transplantation.

In those transplanted, 1 had renal cysts prior to the development of renal dysfunction, 3 children had cysts identified at the time of dysfunction and 1 child developed cysts since stopping CI. 4 children have had an increase in size and number of cysts since stopping CI, at a time of cGFR recovery. The other 4 children have had no change in number or size of renal cysts over time.

Renal cysts development were identified in 10/463 non biliary atresia cases following liver transplant The primary pathology was hepatoblastoma (3) with single cases of tyrosinaemia, autoimmune hepatitis, fulminant hepatitis A, Alagilles syndrome, Wilson's Disease, biliary hypoplasia and cystic fibrosis. 4/10 developed renal dysfunction post transplant, 2 following transplant for hepatoblastoma and one for cystic fibrosis and biliary hypoplasia.

4.6.2 The detection of fibrocystin by immunohistochemistry

To investigate the bile ducts of children with biliary atresia further work was carried out by S. Blair-Reed using immunohistochemical techniques to identify primary cilia and the expression of fibrocystin.

Four micron sections of formalin fixed paraffin embedded liver tissue were cut, placed onto charged slides and heated for 1 hour at 60^0C. After deparaffinisation and rehydration, sections were treated in 0.3% hydrogen peroxide in water to block endogenous peroxidise activity. Antigen retrieval was performed using the ALTER technique, as previously described. Following a brief wash in water, sections were loaded onto a Sequenza^{TM} (Shandon, UK) and washed in TBS/Tween ph7.6. The wild type- C terminus monoclonal antibody (5A) specific for the intracellular domain of fibrocystin. was applied at a 1/200 dilution for 1 hour at room temperature. Sections were then washed in TBS/Tween and visualised with Dako ChemMate EnVision kit (Dako, UK) and Vector NovaRed chromagen (Vector, UK). After washing in water, sections were counterstained in haematoxylin, dehydrated, cleared and mounted.

Fibrocystin staining of liver sections from children with biliary atresia and renal cysts was found to be absent. This may be due to a lack of protein expression by the *PKHD1* gene due to a molecular defect within the gene. This would support the hypothesis of cilial defect in the aetiology of biliary atresia. Alternatively, the absence of fibrocystin in this cohort may be related to loss of expression secondary to advanced biliary disease and could be a non specific marker of liver damage. However, in contrast, fibrocystin was detectable in the other 6 cases of advanced liver disease studied, suggesting either that fibrocystin was expressed in these conditions or that the diseases were not at such an advanced stage of liver damage compared to the biliary atresia cohort.

4.6.3 PKHD1 sequencing in children with biliary atresia and renal cysts

Sequencing the PKHD1 gene as carried out using DHPLC or direct sequencing, by Prof P Harris and revealed a novel frameshifting change c.5476delG (A1826fsX147) mutation in one child which is predicted to result in a truncated protein. In addition two novel missense changes were detected E2124G (6371A>G) and F283L (847T→C). F283L is a conservative change at a highly conserved site, phenylalanine to leucine, as found in chicken and frog, while E2124G is a non-conservative change, but at a poorly conserved site. Extended analysis of the family, who are phenotypically normal (including normal ultrasound scan imaging of the kidneys), revealed a sibling and the father to have the deletion mutation and both missense changes indicating that the missense changes are unlikely to be pathogenic. No likely pathogenic changes to PKHD1 were found in any of the other cases.

A second *PKHD1* mutation was not identified. The single mutation could be a coincidental finding of biliary atresia in a carrier of a *PKHD1* truncating mutation, but with a population frequency of such carriers of ~1 in 200, this seems unlikely. A second, missed, mutation could include a large deletion or an atypical splicing event distant from the splice junctions. Sequencing the promoter region did not identify any variants. An alternative explanation is that a mutation in another gene together with the *PKHD1* mutant allele may result in the biliary atresia plus renal cyst phenotype. Interestingly, in other ciliopathies such as Bardet-Biedl syndrome and nephronophthisis, oligogenic inheritance involving mutations in more than one gene has been reported and may also explain the absence of PKHD1 mutations in the other BA children with renal cysts.

4.6.4 The investigation of nasal cilia

Children with Bardet-Biedl syndrome have anosmia which is likely to be due to anomalies in nasal cilia. I hypothesised that an abnormality in primary cilia may be detected in nasal cilia as a proxy for the cilial involvement of the bile ducts. Although patients with biliary atresia do not report anosmia two children with biliary atresia and renal cysts had their nasal cilia examined. *Prof Chris O'Callaghan carried out the work of my hypothesis. The ultrastructure of nasal cilia was assessed using electron microscopy. The cilial beat cycle, pattern and frequency was analysed by digital high speed video camera. Cilial morphology and function was found to be normal in these two children.*

4.6.5 Summary of results

Cholangiocyte primary cilia play an important role in the normal development and function of bile ducts. Genetic mutations resulting in abnormalities of cilia formation or function cause the clinical syndromes named ciliopathies which have common over lapping clinical features.

The constellation of clinical features forming biliary atresia suggests similarity with ciliopathies. Fibrocystin which is an essential component of cilia has been shown to be absent in biliary atresia patients and supports the hypothesis that biliary atresia may in part be attributed to a cilial defect.

Chapter 5

A novel phenotype of canalicular bile salt transport disorders

Contents

5.1 Introduction to the molecular genetics of canalicular bile salt transport disorders

5.1.1 History

The first phenotype of a canalicular bile salt transport disorder was described in 1969, in an American Amish population named Byler from which the condition was originally named (Byler disease). The children in the original cohort had severe cholestasis in infancy, intense pruritis, steatorrhoea and failure to thrive (Clayton *et al*, 1969). The liver disease progressed to cirrhosis and it was uniformly fatal. Since this initial description, children across the world from all genetic backgrounds have been described with bile salt transport disorders, three genes have been identified *ATP8B1*, *ABCB11* and *ABCB4*, and the importance of regulator genes has been revealed.

5.1.2 Nomenclature

The nomenclature of categorising canalicular bile salt transport may be by disease phenotype, gamma-glutamyl transpeptidase activity (γGT), the affected gene or by protein deficit. This is summarised in table 5.1.1.

Names known by	**Gene**	**Protein**	**γGT level**	**Disease phenotype**
PFIC-1 FIC-1 Byler disease	*ATP8B1*	FIC-1	Low or normal	1.PFIC 2. BRIC 3. ICP (rare) 4. Greenland Eskimo Cholestasis
PFIC-2 BSEP deficiency	*ABCB11*	BSEP	Low or normal	1.PFIC 2. BRIC (van Mil, 2004) 3. ICP 4. Drug induced cholestasis (Lang, 2007)
PFIC-3 MDR3 deficiency	*ABCB4*	MDR3	High	1.PFIC 2. ICP 3. Cholelithiasis 4. Drug induced cholestasis (Lang, 2007)
ICP	*FXR*			1. ICP

Table 5.1.1. The nomenclature of bile salt transport defects which can be classified by name, gene, protein deficit or γGT. Table 5.1.1

5.1.3 *ATP8B1*

This 85.38Kb gene is on chromosome 18q21.31 (53,464,656 - 53,550,037) and consists of 28 coding exons. It was first identified (Bull *et al*, 1998) in children with PFIC using an autozygosity mapping technique.

68 different mutations have been identified to date:-

44 PFIC

19 BRIC

3 ICP

1 Greenland Eskimo cholestasis

1 BA (although this is unlikely to be pathogenic as BA has a raised γGT and therefore there is intact bile salt canalicular transport)

Missense are the most common type (40) followed by splicing, insertions and deletion.

It encodes FIC-1, a 140kD P-type ATPase with 10 transmembrane domains and signature motifs that characterize the non-heavy metal-binding P-type ATPase. It is expressed in several epithelial tissues, the pancreas, and more strongly in small intestine than in the liver which explains the severe diarrhoea which may also be a clinical feature. The expression within the liver localises it to the cholangiocyte. Chaperone proteins CDC50 are required to transport FIC-1 from the endoplasmic reticulum to the apical membrane (Paulusma *et al*, 2008). FIC-1 functions as an aminophospholipid flippase which participates in the movement of phosphatidylserine from the outer to the inner leaflet. This creates asymmetry in the membrane distribution of lipids which is necessary for normal function (Paulusma *et al*, 2008). How this results in the phenotype of FIC-1 disease is not known. Mutation type or location correlates with clinical severity i.e. missense mutations are more common in BRIC whilst frameshift and large deletions are more common in PFIC (Klomp *et al*, 2004). Frankenburg *et al*, 2008, demonstrated that *ATP8B1* through protein kinase C ζ leads to phosphorylation, nuclear localisation and activation of FXR. Through this mechanism, mutations in *ATP8B1*, reduces FXR activation and therefore BSEP expression. This is compounded by the reduced negative feedback from FXR on the ileum so there is an increase in bile salt uptake by enterocytes and increased hepatotoxic bile salts entering the hepatocyte which can not transport them across the apical membrane.

5.1.4 *ABCB11*

This gene on chromosome 2q24 encodes the protein bile salt export pump (BSEP) (Strautnieks *et al*, 1998). 106 different mutations have been identified. Missense mutations are the most common followed by splice sites and deletions. Insertions are rare. The membrane-associated protein encoded by this gene is a member of the superfamily of ATP-binding cassette (ABC) transporters. ABC proteins transport various molecules across extra- and intra-cellular membranes. This protein is a member of the MDR/TAP subfamily. Members of the MDR/TAP subfamily are involved in multidrug resistance. The protein encoded by this gene is the major canalicular bile salt export pump in man. It is suggested that different phenotypes directly correlate with the amount of expression of mature protein that is expressed at the cell surface (Lam *et al*, 2007). *ABCB11* is not expressed in the ileum and this phenotype does not experience diarrhoea.

5.1.5 *ABCB4*

Multidrug resistence 3(MDR3) or ABCB4 encode the hepatobiliary phospholipid transporter, mutations in which cause PFIC3 which differs from PFIC1 and 2 by having a raised γGT (Dixon *et al*, 2000)

5.1.6 Bile acid farnesoid X receptor (FXR)

This nuclear receptor regulates bile acids homeostasis so to reduce hepatocyte toxicity. It binds to DNA response elements in promotor regions of target genes including *ABCB11* and *ABCB4* to activate their transcription so increasing bile acid secretion out of the hepatocyte. FXR also promotes genes for sulphonidation and glucuronidation, reduces import and synthesis of bile salts in the hepatocyste and also reduces ileal bile salt reabsorption (van Mil *et al*, 2007).

5.2 Clinical phenotypes of canalicular bile salt transport disorders

5.2.1 Progressive familial intrahepatic cholestasis (PFIC)

Three different PFIC disorders have been describe corresponding to mutations in different genes.

PFIC-1

The presentation of PFIC-1 is usually within the first 3-6 months of life with conjugated hyperbilirubinaemia which may fluctuate. There is hepatomegaly and splenomegaly develops when there is progression to cirrhosis. Pruritus is very intense and can be refractory to treatment. The absorption of long chain fatty acids is poor leading to malabsorption resulting in steatorrhoea, failure to thrive and symptoms of fat soluble vitamin deficiency such as coagulopathy and rickets.

There are extrahepatic manifestations reflecting areas of protein expression with pancreatitis, diarrhoea and sensorineural hearing loss also occurring.

Liver transplantation is usually required early in childhood but the extrahepatic manifestations will not be alleviated and indeed the diarrhoea may become worse.

Laboratory investigations show a low or normal γGT despite severe cholestasis and the serum cholesterol is also normal. The total serum bile salts are elevated but the concentration of chenodeoxycholic acid in bile is very low. Mutations in *ATP8B1* are also associated with ICP (Mullenbach *et al*, 2005).

Liver histology shows cholestasis and canalicular bile plugs. On electron microscopy there is a characteristic granular appearance. There may also be small duct paucity.

There is a reduction in the protein FIC-1 secondary to mutations in *ATP8B1*.

PFIC-2

The affected child presents with conjugated hyperbilirubinaemia in infancy, hepatomegaly, severe pruritis, steatorrhoea, fat soluble vitamin deficiency and failure to thrive. There are no extrahepatic manifestations. Liver disease progresses to cirrhosis and liver transplantation is usually required in childhood.

The γGT level is low despite cholestasis.

Hepatocellular carcinoma and cholangiocarcinoma have both been described in children under the age of 5 years with PFIC-2 (Knisely *et al*, 2006).

The liver biopsy differs from that of PFIC-1 in that there is inflammation with giant cell hepatitis. There may also be ductular transformation and fibrosis. Immunohistochemistry looking for a deficiency in bile salt export pump (BSEP) expression supports the diagnosis.

The disease is due to a reduction in BSEP which is encoded by the gene *ABCB11*. The disease phenotype is due to the reduction in excretion of bile salts.

PFIC-3

Children with PFIC-3 usually present in infancy but may present later with cholelithiasis or in adulthood with biliary cirrhosis when there has been only partial reduction in protein expression such as in the heterozygous state. The pruritis is intense but the conjugated hyperbilirubinaemia may not be as high as the other types of PFIC. A raised γGT distinguishes PFIC-3 from PFIC1 and PFIC2.

The liver biopsy shows portal fibrosis.

There is progressive cirrhosis and liver transplantation is usually required in childhood.

This is a result of MDR3 deficiency which is encoded by the gene *ABCB4*. Adult onset disease, up to 20% of cases of intrahepatic cholestasis of pregnancy, miscarriages and still births have all been linked to mutations in *ABCB4* (Schneider *et al*, 2007) with the severity of phenotype being influenced by the type of mutation (Wasmuth *et al*, 2007) and whether there is heterozygosity or homozygous for the sequence change (Gotthardt *et al*, 2008).

Canalicular bile salt transport defects do not exclusively present in childhood but may only become evident when bile salt transport mechanisms are overwhelmed e.g. in the presence of increased circulating oestrogen when using the combined oral contraceptive pill or during the third trimester of pregnancy.

5.2.2 Benign recurrent intrahepatic cholestasis (BRIC)

The cholestasis of BRIC can present at any time but most commonly it is as a teenager or in the early twenties. The cholestasis is transient but can last for up to six months. Pruritis and fat soluble vitamin efficiency can occur. It is said to be benign as it does not progress to chronic liver disease however there are case reports of chronic liver disease developing (van Ooteghem *et al*, 2002). Although BRIC may occur with no apparent stimulus there is often a trigger. In women the oral contraceptive pill which suppresses bile salt transport activity reduces the transport to a critical level resulting in the onset of clinical symptoms. Women who develop cholestasis secondary to increased serum oestrogen may be at risk of developing intrahepatic cholestasis of pregnancy. BRIC can develop from mutations in *ATP8B1*, *ABCB11* (Kubitz *et al*, 2006) or *ABCB4*.

5.2.3 Intrahepatic cholestasis of pregnancy (ICP)

Pruritis, often worse on the soles and palms, is the main presenting symptom causing discomfort, insomnia and fatigue. There is elevation of serum alanine aminotransferase (ALT/SGPT) and serum bile acids >10µmol/L is the most sensitive laboratory indicators in confirming the diagnosis of cholestasis. Conventional liver function tests are initially normal and it is in this circumstance that serum bile acid measurement may be needed to confirm the presence of cholestasis. Jaundice occurs in up to 10% of ICP and γGT remains normal in two-thirds. The cholestasis can result in steatorrhoea and fat soluble vitamin malabsorption. There is spontaneous relief of signs and symptoms within 2-3 weeks after delivery (Beuers *et al*, 2006).

In Europe ICP occurs in 0.1-1.5% of pregnancies. Incidence is highest in Chile, Baltic States, Scandinavia and Bolivia with up to 15% of pregnancies being associated with ICP. Over the past decade a trend to lower incidence has been observed in Sweden and Chile possibly due to increased micronutrients such as selenium in an improved diet (Reyes *et al*, 2000).

ICP is not influenced by parity but is 5 times more common in multiple pregnancies. It may be sporadic or run in families. There is seasonal variability with more cases occurring in winter.

Oestrogens and progestagens are naturally cholestatic (Vallejo *et al*, 2006). It is likely that in most instances ICP occurs in individuals who have otherwise subclinical mutations within susceptibility genes for intrahepatic cholestasis whose capacity for efficient biliary excretion is exceded in the hormonal milieu of pregnancy. A hormonal role in the aetiology of ICP is suggested by the observation that symptoms are most prevalent in the 3rd trimester when oestrogen levels are highest. When the hormone levels fall after delivery symptoms tend to improve rapidly. Twin and triplet pregnancies which have higher estrogen levels, have an increased incidence of ICP. Biliary excretion of oestrodiol-17β-D-glucuronide occurs via a canalicular multispecific conjugate export pump (multidrug resistence related protein, MRP2) and exerts its cholestatic effect via a trans inhibition of BSEP from the canalicular side of the hepatocyte membrane. This causes a transient inhibition of BSEP. Oestradiol-17β-D-glucuronide also suppresses the expression of MRP2 (Steiger *et al,* 2000).

There is a significant alteration in the ratio of different progesterone isomers in ICP as compared to other causes of liver disease in pregnancy, the cause and effect of which are not known.

A liver biopsy is rarely undertaken unless there is diagnostic uncertainty.

Intracellular bile pigment and canalicular bile plugs are seen in the absence of any other histological abnormality. Electronmicroscopy shows dilated canaliculi with loss of microvilli.

The risks to the child include (Rioseco *et al, 1994)*:-

1. Increased risk of preterm delivery (19-60%)
2. Fetal distress (22-41%) as indicated by meconium liquor. The inhalation of secretions at the time of delivery with high bile salts has been reported to result in the development of respiratory distress syndrome.
3. Fetal loss (0.4-1.6%). The pathogenesis of fetal complications is not fully understood. In stillborns the post mortem findings are those of acute asphyxia. The fetus is usually well grown and surveillance of the fetal placental unit is normal suggesting that placental insufficiency is not the cause.
4. In those neonates who are born healthy there are no long term complications have previously been reported.

Women who have had ICP may be at risk of cholestasis from use of the oestrogen containing oral contracpetive pill.

There are cases of ICP which can not be attributed to BSEP, FIC1 or MDR3 deficiency which suggest either a non genetic aetiology or there are other as yet unidentified genes for cholestasis (Savander, 2007).

5.2.4 Low γGT cholestasis

In the presence of cholestasis damage to the bile ducts usually results in an increased γGT and therefore a low γGT is a significant finding and can be of diagnostic value.

The differential diagnosis of cholestasis with low γGT is:-

5.2.4.1 PFIC-1 and 2, BRIC and ICP

The clinical features are described above. The diagnosis of BRIC and ICP will be evident from the history. It may be difficult to distinguish between PFIC-1 and 2 initially; however the development of diarrhoea or pancreatitis would implicate PFIC-1.

5.2.4.2 ARC syndrome (OMIM: 208085)

Arthrogryposis, renal dysfunction (Fanconi type tubulopathy) and cholestasis (ARC) syndrome is due to the autosomal recessive inheritance of mutations in *VPS33B* on chromosome 15q26 (Gissen *et al*, 2004). The extended clinical spectrum includes icthyosis, bleeding diathesis, sensorineural deafness, congenital heart disease, diarrhoea, growth failure, cerebral malformations and death is usual before the age of one year. The phenotype can be incomplete with the absence of arthrogryposis (Bull *et al*, 2006). The gene encodes the VPS33B protein which is involved with intracellular vesicular trafficking and vesicular membrane fusion by interacting with soluble *N*-ethylmaleimide-sensitive protein attachment receptor (SNARE) proteins. The bleeding diathesis results from abnormal alpha granule development within abnormally enlarged platelets (Lo *et al*, 2005).

5.2.4.3 Bile acid synthesis disorders

Deficiencies in the enzymes converting cholesterol to the primary bile acids, cholic and chenodeoxycholic acid result in bile acid synthesis disorders some of which present as severe neonatal cholestasis with low γGT. Other presentations may be with the development of liver disease later in life, fat soluble vitamin deficiency or neurological disease. The diagnosis is made by identifying a low level of bile acids in the patient's serum and by fast atom bombardment mass spectrometry of the urine to categorise the defect by detecting the excess metabolites and a lack of the deficient. The genetic defect has been identified in some cases. It is important to diagnose early as the instigation of treatment with bile acid replacement therapy (chenoic acid) can stop the development of neurological symptoms.

> *3β-Hydroxy-*Δ^5*-*C_{27}*-steroid dehydrogenase/isomerase (3βHSD) deficiency*: Mutations in *HSD3B7* on chromosome 16p12-11.2 are associated with 3βHSD. Typically the presentation is with neonatal cholestasis with giant cells on histological examination of the liver. Without treatment there is progressive cirrhosis (Cheng *et al*, 2007).

Δ^4-3-Oxosteroid 5β-reductase deficiency:

This presents with severe cholestasis in infancy with a normal (although occasionally raised) γGT. The over production of Δ^4-3-oxo is hepatotoxic. The liver histology is of giant cell hepatitis with pseudoacinar transformation of the hepatocytes. Electron microscopy is unique with small bile canaliculi containing electron dense material and an absence of microvilli. *AKR1D1* on chromosome 7q32-33 encodes this enzyme (Lemonde *et al*, 2003).

5.3 Neonatal liver failure: a novel clinical phenotype of canalicular bile salt transport disorders

5.3.1 Introduction to neonatal liver failure

Neonatal liver failure requiring liver transplantation is rare and accounts for approximately 2.1% of all paediatric liver transplants. For a liver transplant to be successful the underlying cause of liver failure needs to be solely hepatic as extrahepatic features will not modify. Making a diagnosis is therefore important but can be difficult in the neonatal period when not all clinical features may be manifest. Table 5.3.1 provides a list of conditions which can present in the neonatal period with liver failure.

Group	Causes
Infection	Herpes simplex virus
	Echovirus
	Coxsackie virus A and B
	Parvovirus
	Adenovirus
	Cytomegalovirus (rare)
Metabolic defects	Galactosaemia
	Tyrosinaemia
	Neimann-Pick syndrome
	Urea cycle defects
	Organic acidemias
	Mitochondrial respiratory chain defects
	Bile salt synthesis defects
	Zellweger syndrome
	Neonatal haemochromatosis
Congenital heart defects	Ischaemia and abnormal perfusion of the liver

Table 5.3.1. The differential diagnosis of neonatal liver failure

In all conditions the laboratory investigations will show hepatitis (raised liver enzymes AST and ALT), hepatic synthetic failure (raised prothrombin time and low albumin) and cholestasis. Bile duct involvement leads to a rise in alkaline phosphatase and γGT. Other diagnostic investigations will be specific to each condition. The presence of low γGT in the presence of cholestasis and liver failure is notable and may intimate the diagnosis to be a defect in the biosynthesis of bile acids in the hepatocyte with low levels of bile salts in the circulation. The diagnosis is made by fast atom mass spectrometry of serum and urine. The other conditions which have a low γGT are the bile salt transport defects (BSEP and FIC-1 deficiency) which may present with cholestasis in the neonatal period with progressive fibrosis but have not previously been associated with neonatal liver failure.

γGT is highest in the neonatal period (including premature infants) and declines to reach adult reference range levels by the age of approximately 5 months. A low γGT in a neonate is therefore a significant diagnostic finding

.

Presented here is the investigation of six infants presenting with neonatal liver failure with a low γGT despite severe cholestasis.

5.3.2 Cohort of neonatal liver failure

All babies presented in the newborn period with jaundice, low γGT and synthetic liver failure were studied. A full medical and obstetric history was gained form the mother.

Proband 1:

The mother of proband 1 developed intense pruritus at 38 weeks gestation. The serum bile salts were significantly elevated and foetal distress was seen on the CTG. The baby was well at the time of delivery but developed petechiae and jaundice within a few hours of birth. He was found to have hepatosplenomegaly and an inutero infection suspected. Cytomegalovirus (CMV) was isolated from his urine and the infection was confirmed by blood PCR showing 1,000,000 copies/ml. A rapid rise in serum bilirubin required a double exchange transfusion. Despite the extremely high bilirubin levels the γGT was normal. He was treated with intravenous ganciclovir and resolution of the CMV infection was confirmed by being undetectable by blood PCR. Despite successful treatment of the CMV infection his cholestasis worsened and he developed liver failure. He underwent liver transplantation at the age of seven weeks. There has been no recurrence of disease at 2.5 years follow up and he remains well with the original graft.

Proband 2:

The mother of proband 2 initially had a period of jaundice at the age of 23 years which resolved and no cause was found. During the pregnancy she developed intense pruritus at 34 weeks gestation and became jaundiced at 36 weeks. Unusual for ICP the liver disease did not improve immediately following delivery and she underwent a liver biopsy at 1 month post partum which demonstrated cholestasis and giant cell transformation which is an unusual finding in adult liver disease. The neonate was jaundiced from birth with low γGT and rapidly developed liver synthetic failure. All known causes of neonatal liver failure were excluded necessitating a transjugular liver biopsy at the age of six days which also showed giant cell transformation and cholestasis. Transfusion-transmissible virus (TTV) was isolated from the serum in the child. The liver failure progressed requiring liver transplantation aged two weeks. There has been no recurrence of the disease at 6 years follow up and he remains well with the original graft.

Proband 3:

The mother presented at 37 weeks gestation with pruritus of increasing severity. Elevated bile salts and fetal distress necessitated delivery. She remained well until day 4 when she became lethargic and developed jaundice. Investigations identified cholestasis with low γGT and liver failure. All other known causes of liver failure were excluded. The liver failure progressed requiring liver transplantation aged 10 days. She remains well with the original graft at one year follow up.

Proband 4

This infant was born at term weighing 2400g. The mother had been well throughout pregnancy with no symptoms of ICP. The infant was initially well but by 1 week of age had developed intense cholestasis (bilirubin 476 IU/l and γGT 64IU/l) and liver failure. Despite thorough investigation for causes of liver failure in the neonatal period none were identified. The liver failure deteriorated and she received a liver transplant aged 3 weeks.

She has subsequently developed CMV infection which has been successfully treated. Before transplant she presumably carried antibodies from her mother (IgG CMV positive and IgM negative). She has remained well at follow up 3 months post transplant.

Proband 5

The mother of this infant was generally unwell during pregnancy with malaise and sore throat. She did not have any pruritus or jaundice. All antenatal scans were normal. A female was born at term weighing 2651g (25th centile). At delivery she had a petechial rash, jaundice and massive hepatosplenomegaly. Prior to vitamin K administration the PT 90 seconds and the albumin initially was 28. The bilirubin peaked at 766 IU/l whilst the γGT was 12 IU/l. Cranial imaging showed extensive changes associated with congenital CMV infection with calcification as well as white matter changes. The heart was also affected with a ventricular septal defect, patent foramen ovale and mild pulmonary artery stenosis. A bone marrow aspirate to investigate marrow suppression was consistent with CMV infection. The liver failure and cholestasis stabilised with supportive medical care but did not resolve. The infant succumbed to the neurological sequele aged three months.

Proband 6

This mother presented at in her first pregnancy at 34 weeks gestation with intense pruritus, increasing transaminases and also high blood pressure. This necessitated induction of labour at 35 weeks. A male child weighing 2463g (20-50th centile) was born with Apgar scores of 5 at 1 minute and 8 at 5 minutes. The delivered placenta contained a chorangioma. The infant was initially well but became increasingly jaundiced by 24 hours of age with low γGT and liver failure. He had no splenomegaly or dysmorphic features. He was extensively investigated for causes of cholestasis. An open liver biopsy showed neonatal hepatitis with widespread cholestasis. Electron microscopy showed myelin figures which are associated with viral infection although no viral inclusions were seen. Bile acids were elevated (Tauro-dihydroxycholanoate 0.19 µmol/mmol (reference range 0.01-0.08)) and no abnormal intermediates were identified. The mothers' symptoms resolved following delivery. The infant's cholestasis resolved with medical support.

All the infants underwent histological examination of the native liver (either by biopsy or explant) and also mother of proband 2 and the results are shown in table 5.3.2.

	Appearance of hepatocytes	**Presence of cholestasis**	**Presence of fibrosis**	**Portal inflammation**	**Portal tracts**
Proband 1	Hepatocellular disarray Giant cell transformation	Hepatocellular cholestasis	Pericellular distribution	Minimal	Normal
Proband 2	Giant cell transformation	Hepatocellular cholestasis	Early pericellular	Mild mixed inflammatory cells	Ductular transformation
Mother of proband 2	Giant cell transformation prominent in the perivenular region	Hepatocellular cholestasis	Pericellular distribution	Moderate mixed infiltrate	Normal
Proband 3	Giant cell transformation	Pseudoglandular cholestasis	Pericellular distribution	Minimal	Normal
Proband 4	Giant cell transformation	Hepatocellular cholestasis	Minimal	Mild mixed inflammatory cells	Normal
Proband 5	Hepatocellular disarray. Extreme giant cell transformation	Severe hepatocellular cholestasis	Minimal fibrosis of the portal tracts with fibrosis in the perivenular, perisinusoidal and pericellular regions	Mild mixed inflammatory cell infiltrate	Bile duct paucity, with mild ductular transformation
Proband 6	Hepatocellular disarray with swollen and multinuclear hepatocytes	Canalicular cholestasis	No fibrosis	No inflammation	Normal bile ducts

Table 5.3.2. The liver histological features of all probands and the mother of proband 2

5.4 Molecular genetic investigation of low γGT cholestasis neonatal liver failure

5.4.1 Sequencing *ABCB11*

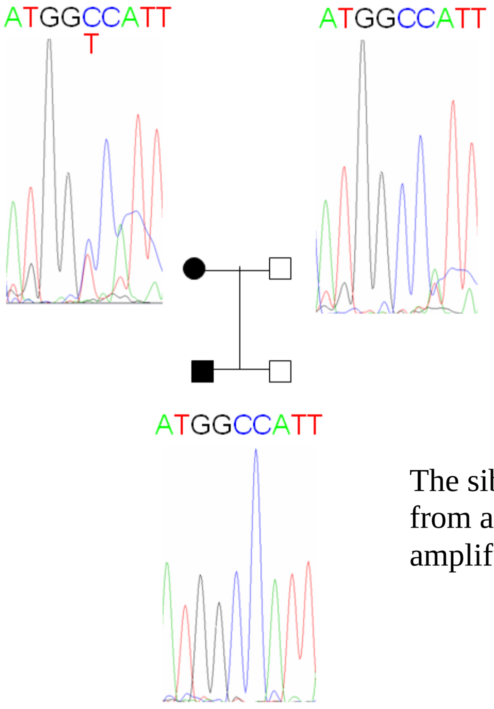
The finding of low serum γGT with raised bile salts led to the investigation of a bile salt transport defect in both mother and child.

ABCB11 was initially chosen for investigation due to the histological feature of giant cell hepatitis which is a feature of PFIC-2. *ABCB11* was also selected as at the time of investigation no extrahepatic features which would suggestive of a defect in *ATP8B1* (diarrhoea or pancreatitis) had become manifest.

The gene was sequenced using the Big Dye sequencing technique on the ABI3700 machine and the results were analysed using Chromas (as described in Chapter 2).

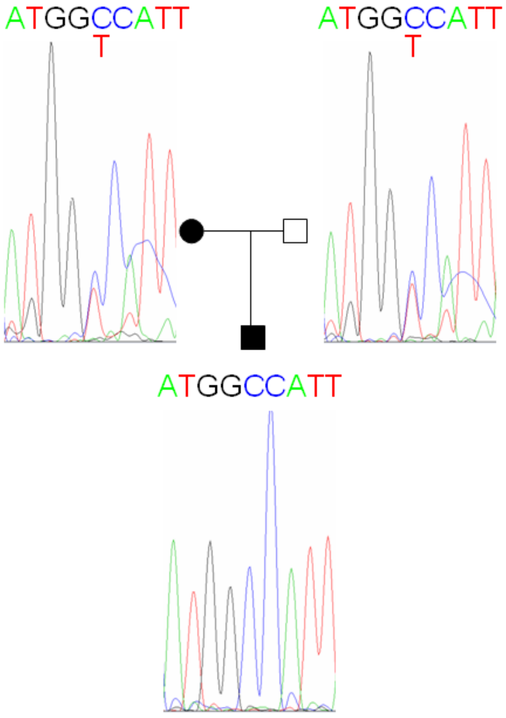
In all probands all the exons were sequenced including 20 bases of the intron / exon boundary to identify splice site changes. The electropherogram results are shown in figure 5.4.1 and a summary of the results are shown in table 5.4.1. Where available the DNA from the parents and siblings was also sequenced and the results are shown with the family pedigree to highlight segregation of sequence variants within each family.

A.Proband 1

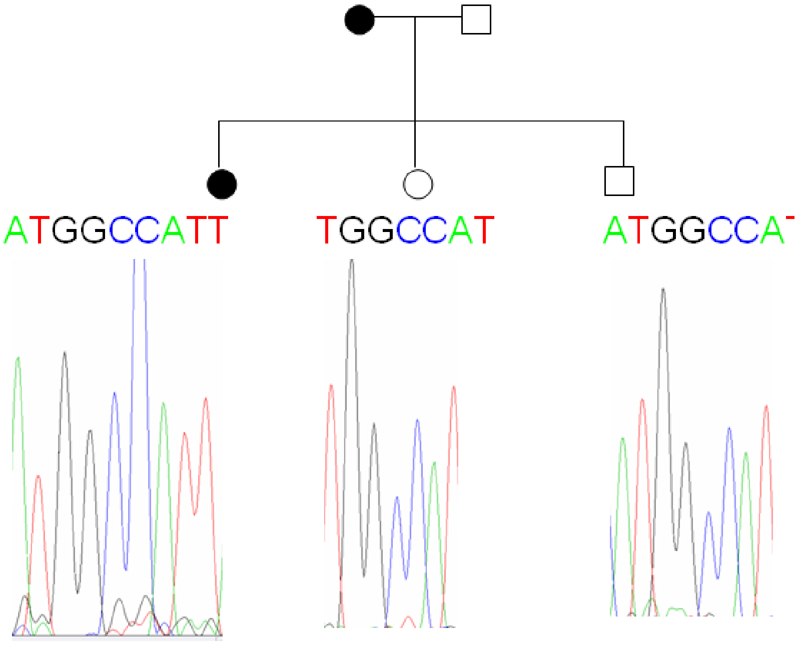


The sibling DNA was taken from a bucal swab and did not amplify

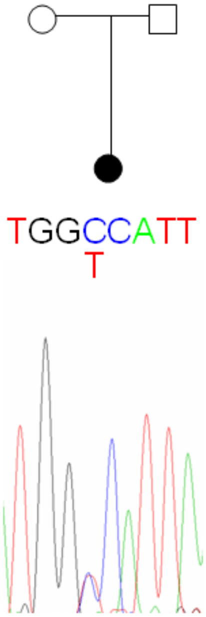
B. Proband 2



C. Proband 3



D. Proband 4



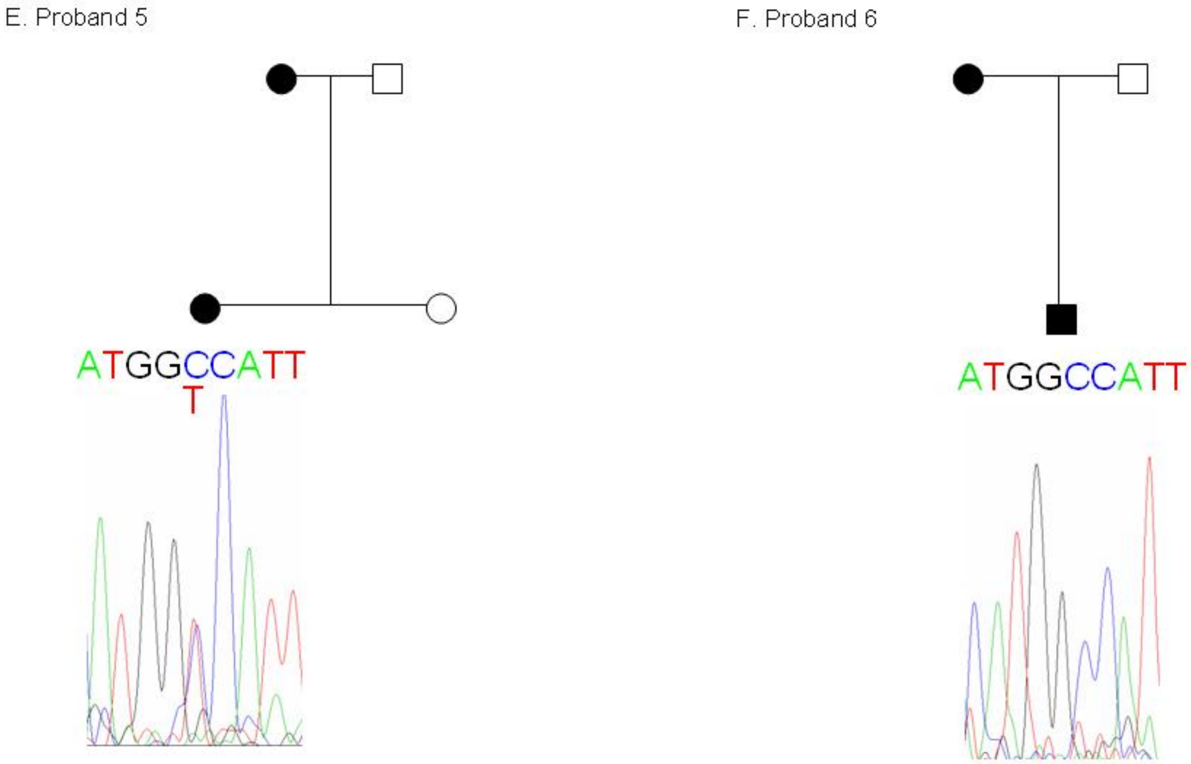


Figure 5.4.1

The pedigree and sequencing results for all the probands are shown in figure 5.4.1

A. The affected child is homozygous for the sequence variant C->T. His sibling was unaffected but his DNA sample was from a bucal swab and did not amplify despite whole genome amplification. The mother who had severe pruritus is heterozygous and the father who is completely asymptomatic is homozygous.
B. The affected child is homozygous for C->T whilst the mother who had severe cholestasis is heterozygous and the unaffected father is also heterozygous
C. The affected child is homozygous for C->T as are the unaffected siblings. The mother suffered with pruritus in the last trimester of the affected pregnancy whilst

she was asymptomatic in the other pregnancies. The father is heterozygous and asymptomatic

D. This neonate had severe liver failure with low γGT. She is heterozygous for the - sequence variant C/T. Her mother did not suffer with intrahepatic cholestasis of pregnancy.

E. The affected neonate was heterozygous for C/T; she also had congenital CMV infection. The mother did not have intrahepatic cholestasis of pregnancy but was generally unwell during pregnancy

F. The affected child is homozygous for the sequence variant C->T. He recovered and did not require liver transplantation

Subject	**Sequence**	**Manifest symptoms**	**Confounding factors**

	variant		
Proband 1	CC	Severe liver failure and cholestasis requiring OLT	CMV (successfully treated)
Mother proband 1	TC	Severe pruritus and elevated bile salts	
Father proband 1	CC	Asymptomatic	
Sibling proband	NA	Asymptomatic	
Proband 2	CC	Severe liver failure requiring OLT	TT virus
Mother proband 2	TC	Severe cholestasis	
Father proband 2	TC	Asymptomatic	
Proband 3	CC	Liver failure requiring OLT	None identified
Mother proband 3	TC	Pruritus in this pregnancy only	
Father proband 3	TC	asymptomatic	
Sibling 1 proband 3	CC	asymptomatic	
Sibling 2 proband 3	CC	asymptomatic	
Proband 4	TC	Liver failure requiring OLT	None identified
Mother proband 4	TC	asymptomatic	
Father proband 4	TT	asymptomatic	
Proband 5	TC	Liver failure which resolved	Congenital CMV
Mother proband 5	TC	Pruritus in pregnancy	
Father proband 5	TT	Asymptomatic	
Sibling proband 5	unknown	Asymptomatic	
Proband 6	CC	Liver failure which resolved	Hypertrophic cardiomyopathy. Unidentified virus?
Mother proband 6	TC	Severe ICP	
Father proband 6	CC	Asymptomatic	

Table 5.4.1 Shows a summary of the sequencing results

The substitution of T for C leads to the amino acid change from valine to alanine at position 444 (V444A) which is a known variant (rs2287622).

Polyphen predicts this to be a benign change.

The 444 position falls within the intracellular component of *ABCB11* in which other pathogenic mutations have been identified which is pictorially shown in figure 5.4.2.

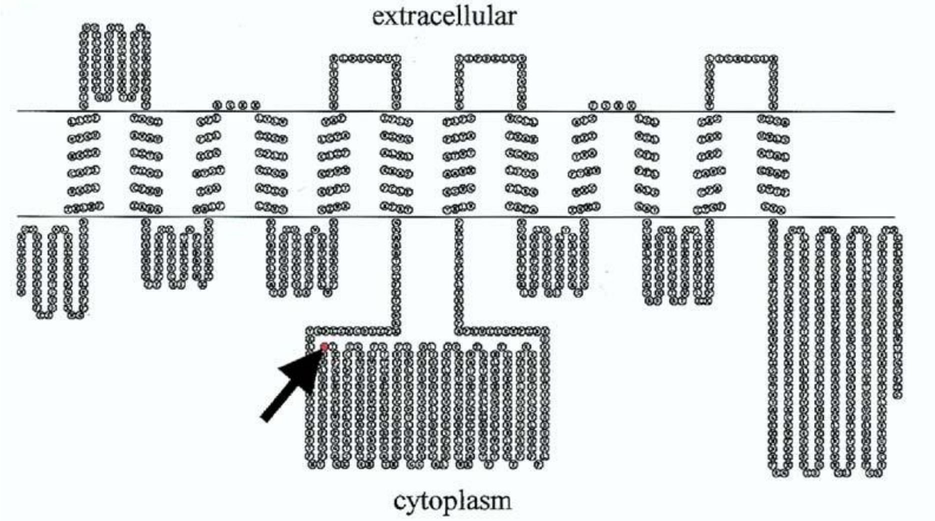


Figure 5.4.2 This topology model of BSEP shows the amino acid position of V444A to be in the nucleotide binding fold in the intracellular domain.

5.4.2 Immunohistochemistry of liver biopsies

In collaboration with Dr Alex Knisely, Kings College London, immunohistochemistry of the proband liver biopsies and liver biopsies from the mothers of proband 1, 2 and 6, were studied for expression of biliary proteins including BSEP and the results are shown in figure 5.4.3

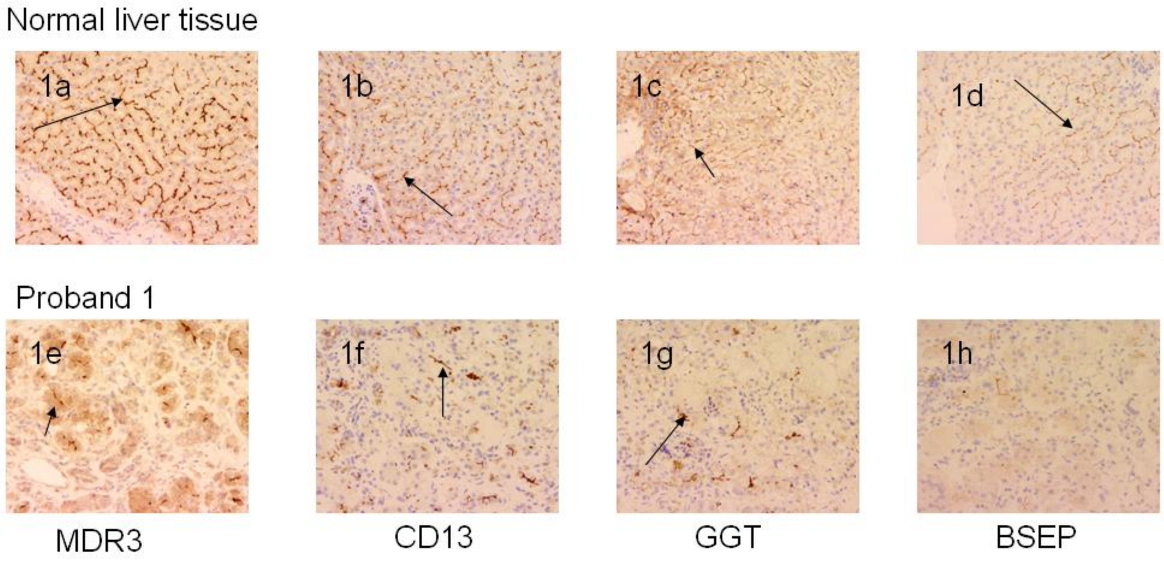


Figure 5.4.3

Images a-d are from normal control liver. Images e-h are taken from sections of liver from proband 1.

a & e show the presence of the protein MDR3 in affected and control liver.

b & f show CD13 expression is present in both livers as is GGT (c & g).

Image d shows a normal distribution of BSEP staining whilst in proband 1 (image h) there is absence of BSEP identified by immunohistochemistry.

5.5 Discussion of results

Reduced expression of BSEP associated with the sequence variant c. C1331T in *ABCB11* has been identified in a cohort of neonates with liver failure and low γGT.

The development of liver failure with cholestasis and a low γGT is unusual and suggests an abnormality in bile salt metabolism – either synthesis or transport. In these children the bile salts in the serum were elevated which excludes a synthetic defect. Mutations in

ABCB11 cause PFIC2, BRIC2 and ICP but there have been no previous reports of association with neonatal liver failure.

The cohort was screened for all known causes of neonatal liver failure. In proband 1 congenital CMV infection was detected. This was successfully treated and the virus cleared. Despite this the liver failure progressed necessitating liver transplantation. CMV is a recognised, although rare, cause of liver failure in neonates and with successful treatment should lead to recovery and no need for liver transplantation. In this child the severity of the liver failure required liver transplantation. TT virus was identified in proband 2. This is a rarely identified virus in neonates and has not previously been reported to result in liver failure and liver transplantation. The significance of the identified TT virus is not known. Congenital CMV infection was also detected in proband 4. The liver failure in this child improved with treatment although she later succumbed to the neurological sequele of the congenital CMV infection. This child had a low γGT suggesting that BSEP expression may also be reduced and she was found to be heterozygous for the sequence variant. In the other probands no other causes of neonatal liver failure were identified.

The sequence change c.T1331C changes the amino acid from valine to alanine (GTC->GCC). Running this sequence change in Polyphen predicts the amino acid change to be benign however studies in pregnant women who develop ICP have shown the sequence variant to be a susceptibility factor for the development of cholestasis with an odds ratio of 3.0 for homozygous CC (Meier *et al*, 2008) and 1.7 for a single allele C (Dixon *et al*, 2009). BRIC and drug induced cholestasis have also been attributed to this V444A polymorphism (Meier, 2006; Kubitz, 2006; Lang, 2007).

The reduced or absent BSEP expression could be attributed to the severity of the liver parenchymal collapse and distorted architecture however other canalicular proteins continued to be expressed suggesting that this is unlikely to be the cause of the lack of BSEP expression. It is likely that the reduced BSEP expression is due to the amino acid change from valine to alanine. This is in keeping with previous findings in ICP cases where there have shown decreased BSEP expression with the V444A variant (Meier, 2006). In ICP with early onset severe symptoms was associated with the identification of more than one bile salt transport mutation with an MDR3 mutation and the polymorphism V444A.

V444A is a common variant in the general population with a heterozygosity rate of 0.48. This means the 25% of the population should be homozygous for alanine. This indicates that although the V444A polymorphism may increase susceptibility to liver disease due to reduced BSEP expression, there needs to be a second pathology (second hit) for symptoms to become manifest.

In this population of neonates three cases had additional causes of liver disease identified – CMV and TT virus. It may be that medical knowledge and technology is thus far unable to detect a second cause in the other affected infants. A second hit would also explain the lack of symptoms in the father of proband 1 and the siblings of proband 3. It should be considered that these homozygous changes will increase their susceptibility to drug induced liver disease and in the female sibling of proband 3 an increased susceptibility to ICP as well as cholestasis secondary to a combined oral contraceptive pill.

5.5.1 Conclusion

The identification of the sequence variant V444A in this cohort of neonates assists in explaining the severity of the liver disease. It is further evidence that this common variant predisposes to cholestatic liver disease as also shown by the association with ICP, BRIC, drug induced cholestasis and now with neonatal liver failure. The phenotype of neonatal low γGT liver failure following a pregnancy complicated by ICP should alert the clinician to the severity of the disease process.

Chapter 6

Conclusion

This thesis has investigated the molecular genetics of paediatric liver disease. The subject of the thesis was chosen due to my clinical work as a paediatric hepatologist in which I identified cohorts of children with liver disease in whom the aetiology remained unknown. In other children in whom an accurate diagnosis could be made due to knowledge of the molecular genetic pathology, accurate counselling, management and investigations of the patient could take place. Therefore I investigated three different presentations of liver disease to ascertain the molecular genetic pathology.

As well as improving the management of the child and family, a greater understanding of the molecular genetics of liver disease may also increase our knowledge of how the hepatocyte and cholangiocyte work and therefore identify pathways and interactions which may provide clues to intracellular therapeutic targets.

Chapter one described the molecular genetic technique known as autozygosity mapping that I used throughout this thesis. It is an extremely effective tool to identify genes associated with disease phenotypes by utilising the large amount of molecular information that can be gained from studying those from consanguineous union.

I have studied a range of different paediatric liver disease conditions which present in different clinical ways and highlight the difficulty of managing a child with liver disease in which a diagnosis is hard to make.

The work on PDI is the main focus of this thesis. Chapter 3 describes the clinical condition, the molecular genetic study of the patients and the successful identification of

TTC37 mutations which are associated with the PDI phenotype. I concentrated my studies on molecular genetic techniques and therefore formed collaborations with experts across the world to work on other research modalities such as immunohistochemistry of the liver and bowel. Although this work was extremely succseful and led to a major publication in Gastroenterology, the time taken to identify the gene allowed no futher investigation to characterise *TTC37* for this thesis.

In chapter 4 I have described the investigation of JATD patients which has enabled it to be newly classified as a ciliopathy. These findings have also led to further investigations of other conditions which may also have primary cilial involvement such as biliary atresia. The study was part of a collaboration which was necessary when studying such a heterozygous condition.Mutations in *IFT80* were only found in one family and highlights the limitations of using autozygosity in a condition which does not have a homogeneous phenotype and it is likely there will be a large number of genes which cause a similar phenotype. New technologies such as whole exome sequencing may facilitate the identification of futher JATD genes and this is an area of research I am exploring.

Chapter 5 describes the identification of a novel phenotype for *ABCB11* and leads to the question what other liver diseases of unknown aetiology can be ascribed to bile salt transport defects.

Towards the end of this period of study it became clear that newer techniques were being developed which were likely to supersede autozygosity mapping as the preferred technique for identifying new genes for rare phenotypes. Whole exome sequencing is such a development and provides information of all sequence changes within the transcribed region of the genome in one single experiment. The information gained is vast and the art will be how to interpret the information so as to identify the correct pathogenic sequence changes.

Despite new technology, all molecular genetic techniques rely on studying cohorts with exactly the same clinical phenotype. The basis of good molecular genetic research therefore will always continue to lie with the patient and the clinician.

Chapter 7

List of References

Ansley SJ, Badano JL, Blacque OE, Hill J, Hoskins BE, Leitch CC, Kim JC, Ross AJ, Eichers ER, Teslovich TM, Mah AK, Johnsen RC, Cavender JC, Lewis RA, Leroux MR, Beales PL, Katsanis N (2003). Basal body dysfunction is a likely cause of pleiotropic Bardet-Biedl syndrome. Nature 425: 628-33

Badano JL, Mitsuma N, Beales PL, Katsanis N (2006). The ciliopathies: an emerging class of human genetic disorders. Annu Rev Genomics Hum Genet. 7: 125-48

Baldari CT, Rosenbaum J (2009). Intraflagellar transport: it's not just for cilia anymore. Curr Opin Cell Biol. [Epub ahead of print]

Balistreri WF, Bezerra JA (2006). Whatever happened to "neonatal hepatitis"? Clin Liver Dis. 10: 27-53

Ballatori N, Christian WV, Lee JY, Dawson PA, Soroka CJ, Boyer JL, Madejczyk MS, Li N (2005). OST alpha-OST beta: a major basolateral bile acid and steroid transporter in human intestinal, renal, and biliary epithelia. Hepatology 42: 1270-9

Barabino AV, Torrente F, Castellano E, Erba D, Calvi A, Gandullia P (2004). "Syndromic diarrhea" may have better outcome than previously reported. J Pediatr. 144: 553-4

Barnes BG (1961). Ciliated secretory cells in the pars distalis of the mouse hypophysis. J Ultrastruct Res 5:453–467

Bennett CL, Christie J, Ramsdell F, Brunkow ME, Ferguson PJ, Whitesell L, Kelly TE, Saulsbury FT, Chance PF, Ochs HD (2001). The immune dysregulation, polyendocrinopathy, enteropathy, X-linked syndrome (IPEX) is caused by mutations of FOXP3. Nat Genet. 27: 20

Berg, M. A., Argente, J., Chernausek, S., Gracia, R., Guevara-Aguirre, J., Hopp, M., Perez-Jurado, L., Rosenbloom, A., Toledo, S. P. A., Francke, U. (1993). Diverse growth hormone receptor gene mutations in Laron syndrome. Am. J. Hum. Genet. 52: 998-1005

Berry S, Dawicki DD, Agarwal KC, Steiner M (1989). The role of microtubules in platelet secretory release. Biochim Biophys Acta. 1012:46-56

Beuers U, Pusl T (2006). Intrahepatic cholestasis of pregnancy – a heterogenous group of pregnancy related disorders. Hepatology 43: 647-9

Bilzer M, Roggel F, Gerbes AL (2006). Role of Kupffer cells in host defense and liver disease. Liver Int. 26: 1175-86

Bischof JM, Chiang AP, Scheetz TE, Stone EM, Casavant TL, Sheffield VC, Braun TA (2006)**.** Genome-wide identification of pseudogenes capable of disease-causing gene conversion. Hum Mutat. 27:545-52

Bittles AH (2008). A community genetics perspective on consanguineous marriage. Community Genet 11: 324-30

Blanchard RA, Furie BC, Jorgensen M, Kruger SF, Furie B (1981). Acquired vitamin K-dependent carboxylation deficiency in liver disease. N Engl J Med. 305: 242-8

Blom E, van de Vrugt HJ, de Vries Y, de Winter JP, Arwert F, Joenje H (2004). Multiple TPR motifs characterize the Fanconi anemia FANCG protein. DNA Repair 3: 77-84

Bomont, P., Cavalier, L., Blondeau, F., Ben Hamida, C., Belal, S., Tazir, M., Demir, E., Topaloglu, H., Korinthenberg, R., Tuysuz, B., Landrieu, P., Hentati, F., Koenig, M. (2000).The gene encoding gigaxonin, a new member of the cytoskeletal BTB/kelch repeat family, is mutated in giant axonal neuropathy. Nature Genet. 26: 370-374

Bookstein C, Xie Y, Rabenau K, Musch M, McSwine R, Rao M, Chang E (1997). Tissue distribution of Na/H exchanger isoforms NHE2 and NHE4 in rat intestine and kidney. Am J Physiol Cell Physiol. 273: 1496-1505

Brunkow ME, Jeffery EW, Hjerrild KA, Paeper B, Clark LB, Yasayko SA, Wilkinson JE, Galas D, Ziegler SF, Ramsdell F (2001). Disruption of a new forkhead/winged-helix protein, scurfin, results in the fatal lymphoproliferative disorder of the scurfy mouse. Nat Genet. 27:68-73

Bull LN, van Eijk MJ, Pawlikowska L, DeYoung JA, Juijn JA, Liao M, Klomp LW, Lomri N, Berger R, Scharschmidt BF, Knisely AS, Houwen RH, Freimer NB (1998). A gene encoding a P-type ATPase mutated in two forms of hereditary cholestasis. Nat Genet. 18: 219-24

Bull LN, Mahmoodi V, Baker AJ, Jones R, Strautnieks SS, Thompson RJ, Knisely AS (2006). VPS33B mutation with ichthyosis, cholestasis, and renal dysfunction but without arthrogryposis: incomplete ARC syndrome phenotype. J Pediatr. 148: 269-71

Bull PC, Thomas GR, Rommens JM, Forbes JR, Cox DW, (1993). The Wilson disease gene is a putative copper transporting P-type ATPase similar to the Menkes gene. Nature Genet. 5: 327-337

Bundey S, Alam H (1993). A five-year prospective study of the health of children in different ethnic groups, with particular reference to the effect of inbreeding. Eur J Hum Genet 1: 206-19

Cali JJ, Hsieh CL, Francke U, Russell DW (1991). Mutations in the bile acid biosynthetic enzyme sterol 27-hydroxylase underlie cerebrotendinous xanthomatosis. J Biol Chem. 266: 7779-83

Cardenas-Rodriguez M, Badano JL (2009). Ciliary biology: understanding the cellular and genetic basis of human ciliopathies. Am J Med Genet C Semin Med Genet. 151: 263-80

Carlton VE, Harris BZ, Puffenberger EG, Batta AK, Knisely AS, Robinson DL, Strauss KA, Shneider BL, Lim WA, Salen G, Morton DH, Bull LN (2003). Complex inheritance of familial hypercholanemia with associated mutations in TJP2 and BAAT. Nat Genet. 34: 91-6

Carruthers L, Phillips AD, Dourmashkin R, Walker-Smith JA (1985). Biochemical abnormality in brush border membrane protein of a patient with congenital microvillus atrophy. J Pediatr Gastroenterol Nutr. 4: 902-7

Carstea ED, Morris JA, Coleman KG, Loftus SK, Zhang D, Cummings C, Gu J, Rosenfeld MA, Pavan WJ, Krizman DB, Nagle J, Polymeropoulos MH, and 26 others (1997). Niemann-Pick C1 disease gene: homology to mediators of cholesterol homeostasis. Science 277: 228-231

Carter BA, Karpen SJ (2007). Intestinal failure-associated liver disease: management and treatment strategies past, present, and future. Semin Liver Dis. 27: 251-8

Chagnon P, Michaud J, Mitchell G, Mercier J, Marion J-F, Drouin E, Rasquin-Weber A, Hudson TJ, Richter A (2002). A missense mutation (R565W) in cirhin (FLJ14728) in North American Indian childhood cirrhosis. Am. J. Hum. Genet. 71: 1443-1449

Chavanas, S., Bodemer, C., Rochat, A., Hamel-Teillac, D., Ali, M., Irvine, A. D., Bonafe, J.-L., Wilkinson, J., Taieb, A., Barrandon, Y., Harper, J. I., de Prost, Y., Hovnanian, A (2000). Mutations in SPINK5, encoding a serine protease inhibitor, cause Netherton syndrome. Nature Genet. 25: 141-142

Chelly, J., Tumer, Z., Tonnesen, T., Petterson, A., Ishikawa-Brush, Y., Tommerup, N., Horn, N., Monaco, A. P (1993). Isolation of a candidate gene for Menkes disease that encodes a potential heavy metal binding protein. Nature Genet. 3: 14-19

Cheng JB, Jacquemin E, Gerhardt M, Nazer H, Cresteil D, Heubi JE, Setchell KD, Russell DW (2003). Molecular genetics of 3beta-hydroxy-Delta5-C27-steroid oxidoreductase deficiency in 16 patients with loss of bile acid synthesis and liver disease. J Clin Endocrinol Metab. 88: 1833-41

Chiang JY (1998). Regulation of bile acid synthesis. Front Biosci. 3: 176-93

Choudhuri S, Klaassen CD (2006). Structure, function, expression, genomic organization, and single nucleotide polymorphisms of human ABCB1 (MDR1), ABCC (MRP), and ABCG2 (BCRP) efflux transporters. Int J Toxicol. 25: 231-59

Ciferri S, Emiliani C, Guglielmini G, Orlacchio A, Nenci GG, Gresele P (2000). Platelets release their lysosomal content in vivo in humans upon activation. Thromb Haemost. 83:157-64

Clayton RJ, Iber FL, Ruebner BH, McKusick VA (1969). Byler disease. Fatal familial intrahepatic cholestasis in an Amish kindred. Am J Dis Child. 117: 112-24

Clotman F, Lannoy VJ, Reber M, Cereghini S, Cassiman D, Jacquemin P, Roskams T, Rousseau GG, Lemaigre FP (2002). The onecut transcription factor HNF6 is required for normal development of the biliary tract. Development 128: 1819-28

Cormier-Daire V, Bonnefont JP, Rustin P, Maurage C, Ogler H, Schmitz J, Ricour C, Saudubray JM, Munnich A, Rötig A (1994). Mitochondrial DNA rearrangements with onset as chronic diarrhea with villous atrophy. J Pediatr. 124: 63-70

Damen G, de Klerk H, Huijmans J, den Hollander J, Sinaasappel M (2004). Gastrointestinal and other clinical manifestations in 17 children with congenital disorders of glycosylation type Ia, Ib, and Ic. J Pediatr Gastroenterol Nutr. 38: 282-7

Dawson PA, Hubbert M, Haywood J, Craddock AL, Zerangue N, Christian WV, Ballatori N (2005). The heteromeric organic solute transporter alpha-beta, Ost-alpha-Ost-beta, is an ileal basolateral bile acid transporter. J Biol Chem. 280: 6960-8

Deleuze JF, Jacquemin E, Dubuisson C, Cresteil D, Dumont M, Erlinger S, Bernard O, Hadchouel M (1996). Defect of multidrug-resistance 3 gene expression in a subtype of progressive familial intrahepatic cholestasis. Hepatology 23: 904-8

Derry, J. M. J., Gormally, E., Means, G. D., Zhao, W., Meindl, A., Kelley, R. I., Boyd, Y., Herman, G. E (1999). Mutations in a delta(8)-delta(7) sterol isomerase in the tattered mouse and X-linked dominant chondrodysplasia punctata. Nature Genet. 22: 286-290

de Vries E, Visser DM, van Dongen JJ, Jacobs CJ, Hoekstra JH, van Tol MJ (2000). Oligoclonal gammopathy in phenotypic diarrhea. J Pediatr Gastroenterol Nutr. 30:349-50

Diamanti A, Basso MS, Castro M, Di Ciommo V, Bracci F, Ferretti F, Pietrobattista A, Gambarara M (2008). Irreversible intestinal failure: prevalence and prognostic factors. J Pediatr Gastroenterol Nutr. 47: 450-7

Dixon PH, Weerasekera N, Linton KJ, Donaldson O, Chambers J, Egginton E, Weaver J, Nelson-Piercy C, de Swiet M, Warnes G, Elias E, Higgins CF, Johnston DG, McCarthy MI, Williamson C (2000). Heterozygous MDR3 missense mutation associated with intrahepatic cholestasis of pregnancy: evidence for a defect in protein trafficking. Hum Mol Genet. 9: 1209-17

Dixon PH, van Mil SW, Chambers J, Strautnieks S, Thompson RJ, Lammert F, Kubitz R, Keitel V, Glantz A, Mattsson LA, Marschall HU, Molokhia M, Moore GE, Linton KJ, Williamson C. Contribution of variant alleles of ABCB11 to susceptibility to intrahepatic cholestasis of pregnancy. Gut. 58: 537-44

Donner MG, Schumacher S, Warskulat U, Heinemann J, Häussinger D (2007). Obstructive cholestasis induces TNF-alpha- and IL-1 -mediated periportal downregulation of Bsep and zonal regulation of Ntcp, Oatp1a4, and Oatp1b2. Am J Physiol Gastrointest Liver Physiol. 293: G1134-46

D'souza S, Garcia-Cabado A, Yu F, Teter K, Lukacs G, Skorecki K, Moore H, Orlowski J, Grinstein S (1998). The epithelial sodium-hydrogen antiporter Na/H exchanger 3 accumulates and is functional in recycling endosomes. J Biol Chem. 273: 2035-2043

Dudeja P, Rao D, Syed I, Joshi V, Dahdal R, Gardner C, Risk M, Schmidt L, Bavishi D, Kim K, Harig J, Goldstein J, Layden T, Ramaswamy K (1996). Intestinal distribution of human Na/H exchanger isolforms NHE1, NHE2 and NHE3 mRNA. Am J Physiol Gastrointest Liver Physiol. 271: 483-493

Dweikat I, Sultan M, Maraqa N, Hindi T, Abu-Rmeileh S, Abu-Libdeh B (2007). Tricho-hepato-enteric syndrome: a case of hemochromatosis with intractable diarrhea, dysmorphic features, and hair abnormality. Am J Med Genet A. 143: 581-3

Egritas O, Dalgic B, Onder M (2009). Tricho-hepato-enteric syndrome presenting with mild colitis. Eur J Pediatr. 168: 933-5

Elferink RO, Groen AK. (2002). Genetic defects in hepatobiliary transport. Biochim Biophys Acta. 1586: 129-45

Fabre A, André N, Breton A, Broué P, Badens C, Roquelaure B (2007). Intractable diarrhea with "phenotypic anomalies" and tricho-hepato-enteric syndrome: two names for the same disorder. Am J Med Genet A. 143: 584-8

Fabre A, Roquelaure B, Lacoste C, André N, Sarles J, Breton A, Martinez-Vinson C, Cezard JP, Colomb V, Goulet O, Levy N, Badens C (2009). Exclusion of EGFR, HRAS, DSP, JUP, CTNNB1, PLEC1, and EPPK1 as functional candidate genes in 7 families with syndromic diarrhoea. J Pediatr Gastroenterol Nutr. 48: 501-3

Fabris L, Cadamuro M, Libbrecht L, Raynaud P, Spirlì C, Fiorotto R, Okolicsanyi L, Lemaigre F, Strazzabosco M, Roskams T (2008). Epithelial expression of angiogenic growth factors modulate arterial vasculogenesis in human liver development. Hepatology 47: 719-28

Fell JM, Miller MP, Finkel Y, Booth IW (1992). Congenital sodium diarrhea with a partial defect in jejunal brush border membrane sodium transport, normal rectal transport, and resolving diarrhea. J Pediatr Gastroenterol Nutr. 15: 112-6

Figus A, Angius A, Loudianos G, Bertini C, Dessi V, Loi A, Deiana M, Lovicu M, Olla N, Sole G, De Virgiliis S, Lilliu F (1995). Molecular pathology and haplotype analysis of Wilson disease in Mediterranean populations Am. J. Hum. Genet 57: 1318-1324

Fliegauf M, Omran H (2006). Novel tools to unravel molecular mechanisms in cilia-related disorders. Trends Genet. 22: 241-5

Fontaine O, Garner P, Bhan MK (2007). Oral rehydration therapy: the simple solution for saving lives. BMJ 334: s14

Frankenberg T, Miloh T, Chen FY, Ananthanarayanan M, Sun AQ, Balasubramaniyan N, Arias I, Setchell KD, Suchy FJ, Shneider BL (2008). The membrane protein ATPase class I type 8B member 1 signals through protein kinase C zeta to activate the farnesoid X receptor. Hepatology. 48: 1896-905

Garrod AE, Hele TS (1905). The uniformity of the homogentisic acid excretion in alkaptonuria. J Physiol 33: 198-205

Garver WS, Francis GA, Jelinek D, Shepherd G, Flynn J, Castro G, Vockley CW, Coppock DL, Pettit KM, Heidenreich RA, Meaney FJ (2007). The National Niemann-Pick C1 Disease Database: report of clinical features and health problems. Am. J. Med. Genet 143: 1204-1211

Gawenis L, Stein X, Shull G, Schultheis P, Woo A, Walker N, Clarke L (2002). Intestinal NaCl transport in NHE2 and NHE3 knockout mice. Am J physiol Gastrintest Liver Physiol. 282: 776-784

Girault D, Goulet O, Le Deist F, Brousse N, Colomb V, Césarini JP, de Potter S, Canioni D, Griscelli C, Fischer A, *et al* (1994). Intractable infant diarrhea associated with phenotypic abnormalities and immunodeficiency. J Pediatr. 125: 36-42

Gissen P, Johnson CA, Morgan NV, Stapelbroek JM, Forshew T, Cooper WN, McKiernan PJ, Klomp LW, Morris AA, Wraith JE, McClean P, Lynch SA, Thompson RJ, Lo B, Quarrell OW, Di Rocco M, Trembath RC, Mandel H, Wali S, Karet FE, Knisely AS, Houwen RH, Kelly DA, Maher ER (2004). Mutations in VPS33B, encoding a regulator of SNARE-dependent membrane fusion, cause arthrogryposis-renal dysfunction-cholestasis (ARC) syndrome. Nat Genet. 36: 400-4

Goodarzi MO, Xu N, Cui J, Guo X, Chen YI, Azziz R (2008). Small glutamine-rich tetratricopeptide repeat-containing protein alpha (SGTA), a candidate gene for polycystic ovary syndrome. Hum Reprod. 23: 1214-9

Gotthardt D, Runz H, Keitel V, Fischer C, Flechtenmacher C, Wirtenberger M, Weiss KH, Imparato S, Braun A, Hemminki K, Stremmel W, Rüschendorf F, Stiehl A, Kubitz R, Burwinkel B, Schirmacher P, Knisely AS, Zschocke J, Sauer P (2008). A mutation in the canalicular phospholipid transporter gene, ABCB4, is associated with cholestasis, ductopenia, and cirrhosis in adults. Hepatology. 48: 1157-66

Goulet OJ, Brousse N, Canioni D, Walker-Smith JA, Schmitz J, Phillips AD (1998). Syndrome of intractable diarrhoea with persistent villous atrophy in early childhood: a clinicopathological survey of 47 cases. J Pediatr Gastroenterol Nutr. 26: 151-61

Goulet O, Salomon J, Ruemmele F, de Serres NP, Brousse N (2007). Intestinal epithelial dysplasia (tufting enteropathy). Orphanet J Rare Dis 2:20

Grizot S, Fieschi F, Dagher MC, Pebay-Peyroula E (2001). The active N-terminal region of p67phox. Structure at 1.8 A resolution and biochemical characterizations of the A128V mutant implicated in chronic granulomatous disease. J Biol Chem. 276: 21627-31

Guyton (1991). Textbook of Medical Physiology 8th Edition.W.B.Saunders Company, Philadelphia

Hadj-Rabia S, Baala L, Vabres P, Hamel-Teillac D, Jacquemin E, Fabre M, Lyonnet S, De Prost Y, Munnich A, Hadchouel M, Smahi A (2004). Claudin-1 gene mutations in neonatal sclerosing cholangitis associated with ichthyosis: a tight junction disease. Gastroenterology 127: 1386-90

Hartman H, Smith TF (2009). The evolution of the cilium and the eukaryotic cell. Cell Motil Cytoskeleton. 66: 215-9

Hayward, C. P. M., Cramer, E. M., Kane, W. H., Zheng, S., Bouchard, M., Masse, J.-M., Rivard, G. E (1997). Studies of a second family with the Quebec platelet disorder: evidence that the degradation of the alpha-granule membrane and its soluble contents are not secondary to a defect in targeting proteins to alpha-granules. Blood 89: 1243-1253

Heinz-Erian P, Müller T, Krabichler B, Schranz M, Becker C, Rüschendorf F, Nürnberg P, Rossier B, Vujic M, Booth IW, Holmberg C, Wijmenga C, Grigelioniene G, Kneepkens CM, Rosipal S, Mistrik M, Kappler M, Michaud L, Dóczy LC, Siu VM, Krantz M, Zoller H, Utermann G, Janecke AR (2009). Mutations in SPINT2 cause a syndromic form of congenital sodium diarrhea. Am J Hum Genet. 84: 188-96

Höglund P, Haila S, Socha J, Tomaszewski L, Saarialho-Kere U, Karjalainen-Lindsberg ML, Airola K, Holmberg C, de la Chapelle A, Kere J (1996). Mutations of the Down-regulated in adenoma (DRA) gene cause congenital chloride diarrhoea. Nat Genet 14: 316-9

Hogrefe C, Römermann D, Friedrich B (1984). Alcaligenes eutrophus hydrogenase genes (Hox). J Bacteriol 158: 43-8

Holt JA, Luo G, Billin AN, Bisi J, McNeill YY, Kozarsky KF, Donahee M, Wang DY, Mansfield TA, Kliewer SA, Goodwin B, Jones SA (2003). Definition of a novel growth factor-dependent signal cascade for the suppression of bile acid biosynthesis. Genes Dev. 17: 1581-91

Holmsen H (1994). Significance of testing platelet functions in vitro. Eur J Clin Invest. 24 Suppl 1:3-8

Holzinger A, Maier EM, Bück C, Mayerhofer PU, Kappler M, Haworth JC, Moroz SP, Hadorn HB, Sadler JE, Roscher AA (2002). Mutations in the proenteropeptidase gene are the molecular cause of congenital enteropeptidase deficiency. Am J Hum Genet. 70: 20-5

Hsu JC, Laz T, Mohn KL, Taub R (1991). Identification of LRF-1, a leucine-zipper protein that is rapidly and highly induced in regenerating liver. Proc Natl Acad Sci U S A. 88: 3511-5

Hunter MP, Wilson CM, Jiang X, Cong R, Vasavada H, Kaestner KH, Bogue CW (2007). The homeobox gene Hhex is essential for proper hepatoblast differentiation and bile duct morphogenesis. Dev Biol 308: 355-67

Imbach, T., Burda, P., Kuhnert, P., Wevers, R. A., Aebi, M., Berger, E. G., Hennet, T. (1999). A mutation in the human ortholog of the Saccharomyces cerevisiae ALG6 gene causes carbohydrate-deficient glycoprotein syndrome type-Ic. Proc. Nat. Acad. Sci. 96: 6982-6987

Johnson GC, Esposito L, Barratt BJ, Smith AN, Heward J, Di Genova G, Ueda H, Cordell HJ, Eaves IA, Dudbridge F, Twells RC, Payne F, Hughes W, Nutland S, Stevens H, Carr P, Tuomilehto-Wolf E, Tuomilehto J, Gough SC, Clayton DG, Todd JA (2001). Haplotype tagging for the identification of common disease genes. Nat Genet. 29: 233-7

Kaback MM, Rimoin DL, O'Brien JS (1977). Tay-Sachs Disease: Screening and Prevention. New York: Alan R. Liss (pub.)

Katsanis N, Ansley SJ, Badano JL, Eichers ER, Lewis RA, Hoskins BE, Scambler PJ, Davidson WS, Beales PL, Lupski JR (2001). Triallelic inheritance in Bardet-Biedl syndrome, a Mendelian recessive disorder Science 293: 2256-9

Kamiya A, Inoue Y, Gonzalez FJ (2003). Role of the hepatocyte nuclear factor 4alpha in control of the pregnane X receptor during fetal liver development. Hepatology 37:1375-84

Kiela PR, Laubitz D, Larmonier CB, et al (2009). Changes in mucosal homeostasis predisposes NHE3 knockout mice to increased susceptibility to DSS-induced epithelial injury. Gastroenterology 137: 965-75

Klomp LW, Vargas JC, van Mil SW, Pawlikowska L, Strautnieks SS, van Eijk MJ, Juijn JA, Pabón-Peña C, Smith LB, DeYoung JA, Byrne JA, Gombert J, van der Brugge G, Berger R, Jankowska I, Pawlowska J, Villa E, Knisely AS, Thompson RJ, Freimer NB, Houwen RH, Bull LN (2004). Characterization of mutations in ATP8B1 associated with hereditary cholestasis. Hepatology. 40: 27-38

Klammt S, Mitzner S, Strange J, Brinkmann B, Drewelow B, Emmrich J, Liebe S, Schmidt R (2007). Albumin-binding function is reduced in patients with decompensated cirrhosis and correlates inversely with severity of liver disease assessed by model for end-stage liver disease. Eur J Gastroenterol Hepatol. 19: 257-63

Knisely AS, Strautnieks SS, Meier Y, Stieger B, Byrne JA, Portmann BC, Bull LN, Pawlikowska L, Bilezikçi B, Ozçay F, László A, Tiszlavicz L, Moore L, Raftos J, Arnell H, Fischler B, Németh A, Papadogiannakis N, Cielecka-Kuszyk J, Jankowska I, Pawłowska J, Melín-Aldana H, Emerick KM, Whitington PF, Mieli-Vergani G, Thompson RJ (2006). Hepatocellular carcinoma in ten children under five years of age with bile salt export pump deficiency. Hepatology. 44: 478-86

Koh ET, Chan DK, Ho LY, Yeo CL (1998). Congenital sodium diarrhoea in an Indian girl. Singapore Med J. 39: 468-70

Kubitz R, Keitel V, Scheuring S, Köhrer K, Häussinger D (2006). Benign recurrent intrahepatic cholestasis associated with mutations of the bile salt export pump. J Clin Gastroenterol. 40: 171-5

Kullak-Ublick GA, Stieger B, Meier PJ (2004). Enterohepatic bile salt transporters in normal physiology and liver disease. Gastroenterology. 126: 322-42

Kuokkanen M, Kokkonen J, Enattah NS, Ylisaukko-Oja T, Komu H, Varilo T, Peltonen L, Savilahti E, Jarvela I (2006). Mutations in the translated region of the lactase gene (LCT) underlie congenital lactase deficiency. Am J Hum Genet. 78: 339-44

Lam P, Pearson CL, Soroka CJ, Xu S, Mennone A, Boyer JL (2007). Levels of plasma membrane expression in progressive and benign mutations of the bile salt export pump (Bsep/Abcb11) correlate with severity of cholestatic diseases. Am J Physiol Cell Physiol. 293: C1709-16

Lamaigre F (2008). Notch signaling in bile duct development: new insights raise new questions. Hepatology. 48: 358-360

Landers ES, Botstein D (1987). Homozygosity mapping: a way to map human recessive traits with the DNA of inbred children. Science. 236: 1567-70

Landers MC, Schroeder TL (2003). Intractable diarrhea of infancy with facial dysmorphism, trichorrhexis nodosa, and cirrhosis. Pediatr Dermatol. 20: 432-5

Lang C, Meier Y, Stieger B, Beuers U, Lang T, Kerb R, Kullak-Ublick GA, Meier PJ, Pauli-Magnus C (2007). Mutations and polymorphisms in the bile salt export pump and the multidrug resistance protein 3 associated with drug-induced liver injury. Pharmacogenet Genomics. 17: 47-60

Lee CS, Friedman JR, Fulmer JT, Kaestner KH (2005). The initiation of liver development is dependent on Foxa transcription factors. Nature 435: 944-7

Lefebvre P, Cariou B, Lien F, Kuipers F, Staels B (2009). Role of bile acids and bile acid receptors in metabolic regulation. Physiol Rev 89: 147-191

Lemna WK, Feldman GL, Kerem B, Fernbach SD, Zevkovich EP, O'Brien WE, Riordan JR, Collins FS, Tsui CL, Beaudet AL (1990). Mutation analysis for heterozygote detection and the prenatal diagnosis of cystic fibrosis New Eng. J. Med. 322: 291-296

Lemonde HA, Custard EJ, Bouquet J, Duran M, Overmars H, Scambler PJ, Clayton PT (2003). Mutations in SRD5B1 (AKR1D1), the gene encoding delta(4)-3-oxosteroid 5beta-reductase, in hepatitis and liver failure in infancy. Gut 52: 1494-9

Li A, Davila S, Furu L, Qian Q, Tian X, Kamath PS, King BF, Torres VE, Somlo S (2003).Mutations in PRKCSH cause isolated autosomal dominant polycystic liver disease. Am J Hum Genet. 72: 691-703

Lo B, Li L, Gissen P, Christensen H, McKiernan PJ, Ye C, Abdelhaleem M, Hayes JA, Williams MD, Chitayat D, Kahr WH (2005). Requirement of VPS33B, a member of the Sec1/Munc18 protein family, in megakaryocyte and platelet alpha-granule biogenesis. Blood. 106: 4159-66

Lu TT, Makishima M, Repa JJ, Schoonjans K, Kerr TA, Auwerx J, Mangelsdorf DJ (2000). Molecular basis for feedback regulation of bile acid synthesis by nuclear receptors. Mol Cell. 6: 507-15

Maas, N. M. C., de Putte, T. V., Melotte, C., Francis, A., Schrander-Stumpel, C. T. R. M., Sanlaville, D., Genevieve, D., Lyonnet, S., Dimitrov, B., Devriendt, K., Fryns, J.-P., Vermeesch, J. R (2007). The C20orf133 gene is disrupted in a patient with Kabuki syndrome. J. Med. Genet. 44: 562-569

Martín MG, Turk E, Lostao MP, Kerner C, Wright EM (1996). Defects in Na+/glucose cotransporter (SGLT1) trafficking and function cause glucose-galactose malabsorption. Nat Genet. 12: 216-20

Masyuk AI, Masyuk TV, Splinter PL, Huang BQ, Stroope AJ, LaRusso NF (2006). Cholangiocyte cilia detect changes in luminal fluid flow and transmit them into intracellular Ca2+ and cAMP signaling. Gastroenterology. 131: 911-20

Mazariegos GV, Soltys K, Bond G, Girnita A, Machaidze Z, Jaffe R, Green M, Martin D, Zeevi A, Squires R, Perez G, Abu-Elmagd K, Sindhi R (2008). Pediatric intestinal retransplantation: techniques, management, and outcomes. Transplantation. 86: 1777-82

McCollum D, Monosov E, Subramani S (1993). The pas8 mutant of Pichia pastoris exhibits the peroxisomal protein import deficiencies of Zellweger syndrome cells--the PAS8 protein binds to the COOH-terminal tripeptide peroxisomal targeting signal, and is a member of the TPR protein family. J Cell Biol. 121: 761-74

McDaniell R, Warthen DM, Sanchez-Lara PA, Pai A, Krantz ID, Piccoli DA, Spinner NB (2006). NOTCH2 mutations cause Alagille syndrome, a heterogeneous disorder of the notch signaling pathway. Am J Hum Genet. 79: 169-73

Meier PJ (1995). Molecular mechanisms of hepatic bile salt transport from sinusoidal blood into bile. Am J Physiol. 269: G801-12

Meier Y, Zodan T, Lang C, Zimmermann R, Kullak-Ublick GA, Meier PJ, Stieger B, Pauli-Magnus C (2008). Increased susceptibility for intrahepatic cholestasis of pregnancy and contraceptive-induced cholestasis in carriers of the 1331T>C polymorphism in the bile salt export pump. World J Gastroenterol. 14: 38-45

Meng X, Lu X, Morris CA, Keating MT (1998). A novel human gene FKBP6 is deleted in Williams syndrome. Genomics. 52: 130-7

Mercer, J. F. B., Livingston, J., Hall, B., Paynter, J. A., Begy, C., Chandrasekharappa, S., Lockhart, P., Grimes, A., Bhave, M., Siemieniak, D., Glover, T. W (1993). Isolation of a partial candidate gene for Menkes disease by positional cloning. Nature Genet. 3: 20-25

Messenger A, Dawber R (1997). The physiology and embryology of hair growth. Diseases of the hair and scalp. BlackwellScience Inc

Mighell AJ, Smith NR, Robinson PA, Markham AF (2006). Vertebrate pseudogenes. FEBS Lett. 468:109-14

Moszynski P (2006). Unicef warns of "epidemic" of childhood malnutrition. BMJ 332: 1051

Müllenbach R, Bennett A, Tetlow N, Patel N, Hamilton G, Cheng F, Chambers J, Howard R, Taylor-Robinson SD, Williamson C. ATP8B1 mutations in British cases with intrahepatic cholestasis of pregnancy. Gut. 54: 829-34

Muller M, Jasmin J, Monteil R, Loubiere R (1991). Embryology of the hair follicle. Early human development. 26: 159-166

Müller T, Hess MW, Schiefermeier N, Pfaller K, Ebner HL, Heinz-Erian P, Ponstingl H, Partsch J, Röllinghoff B, Köhler H, Berger T, Lenhartz H, Schlenck B, Houwen RJ, Taylor CJ, Zoller H, Lechner S, Goulet O, Utermann G, Ruemmele FM, Huber LA, Janecke AR (2008). MYO5B mutations cause microvillus inclusion disease and disrupt epithelial cell polarity. Nat Gen 40: 1163-508

Müller T, Wijmenga C, Phillips AD, Janecke A, Houwen RH, Fischer H, Ellemunter H, Frühwirth M, Offner F, Hofer S, Müller W, Booth IW, Heinz-Erian P (2000). Congenital sodium diarrhea is an autosomal recessive disorder of sodium/proton exchange but unrelated to known candidate genes. Gastroenterology 119: 1506-13

Nagaki M, Moriwaki H (2008). Transcription factor HNF and hepatocyte differentiation. Hepatol Res 38: 961-9

Nigg EA, Raff JW (2009). Centrioles, centrosomes, and cilia in health and disease. Cell. 139: 663-78

Norio R (2003). The Finnish Disease Heritage III: the individual diseases. Hum Genet. 112: 470-526

Nurden, A. T., Nurden, P (2007). Inherited thrombocytopenias Haematologica 92: 1158-1164

Oda T, Elkahloun AG, Pike BL, Okajima K, Krantz ID, Genin A, Piccoli DA, Meltzer PS, Spinner NB, Collins FS, Chandrasekharappa SC (1997). Mutations in the human Jagged1 gene are responsible for Alagille syndrome. Nat Genet. 16: 235-42

Oelkers P, Kirby LC, Heubi JE, Dawson PA (1997). Primary bile acid malabsorption caused by mutations in the ileal sodium-dependent bile acid transporter gene (SLC10A2). J Clin Invest. 99: 1880-7

Ohura T, Kobayashi K, Tazawa Y, Nishi I, Abukawa D, Sakamoto O, Iinuma K, Saheki T (2001). Neonatal presentation of adult-onset type II citrullinemia. Hum. Genet. 108: 87-90

Paulusma CC, Folmer DE, Ho-Mok KS, de Waart DR, Hilarius PM, Verhoeven AJ, Oude Elferink RP (2008). ATP8B1 requires an accessory protein for endoplasmic reticulum exit and plasma membrane lipid flippase activity. Hepatology. 47: 268-78

Paznekas, W. A., Boyadjiev, S. A., Shapiro, R. E., Daniels, O., Wollnik, B., Keegan, C. E., Innis, J. W., Dinulos, M. B., Christian, C., Hannibal, M. C., Jabs, E. W (2003). Connexin 43 (GJA1) mutations cause the pleiotropic phenotype of oculodentodigital dysplasia. Am. J. Hum. Genet. 72: 408-418

Pedersen LB, Rosenbaum JL (2008). Intraflagellar transport (IFT) role in ciliary assembly, resorption and signalling. Curr Top Dev Biol. 85:23-61

Pennarun G, Escudier E, Chapelin C, Bridoux A-M, Cacheux V, Roger G, Clement A, Goossens M, Amselem S, Duriez, B (1999). Loss-of-function mutations in a human gene related to Chlamydomonas reinhardtii dynein IC78 result in primary ciliary dyskinesia. Am. J. Hum. Genet. 65: 1508-1519

Phillips MJ, Azuma T, Meredith S-LM, Squire JA, Ackerley CA, Pluthero FG, Roberts EA, Superina RA, Levy GA, Marsden, PA (2003). Abnormalities in villin gene expression and canalicular microvillus structure in progressive cholestatic liver disease of childhood. Lancet 362: 1112-1119

Price, V. H., Odom, R. B., Jones, F. T., Ward, W. H (1980). Trichothiodystrophy: sulfur-deficient brittle hair. In: Brown, A. C., Crounse, R. G. Hair, Trace Elements, and Human Illness. New York: Praeger Publ. (pub.)

Reifen RM, Cutz E, Griffiths AM, Ngan BY, Sherman PM (1994). Tufting enteropathy: a newly recognized clinicopathological entity associated with refractory diarrhea in infants. J Pediatr Gastroenterol Nutr 18: 379-85

Reyes H, Baez M, Ganzalez M, Hernandez I, Palma J, Ribalta J (2000). Selenium, zinc and copper plasma levels in intrahepatic cholestasis of pregnancy, in normal pregnancies and in healthy individuals. J Hepatol 32: 542-9

Rhodes CJ, Brennan SO, Hutton JC (1989). Proalbumin to albumin conversion by a proinsulin processing endopeptidase of insulin secretory granules. J Biol Chem. 264: 14240-5

Riordan JR, Rommens JM, Kerem B, Alon N, Rozmahel R, Grzelczak Z, Zielenski J, Lok S, Plavsic N, Chou JL, et al (1989). Identification of the cystic fibrosis gene: cloning and characterization of complementary DNA. Science 245: 1066-73

Rioseco A, Ivankovic M, Manzur A, Hamed F, Kato S, Parer J (1994). Intrahepatic cholestasis of pregnancy: a retrospective case-control study of perinatal outcome. Am J Obstet Gynecol 170: 890-5

Rizzo G, Renga B, Mencarelli A, Pellicciari R, Fiorucci S (2005). Role of FXR in regulating bile acid homeostasis and relevance for human diseases. Curr Drug Targets Immune Endocr Metabol Disord. 5: 289-303

Russell D (2003). The enzymes, regulation, and genetics of bile acid synthesis. Annu Rev Biochem. 72: 137-174

Saarela T, Similä S, Koivisto M (1995). Hypercalcemia and nephrocalcinosis in patients with congenital lactase deficiency. J Pediatr. 127: 920-3

Sadler TW (1990). Langman's Medical Embryology. 6th edition. Williams and Wilkins, Maryland

Salonen R, Norio R (1984). The Meckel syndrome in Finland: epidemiologic and genetic aspects. Am. J. Med. Genet 18: 691-698

Salvestrini C, McGrath JA, Ozoemena L, Husain K, Buhamrah E, Sabery N, Leichtner A, Rufo PA, Perez-Atayde A, Orteu CH, Torrente F, Heuschkel RB, Thomson MA, Murch

SH (2008). Desquamative enteropathy and pyloric atresia without skin disease caused by a novel intracellular beta4 integrin mutation. J Pediatr Gastroenterol Nutr. 47: 585-91

Sanyal S, Båvner A, Haroniti A, Nilsson LM, Lundåsen T, Rehnmark S, Witt MR, Einarsson C, Talianidis I, Gustafsson JA, Treuter E (2007). Involvement of corepressor complex subunit GPS2 in transcriptional pathways governing human bile acid biosynthesis. Proc Natl Acad Sci U S A. 104(40):15665-70

Sashiyama H, Shino Y, Sakao S, Shimada H, Kobayashi S, Ochiai T, Shirasawa H (2002). Alteration of integrin expression relates to malignant progression of human papillomavirus-immortalized esophageal keratinocytes. Cancer Lett. 177: 21-8

Savander M, Ropponen A, Avela K, Weerasekera N, Cormand B, Hirvioja ML, Riikonen S, Ylikorkala O, Lehesjoki AE, Williamson C, Aittomäki K. Genetic evidence of heterogeneity in intrahepatic cholestasis of pregnancy. Gut. 52: 1025-9

Schneider G, Paus TC, Kullak-Ublick GA, Meier PJ, Wienker TF, Lang T, van de Vondel P, Sauerbruch T, Reichel C (2007). Linkage between a new splicing site mutation in the MDR3 alias ABCB4 gene and intrahepatic cholestasis of pregnancy. Hepatology. 45: 150-8

Setchell KD, Schwarz M, O'Connell NC, Lund EG, Davis DL, Lathe R, Thompson HR, Weslie Tyson R, Sokol RJ, Russell DW (1998). Identification of a new inborn error in bile acid synthesis: mutation of the oxysterol 7alpha-hydroxylase gene causes severe neonatal liver disease. J Clin Invest. 102: 1690-703

Silengo M, Valenzise M, Spada M, Ferrero GB, Ferraris S, Dassi P, Jarre L (2003). Hair anomalies as a sign of mitochondrial disease. Eur J Pediatr. 162:459-61

Sing CF, Risser DR, Howatt WF, Erickson RP (1982). Phenotypic heterogeneity in cystic fibrosis Am. J. Med. Genet 13: 179-195

Sivagnanam M, Mueller JL, Lee H, Chen Z, Nelson SF, Turner D, Zlotkin SH, Pencharz PB, Ngan BY, Libiger O, Schork NJ, Lavine JE, Taylor S, Newbury RO, Kolodner RD, Hoffman HM (2008). Identification of EpCAM as the gene for congenital tufting enteropathy. Gastroenterology 135: 429-37

Skrede S, Björkhem I, Kvittingen EA, Buchmann MS, Lie SO, East C, Grundy S (1986). Demonstration of 26-hydroxylation of C27-steroids in human skin fibroblasts, and a deficiency of this activity in cerebrotendinous xanthomatosis. J Clin Invest. 78: 729-35

Smith MP, Cramer EM, Savidge GF (1997). Megakaryocytes and platelets in alpha-granule disorders. Baillieres Clin Haematol. 10:125-48

Smushkin G, Vella A (2010). Genetics of type 2 diabetes. Curr Opin Clin Nutr Metab Care. Epub May 13 2010

Splinter PL, Lazaridis KN, Dawson PA, LaRusso NF (2006). Cloning and expression of SLC10A4, a putative organic anion transport protein. World J Gastroenterol.12: 6797-805

Stankler L, Lloyd D, Pollitt RJ, Gray ES, Thom H, Russell G (1982). Unexplained diarrhoea and failure to thrive in 2 siblings with unusual facies and abnormal scalp hair shafts: a new syndrome. Arch Dis Child. 57: 212-6

Steiger B, Fattinger K, Madon J, Kullak-Ublick G, Meier P (2000). Drug- and estrogen-induced cholestasis through inhibition of the hepatocellular bile salt export pump (Bsep) of a rat liver. Gastroenterology 118: 422-30

Strautnieks SS, Bull LN, Knisely AS, Kocoshis SA, Dahl N, Arnell H, Sokal E, Dahan K, Childs S, Ling V, Tanner MS, Kagalwalla AF, Németh A, Pawlowska J, Baker A, Mieli-Vergani G, Freimer NB, Gardiner RM, Thompson RJ (1998). A gene encoding a liver-specific ABC transporter is mutated in progressive familial intrahepatic cholestasis. Nat Genet. 20: 233-8

Trauner M, Boyer JL (2003). Bile salt transporters: molecular characterization, function, and regulation. Physiol Rev. 83: 633-71

Tsujii H, König J, Rost D, Stöckel B, Leuschner U, Keppler D (1999). Exon-intron organization of the human multidrug-resistance protein 2 (MRP2) gene mutated in Dubin-Johnson syndrome. Gastroenterology 117: 653-60

Turkel S B, Diehl E J, Richmond J A (1985). Necropsy findings in neonatal asphyxiating thoracic dystrophy. J. Med. Genet 22: 112-118

Vabres, P., Lacombe, D., Rabinowitz, L. G., Aubert, G., Anderson, C. E., Taieb, A., Bonafe, J.-L., Hors-Cayla, M.-C (1995). The gene for Bazex-Dupre-Christol syndrome maps to chromosome Xq. J. Invest. Derm. 105: 87-91

Vallejo M, Briz O, Serrano MA, Monte MJ, Marin JJ (2006). Potential role of trans-inhibition of the bile salt export pump by progesterone metabolites in the etiopathogenesis of intrahepatic cholestasis of pregnancy. J Hepatol. 44: 1150-7

Van Mil SW, Milona A, Dixon PH, Mullenbach R, Geenes VL, Chambers J, Shevchuk V, Moore GE, Lammert F, Glantz AG, Mattsson LA, Whittaker J, Parker MG, White R, Williamson C (2007). Functional variants of the central bile acid sensor FXR identified in intrahepatic cholestasis of pregnancy. Gastroenterology. 133: 507-16

van Mil SW, van der Woerd WL, van der Brugge G, Sturm E, Jansen PL, Bull LN, van den Berg IE, Berger R, Houwen RH, Klomp LW. Benign recurrent intrahepatic cholestasis type 2 is caused by mutations in ABCB11.Gastroenterology. 127: 379-84

van Ooteghem NA, Klomp LW, van Berge-Henegouwen GP, Houwen RH (2002). Benign recurrent intrahepatic cholestasis progressing to progressive familial intrahepatic cholestasis: low GGT cholestasis is a clinical continuum. J Hepatol. 36: 439-43

Varga M, Obrist P, Schneeberger S, Mühlmann G, Felgel-Farnholz C, Fong D, Zitt M, Brunhuber T, Schäfer G, Gastl G, Spizzo G (2004). Overexpression of epithelial cell adhesion molecule antigen in gallbladder carcinoma is an independent marker for poor survival. Clin Cancer Res 10: 3131-6

Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, Smith HO, Yandell M, Evans CA, Holt RA, Gocayne JD, Amanatides P, Ballew RM, Huson DH, Wortman JR, Zhang Q, Kodira CD, Zheng XH, Chen L, Skupski M, Subramanian G, Thomas PD, Zhang J, Gabor Miklos GL, Nelson C, Broder S and 248 more. (2001). The sequence of the human genome. Science. 291: 1304-51

Verloes A, Lombet J, Lambert Y, Hubert AF, Deprez M, Fridman V, Gosseye S, Rigo J, Sokal E (1997). Tricho-hepato-enteric syndrome: further delineation of a distinct syndrome with neonatal hemochromatosis phenotype, intractable diarrhea, and hair anomalies. Am J Med Genet. 68: 391-5

Vulpe, C., Levinson, B., Whitney, S., Packman, S., Gitschier, J (1993). Isolation of a candidate gene for Menkes disease and evidence that it encodes a copper-transporting ATPase. Nature Genet. 3: 7-13

Walker, D. C., McCloskey, D. A., Simard, L. R., McInnes, R. R (1990). Molecular analysis of human argininosuccinate lyase: mutant characterization and alternative splicing of the coding region. Proc. Nat. Acad. Sci. 87: 9625-9629

Ward CJ, Hogan MC, Rossetti S, Walker D, Sneddon T, Wang X, Kubly V, Cunningham JM, Bacallao R, Ishibashi M, Milliner DS, Torres VE, Harris PC (2002). The gene mutated in autosomal recessive polycystic kidney disease encodes a large, receptor-like protein. Nat Genet. 30: 259-69

Wasmuth HE, Glantz A, Keppeler H, Simon E, Bartz C, Rath W, Mattsson LA, Marschall HU, Lammert F (2007). Intrahepatic cholestasis of pregnancy: the severe form is associated with common variants of the hepatobiliary phospholipid transporter ABCB4 gene. Gut. 56: 265-70

Watson JD, Crick FH (1953). The structure of DNA. Cold Spring Harb Symp Quant Biol. 18: 123-31

Weinman SA, Carruth MW, Dawson PA (1998). Bile acid uptake via the human apical sodium-bile acid cotransporter is electrogenic. J Biol Chem. 273: 34691-5

Weinstein LS, Shenker A, Gejman PV, Merino MJ, Friedman E, Spiegel AM (1991). Activating mutations of the stimulatory G protein in the McCune-Albright syndrome. New Eng. J. Med. 325: 1688-1695

Wildin RS, Ramsdell F, Peake J, Faravelli F, Casanova JL, Buist N, Levy-Lahad E, Mazzella M, Goulet O, Perroni L, Bricarelli FD, Byrne G, McEuen M, Proll S, Appleby M, Brunkow ME (2001). X-linked neonatal diabetes mellitus, enteropathy and endocrinopathy syndrome is the human equivalent of mouse scurfy. Nat Genet. 27: 18-20

Woods CG, Cox J, Springell K, Hampshire DJ, Mohamed MD, McKibbin M, Stern R, Raymond FL, Sandford R, Malik Sharif S, Karbani G, Ahmed M, Bond J, Clayton D, Inglehearn CF (2006). Quantification of homozygosity in consanguineous individuals with autosomal recessive disease. Am J Hum Genet 78: 889-96

Yu L, Li-Hawkins J, Hammer RE, Berge KE, Horton JD, Cohen JC, Hobbs HH (2002). Overexpression of ABCG5 and ABCG8 promotes biliary cholesterol secretion and reduces fractional absorption of dietary cholesterol. J Clin Invest. 110: 671-80

Zappu A, Magli O, Lepori MB, Dessì V, Diana S, Incollu S, Kanavakis E, Nicolaidou P, Manolaki N, Fretzayas A, De Virgiliis S, Cao A, Loudianos G (2008). High incidence and allelic homogeneity of Wilson disease in 2 isolated populations: a prerequisite for efficient disease prevention programs. J Pediatr Gastroenterol Nutr 47:334-8

Zhou F, You G (2006). Molecular insights into the structure-function relationship of organic anion transporters OATs. Pharm Res. 24: 28-36

Zhu QS, Xing W, Qian B, von Dippe P, Shneider BL, Fox VL, Levy D (2003). Inhibition of human m-epoxide hydrolase gene expression in a case of hypercholanemia. Biochim Biophys Acta. 1638: 208-16

K.W. Zimmerman. Arch. Mikrosk. Entwickl.Mech. 52 (1898), pp. 552–706

Chapter 8

Appendix

Contents

8.1 Proformas used to ascertain clinical details of patients referred from collaborators

Proforma for Phenotypic diarrhoea of Infancy clinical details

Family history with family tree, consanguinity details, other affecteds in the family and ethnicity
Gestational and birth weight
Need for parenteral nutrition
Bowel biopsy/histology reports (if possible pictures)
Liver disease / ultrasound findings / histology (if possible pictures)
Facial features
Hair findings on microscopy / trichorrhexis nodosa (if possible pictures)
Cardiac anomalies
Spinal / skeletal anomalies
Immunodeficiency
Platelet anomalies / thrombocytosis
Unexplained anaemia
Renal anomalies (protein / creatinine ratio)

Development
Skin or hair pigment changes
Eyes signs
Central venous catheter difficulties
If died – age and mode of death

Address to send DNA to is:

Dr Jane Hartley
Medical and Molecular Genetics
C/O Regional Genetics Department
Birmingham Women's Hospital
Metchley Park Road,
Edgbaston
Birmingham
B15 2TG
UK

I would be very grateful if you could let me know by email when the samples are sent so that I can ask to the lab to look out for them for me.

[redacted]

Many thanks for your help with this important research project. If you have any queries then please do not hesitate to contact me by email or on my mobile: +[redacted]

Jane

Dr Jane Hartley
Consultant Paediatric Hepatologist
MRC Clinical Research Fellow

Clinical proforma for PFIC patients - to accompany DNA sample

Patient name:

Patient date of birth:

Male/female:

Ethnicity:

Demographics	
Birth weight	
Gestational age	
Age at presentation	
Age at diagnosis	
Family pedigree/ family tree	Please attach a family pedigree / tree. This should also indicate family members with **any** disease or illness
Parental consanguinity	Yes No First cousin Second cousin Other – please specify…………………..
Familial illnesses	
Current age	
Current centile weight	3rd 25th 50th 75th 95th
Current centile height	3rd 25th 50th 75th 95th

Clinical features	**Yes/no answers**
Hepatomegaly	Yes No
Splenomegaly	Yes No
Portal hypertension	Yes No
Dysmorphic features	Yes No
Continuous or intermittent jaundice	Continuous Intermittent
Diarrhoea	Yes No
Steatorrhoea	Yes No
Xanthomas	Yes No
Pruritis none, mild, severe	None mild severe
Failure to thrive (persistently less then 3rd percentile or crossed 2 percentile lines on growth chart)	Yes No
Poor feeding (nutritional support required eg Nasogastric tube feeding)	Yes No
Hearing loss	Yes No
Rickets on x-ray	Yes No
Episodes of pancreatitis?	Yes No

Gallstones?	Yes No
Hepatocellular carcinoma?	Yes No
Renal dysfunction or tubular defect?	Yes No
Bleeding tendency	Yes No

Biochemistry (current)	
GGT low, normal or high	Low Normal High
Transaminases	Normal Raised
Bilirubin (maximum)	Normal Raised
Cholesterol	Normal Raised
Triglycerides	Normal Raised
Vitamin A (serum level)	Low Normal High
Vitamin D (serum level)	Low Normal High
Vitamin E (serum level)	Low Normal High
Prothrombin time	Normal Raised

Management	
Medication?	
Biliary diversion?	
OLT and age at transplant	
New clinical features post transplant ie diarrhoea	

Consented for research and publications Yes / No

It would be essential to have histology blocks of the liver biopsy from each of the patients.

Histology blocks enclosed Yes / No

DNA is also required from as many family members as possible

Name	Date of Birth	Relationship with affected child
		Affected child
		Father
		Mother
		Sibling

If there is anything not on this form that you feel is important?

8.2 Microsatellite markers

Microsatellites used to look for linkage to region on chromosome 5, 19 and 16 in chapter 3

D5S428F	6CACACATACACACACTCATACAC
D5S428R	GGAGCATTTTAGTAGATATTCACAG
D5S401F	6CTAAACGATCCCCAATGTCT
D5S401R	AGCTATTTTGGTTTTTCTGTTGA
D5S1725F	6TGTACTTCAGGCTACCCTGC
D5S1725R	CCAGAGAAAGAAAACCAATAGG
D5S1463F	6ATTAGCCAGTCATTTAAAAATCG
D5S1463R	ATTAAATACATACAGGTGTGTGCG
D5S815F	6TGGTATACTTGTGTAGCAAATTACA
D5S815R	TGCCATGATTGTTAAGTTTCC
D5S2100F	6TTAATTGAGANGCAATAGTGATAAT
D5S2100R	GCTGGGAAGTCATATACTAGATTTG
D5S644F	6ACTAACTGGTAGATCAATGTGC
D5S644R	TTGGATTTGCTAAGACTGTG
D5S1462F	6CTTTCCCTCTCTTCCCACAT
D5S1462R	CCTGGGGGTAAGAGGAGATA
D5S1503F	6AAGACCTATTAGGAGGCAGATG
D5S1503R	AAATCCCTGTACCAAGTCCC
D5S495F	6GCATAAGTTTTGCTAGGGGA
D5S495R	TTTAACCCCATCACAAGTTG
D5S409F	6GAGGGGATGAAGTGTGGATAAAC
D5S409R	CTGTAGGATGGCAGTGCTCTTAG

D5S433F	6TGTAAGACATACTCTCTATCACCC
D5S433R	TCAGACATCCATCTGTGTG
D5S460F	6GCTATTAGCTGCCAGAATGTT
D5S460R	TGGTCTGTTTTTGACTCAGAAA
D5S485F	6GATCCAATCAATCCAATG
D5S485R	GACACAGCATACAGAATGAA

D5S475F	6CAGATACCGCTACTCTTATCCTACC
D5S475R	TCTGGTTGGTGCTGAATGAA
D5S1466F	6GTATCAGAACTTCATGTTGTACACC
D5S1466R	GGCACCTAGGTTTGTTCTGA
D5S2501F	6TGATTACTCTGAGGAAGAAGGC
D5S2501R	TTGAAATGGGCACAGAAATT
D5S2027F	6ACTTGGCAGATTTTCCACTC
D5S2027R	CACCTCATTGACTGGGAC
D5S2065F	6CAGCCTCATTGTTTATTGACAG
D5S2065R	AATGGCATAGTTTTGGCTC
D5S2055F	6GTCTACCCCATCTCTGAACC
D5S2055R	GATCCCACTGCAGGCT
D5S494F	6GCTTTCACGAAGGTAGATATTG
D5S494R	CAGGCTAGGCAGATTACAGA
D5S471F	6TTTTCACACATTTTCCCAGC
D5S471R	AAAACTTCATTTACAAAAACAGGAG

D19S714F	6ATGCCCTCTTCTGTCTCTCC
D19S714R	GCAGAGAATCTGGACATGCT
D19S898F	6CCAGGAGGTCAAGGCTGC
D19S898R	AGCTCACTCTGCCCATTTCC
D19S560R	6GAAATATAGCGAGACCC
D19S560R	GTCTCACCACAAAAAATG
D19S568F	6TGAGTCTGCTGAGACCAAAGTTAG
D19S568R	ATAATGTAGCCTTGTCCTGGAATAG
D19S433F	6CCTGGGCAACAGAATAAGAT
D19S433R	TAGGTTTTTAAGGAACAGGTGG
D19s414F	6TGAGTCTGCTGAGACCAAAGTTAG
D19s414R	ATAATGTAGCCTTGTCCTGGAATAG
D19s245F	6TGAGTCTGCTGAGACCAAAGTTAG
D19s245R	ATAATGTAGCCTTGTCCTGGAATAG

D16S753F	6CAGGCTGAATGACAGAACAA
D16S753R	ATTGAAAACAACTCCGTCCA
D16S3183F	6GCCCATCTAGCCAACTAAATC
D16S3183R	CCCGATCTCACAATTTTGCAG
D16S3232F	6AATGCAATATACAAAAACTCACCC
D16S3232R	ATTTTAAACAATTTTGTCATCACCA
D16S3321F	6TCATGAGGCTACAGGCACAG
D16S3321R	AGAGCTTCCTGCCTAGTCCC
D16S2964F	6CTTCCCAGATTTGCTGACTTG

D16S2864R	CTTTCCCTATCCTGTTGAATGC
D16S3409F	6GGCAGAGGTTGAAGAGAGCTG
D16S3409R	CAGTAGAACACAGGCTAGAG
D16S746F	6CTCTAGCCTGGGTGACAAG
D16S746R	GGCAACAAGAGTTAAACTCC
D16S3105F	6GATCTTCCCAAAGCGCC
D16S3105R	TCCCGTCAGCCAAGCTA
D16S3044F	6ATACTCACTTTTAGACAGTTCAGGG
D16S3044R	GGCTCAGTTCCTAACCAGTTC

Microsatellites used to look for linkage to ARC syndrome in chapter 3

D15S996F	6GAAGGATGGTTTGAGCCC
D15S996R	ACTTAGGAATAATCATTACTGGCAT
D15S127F	6CCAACCACACTGGGAA
D15s127R	AACAGTTGCCCACGGT
D15S158F	6CAGGAGACCTCCAAACACA
D15S158R	TTTCAGCCAAGAAGCACG
D15S963F	6CAGCAAATCTCTGGAGCAC
D15S963R	GCAACTTCTTGTAAGTAGCCTAGC

Microsatellites used to look for linkage to intraflagellar transport proteins complex B in chapter 4

D15S1510F	6AGGGTCCAATTTCAACATGA
D15S1510R	CCAATGGAGCTGAAAGTCAT
D15S206F	6AACTTTTGGTGCTGAGGGC
D15S206R	CTGGCACTTGCTGGTGATA

D20S1121F	6ACTTTTTCTCCTTCAAGAGTCACC
D20S1121R	TGCACTTTCATTCACATGCA
D20S150F	6CCCGTAATCCCAGCTACTC
D20S150R	CAGCCCCACGTAGTCACCT

D3S1587F	6TACAGTTCTATAAGGGCAGCC
D3S1587R	AGGGAGACAGAGTGATGGATT
D3S621F	6ATACCCATGTTCACTGCACC
D3S621R	CACTTAGCACGTTTTCAAGG
D3S548F	6CTTCCAGGTCCAAGAGTG
D3S3548R	CAAAGGCAGCAGAATATG

D14S66F	6CTTATAATCAACACCTGCCTTC
D14S66R	TTTAGACTAGGGATGCAATCC
D14S980F	6CTGGGCAACAAGAGTGG
D14S980R	GAAGCGGGACAATTCTCTAAG
D14S274F	6TGAACTTTGGGCACCCT
D14S274R	TCTGACAAACCAGCAAATGA
D14S696F	6CAGCACAATGCAGGAGGAC
D14S696R	GAGAGAGGGTTTTAAGCCAAG
D14S1038F	6GATCCATCTTAGCCATTAAGG
D14S1038R	CAGTCAGGTGTCCATCTAAAAC
D14S994F	6GGCAGACAGGGCTAGAA
D14S994R	CGTTATCAGATGTAAGAGACTCCAG

D12S1300F	6CCTCACACAATGTTGTAAGGG
D12S1300R	TGTAACATCCGTGATTAAAATAGC
PAHF	6CTCTGTTCCCAGTGCTTCAC
PAHR	TGCAGATTCAGGAAGCACG
D12S353F	6AATTCATTGGGAGGGCA
D12S353R	TGGCAAATCGGAGAGC
D12S84F	6TGGGGTAGAGTTCTTATCTGG
D12S84R	AAATATGTCTCTAGGCTAATGGC
D12S1583F	6AGCCAGGAGTTGGAGAC
D12S1583R	CCCCTTGTTTACTTCCAG
D12S79F	6TTGGACTGAACTGAGATGCC
D12S79R	TATGTGCACCCAGACTACCA
D12S2070F	6GGGTCAGCGAATATTTCCTT
D12S2070R	TGGCTGACAGAGCCTAAAGT

8.3 Oligonucleotide primers for sequencing

Primers used in chapter 3

CAST X1F	CTCTCTCCCTGGCAGGAC
CAST X1R	GACCCGGTACCCCATCG
CASTX2 F	CATTTAGCCTTTGGGGATTAAAG
CAST X2R	GGCAAACTTCACAGTTTAAAAGC
CAST X3 F	AAGTTTGCATTTGACTACATTGG
CAST X3 R	AGGCCCCACCCACTAATC
CAST X4 F	GGAGCTGAAGGACCATGTAG
CAST X4 R	TTTGTTTTGCTTTGCAGCAC
CAST X5 F	TGGAGTTATACACATATGCATCGAG
CAST X5 R	TGATGGGCAGCTAAGCCTAC
CAST X6 F	TCAAGCAAAGTGTAACAAACTTAACTG
CAST X6 R	CACAGCCATTCAAAAGGGTC
CAST X7 F	CCTAAGCGGGCTTTTCCTG
CAST X7 R	GGATGGCACCATTTTAGCTG
CAST X8 F	GAACCACAGACAGCACAACTG
CAST X8 R	GTGAACAGGGGAACCGTATC
CAST X9 F	GAGGGTGAACTGGCAGATG
CAST X9 R	GATAGGTTATTTGAGCATACGTTTC
CAST X10 F	CTGTAGAGCGGATGTTCAGTG
CAST X10 R	GCCCACAAGAAATCAGATTATG
CAST X11 F	GAACTTTTGGTGTTTGGTGG
CAST X11 R	ATGACCATGGCCTTATTTGC
CAST X12 F	GGAATAGTGCTTTAGTTTCAGAATATG
CAST X12 R	AATGACAATTTCACAAGCACC
CAST X13 F	AACCCACAACGGGCAATAC
CAST X13 R	AATTGATAGCCCTGTTCCCC
CAST X14-15 F	CATTCTTCCACTTGAGTGCTTC
CAST X14-15 R	GTGCTACTGGCAGTTTCCTG
CAST X16 F	TTTTACAAAGAAGTAAATGGATTGTTC
CAST X16 R	TTCGGTTTTGTAATGTGCAG
CAST X17 F	GGAAGTGATATCATCTCATTATGTGC

CAST X17 R	AACAAAACAAGATGCAAAGGC
CAST X18F	TCCCAGAAACAAACTTGATGG
CAST X18 R	TGAATTCCAGACTGTCAAAATTATC
CAST X19 F	GCTTTGGGTGTTAAGTATGGC
CAST X19R	TGGCATATTGGGCTTTCTAAC
CAST X20 F	GTAGCGGAGTGTCTTGTTGG
CAST X20 R	CAGCAGAGAGTGACAGGAGG
CAST X21 F	TGATATGCCTTTCTGAATTGATAG
CAST X21R	GGGGAGAAGGTGATCAATG
CAST X22 F	CAGGCATGAATTTTATGGAGAG
CAST X22 R	TCCCAAGTACCTTTGTTCTGTG
CAST X23-24 F	AACAGCAAGTATAACCTGTAGACCC
CAST X23-24 R	GATGGCACCACTGTACTCCC
CAST X25 F	AGTTAAGTGATGGCATTGTGC
CAST X25 R	TCTTTCTAGTTGTGGCTGCG
CAST X26 F	TGCCTCTGATACAGTTTGGC
CAST X26 R	TGACAGAGTCAAGACCCAAGG
CAST X27 F	AGTCATCTGTGTTGTTCTGTTGC
CAST X27 R	ATATTTACCTGCTGGTCGGC
CAST X28 F	TCTCTACTGCCTAAACCTAAGTAAACC
CAST X28 R	ATGCCAAATACCCCTATGGC
CAST X29 F	TATTGCATTTTGCCTTCTGC
CAST X29 R	GCCTGGGCTATAGAGTAAGACC
CAST X30 F	TTTAACTGCAGCCTAGTTAATTGC
CAST X30 R	TGATAAACCCTACTCTGTGTAACTGC
CAST X31 F	TCCCCTAATTCTGAGAAGGC
CAST X31 R	AAATTGGCAGAAGATACAGTTGTC
SLCO4C1 F1 F	GAGAGGGACCTGGCTCTG
SLCO4C1 F1 R	CAACTGGTGCTATTCCTTGTTG
SLCO4C1 F2 F	GCCTGATTTCATCAAGCTACG
SLCO4C1 F2 R	GCAAATCCAGTAGTAATTAAGGCTTC
SLCO4C1 F3 F	GCTTTCCAAAACATTTACCAGG
SLCO4C1 F3 R	CCAGCTTTAGCTTCAAAACC
SLCO4C1 F4 F	CCTGTCTGTGGAGATGGAG
SLCO4C1 F4 R	GCCATGGCAATGTGTGTTC

SLCO6A1 F1 F	GTTGGCCCAGGCAGGAC
SLCO6A1 F1 R	TCGCAAATATCTTCAATTCCTAC
SLCO6A1 F2 F	CGATATTTCATCTGGCCTGG
SLCO6A1 F2 R	GGCATATGAGCACTGGATTC
SLCO6A1 F3 F	CCAAACAATATGCCAGGTTC
SLCO6A1 F3 R	GGGTCTGGCATCAATAAAATC
SLCO6A1 F4 F	ACCTCACGGCTCCTTGC
SLCO6A1 F4 R	CCTCTTCTGAAAGTTGGGC
SLCO6A1 1F	AGTCCTCCGACGCCCTC
SLCO6A1 1R	CTCCGCCATACTCCCTCTC
SLCO6A1 2F	TTGTAGAGGTTAAGGCCTATTTATG
SLCO6A1 2R	GTTCCTTGAGCCAAGGTTTG
SLCO6A1 3F	TACTGCACATTTGTTTTATGTTTG
SLCO6A1 3R	TTATCATAAGGGCCAATAGTACTTC
SLCO6A1 4F	GCTTGGTTCCCACTTCTGAG
SLCO6A1 4R	TGAATAAAGCCAGTGTGGTATG
SLCO6A1 5F	CTCTGTTTGATTATGGTTTTGTTG
SLCO6A1 5R	TAAGAACCATGGTGGAAGAAAC
SLCO6A1 6F	TTTGCCATTTCAGGATTAATTG
SLCO6A1 6R	TTGAACCCTTTATGGGTATTTTC
SLCO6A1 7F	GATCATTAGCATTTGGTTCTGTC
SLCO6A1 7R	CACTATGATGGCAAAGTGCTC
SLCO6A1 8F	TGTTTAGAGCAAAGGAAATTGTATC
SLCO6A1 8R	TTGCCTGAGGAATATGCTGTAG
SLCO6A1 9F	TTTCAATAAGCATATTTTATGGATCTC
SLCO6A1 9R	GAACATTTCACTAAAGAGGAAGAAGC
SLCO6A1 10F	TTCATTTCTAAATAATACATAGCAGGC
SLCO6A1 10R	TCTGCATTTGCTACCATGAAAC
SLCO6A1 11F	AAAATGATTAGCCTTGAATCTAGC
SLCO6A1 11R	CCCTTTCACTCGCTTTGTG
SLCO6A1 12F	TGTGTGTGCATTACATGTTTAAGAG
SLCO6A1 12R	CAGGTTAAGCCTCCTTTACTCATC
SLCO6A1 13F	AATAGCACACCTTGCACTTGG
SLCO6A1 13R	TCCAAATAGCAAAGGCCAAC
LIX1 EX1 F	GCAAGAATTCAGGCATGAGG

LIX1 EX1 R	GGATGTGACTTGACGTTTAGGAC
LIX1 EX2 F	TTCACAGCTAAAAGCCCCTC
LIX1 EX2 R	TCCTTTCTGGTTGTTTTCATTC
LIX1 EX3 F	TGATGCAGGAGTGGAATGG
LIX1 EX3 R	AACATCTCTCAAGATTCAAATGC
LIX1 EX4 F	TCTACTGGAGACTTGGTTGTGAAC
LIX1 EX4 R	CAAAAGATTATTGGGTCGGG
LIX1 EX5 F	CTGCTCAGCTCTCACCAATG
LIX1 EX5 R	TGGGAAGATGATAAGCCACC
LIX1 EX6 F	ACGGGTGCCTTTGTTGC
LIX1 EX6 R	GCCTCTGAGGACCTCTGCAC
RIOK2 EX1 F	AATAACGACCTGCCTTACGC
RIOK2 EX1 R	CTTGCAGGAACAGGAAGAGG
RIOK2 EX2 F	CGCAGGGTTTTGATAGACAAG
RIOK2 EX2 R	TTGGAACACTTGCTCTGGTG
RIOK2 EX3 F	AAATCTTTCTTATGACTTGAACTTACG
RIOK2 EX3 R	AAGAGGAAAGGTTTCTGAAGGG
RIOK2 EX4 F	TGTTTTAACTGATTGGAATAAACG
RIOK2 EX4 R	GCCTTTGAGACACATGTATTATTTG
RIOK2 EX5 F	TGTGGCTAATTCTACGCAGTG
RIOK2 EX5 R	TCATATTGTCAGAATGAAACTGGG
RIOK2 EX6 F	GCATACATACTTTCGTTAAAACCG
RIOK2 EX6 R	GCAACAAGCAAAACTCCG
RIOK2 EX7 F	AAGGAGGAGCAATACCATTAACC
RIOK2 EX7 R	CAGAGCAATGACATCTAACAGTG
RIOK2 EX8_1 F	TGACAGCCAGCATATACCTTTC
RIOK2 EX8_1 R	TCTGACCATCTTGCCTGTTG
RIOK2 EX8_2 F	AAGAATCAGAGGGCTGCTATTG
RIOK2 EX8_2 R	TTTCATTTTATACATCAAGGTGGC
RIOK2 EX9 F	AGATGAATTGTGGGGAGGTG
RIOK2 EX9 R	TAATGGTGGCAGGACACAAC
RIOK2 EX10 F	GGACACTAGGTGGCACTATGG
RIOK2 EX10 R	AATGGACTGCTTTGGCAGG
FLJ35946 EXON1F	TTTCAGCCAATCACATCTGC
FLJ35946 EXON1R	AGGGCGGTATCAAGTGTCTC

FLJ35946 EXON2F	TTTCCACTCGTGACTTTCTATGA
FLJ35946 EXON2R	TGCAGCAATAAAAAGTGAATTG
FLJ35946 EXON3F	CACATCCACCCTGTTTCTTG
FLJ35946 EXON3R	GGCTCAGGAAATGAAACTGC
LOC441066 X1F	ATTCCCGTTTGTGTGACCC
LOC441066 X1R	AAGCCAGGGCTACCAGATG
LOC441066 X2F	GGTCTTCCTGCTGTAGGGC
LOC441066 X2R	GGAAGGCTAAAGTGGAAGGTG
LOC441066 X3F	GGGTGGTTCTCCAGGGTC
LOC441066 X3R	GGCAGATTCCAGGAAGGAC
LOC728104 EX1 F(2)	GGAGGAACCGGTGCTCAGAGA
LOC728104 EX1 R(2)	GATCTCGGATAGCACACGTCC
LOC728104 EXON2F	TTTTGGTTTTCTTCCCCTCA
LOC728104 EXON2R	CAGAGAGGGTCAAATCATTGAA
LOC728104 EXON3F	GGGGTAATAAGGTGGTTGAAAA
LOC728104 EXON3R	GCAAATGCAGACCCTGAGAT
TMEM157 EXON1F	CCACCGCGCCTATGGTCCC
TMEM157 EXON1R	GGGAAGGGGGAAAGGGTCAC
TMEM157 EXON2F	CCTAGGTGGTAAATTTGTTGC
TMEM157 EXON2R	CAGAACTGAAGAAACTGTGTTGG
TMEM157 EXON3F	CCTTCCCCTCCAAATCTTTC
TMEM157 EXON3R	CCCCCCCAACATTGATATAG
ST8SIA4 EXON1F	CGCGACTATCTCCCCAAAACG
ST8SIA4 EXON1R	CTAACCATCACTCTACCCTC
ST8SIA4 EXON2F	GTATTTCCTTCTAACTTGTAGGG
ST8SIA4 EXON2R	CCAGTGTTGAATACAAGTTTGCC
ST8SIA4 EXON3F	CCACCTAAGATTCATAGCTTACC
ST8SIA4 EXON3R	CTAACAGTTTTCCACCCCC
ST8SIA4 EXON4F	GCACGTGCGGAGACTTATTG
ST8SIA4 EXON4R	CGGAAACATATATCCATTTGGAG
ST8SIA4 EXON5F	GATACCTTTTAAGTTTTCTGC
ST8SIA4 EXON5R	GATGCTGAAACCCAGCCGTG
PAM X1F	ATGAAGTAGCGGCTGCTGG
PAM X1R	CAAAGAACATCAACCCCGTC
PAM X2F	TCATAGAATTCCTTTCTTTCCCC

PAM X2R	TGCCTACGTTTAAAGAAGTTCC
PAM X3F	TTTTACAGTCATAGAAGCCACG
PAM X3R	ATTTCAAATCACTGCCCTCC
PAM X4F	CTTTGGTGCACAGAAGTGTAAC
PAM X4R	TTTGCACTTGTAAAACACATTTTC
PAM X5F	ACCATGGGGAGCAATTTATG
PAM X5R	CACTGTCAACACTGGTTAGAAAGG
PAM X6F	TTTGAGATTAACTCCATTTTAGATGC
PAM X6R	GAGGGTGTGGGCAATTC
PAM X7F	GAACCATACAGATGAAACATAACTTG
PAM X7R	AAACTTCCCACTCTAAGTTCGC
PAM X8F	GGAACCATTTAGAGTCTGCCTTC
PAM X8R	TGGAAAAGTAATAAACCACTGGC
PAMX9-10F	AACCTCAGTCCCTAGCATCTTG
PAM X9-10R	CAGCTCAGTTATACAGTTAAGGCAG
PAM X14F	TCCTATGCTTAAAAGTCTGAGTGC
PAM X14R	CCCTTCAGGTTTTGCTAACG
PAM X17F	TTCTGTCACATATGCCTTGAGAG
PAM X17R	AAAACCCATACAACTCCAGAGC
PAM X18F	CTCCAATTGCTTGTGAAAGTG
PAM X18R	GAGTGTTGGCCAATATTTAGTTCC
PAM X19F	CCTTATCTTTCCCCTGGCTC
PAM X19R	CAAAACTGAGTTAGAACACACTAACAG
PAM X20F	GCCAATTGTAGGGACCAGG
PAM X20R	AACACTCCCATCAACCTTGC
PAM X21F	TTGTTTTGTTGTGGGCTTTG
PAM X21R	TCATATGCTTTTAAGCAATCAAAC
PAM X22F	AGTGGTTTGTATTGAACTCTTTCC
PAM X22R	TTTTAAGTTGCAATCAAGGGC
PAM X23F	TTGATTTGACTGGGTAGGGG
PAM X23R	GGGGCAATATTTGGAACATAG
PAM X24F	GAAGTTGAGTGCATGGGTTG
PAM X24R	AAGACATGTGCTTGTTGACTTTG
PAM X25F	CTTTCCCATCCCCACCTTC
PAM X25F	GGAAAGGAATCTGACATTCTGG

LOC134505 X1F	GAGGACACCGTGCCAGTC
LOC134505 X1R	TCCAGCAGGTCCTTGAGATAC
LOC134505 X2F	CCCGTGTAAGTGCATGATTG
LOC 134505 X2R	AAAGGACTCAACTTACCCATCAG
FLJ20125 EXON2F	AATTCTGGGAATTCTCGGTTC
FLJ20125 EXON2R	TCCACAAGGAAGAATGTGGG
FLJ20125 EXON3F	TGGTTTAAAGAGAAGGCTATGATTC
FLJ20125 EXON3R	GAGAGGTGATACCCAGAATAAATC
FLJ20125 EXON4F	TCCAGAAGTTTCTCCTTCAAAAC
FLJ20125 EXON4R	TATAATGCAGAGGGCACTGG
FLJ20125 EXON5F	CTTGGGATTTTGTTCCTCAC
FLJ20125 EXON5R	CCAATAGTTGCATCTTCCACC
FLJ20125 X5-6F	CCGAGATTTTGCCACTGC
FLJ20125 X5-6R	CAGCATTTCATAGCTAGTAAACCAAC
FLJ20125 X7F	ATTCTCATTATTTGCATTTACCTTC
FLJ20125 X7R	AACAAAAGAGGGGAGAAAGG
FLJ20125 X8F (201)	CTTGGGATTTTGTTCCTCAC
FLJ20125 X8R (201)	CGTATACTGATATTCAAATGTGGC
HISPPD1EXON 1 F	CATATGTTTGTCAAATTGCTAAG
HISPPD1EXON 1 R	CGGTTCAACCAGAACTCAC
HISPPD1EXON 2 F	GGAAAACAAATGAAGATAGCTG
HISPPD1EXON 2 R	GTCCATGAAAGTGAGTTAATG
HISPPD1EXON3 F	CTTTTATGGGAGGTGGCATC
HISPPD1EXON3 R	CACCTGATGCCATGTCAGTG
HISPPD1EXON4 F	GGCTTTTAATGCTTAGGTAATG
HISPPD1EXON4 R	TGTAAGAATGTGGCCCAAAAG
HISPPD1EXON5 F	CCTCGATAATAATGTGGTGTCC
HISPPD1EXON5 R	AAATGATGCCATGCTTTTCTAAC
HISPPD1 FRAG2 F	GATTGAAGGGGAAGATCATGTAG
HISPPD1 FRAG2 R	TCCATAAGAAGCTGTCGTGC
HISPPD1 FRAG3 F	AGCTGTTATACGTCATGGGG
HISPPD1 FRAG3 R	GGAGTAAGCTTTTCATAATCTTC
HISPPD1 FRAG4 F	TGGAAGGAGAGCTTACACCC
HISPPD1 FRAG4 R	GGCACCATAGCGAAGAATAGAC
HISPPD1 FRAG5 F	CGCTCAGACCTTCAGAGGAC

HISPPD1 FRAG5 R	TCCACAAGAGTTCTTGGTGTTC
HISPPD1 FRAG6 F	AGTGTGTCTAGCCCAGAGGG
HISPPD1 FRAG6 R	TCATGACAACAGTTCCTTACCTG
HISPPD1 24F	TGGCTCTGTTGTATCTAACTTTGC
HISPPD1 24R	AAAACAGAAGGATTTACTCTGATGTG
HISPPD1 25F	GTCTGTCTGGAACAGGCTGC
HISPPDI 25R	CATGCTTTCACCAAAGAGTTTTAC
HISPPD1 26F	TCCCAATGTGAATGGAAGAAC
HISPPD1 26R	TACTGCTCATGGGTGTGTGC
HISPPDI 27F	CCTAAACTGTCCAGTTCTAAGCAG
HISPPD1 27R	CAGATAGTATGTAGTTGCCCAGC
HISPPDI 28F	TGAATTTACCTTGACCATTTTCTTC
HISPPD1 28R	CGCTCTAGGGAAGTGCAAAG
HISPPD1 29F	CAAGAATTGTTTTGTCAATCAGC
HISPPD1 29R	TCATGACAACAGTTCCTTACCTG
HISPPDI EX18 F	GCGTATGCCTAAAGTAACTCTC
HISPPD1 EX18 R	CCTTTTTATCCACCAACAGAC
LOC90355 EXON3.1F	GCCTGGCCCAAGTCTCTG
LOC90355 EXON3.1R	TGGAATCTCGTATGGCTGG
LOC90355 EXON3.2F	ATCCTTACGCTCCAAACAGC
LOC90355 EXON3.2R	CATGTACAGAGTACAGCAGTGGG
CETN3 EXON 1F	GTCTTGCTGCCTTGGGTAGG
CETN3 EXON1R	CCCTTCCACACACACCCTC
CETN3 EXON2F	TTGAAATTTAGAAGGTAATTATTGGC
CETN3 EXON2R	TGCCTTCATAGTTCCAAAACC
CETN3 EXON3F	CTGTAAACACTTGTTGGTTACTTTG
CETN3 EXON3R	AAAACTATAAATCTCAGAGCAAAGCC
CETN3 EXON4F	TGAATGCCTTTGTGATTATGC
CETN3 EXON4R	TCTTGCAAGTCATTTGGTTTTAG
CETN3 EXON5F	AAAATGGTAGTCGTGGATTCTG
CETN3 EXON5R	TTTCACATGGCTCCAGGC
LOC153363 X1_1F	GCCGCAGTGCCCTGTGTGTG
LOC153363 X1_1R	CGAACTGGTAGAGGCCGCCG
LOC153363 X1-2F (2)	GCCTCTTGCAGGACCGAGAGGC
LOC153363 X1_2R	GGCCACGCAAGTCTGCAAC

LOC153363 X2_1F	CCTTCCCAATTTCTTTGGAAGC
LOC153363 X2_1R	CGAAAAAGCCTTTCAGCACC
LOC153363 X2_2F	GGACAGAGGTCTGGTAGAG
LOC153363 X2_2R	GCACTTCATGAAATGCTCAC
POLR3GX2F	TGTGCTACTTAAGGGGTGCAG
POLR3GX2R	TGAAAACTAATTCTTTACTGGTTCAC
POLR3GX3F	GGCCAATAGGAAAATTACGC
POLR3GX3R	TGGTTCTCTAGACATAAGAATGGG
POLR3GX4F	GGAATGGCTTAAGTTCAGGG
POLR3GX4R	AACCAGACTCAGTTTTCCTATGG
POLR3GX5F	ACTTTATTGTCATTTTCTGGACAAC
POLR3GX5R	ATAACTTTGGGTTGACAAATCC
POLR3GX6F	CAGCATACGCAAAGACAAGG
POLR3GX6R	AAACAGAGCAAAACGAGGTAGG
POLR3GX7F	AAGCAGGGAAGTCAGCAGTC
POLR3GX7R	TTCCCATCATCTTGGTTTCTG
POLR3GX8F	GCAGAGACTTTTAGGCCTGG
POLR3GX8R	GGTAACAACTATTTGTCC
LYSMD3 EX1 F	TTTAACATTATGGCAGGGAGG
LYSMD3 EX1 R	ACAGAATTTATTTCAGTGTGCG
LYSMD3 EX2_1 F	AGAATGAATGACTATTGCTGCC
LYSMD3 EX2_1 R	TTTATGATCATCTTGTTGGCTG
LYSMD3 EX2_2 F	TGTTAGTCATCATTCAACAGTGGAC
LYSMD3 EX2_2 R	TTAAAAGTTTTGAAGGGGTGTG
LYSMD3 EX2_3 F	TGGTGTGTGAATGTTTCAAGC
LYSMD3 EX2_3 R	CCTCAAAGGCAGTTTTGACC
LYSMD3 EX2_4 F	CATTTAATGCCTTACTGCTTTATG
LYSMD3 EX2_4 R	TGTCCTGTCTTTTCCTGATGG
LYSMD3 EX2_5 F	TACCACCTCATTTTCTGGGC
LYSMD3 EX2_5 R	TTAGTTATACAAGGTAATTTGCATTTG
LYSMD3 EX2_6 F	GGGTTCGATTTCCAAAATATG
LYSMD3 EX2_6 R	CCAAGCAGGTAGAGTCCAGG
ARRDC3 X1F	AAAATATCGATCAGTGTTAAGTGAAG
ARRDC3 X1R	GAATCACGTTAATTCAGAAAACTG
ARRDC3 X2F	TAAAGTTATGCCACCCTGCG

ARRDC3 X2R	GCTTTCTATAATTTTGCCACGC
ARRDC3 X3F	TTTGATACCTCATTGGAACATAAAC
ARRDC3 X3R	GACAACTGCATAGTTTAGGTGATTG
ARRDC3 X4F	CAGATTGCTTTCTTTGCTGC
ARRDC3 X4R	CGGGAAGAGCAGTTCTCAATC
ARRDC3 X5F	GCTTTAAGAATCTTGACTTCTTTTG
ARRDC3 X5R	CCTGGGTGACAGGTGTTTC
ARRDC3 X6F	TGACATAGTTTTGTCTTACATCCAG
ARRDC3 X6R	TCGAACAGCATGTTCTTATGC
ARRDC3 X7F	TGTGCCCCTAATATTCAGTTTTG
ARRDC3 X7R	ATGCCTGCAATGCTATTTCC
ARRDC3 X8F	TTGTAGTCATTTGCTACATTCTTCC
ARRDC3 X8R	GCCGGAAGAGATACAGTTCG
C5ORF21 X1F	TCTCGTGGAATAGATAGGTTTTGAG
C5ORF21 X1R	AATATTCCTACACCTTTATTAGGCAC
C5ORF21 X2F	CAAGGTGCTTATGTTAATACTGTGC
C5ORF21 X2R	TTTCACCATGCAAGGATGTC
C5ORF21 X3F	TGTTTGCATCATAGTGACATCTG
C5ORF21 X3R	AAAGAGTGCATAAACTGCTGC
C5ORF21 X4F	GCAATAACTTGCAACAATTTACC
C5ORF21 X4R	AAACTTATCACAGTCAACCCAAAAC
C5ORF21 X5F	AACCATTTAATACTTTGTAACAGCC
C5ORF21 X5R	TGCCACTTTGTGTTTCTTTG
C5ORF21 X6F	AAAATAAAATCATTATCTTGCCTAAAG
C5ORF21 X6R	GAATGATGCGTTTTGTAATTCTG
C5ORF21 X7F	CATATGCTGTGTATGCAAGC
C5ORF21 X7R	GGAGAGGGGATAAACCTAC
C5ORF21 X8F	TCTGACTAAGCAGGCTGCAC
C5ORF21 X8R	GGGGCTCAAGATATAAAGGC
C5ORF21 X9F	TGAATAATTACCGACTGTTTACATAAG
C5ORF21 X9R	TGATAGTGGATTCTACCCATGC
C5ORF21 X10F	CTAGTTGTGCCCTGGGAGTC
C5ORF21 X10R	TCGACACAAGCAGTGAGGAG
C5ORF36 EX3 F	AATTTCTGCCCACCCCTATC
C5ORF36 EX3R	GAGCAGGATACAAATAGACCAGG

C5ORF36 EX4F	TTTCAGGGCCTGTGATATTATG
C5ORF36 EX4R	TGAATGAGTTGTATTCCCTGTTTAC
C5ORF36 EX5_1F	GATAATCAGCTAGGTTTGAAATGAAC
C5ORF36 EX5_1R	TTTCACCATTGAAGATGGAGC
C5ORF36 EX 5_2F	CAGTGCTTACAACAACTCTTGTTTC
C5ORF36 EX5_2R	CATGTTTGCAGTGAGATTTAACC
ANKRDX2-3F	TTTTAACAGCTGTGCCTAACATC
ANKRDX2-3R	ATGCTATCTCTCCAGAGACTAAAATAC
ANKRDX4F	TGTTAGCAGCATAGCTTAAGAGTC
ANKRDX4R	AATCACCATGGCTCATTTGG
ANKRDX5F	AAGAAATTGCAGCCGTGTTC
ANKRDX5R	AAGTAATTGAGAGCCACTGCC
ANKRDX6F	GAGGGTGCATTTTACCTACTTG
ANKRDX6R	CACATGGTTCAAACACCAGC
ANKRDX7F	CTTGTGTTGATTGTGTTAAGGC
ANKRDX7R	CAGCCTGGTAACAGAGCAAG
ANKRDX8F	CTGTCAATGGATGAGATTTTCTG
ANKRDX8R	TCAGGCAAAACAAAATTCACTG
ANKRDX9F	TGTGAGGATCTTTGCTGTGG
ANKRDX9R	ACTTCGGGGCATTGATGTAG
ANKRDX10F	TTCCAAGGTTTTGGGGAAG
ANKRDX10R	CGTTACCTTTTATTGTGCATACAAC
ANKRDX11F	CAGCAAGGCCAGGAATTATC
ANKRDX11R	TTTTCCCCAAAACAAAATTATC
ANKRDX12F	GGGTTACTGTTTTGGTTTTAGACTG
ANKRDX12R	TTTGCCTTTTCATAAGTTTTGC
ANKRDX13F	TTCCAGCAGTATTCTAGAAGGAG
ANKRDX13R	GGCATATAAACCTTTACTCCCAC
ANKRDX14F	TGACTTTATTGAACTGGGGC
ANKRDX14R	GGAACCAGTCTCCAAGGACC
ANKRDX15F	GGGAGAGAATTAGTTAGCTAGGGTTC
ANKRDX15R	ATCTATGTGTGCGTGCATCC
ANKRDX16F	TTTGTGGATTAAAATCATGTGG
ANKRDX16R	TACAGTGAGCCGCGATTG
ANKRDX17F	GCAACCATAGTGATGCAACC

ANKRDX17R	TCTTTTGGGAAGCACTTAGAAAG
ANKRDX18F	GTATGCTGGCACGCCTTG
ANKRDX18R	GGGCAAATGTGTGAAACCTG
ANKRDX19F	GGGGCCAGATCTCCTAGTTC
ANKRDX19R	TTGGGTATTATCAAAATATGGACAC
ANKRDX20F	TGAAGCACCTAGAAAGTCACAC
ANKRDX20R	TGCTATCTACCTCTGCAGGATTC
ANKRD32 X21F	AAGGAGAAAAGAGCTGTTAAGCTG
ANKRD32 X21R	TCCACAGTAAGTACCAATGCAAAC
FAM81B X1F(2)	GCTTCTAAATCCCAGGGATTTC
FAM81B X1R(2)	CTCTGTTGTGCCCAAGCC
FAM81B EX2F	TGAAAGCCTCCCTTTAAGCC
FAM81B EX2R	ACAAACAGGAAATGTGGGAG
FAM81B EX3F	AACTGAGCTGATTATCTGGTCTATG
FAM81B EX3R	TTTTGCGTGTGAGATTTGTC
FAM81B EX4F	GCTCCTCCAGCTCATTCTTG
FAM81B EX4R	CAAGCAATGCTGCCAAAATAC
FAM81B EX5F	CCCCTGCTTTAAATGATGTTTC
FAM81B EX5R	AGGAATGCGAATTCACTGGC
FAM81B X6F (2)	GCCATCAAATGAATGTGCAGG
FAMB1B X6R (2)	CCTGATTGGGCATTTTGATAAGC
FAMB1B X7F (2)	GTTTCCCGGCTCCTTAAAGG
FAMB1B X7R (2)	GCCAATGCTTTCATTAGGC
FAM81B EX8F	GACACACTGAGTCTCGTCACC
FAM81B EX8R	CCCAGACCTTTACTCACTGC
FAM81B EX9F	CAGTTTTAGATACAAAGGAACTTGG
FAM81B EX9R	GATTGATCCCTGGCCATTAG
FAM81B EX10F	GCAGGACTGCCGATTAGC
FAM81B EX10R	CATTAGAAGCAGTAAACATCCCAG
ARSK EX1 F	AGAGAACGCCAGAGGGAGG
ARSK EX1 R	CCAAAGCCTGCAGCACAG
ARSK EX2 F	TGTACCAGATACTGGAAACACCC
ARSK EX2 R	AACCCCAATAGCCAATATATACC
ARSK EX3 F	TGGACTGCTCTCTCAGAGTAATTG
ARSK EX3 R	TGTACACGCTATGGCAGAGG

ARSK EX4F(2)	CCATGTTGTTTTCTTAGTGTATGC
ARSK EX4R(2)	GAGCAAATACCAATAAGCCAC
ARSK EX5 F	TCATCATACATACTTATTGGAGGAAAG
ARSK EX5 R	TGTAAGAAAGTCACCAAGCAGC
ARSK EX6F(2)	GGGTAGAGTGGTTATAAAC
ARSK EX6R(2)	GGACCTCTGACTACTAACTG
ARSK EX7 F	GTGAATGATGCTATTTGTGAATG
ARSK EX7 R	GAGCAATGCTACATAGGGCAG
ARSK EX8 F	AACATAGGGTGTTTAACTCCTGG
ARSK EX8 R	TGACCTTGTGATCTGCTTTCC
RFESD EX2F	GCATTGGAATTTATGTGGCC
RFESD EX2R	GAGATGTCAAGTCAACAACC
RFESD EX3F	GGCTCCTGCCTGTTTACC
RFESD EX3R	GCCTGGAGTACACAGTTGGC
RFESD EX4-5F	CCGAACTATTCATTCTTAGGG
RFESD EX4-5R	GCATTAAAATATAGTTCAAACC
RFESD EX6F	GGTTCTTCTGCAGCAGACTC
RFESD EX6R	CCTTTAAAACTGAAGCAGAG
SPATA9 EX1-2F	TTCCAGCAGTGAGCTGTGAG
SPATA9 EX1-2R	TGAGGCAAAATTAGTGTTTTCTAAAG
SPATA9 EX3F	GCATGACACCATGCACAATG
SPATA9 EX3R	TCCAAGTATATCTGTTTTACACTGGAC
SPATA9 EX4F	GCCCAGATAGCTCTAAAGTAGATGAC
SPATA9 EX4R	TTGTGTAACGAGACTGTGAATGG
SPATA9 EX5F	TGTAGATTAAATGACTTTATGTGGAAC
SPATA9 EX5R	GCAGAGCAATTCAGAATATGTAAAC
RHOBTB3 X1-2F	CGGATTGCGGGTGAACTC
RHOBTB3 X1R(2)	GCGGGGGACAGCGCGCGGCG
RHOBTB3 X2F(2)	CGCCGCGCGCTGTCCCCCGC
RHOBTB3 X1-2R	GGACACTCCACACTCACGG
RHOBTB3 X3F	ATTTCAGGATTGATTGGACG
RHOBTB3 X3R	ATAGCAGCTGGCATTGAACC
RHOBTB3 X4F	TCTTTGTTGTTTAATGTTCAGATTG
RHOBTB3 X4R	CAAAATACCTAAAGAGGCAACC
RHOBTB3 X5F	TTGACAATTGTGACTTCAGATGC

RHOBTB3 X5R	TGCTCAAGTTTGAGAATAATTGC
RHOBTB3 X6F	AAATTGTTGGGTCTGTAAACCTG
RHOBTB3 X6R	CACAGGCACATTTCATTTCC
RHOBTB3 X7F	TTAGGGGAAGTTAAACTGATCG
RHOBTB3 X7R	CACACAATCTAAGATACAAATAACCC
RHOBTB3 X8F	GCATAATTTTCCAAATTCCTAGC
RHOBTB3 X8R	CATTCCTTCACACCAAACATTC
RHOBTB3 X9F	GGAAAATTGCTGGAATAACCC
RHOBTB3 X9R	TAACAAAATGGGTGGCAAGG
RHOBTB3 X10F	ATGGTGGATGCTACGGTCAG
RHOBTB3 X10R	TTGTGAAGGCACTTTAGCCTC
RHOBTB3 X11F	ATGAATATGGGCTCTCCGTG
RHOBTB3 X11R	TTTCCTCAATTACCACACTGC
RHOBTB3 X12F	TCAATAGCTGATATCTTAATTGATCC
RHOBTB3 X12R	CCTGGAGGTATCCCATTCTTC
GLRX EX1 F	ATACCCCAAGCAATACCTGC
GLRX EX1 R	TGACAAGAAAAGGCAATCCG
GLRX EX2 F	GAAAAGTTTGGAATCACGGG
GLRX EX2 R	CACGACAGAAGAATTCCACG
FIS X1-2F	TGTCAAAACAGACCACTGGG
FIS X1-2R	GGACAACTAGCAGGAAGTGGAG
FIS X3F	AATAGCACTGGTGTCGGTGG
FIS X3R	TTTTATGGTTATCAGTTTCACAGC
FIS X4.1F	TAAGGGCTGGTTTGCGTAAC
FIS X4.1R	TGTGAAATACTCAAATCCCTTCAG
FIS X4.2F	CAGGATCTGCCATGACTCAG
FIS X4.2R	TCCTTTCCAACCTTTCCAAC
FIS X4.3F	GGCAGAATTCAATATTGGGG
FIS X4.3R	CCTTATCTAGTCCCCAGAGGC
ELL2 EX1F(2)	CTGCAGAAGCCCAAGCAGCCG
ELL2 EX1R(2)	GCCTCTCTGAGCCCAGCCTG
ELL2 EX11-12F	CCTATACTCTTGCTGAATGCATG
ELL2 X11R	GAGATTGCAGAATTTACAAC
ELL2 X12F	GTTGTAAATTCTGCAATCTC
ELL2 EX11-12R	CTTTCAACAGCCAAAGTTTCACC

MIRN583 X1F	CAACAAAACCCCATGATACAAG
MIRN583 X1R	TTGGTAAGTGGGAGCTCAGG
PSCK1 EXON1 F	GGGAGTCTGTCTGGCTTTTC
PSCK1 EXON2 F	TCAGATACCAGGAATGGGC
PSCK1 EXON3 F	GCAGAGCTGCCTGAACTC
PSCK1 EXON4 F	GTGCCCATTTTGTAGCAAGC
PSCK1 EXON5 F	AGCTATCAGCCAGGATGGG
PSCK1 EXON6 F	CCTATGCCCCATTAATTCATTC
PSCK1 EXON7 F	GCTGGGTTACAGAAATTGGG
PSCK1 EXON 8 F	CACACATGTAGTCGTTGGGG
PSCK1 EXON9 F	ATAGCAGATGCAGCCTTTGG
PSCK1 EXON10 F	CACTTGCTTTTGCAAGGAAAC
PSCK1 EXON11 F	GCAGATATGCATAAAATGCAAG
PSCK1 EXON12 F	TCAGTTATCAGATGCTAGAGTGTATCC
PSCK1 EXON 13 F	TCACATGAAATGCACAGAATC
PSCK1 EXON14 F	CGTTCCTGTGGTAGGGTTG
PSCK1 EXON1 R	CTCCCACTTGGAAGACCG
PSCK1 EXON2 R	CCACATTTGCAAAATGCTTC
PSCK1 EXON3 R	GCTCCCTATGAAATTCTCCATC
PSCK1 EXON4 R	GCATCCCCTTGAAAAGAAGG
PSCK1 EXON5 R	TGTTGTGCATTAACATTTCCTG
PSCK1 EXON6 R	CATTTACAATGGGTGATGTAAGC
PSCK1 EXON7 R	TGAAGAATGATTTGGCCCTC
PSCK1 EXON8 R	AGGGGAATCATTTCACATGC
PSCK1 EXON9 R	TGCAGTATTCCAAAATTTCTCC
PSCK1 EXON10 R	AATCAGGCAGAATGGCAAAC
PSCK1 EXON11 R	CAAAAGCAATCAGTTATTTGAATC
PSCK1 EXON12 R	AAGGGGTACAATTCTTTAGGGC
PSCK1 EXON13 R	TTCCTGCTTGGAAGTTGGG
PSCK1 EXON 14 R	TTCAGACACAGGCAAAATCG
LNPEP X1F	CTCGGAGTAGGAAGCTCGG
LNPEP X1R	GACTGCCCACGACCCTC
LNPEP X2-1F	CAGTAGCTCTGAGATGGAAATAAGTC
LNPEP X2-1R	TTAGGTTCGGGTGTAGGCTG
LNPEP X2-2F	CTGCCCAGATGTACCTTTACC

LNPEP X2-2R	TCTCTGGTGTGGACAAAGGG
LNPEP X3F	GGTGCCTTGTACAGTAGGTGC
LNPEP X3R	TCTGGCAAAGAGGAATTTAAGTG
LNPEP X4F	TTGTTTGTTATTGCCATTTATTCC
LNPEP X4R	CCAGGAATCGTTAAAGCAAGAG
LNPEP X5F	TGCTAGTCTTGACAGTGCTACTTG
LNPEP X5R	CAAACAACTTGGGCAGAACC
LNPEP X6F	CTTCCTCAGCTGACATGTGG
LNPEP X6R	AAATGCTAACAATAACTCCAGGAC
LNPEP X7F	CTTGCTTGTTTCCCATTCAG
LNPEP X7R	TCAAGTTTGGCATACCTTCC
LNPEP X8F	TCCTGAAGTTTAGTGACACTGGTC
LNPEP X8R	CACACAGCTTGCTCACATACG
LNPEP X9F	ACAGTTGAACACAGTCCACATC
LNPEP X9R	TTGACATAAAATTTAGTGCTCCATTC
LNPEP X10-11F	TTGCTCATTTACCTAGATTAGAAGTG
LNPEP X10F	GGAAAAGAGGGAAGGGTACAG
LNPEP X10R	GGGATAATCTTGTTAGGTTTCAC
LNPEP X10-11R	TTGCTCATTTACCTAGATTAGAAGTG
LNPEP X11F	CATATGTAACTAATTTTCC
LNPEP X11R	GCACACTGGAAAGTGTAATC
LNPEP X12F	CCTTCACTAACGTGAAAATAACTG
LNPEP X12R	TTTCCCAATCCTCCAAAATG
LNPEP X13F	AGCATGCATAGACTTCTTCTGTG
LNPEP X13R	TTACTGGCTGAAAATGGAACC
LNPEP X14F	TTACAGGCACGAGACACTGC
LNPEP X14R	GATCCAGGATGAGAATACCCTG
LNPEP X15F	GAAACTCCACCACAGCCAAC
LNPEP X15R	GGGGAAAAGAGATCCCCTC
LNPEP X16F	GCAACAGTTATTTCTCAATTTGC
LNPEP X16R	TCTGTGAGCCAAAACACTGG
LNPEP X17F	CACTCAGGAACAACGCTTTAC
LNPEP X17R	ACTAAAACTGAGCAGCCTGG
LNPEP X18F	GCAGCACAGCCATCACTAAG
LNPEP X18R	TGGCTTTACCTTGGCTTTTG

LRAP X1-1F	TGCAGTGCCATGAAGAACTAC
LRAP X1-1R	GATTTCAAGATCTTTGCTGTGC
LRAP X1-2F	TCTTTGTCCACCCCAATCTC
LRAP X1-2R	GAAGTCTGCTCCTGGCAAAG
LRAP X2F	CCCAGGTTTGATAAGCTGTTTG
LRAP X2R	ATGGCTGAAATCCTGAAAGC
LRAP X3F	TGCCATAAGTCATAGGCATGG
LRAP X3R	CAAAAGTTCATTGCTCTCAGG
LRAP X4F	AATGGAATTATGCCAGGGAG
LRAP X4R	GGAAGAGATCAAAGCAAGTTTTATG
LRAP X5F	AAAGGATCATAGTGTATTTGAGCAG
LRAP X5R	TTGTTTCAGGAAGATTGGTCG
LRAP X6F	CAGAGAGAAAAGGGTAGCAAAG
LRAP X6R	TCCAGTTCTGTGATCTGTTTCG
LRAP X7F	CATCTTACTAAACCTACTATG
LRAP X7R	GGCACCAGTGACTTAAACTGCTG
LRAP X8F	GCTTCCTTTAAAAACATACC
LRAP X8R	GCAGAATAAATATGGTTGACTGAC
LRAP X10F	TGTCCCTCTGTCTTTGGATTG
LRAP X10R	AAGGAAATAACGAGGAAACTGAG
LRAP X11F	TGGGTACTTAGGTGCCTTTTAC
LRAP X11R	TTATTGTATCGTTCCTTCAAATAGC
LRAP X12F	TCCCCAAACTTCATCTTCCC
LRAP X12R	TTGAAGTCTAGCATATCTGGTTACG
LRAP X13F	AAGGCAATAGGTCCAGAGAGG
LRAP X13R	GCTCTTAGGATTGGCAATGG
LRAP X14F	ACTGTTTGGGGAAAGATTGG
LRAP X14R	AGGGTTTGTAATGCCCAGC
LRAP X15F	AAAATGCCTTCAACTTGGTG
LRAP X15R	GCCTTTAAAACTGGTGGCTG
LRAP X16F	TGGTGGCTTGAGTTAATGTCC
LRAP X16R	TTGAGGCAACTGAGACATCTG
LRAP X17F	TTTCACAGCTTTAGAATTGTACTGG
LRAP X17R	GGTTTTAATCCTGTCCCTTCC
LRAP X18F	CCTGTGTGTATGTAATTTGCAAGC

LRAP X18R	TGAGCTTCACTGGTGTTACCC
ELL2 EX1F(2)	CTGCAGAAGCCCAAGCAGCCG
ELL2 X1R (3)	GCCAGCCACGGCTCCGCCGGC
ELL2X2F	GCCTGGCCCAAATAACATAG
ELL2X2R	AAAGGGTCAACTGAAAACGG
ELL2X3F	CATGTTTGCCAGATTCCTTG
ELL2X3R	TTGAACTGTTGTGATTATATCCCC
ELL2X4F	TTGCAAACTTAGTTGGGCTG
ELL2X4R	TCACCATGTTATCCAAAATGTTC
ELLRX5F	TTGTGGAAACTGAGGTGCTC
ELL2X5R	GTCCCAAATAAAGACAATGCTC
ELL2X6-7F	AAGGTTGCCTGTTTTAGGTTAGG
ELL2X6-7R	CTCCGGGTCTCATTTCCTTC
ELL2X8_1F	GCTTAAGCATGAAGAAAGCG
ELL2X8_1R	CATGATAGCATGCACTTCTGG
ELL2X8_2F	TTTCTCCTTTTCCCTCCCC
ELL2X8_2R	GTCATGCTGAGGCTGCAC
ELL2X9R	CAGGCATTTGATAACCTTCG
ELL2X9R	AGGGTGGCTCCAGGACTC
ELL2X10F	AGTGGCAGAGTGGCATTAAAC
ELL2X10R	AGGGCAGGGAAAGAAAGAAG
ELL2 EX11-12F	CCTATACTCTTGCTGAATGCATG
ELL2 X11R	GAGATTGCAGAATTTACAAC
ELL2 X12F	GTTGTAAATTCTGCAATCTC
ELL2 EX11-12R	CTTTCAACAGCCAAAGTTTCACC
RGMBX2F	GTGTTCCGAAGTTCAGGGG
RGMBX2R	GAAAGGAAAGGAGAGAAAATAAGG
RGMBX3F	CTTTGTGTCTTCTCTTCCGC
RGMBX3R	GCTATTTGGGAACAAAGCCC
RGMBX4F	TTCTCATGTAGCTCAGGAAGTTATG
RGMBX4R	AGGCCCTTTATAGCACATTTC
RGMBX5_1F	GGTTACCCCGTTCTGCTTC
RGMBX5_1R	TGTTGGCAGTCTCCAGTGTG
RGMBX5_2F	CATCCGTATGCCTGAAGACC
RGMBX5_2R	GGTGTCCCAGGACAAACATC

NUDT12X2F	GGGCATATAAATTGCGTTGC
NUDT12X2R	GACTCCTCATAAATTTGCTTTGC
NUDT12X3F	TGAAATAAATACGGAGAAAGTGTTTAC
NUDT12X3R	GGAAAGTGCAAATACATACAACTCC
NUDT12X4F	AAAACCCCTGATCACACAGC
NUDT12X4R	AAAGGAAAACTCTTTAGCTCCTCAC
NUDT12X5F	TCAGCTACATTAGGGCAGAAAG
NUDT12X5R	AAAGCAGACCAAATTACTAGGAAAAG
NUDT12X6F	TTGAACCATTCAGGACACAAAG
NUDT12X6R	TGCTCTGTGACCAAACCAATAC
NUDT12X7F	TGGATGTTGTCACTTTCTACTGTTG
NUDT12X7R	CCAACATCGTATTTTGTGTTGTG
CHD1 FRAG1 F	TTTATATTCATCTTTTAAACTGGGAAG
CHD1 FRAG1 R	ACATTAACAGTTGCTTGGCG
CHD1 FRAG2 F	CAAAGTCAAAAGCAGAAAACCTC
CHD1 FRAG2 R	TTCAGGGCTACAAACCTTGG
CHD1 FRAG3 F	TCCAATCAAAAGTCAGCAGC
CHD1 FRAG3 R	TCTTCTTCAAAATCTTCCCAGG
CHD1 FRAG4 F	GTGTTGATGAAGCACACCG
CHD1 FRAG4 R	ACAACAGTGTCAGCAGAGGC
CHD1 FRAG5 F	TCCCCTTTCAAAGATTAGATGG
CHD1 FRAG5 R	TCTCCTACTTCTACTGCGCC
CHD1 FRAG6 F	GAATTGGGAGGAAATTATTCCAG
CHD1 FRAG6 R	TGATGAGGTAGTCTGCACGG
CHD1 FRAG7 F	TGGGGCAAAGAAGATGATTC
CHD1 FRAG7 R	CGTGAGGATTCAAGTTGCTG
CHD1 FRAG8 F	TCCTGAACAAATTAAGCAATGG
CHD1 FRAG8 R	CTGTGTTGGTTTATGATATATGGC
CHD1 EX7(2) F	GCTGTGGGATCTCAGGTGGATC
CHD1 EX7(2) R	GCTGTGGACAAAGACAGCAAC
CHD1 EX14(2) F	CATGCATCTTCTTGGGAGGGAG
CHD1 EX14(2) R	GTAGAGAAGGTGGGGGCACC
TCC37 X4F	CTTTTCCAGGGTCCCTTTTC
TTC37 X4R	ACCACTTTTGCTGTTTGGAC
TTC37 X5R	TCCTGACTCATGCTAGACTTTCTG

TTC37 X5R	AGTTTAATGTTGAAACCTGTATTTTG
TTC37 X6F	ATATTTAGAGGCAAATCTCAGACAG
TTC37 X6R	AAAGCTTCGCAATTCCCAG
TTC37 X7F	GGATTGACTAATTTTGGAACTTTG
TTC37 X7R	ACCACAACATCTGCCTCCTG
TTC37 X8F	ATGGGGTTCTATGCTTTTGC
TTC37 X8R	AAATGACAATGATCAATTATGTTTCC
TTC37 X9-10F	TGCAGTTAATTGAACGTGTAGTTTG
TTC37 X9-10R	AACAACACCTTAAACATGGTAGACAC
TTC37 X11F	GGGTTTGGGCATTTAGAAAC
TTC37 X11R	GGCTCAAAGATAAATGCCTACTG
TTC37 X12F	TGGATGGCAAGCACTTGTAG
TTC37 X12R	TGTCAATATTCCAAATGCAGG
TTC37 X13F	TTGTTTTGGAAATTTATGCTTATTG
TTC37 X13R	AGGCAGAGAATGGTTGTTGC
TTC37 X14-15F	AACATTGCTGTGGTTGTGTTG
TTC37 X14-15R	AAAATCAAATCTTCATCAAGCC
TTC37 X16F	TTGGCACCTCTGAATCACTC
TTC37 X16R	CAAAACCTGACATATGAAAGATAAAG
TTC37 X17F	AGGGACAGCTTTCCTGTCTG
TTC37 X17R	AAAAGACAGAACGCTCATGTAAAG
TTC37 X18F	TTGAAACTGTAATTTGATGTATACCAG
TTC37 X18R	TCTCAAAATACAACCTGGTAAGAGAC
TTC37 X19F	GAGATGGACACCTGTCCTGG
TTC37 X19R	TCATATGGATGCTTGCTGTTG
TTC37 X20F	AAAGTGCTTTAAAATAACACCAATAAC
TTC37 X20R	TGTATGTTATAGAATTGGCTTGTCTG
TTC37 X21-22F	AATCAGGGAGATAAACTTCATGC
TTC37 X21-22R	CCAAAACAACTTCGAGAAAAGC
TTC37 X23-25	TTGGATGAAGTAAATAGTGTTTTCAC
TTC37 X23-25	CATTTCTCCACAAGAAGCTGAG
TTC37 X26F	AAATCTGTTGCACCCAGGAG
TTC37 X26R	AGCCCTCAAATGCTATCCG
TTC37 X27F	GCTCAGCTAGTAAACCTTTGGTTC
TTC37 X27R	TAAGCCACTATGCCTAGCCC

TTC37 X28F	TTAACAATTGACTTAAAGAACACTGAG
TTC37 X28R	AGACTGCTTCGCTGTTGC
TTC37 X29F	GCCACTATTCTTAACAGTGTGAGTG
TTC37 X29R	AGGAACCCTTCCCCTCATC
TTC37 X30F	TTCAGCTAAAGAGGAAGGTAACAG
TTC37 X30R	CAGTCCCATAGCTTTCACCAG
TTC37 X31F	GTCCGGCCCAACTTAGTG
TTC37 X31R	TTTTAATAAACATTGCAACCCC
TTC37 X32F	GCTTCATCAATCAGGATGGC
TTC37 X32R	GCTTTCTCCAATTCTCAACCC
TTC37 X33F	AGAATGGCGTGAACCTGG
TTC37 X33R	AGTCCCAATGTCAAATGCAAC
TTC37 X34F	GGTTGCTTGTACATTTTGTGG
TTC37 X34R	TCTTGCAGAAGAATAATGCCATAG
TTC37 X35-36F	TCCAGTGATTTAGTGGAAAGGTC
TTC37 X35-36R	CACCCAAACTCCTGTTGTAGAAG
TTC37 X37F	GGAAAGATGAATGTAGCAATGTG
TTC37 X37R	TGCTGAAGCCGTTTGTATTC
TTC37 X38F	TTTTAGCAACCAGTTACTCCTCTAAG
TTC37 X38R	CACCGCATCTAGCCTCATAC
TTC37 X39F	TTTACAGGCTGTTCATTGTTTTG
TTC37 X39R	AAATGGAGCATAGCCCTGAC
TTC37 X40F	TGTTGGAAGTAGGGGTAACAGAC
TTC37 X40R	AAGCCTTCACGTCAAAGAAAG
TTC37 X41F	AACCTGGGAGGCAGAGC
TTC37 X41R	CTGGAACCAGTTCTTCCAAAG
TTC37 X42F	TGGTACCTAGTGCATGATAAACAC
TTC37 X42R	AACCCATGGTAAGATACAAACTG
TTC37 X43F	CCCTTGGATTACTTGGCTTC
TTC37 X43R	GGGTAGAAGTCTTGATTCATTGC
TTC37 FRAG X10F	GGCAGGAACAAGGTGCAG
TTC37 FRAG X10R	GATGATACCATCCACTTGTGC
TTC37 FRAG X17F	GCTGCAAAGATTATGGAAGACC
TTC37 FRAG X17R	GTTGTGTAGCCTCCTCTGC
TTC37 FRAG X29F	GCTCTTATTGCTGAGGCAG

TTC37 FRAG X29R	CATAGGCTTTGCTGCTCTC
TTC37 X41(1)F	CAGAGCGAGACTCCGTCTC
TTC37 X41(2)F	GAATTGAATTGAATACTTTTC
TTC37 X41(1)R	GGTCTGTTACAAACTTTAAAG
TTC37 X41(2)R	GGATTTCTGGAACCAGTTCTTCC
TTC37 FRAGX29 (1)F	GCAAGCTCATGAGGCTTTC
TTC37 FRAGX29 (2)F	GACACCATGGATCTCTTCAG
TTC37 FRAGX29 (1)R	GCCAATGCAAAACCTATGATGTC
TTC37 FRAGX29 (2)R	GTTATTGCCAGAGCTGTCAAG

Primers used in chapter 4

DNM1EX1F	ACTACGACACCCATGATGCC
DNM1EX1R	GACGACGTTTGCAGACTAGC
DNM1EX2_3F	ACAGGTACCCCTGGGACAG
DNM1EX2_3R	AGTCTGCAGGAGGAGCGAG
DNM1EX4F	ACCCCGCCCTATTCTTTAAC
DNM1EX4R	TCAGAGGCTCCCCTTTACAC
DNM1EX5_6F	TGCATGGAATTGTGTGTGAG
DNM1EX5_6R	AGGCTAACCTCTGCTGTTCAC
DNM1EX7_8F	CATCCAAGTCCCTTCCTGG
DNM1EX7_8R	CGCAGAGAGAGTCCTGAGTG
DNM1EX9F	CCCTGTCTTGACCTCCCAG
DNM1EX9R	ACTTCATCACTCCCTGTGCC
DNM1EX10F	CAGAGTTCGTGGGCTTGG
DNM1EX10R	CTGGCAGCAGTGCAGGG
DNM1EX11F	TGGACTCGTGGGGTGTG
DNM1EX11R	AGGAGTGGAAGGCGCAG
DNM1EX12_14F	TACTTTTCTCCCAGCTTGCC
DNM1EX12_14R	GTTGCAGATCCCCACTGC
DNM1EX15F	TTTCGAATGCCTCTCTTTGC
DNM1EX15R	GTCTGAGCCTGACTCTGTCC
DNM1EX16F	GGGGACTCTTAGAGAGTGAGGTG
DNM1EX16R	GGATCCCTGTACCCAGCTTC
DNM1EX17_18F	GGCATGCCCTTAACCCC
DNM1EX17_18R	GTCCCTTGCTTGAAAACCTG
DNM1EX19F	ATTCTGGGTACCCTTGGAGG
DNM1EX19R	AACATGGCCACCTGCATC
DNM1EX20F	AAGCCACACCAGCTCACAG
DNM1EX20R	CTCTAAGCTCCGCCCCTC
DNM1EX21F	CTTAGGGGTGCGGCAGG
DNM1EX21R	CGATTGCCTGGGATAGAAAC
DNM1EX22F	TTGCCTTACCAGCTCTCTCC

DNM1EX22R	CAAGTCACAGAACAGCCGTC
KIAA0586 EX2 F	GGGACTTGTTTTGTGACCAAC
KIAA0586 EX2 R	AAACTAGACTTATATCTTACGTTTGGG
KIAA0586 EX3 F	GAATCATGCATATACAGTTTCTAAAGC
KIAA0586 EX3 R	AAACAAGGCATGCTGCTAATC
KIAA0586 EX4 F	AATACACAGTCCTGCCCTCC
KIAA0586 EX4 R	ACAAAGCTCTGGATATGAAAGG
KIAA0586 EX5 F	TTCGATCGTGGAATGCTATG
KIAA0586 EX5 R	GTGTCTAGCACCATGAGAAATG
KIAA0586 EX6 F	AACCACCTTCGGCAGATTAC
KIAA0586 EX6 R	AACACACAACCACAAATGGG
KIAA0586 EX7 F	CTGCCTTTCATTTCTGTGACTC
KIAA0586 EX7 R	TTACCCGCATTACAGGAGAAG
KIAA0586 EX8 F	GCTGCTACTACAGTGGATACTTGC
KIAA0586 EX8 R	AGCACAAAGAAAATTCACAGC
KIAA0586 EX9 F	CACCGCAGCTTCAGTAGTTC
KIAA0586 EX9 R	AAGAAGCTTGATTTTCCCAGC
KIAA0586 EX10 F	AGATGATCCACCCACCTCAG
KIAA0586 EX1O R	TTTTGGCTAGCCTCATTTTATC
KIAA0586 EX11 F	AAGACTGTATCAAAGATTTGTGTAATG
KIAA0586 EX11 R	TTCTTCATCCTTCCTCCTAATTTC
KIAA0586 EX12 F	ATGCCTGGCCCACATTAG
KIAA0586 EX12 R	TGTGGATCCAAGAAGATATGTACC
KIAA0586 EX13 F	AGATCTGATCCAAGTCCTGCTC
KIAA0586 EX13 R	TGTAATGCCTGTAAATTGGAATC
KIAA0586 EX14 F	CATTCCCCACTCCAAAAGATAC
KIAA0586 EX14 R	AAACACCTTCACATTCGCTTC
KIAA0586 EX15 F	AAGGGCTGTTGAAAACAGAG
KIAA0586 EX15 R	TTTTCCATCCCCAAGCATC
KIAA0586 EX16 F	TTTCTGTGAAGAGCTGTTGTATTTC
KIAA0586 EX16 R	CTGAGCCACACAGTGAAACC
KIAA0586 EX17 F	TCCAGATGGTATATTAGACTTTTCCC
KIAA0586 EX17 R	GAAGCAATATAATGGATTTCCG
KIAA0586 EX18 F	AAAACTTAAAGGGCATGCTTTG
KIAA0586 EX18 R	TCATTATGATAAAAGGCTTTAAATGG
KIAA0586 EX19 F	TGGGGACTAGCTTGGATTTC
KIAA0586 EX19 R	CCAAAGATAGTCAACAACCTGG
KIAA0586 EX20 F	TCACCTAGTTCCTTGCCTATAGTTG
KIAA0586 EX20 R	TTTCAGCCACATTTTAATACTTTCC
KIAA0586 EX21 F	CAACATGTTTGTGGCATTGG
KIAA0586 EX21 R	TGGAGAATTTACTTGACCTCTGG
KIAA0586 EX22 F	CTTTTAAGAATGAAAGTTCGGC
KIAA0586 EX22 R	GCCCAGAACTTCAAGACCAG
KIAA0586 EX23 F	CTGCAGGCCAATCGTAGAC
KIAA0586 EX23 R	TGCCTGGCCTCAATTATAGAC
KIAA0586 EX24 F	AAAATGGTAAGGGATATCTGAAGC
KIAA0586 EX24 R	TGTTTTCAATGGATCACATAAGTC

KIAA0586 EX25 F	CAAAATGAATACAGGCCTTCAG
KIAA0586 EX25 R	AACTAAATTGCATTGAGTCCCC
KIAA0586 EX26 F	TATGCTATGGCTTTATGGTCAG
KIAA0586 EX26 R	AAACCTGATCAGTGGCAACC
KIAA0586 EX27 F	TTTGAGGCAAAGATAGAGGG
KIAA0586 EX27 R	AAGGTGGTTGGAACAGATGG
KIAA0586 EX28 F	CCTTCCAAGGAGTCAAAGGG
KIAA0586 EX28 R	AAAGAAATTTTGCAAGCCCC
KIAA0586 EX29 F	GCAGGTGAGAAATATTGGTGG
KIAA0586 EX29 R	ACCACGCTAAGGCAACAGTG
KIAA0586 EX30 F	CACCTATTTATTGGTAGAAGTGCG
KIAA0586 EX30 R	AAACCTAGAGAGCAGTGGGATG
KIAA0586 EX31 F	GGGAGTGCATTTGGTTTGTG
KIAA0586 EX31 R	GCTGAGTTTGTAACCTGGGG
KIAA0586 EX32 F	TCGCAAAGCATAAAGCACG
KIAA0586 EX32 R	GGGTTTTAAGCCAAGGTAAAATG
IFT81 X2F	CCAGGACTTTTGTTGCCAG
IFT81X2R	AATGTGCCCAAGACTCACAG
IFT81X3F	AGCAGAACAATTTTGTCAGTGAG
IFT81X3R	TGATTGCAGTATATGAAGGTTTGG
IFT81X4F	GCTGTCTCTCTTCAAAGTTTACCG
IFT81X4R	TGCCACTTGGAACTACAAAGAC
IFT81X6-7F	CCTGCATAAGAAGCCCTGG
IFT81X6-7R	TGTTAGCCACTGCACCCG
IFT81X8F	TTTTCATTGCTTGCCTAAGTG
IFT81X8R	TGATGTTCCAAAATGCTGAGAC
IFT81X9F	CAAATATGCCGTATCCAGGTC
IFT81X9R	TGGGATACTACAGAAAGCAGTTC
IFT81X10F	AAATTCAAAGCTGGTTTGGC
IFT18X10R	TGGAGATTAGCAGTGACAGAAC
IFT81X11F	TTGGTTAAGAGAGGCCAAGC
IFT81X11R	AACGTCCCAATATTAACATTCAC
IFT81X12F	ACGTGTGACATGTTGGGAAG
IFT81X12R	GCTGTAACCCATAAACCTCCTG
IFT81X13F	CAATGAAGATTGTCCTTTAGGATTC
IFT81X13R	AACACAAAACACCTGCCTTC
IFT81X14F	TGTTTGGGACTGAAGTAATTTGC
IFT81X14R	CCCAAGTAAAATATCCAGGGTTC
IFT81X15F	GTTTCAGAGCCATGTTTGGG
IFT81X15R	CTAGGCTCAAGCAATCCTCC
IFT81X16-17F	CATCGTCGCCATATCTGTCTC
IFT18X16-17R	TTCCTACTTAGCAGGGCTAAAATATC
IFT81X18F	TGGGTTTCATTCTGACTCCAG
IFT81X18R	TGAATGAACAATTTCCCCTC
ODF2EXON5-6F	GCCCAGTTTACTCATCCACAG
ODF2EXON5-6R	GTGAGCCTACCTCCCACTTG
ODF2EXON 7F	ACTCGTGGTCAAGGTTCTGC

ODF2EXON 7R	GAAATCCACCTGGGTGCTC
ODF2EXON 8F	TGTGGCATAGAGGAAGGACC
ODF2EXON8R	ATTATTGCCCAGGGCTCAAG
ODF2EXON 9F	GATCTCTGGGCATGCGG
ODF2EXON9R	ACGCGTAAAGACCCTGCC
ODF2EXON 10F	GGCCTGCTTCTTCCCTTC
ODF2EXON10R	GAAAGGCTCAGAGAGGATGG
ODF2EXON 11F	AAATGTAGACATCTGGGTTTATTGC
ODF2EXON11R	TAGTGGAACGACGCTCTTGC
ODF2EXON12F	CAAATCTCTGTCCTCCCTGG
ODF2EXON12R	GAACCCAGCTCTCTGCATATC
ODF2EXON 13F	ACGGAGCAAGACTCCGTATG
ODF2EXON13R	CTACCATCAAATTGGCAGGC
ODF2EXON 14F	CTCCTTCACCCTCTGCTTTG
ODF2EXON14R	CACCCAGCCCTCCTCAG
ODF2EXON 15F	GGCAGCCTCACCACTGAC
ODF2EXON15R	CTTCCCTGAGCCACTAGGC
ODF2EXON 16F	GTCCCACACCACACAAAGAC
ODF2EXON16R	TGGTCAAATCCAGGCTTTTC
ODF2EXON 17-18F	GTAGGAGGTCCACCTGAAGC
ODF2EXON17-18R	CAAGGGGAACAGACTCACAG
ODF2EXON19F	CCACTAGGAGCCAGGTCTTTC
ODF2EXON19R	CAGCTCATCTTGGTATGCGG
ODF2EXON 20F	ACAGGTCATAGGGAAAGGGG
ODF2EXON20R	TGGCTGAAAGTCTCAGGTCC
ODF2EXON 21F	GTGGAACATGACAGGTTGGG
ODF2EXON21R	GCTGAGAGTTTAGGTGATTTGC
ODF2EXON 22F	TCTCTTTGCCAAGTGTTCCC
ODF2EXON22R	GCCCCACAGAAGCTAGGTC
ODF2EXON 23F	AAGAAAGTCCCTATGTCTCTATGG
ODF2EXON23R	GGCAACTCTGGAGTTCCTATG
IFT52X2F	TCCTAGATAGGTTTCAGTTGTATACCC
IFT52X2R	TTTAAATTGAGCTTCTCCCACC
IFT52X3F	GCCTCTGCTTCCAAATAGTTAAG
IFT52X3R	ATCCGGTATTTCAAAGCTCAG
IFT52X4-5F	TGGTATTTCGGGCAAAAGTC
IFT52X4-5R	AAATGATCCATCCACCTTGG
IFT52X6F	TGCCCTTTCTTCTTTCTTGC
IFT52X6R	GGGGAGGAGCATGGACTTAC
IFT52X7F	AATGAATCAAAGAGTGTTACTCAGC
IFT52X7R	TCCACTGTTTCCCCATAAGC
IFT52X8F	TGGGTACTAGGATACATGAGATGC
IFT52X8R	TTTAAATGAGCAAATAAACTTACCAG
IFT52X9F	TCTTGATTTAATTCTCTAGTCCTGACC
IFT52X9R	TTGAGCTGTGAGAAGTGGTACTG
IFT52X10F	GACCTGAGACAGTGGAGACTTG
IFT52X10R	TAATTCCCGTTCCCAGGTTC

IFT51X11F	CACTGTATTTCATGGAGAACAACAC
IFT52X11R	CCTGCTGACAAGATAAGGGC
IFT52X12F	TTTGGACAAGCTCTTTCTTGAG
IFT52X12R	AGCCTGGGCAACAAGAG
IFT52X13F	GTCTGTACTTTGGGAAGGGC
IFT52X13R	CGTCATAAAACATGGCAAGC
IFT52X14F	CAGCCCTAAAGCCAGATATTATG
IFT52X14R	AGCTCATGGAGGAAAGAGTG
IFT80X2F	TCCCTTGTTAGGATTGGCAC
IFT80X2R	GCAAACGGTTCACTTTGGAG
IFT80X3F	TTGAAACCATTTGACCCTATTTTAC
IFT80X3R	TGCAGATCCACAAATACCAG
IFT80X4F	AATGTTAATATTTGCCTTTATGCC
IFT80X4R	CAATGAAAGACACTATTGCTAAATG
IFT80X5F	AAAGAATATGCGGTAGATACTGTTG
IFT80X5R	CCAAAGTCAGTCTCAACCACC
IFT80X6F	AAGAGAAGAAATGAAATGACACCTC
IFT80X6R	GCAAGTGCCCTCTAGTTTAGC
IFT80X7F	CCCTGAGAAGATGACAAAGG
IFT80X7R	TCCATGCTTTAAAATACCACTTAAC
IFT80X8F	TTGCTAATATATAAGGTGCTGATTTC
IFT80X8R	CCATGTGTGTTGTGCATTACC
IFT80X9F	CATGTGCCTGATTTAAAACTTCC
IFT80X9R	AGCTTTGGGCATTCAGG
IFT80X10F	TCTGTATCTATGCTTTGTCAGTTGG
IFT80X10R	CACTTTCTCTCACTTCCGGG
IFT80X11F	TGACTTGCCCAAACTCACAG
IFT80X11R	CCTGTGTTGCTCTCATTCTCTC
IFT80X12F	AACAACCAATGCTGGCTTTC
IFT80X12R	AAGTTCAAACTCTCTACCAACAGC
IFT80X13F	AATTGGATGCTGCTGCTCTC
IFT80X13R	CCAGTGCTTAGTCTTTAACTGTGAC
IFT80X14F	TTCAGTGCAGCTTTTGAACG
IFT80X14R	GACCTTGATCTTGCCACTCC
IFT80X15F	ATGGGATTTTGTTTGTCAGG
IFT80X15R	TGTAAAACACCTTGCAGGATTG
IFT80X16F	GTATTTCACGGTCCACTCCC
IFT80X16R	TCATTTTGCACAAAGGTAAACTAC
IFT80X17F	ATGGCCACAATGAAATTCCC
IFT80X17R	CCACTAAACAGTAGTATGTGGGC
IFT80X18F	TGCAGCAATTGGTGAAGTAAG
IFT80X18R	AAATTCTTCTGTGGTGTTAAGAAAC
IFT80X19F	AGAGGCCATGTTTGATCTGG
IFT80X19R	TCAAAATATGCCTGCAAGATTC
IFT80X20-1F	AATCTTTTGGAAGTCACACAGTC
IFT80X20-1R	GAGCTGTTTCCATCTTTTGC
IFT80X20-1F	ATTTGACCTCTAAATTGTAAAATGC

IFT80X20-2R	CACAAATTCAGACTTCAGGGC
LMX1BX1F	ATAGCAACAGGTCCCGAGTC
LMX1BX1R	GTCCACAGCCGGACGAC
LMX1BX2F	GGACTGGGACGGACTAGC
LMX1BX2R	AGCTCTCGGAACCCTTGG
LMX1BX3F	CCTCTGGGAGGGACTTCTG
LMX1BX3R	ACTCCCCTCCAGGACACC
LMX1BX4-5F	CGAAGGGGACAAGGCTG
LMX1BX4-5R	GTAGTCTTCTGGCTGCCCC
LMX1BX6F	GCAGCCAGAAGACTACGGTC
LMX1BX6R	CCTCTGCCCCAGCTCAC
LMX1BX7F	AAGGAGATCAGAAGGGGAGG
LMX1BX7R	CCTAGGGCAGCAGGTGG
LMX1BX8F	ACAGCCTACAGGGCAAACAG
LMX1BX8R	GTAGTCTGTGCGGAGAGCAG
WDR34X1F	GACCAATTCAAACATGGCG
WDR34X1R	ACTGTAAAGGCAAACGGTGG
WDR34X2F	ACAGCAGCAAGGAGGGG
WDR34X2R	TGGGAAGGAAAACAAGAGGAG
WDR34X3F	TAGGTGGGGTTTGGTTAGGC
WDR34X3R	AGACAAGGCCCAAGGAGG
WDR34X4F	GTGATAGGAGCCCCACCC
WDR34X4R	ATGGAAAGCTCCCTTGTGC
WDR34X5F	GTGTTTGGCCAGGAAGTCAG
WDR34X5R	CTCTCCCCAGCTCTGCC
WDR34X6F	TTTCCTGGAAGAGAGGGAGG
WDR34X6R	AGGGGAGATCACAGCAAGTC
WDR34X7F	TGTTCTCTGGCTCTGTTGGG
WDR34X7R	CAGGCTAGAGACCCGCAG
WDR34X8F	CACCACCCAGACCCCAG
WDR34X8R	AGAACCCCAGGTACAAGCAG
WDR34X8F	CTGCCATGTGGTTGGTCTC
WDR34X9R	ATTTGGCTTGCGTCAGAAAC
C9ORF74X1F	CACCGCCTCTTCCGTCC
C9ORF74X1R	CACCGCCTCTTCCGTCC
C9ORF74X2F	CCTCAGTTTTCTGTTTCCGC
C9ORF74X2R	CCATGGACCAACTCCTGTTC
C9ORF74X3F	GTGGGCAGGTGCTATTTGG
C9ORF74X3R	AAGGCCAGGACTGAACCC
C9ORF74X4-5F	AGCTCTCAGCCTCTCTTCCTC
C9ORF74X4-5R	GGATGTCTGATTGCTTCGTTC
C9ORF16X1F	CCCGTGTTCTATCCGCC
C9ORF16X1R	ATGGGTGGGCTCTTTGTTC
C9ORF16X2F	TTCCAGACGGTGACACTGAG
C9ORF16X2R	GGCAAGTGGGTGGGTTG
ZNF79X1F	ATAGACCCTTACGCCCAGAG
ZNF79X1R	CTGATTCGGCTCACACAGC

ZNF79X3F	AACTGCCAACGTTAACCACC
ZNF79X3R	AAGCAAGTGCCACAGATTTTC
ZNF79X4F	CTCTCCTGAGTTCTGGGTGG
ZNF79X4R	ATTTAAAGGCGCTTTTGCTC
ZNF79X5F	ACCCATGTGTAGAGATGCCC
ZNF79X5R	TGTTCCTAGTTACTCTCCGGC
ZNF297X3_1F	TTATCTTTGGGCTGGAATTTG
ZNF297X3_1R	CCATCTCTCAGTTCTTGGGC
ZNF297X3_2F	ACCACCAGTCACCAAGCAG
ZNF297X3_2R	TTCCCATCTGACCTCTCTCC
ZNF297X3_3F	GGGGAGAAGAAAGTGGAAGC
ZNF297X3_3R	GCCCAGATTTGCATTAATTACC
C9ORF88X1F	CTCAGCAGCAAACAGCTCTC
C9ORF88X1R	TGACAGAGCAAAACCTCCTC
C90RF88X2F	CTAGGGGAAGGATCGTGATG
C9ORF88X2R	AGGCCACATGGCGAGAAG
C9ORF88X3F	CAGCTTCATGGGGAGCAC
C9ORF88X3R	GCACACATGGAGAAAGCG
C9ORF88X4F	CCACACTGAGTAGGGCCTTG
C9ORF88X4R	ATGGGGCTGAAAGCAATG
C9ORF88X5F	CTCTGACATCCTCTGGTGGAG
C9ORF88X5R	TTCAGCAATAACCCAGGTCC
C9ORF88X6F	CTTGGTAGAGGCCACTGGG
C9ORF88X6R	GCTGGGTCCTTGGGTGAC
C9ORF88X7-8F	GCTTGTGGGAATGCAGAGAG
C9ORF88X7-8R	AAAAGAAAGTCCACCTTGGC
C9ORF88X9F	CTCCAGCACCTGGGCTC
C9ORF88X9R	TCTGGGACAGACCAGACCTAC
C9ORF88X10F	GACCCAGGCAAAGTGTGTG
C9ORF88X10R	CCTGTGGAAGGCAGTGGAC
C9ORF88X11-12F	AGGAAGGGGCTGTGGTC
C9ORF88X11-12R	GTCAGGGCCAAGGGTAGAG
C9ORF88X13F	AGGCCTCCCTCTACCCTTG
C9ORF88X13R	TCAGCTGCCACCACTTCTC
C9ORF88X14_1F	TTGAGCAGAGTCCAGCACAG
C9ORF88X14_1R	CTCTCAGGCCGCAGACC
C9ORF88X14_2F	TTTGAGGCTAGCCCTGAGTC
C9ORF88X14_2R	GAGCTGAGCCTGCCTGG

Primers used in chapter 5

ABCB11X2F	CAAATTGTTCTTTCGTTTGGC
ABCB11X2R	TGCTCCTTGAAACTTGACCAG
ABCB11X3F	TGAGCAGGAAGAAAGAAAAGG

ABCB11X3R	CCTAGAAGGGATATTCCAAAAGG
ABCB11X4F	GCCAGTGGGGATTTTCTTTC
ABCB11X4R	AACACTCCCCTCATGATCTAAAC
ABCB11X5F	AGTCCTCCTACCTCTCCTGC
ABCB11X5R	TCAGCCAGTAAAATCCCCTC
ABCB11X6F	AATCTCTGGTGGCTTGATCC
ABCB11X6R	GTGGCAACACATTGCATCTC
ABCB11X7F	CCCCTTTTCTCAACTGTTGTATTG
ABCB11X7R	AATTTAGAAACAAGGGTTTTATTATCC
ABCB11X8F	GAGAGATGGGAATGTTTAAAAGG
ABCB11X8R	TCAGGAAAAGGGACTCAAGC
ABCB11X9F	GACAGACTGACTTACCTAATTTCTTGG
ABCB11X9R	CCGCTTTGCACAAACTGAG
ABCB11X10F	TCCCTGAAGCTGCTCTGTG
ABCB11X10R	CCTGAAGGCACCAAAGTAATAAAC
ABCB11X11F	TGCGTTAACATGGAAGACCC
ABCB11X11R	GAGTTCATTCTGTGCCCCAC
ABCB11X12F	GCAGAGATACGCCAAAGATG
ABCB11X12R	GGAAACAGAGTCAGGCTTCAG
ABCB11X13F	AAGCATCTGCACCTGTAGCC
ABCB11X13R	CTGCCATTTGCACTTTACTG
ABCB11X14F	TGCCCATTGGTCAAGTATG
ABCB11X14R	CTAAAACATGGCTTAAGAATTTAATG
ABCB11X15F	TTATAGTGGATCACTGTCAGAAGC
ABCB11X15R	AGCAGCACAAGCATTTCCAC
ABCB11X16F	TGATGCAAAGGTCAGTGTCAG
ABCB11X16R	CATAGAAAACCGTAAAGCACTATAGAC
ABCB11X17F	GTAAAGAATTCTACTTGGATATGGTTC
ABCB11X17R	CAGAGTTTCCTTGTTGTACCTGAG
ABCB11X18F	ACACCAGTTGATCCTGCTCC
ABCB11X18R	AAAGGGTACCCAACAGTCCC
ABCB11X19F	TGTGAATGCCAAAGGATCTG
ABCB11X19R	CATGAAAACAAAGAGCGGAC
ABCB11X20-21F	CCCACCAGAATGATACATTTCC
ABCB11X20-21R	ATCCCACTGGTCCCTATTCC

ABCB11X22F	TGGTAATTGGTAAAAGCGACTG
ABCB11X22R	AACAGTTTGTCTGATAGCCACTC
ABCB11X23F	GCCACTGAAATGTCACGAAAG
ABCB11X23R	CAGAACCAGGCTATTCCTTCC
ABCB11X24F	ATCACACCAACCACGCC
ABCB11X24R	CAACCTTACCCCTCATCAATAC
ABCB11X25F	AAACTCAAGATTTAGGTGTGTTTTC
ABCB11X25R	AGGGGTTGGAAATACTCTGC
ABCB11X26F	AAGCAAACCAAATGTCCTGC
ABCB11X26R	TGCTCAACCTGTACACTCTGG
ABCB11X27F	GAGTTCAGTACAGCACAGGAGC
ABCB11X27R	TTGAAAATAGTGCCATTTTATTAAGG
ABCB11X28F	TCATGATGGTGGGCTGG
ABCB11X28R	AAAACAATCCCAGCAATCCC

8.4 The nucleotide and amino acid sequences of *TTC37*

This is the coding sequence of TTC37. The exons are coloured alternately blue and black.

Nucleotide sequence

ATGTCCAGCAAGGAAGTGAAGACTGCTCTAAAAAGTGCTAGAGATGCAATCAGAAACAAAGA

ATACAAAGAAGCTTTGAAACACTGTAAG

ACAGTGTTAAAGCAAGAGAAAAATAACTATAATGCCTGGGTTTTTATTGGCGTTGCTGCAGCT

GAACTAGAACAACCTGATCAGGCCCAGAGTGCCTATAAAAAAGCTGCTGAATTAGAG

CCAGACCAATTACTAGCTTGGCAG

GGGTTAGCAAACTTGTATGAGAAATATAATCACATAAATGCTAAGGATGACTTGCCTGGTGTT

TACCAAAAGCTCCTGGATCTTTATGAGAG

TGTTGACAAGCAGAAGTGGTGTGATGTCTGCAAGAAACTTGTGGATCTATATTACCAAGAAAA

GAAACACCTAGAG

GTGGCTCGAACATGGCACAAGTTGATAAAAACACGGCAGGAACAAGGTGCAGAAAATGAAG

AGCTTCATCAACTATGGAGAAAATTGACTCAGTTCCTGGCTGAAAGTACAGAGGACCAGAAT

AATGAAACTCAGCAATTG

CTTTTTACTGCTTTTGAGAATGCACTGGGATTATCAGATAAGATTCCTAGTGAAGATCACCAA

GTACTTTATAGGCATTTCATTCAGAGTTTATCCAAA

TTTCCTCATGAGTCTGCTAGATTGAAGAAGGCCTGTGAAGGAATGATAAACATCTATCCTACT

GTACAGTATCCATTAGAAGTGCTTTGTTTGCATTTAATTGAATCAG

GAAATCTTACTGATGAGGGGCAGCAGTATTGTTGTAGATTAGTGGAAATGGATTCAAAAAGTG

GTCCAGGCCTCATTGGCTTAGGCATTAAAGCATTACAAGACAAAAAGTATGAAGATGCTGTTA

GGAACCTAACAGAAG

GGTTAAAGGAAAGCCCTGTCTGCACAAGTGGATGGTATCATCTGGCAGAAGCCCAAGTCAAA

ATGCATAGACCTAAAGAAGCTGTTCTTTCATGCAGTCAAG

CTCTGAAGATCGTAGATAATCTTGGTGCGTCTGGTAACAGTCTTTATCAGAGGAATCTTTGTCTTCATTTGAAAGCAGAGGCTTTGATTAAACTCTCAGATTATGACTCTTCAGAGGAAGCAATTCGTACGCTTGATCAG

ATTTCTGATGCAGATAATATCCCAGGACTTTTGGTTCTCAAAAGCTTGGCCTATCGGAACAAAGGTTCATTTGATGAAGCTGCAAAG

ATTATGGAAGACCTTCTCTCTTCTTACCCTGACCTAGCTGAAGTTCATGCCCTTGAGGCTTTGATTCATTTCACCAAAAAGGACTATCTACAAGCAGAAAAATG

TTTTCAGAGAGCTCTTGAGAAAGATACCGAAGTTGCAGAATATCATTACCAACTTGGATTAACATACTGGTTCATGGGTGAAGAGACAAGAAAAGATAAAACAAAGGCTCTTACCCACTTTCTGAAG

GCTGCAAGACTGGATACATATATGGGCAAAGTTTTCTGCTATTTAGGTCATTATTATAGAGACGTAGTGGGAGATAAAAACAGAGCTCGTGGATGTTATAGGAAAGCCTTTGAATTAGATGACACTGATGCTGAATCTGGAGCTGCAGCAGTTGACCTAAGTGTGGAGCTTGAAGATATG

GAAATGGCTTTAGCTATCCTAACAACAGTAACTCAAAAGGCAAGTGCTGGAACGGCAAAATGGGCCTGGCTTAGGCGAGGACTATACTATTTGAAAGCTGGTCAGCATTCTCAAGCAGTGGCTGA

TTTACAGGCAGCATTAAGAGCAGACCCAAAGGACTTCAATTGTTGGGAATCGTTAGGAGAAGCATACTTAAGCAGAGGAGGCTACACAACAGCCTTGAAGTCCTTCACAAAAGCCAGTGAGCTGAACCCAGAATCCATATACAGTGTGTTTAAGGTTGCAGCAATACAGCAAATCCTAGGCAAATATAAGGAGGCTGTAGCTCAATACCAGATGATCATTAAAAAGAAAGAAGATTATGTGCCTGCTTTAAAAG

GTTTGGGTGAATGCCATCTTATGATGGCAAAAGCAGCTCTAGTTGATTATCTTGATGGAAAAGCCGTAGACTACATAGAAAAAGCACTGGAATATTTTACTTG

TGCTCTACAGCATCGAGCTGATGTGTCCTGCCTCTGGAAGCTAGCTGGGGATGCTTGTACCTGTCTGTATGCTGTCGCACCATCTAAAGTGAATGTTCATGTTTTAGGAGTCCTTCTAGGTCAGAAAGAAGGAAAACAAGTATTAAAGAAAAATGAGCTCCTCCACCTTGGAGGAAG

GTGTTATGGTCGTGCATTAAAACTGATGTCTACATCTAATACATGGTGTGACCTTGGAATTAATTATTATCGCCAAGCACAACATCTAGCAGAAACAGGCAGCAACATGAATGATCTTAAGGAGTTGCTGGAGAAATCTTTACAT

TGTCTGAAAAAAGCAGTGAGACTCGACAGTAATAATCACTTATACTGGAATGCTCTTGGTGTG

GTTGCATGTTACAGTG

GTATTGGAAATTATGCCCTTGCTCAGCACTGTTTCATCAAATCAATCCAGTCAGAACAAATT

AATGCTGTTGCATGGACCAACTTGGGAGTGTTATACCTCACAAATGAAAACATTGAG

CAAGCTCATGAGGCTTTCAAAATGGCTCAATCCCTTGATCCATCTTATTTAATGTGCTGGATTG

GACAA

GCTCTTATTGCTGAGGCAGTTGGAAGTTATGACACCATGGATCTCTTCAGGCACACTACAGAA

CTAAATATGCAT

ACTGAAGGAGCATTAGGTTATGCGTATTGGGTCTGCACAACATTGCAAGATAAAAGCAACAG

AGAAACAGAGCTGTACCAGTACAACATCCTCCAGATGAATGCTATTCCAGCAGCACAAGTTAT

TTTGAATAAATATGTAG

AAAGAATTCAGAATTATGCCCCAGCTTTCACAATGTTGGGTTACTTAAACGAACATCTACAAC

TGAAAAAGGAAGCAGCAAATGCATACC

AAAG

GGCAATTTTGTTGTTACAGACTGCAGAAGACCAAGATACTTACAATGTTGCAATAAGAAATTA

CGGCAGATTGTTATG

TTCCACTGGTGAATATGATAAAGCTATCCAGGCTTTTAAGTCAACACCCCTTGAAGTGTTAGA

AGACATCATAGGTTTTGCATTGGCTTTATTCATGAAGGGGCTTTATAAAGAGAGCAGCAAAG

CCTATGAGAGAGCCTTGTCTATTGTTGAATCGGAGCAAGACAAAGCCCATATCTTGACAGCTC

TGGCAATAACTGAATATAAACAAGGAAAAACGGATGTAGCCAAGACATTGCTATTTAAATG

CTCTATCTTAAAGGAACCAACCACAGAAAGCCTTCAAGCCCTGTGTGCTCTAGGGTTGGCAAT

GCAGGATGCTACACTGTCAAAAGCAGCACTTAATGAGTTACTGAAGCACATCAAACACAAAG

ACAGTAATTATCAGAGGTGCCTTCTTACATCAGCGATTTATGCACTCCAAGGCCGCAGTGTGG

CTGTGCAAAAACAAATATCTAAAGCTGTTCACAG

CAACCCTGGTGACCCTGCTCTTTGGTCTCTGTTGTCTCGAGTTGTTGCACAGTATGCTCAACGA

AATGCAAAG

GGAGGTGTTGTAGCAGGAAATGTGGCTCATATTCTGGACTCAAATCATGGAAAG

AAGGCATTACTGTACACTGCGGTAAATCAGTTGGCTATGGGAAGCAGTTCAGCAGAAGATGA

AAAAAATACTGCACTAAAGACCATTCAGAAGGCAGCTCTCCTTTCTCCAG

GTGATCCTGCTATCTGGGCTGGGCTAATGGCAGCCTGTCACGCTGATGATAAACTGGCCTTAG

TGAACAACACTCAGCCAAAGAGGATAGATTTATACTTGGCACTGTTATCTGCTGT

TTCTGCTTCAA

TTAAAGACGAAAAATTCTTTGAAAATTACAACCAGTCCCTTGAAAAGTGGTCTCTCTCA

CAAGCTGTCACTGGTCTAATAGACACAGGAAGAATATCTGAAGCTGAAACTCTCTGCACAAA

G

AATTTAAAAAGTAACCCTGATCAGCCAGCCGTTATCTTACTTTTGAGACAAGTTCAGTGTAAA

CCACTCCTGGAGTCACAAAAGCCACTCCCAGATGCTGTACTTGAAGAACTACAAAAAACAGTC

ATGTCCAACTCAACCTCTGTTCCAGCTTGGCAG

TGGCTGGCACATGTGTATCAATCCCAAGGAATGATGAGAGCTGCAGAGATGTGTTACAGAAA

GAGTCTACAATTGGCATCCCAACGGGGCAGTTGGAGTGGGAAGCTCTCAAGTCTGTTGAGACT

AGCACTACTTGCATTAAAAGTCTGTATG

GCTAACATTTCCAATGATCACTGGCCATCTTTGGTTCAAGAGGCTACAACTGAGGCCTTGAAG

CTTTGCTTTTGTCCACTGGCTGTTCTTTTACAAGCTTTGTTACAATTCAAACGCAAAATGGGGG

CAAG

AGAGACACGGCGTCTTTTGGAAAGAGTGGTATATCAGCCTGGGTATCCCAAATCTATTGCATC

AACTGCACGTTGGTACCTACTGAGACACTTATATGCCAAAGATGACTATGAGCTTATTGAC

GTGCTGGTAAACAATGCCAAAACTCATGGAGATACAAGAGCATTGGAACTGAATCAGAGATT

GTCCTCACAATAA

Amino acid sequence

The amino acid sequenced is alternate blue and black to denote individual exons

MSSKEVKTALKSARDAIRNKEYKEALKHCK

TVLKQEKNNYNAWVFIGVAAAELEQPDQAQSAYKKAAELEPDQLLAWQ

GLANLYEKYNHINAKDDLPGVYQKLLDLYES

VDKQKWCDVCKKLVDLYYQEKKHLE

VARTWHKLIKTRQEQGAENEELHQLWRKLTQFLAESTEDQNNETQQL

LFTAFENALGLSDKIPSEDHQVLYRHFIQSLSK

FPHESARLKKACEGMINIYPTVQYPLEVLCLHLIESG

NLTDEGQQYCCRLVEMDSKSGPGLIGLGIKALQDKKYEDAVRNLTE

GLKESPVCTSGWYHLAEAQVKMHRPKEAVLSCSQA

LKIVDNLGASGNSLYQRNLCLHLKAEALIKLSDYDSSEEAIRTLDQ

ISDADNIPGLLVLKSLAYRNKGSFDEAAK

IMEDLLSSYPDLAEVHALEALIHFTKKDYLQAEK

CFQRALEKDTEVAEYHYQLGLTYWFMGEETRKDKTKALTHFLK

AARLDTYMGKVFCYLGHYYRDVVGDKNRARGCYRKAFELDDTDAESGAAAVDLSVELEDM

EMALAILTTVTQKASAGTAKWAWLRRGLYYLKAGQHSQAVAD

LQAALRADPKDFNCWESLGEAYLSRGGYTTALKSFTKASELNPESIYSVFKVAAIQQILGKYKEAVAQYQMIIKKKEDYVPALK

GLGECHLMMAKAALVDYLDGKAVDYIEKALEYFTC

ALQHRADVSCLWKLAGDACTCLYAVAPSKVNVHVLGVLLGQKEGKQVLKKNELLHLGG

RCYGRALKLMSTSNTWCDLGINYYRQAQHLAETGSNMNDLKELLEKSLH

CLKKAVRLDSNNHLYWNALGVVACYS

GIGNYALAQHCFIKSIQSEQI

NAVAWTNLGVLYLTNENIE

QAHEAFKMAQSLDPSYLMCWIGQ

ALIAEAVGSYDTMDLFRHTTELNMH

TEGALGYAYWVCTTLQDKSNRETELYQYNILQMNAIPAAQVILNKYVE

RIQNYAPAFTMLGYLNEHLQLKKEAANAYQ

RAILLLQTAEDQDTYNVAIRNYGRLLC

STGEYDKAIQAFKSTPLEVLEDIIGFALALFMKGLYKESSK

AYERALSIVESEQDKAHILTALAITEYKQGKTDVAKTLLFKC

SILKEPTTESLQALCALGLAMQDATLSKAALNELLKHIKHKDSNYQRCLLTSAIYALQGRSVAVQKQISKAVH

SNPGDPALWSLLSRVVAQYAQRNAK

GGVVAGNVAHILDSNHGK

KALLYTAVNQLAMGSSSAEDEKNTALKTIQKAALLSPG

DPAIWAGLMAACHADDKLALVNNTQPKRIDLYLALLSAVSAS

IKDEKFFENYNQSLEKWSLSQAVTGLIDTGRISEAETLCTK

NLKSNPDQPAVILLLRQVQCKPLLESQKPLPDAVLEELQKTVMSNSTSVPAWQ

WLAHVYQSQGMMRAAEMCYRKSLQLASQRGSWSGKLSSLLRLALLALKVCM

ANISNDHWPSLVQEATTEALKLCFCPLAVLLQALLQFKRKMGA

RETRRLLERVVYQPGYPKSIASTARWYLLRHLYAKDDYELID

VLVNNAKTHGDTRALELNQRLSSQ

8.5 Manuscripts and abstracts derived from the work contributing to this thesis

- Mutations in *TTC37* cause tricohepatoenteric syndrome (phenotypic diarrhoea of infancy). GASTROENTEROLOGY 2010;138:2388–2398
- Clinical phenotype and autozygosity mapping of phenotypic diarrhoea of infancy. British Society of Human Genetics 2006. Poster presentation

  &

  Clinical phenotype and autozygosity mapping of phenotypic diarrhoea of infancy. American Society of Human Genetics 2006. Poster Presentation

- IFT80, which encodes a conserved intraflagellar transport protein, is mutated in Jeune asphyxiating thoracic dystrophy. Beales PL, Bland E, Tobin JL, Bacchelli C, Tuysuz B, Hill J, Rix S, Pearson CG, Kai M, Hartley J, Johnson C, Irving M, Elcioglu N, Winey M, Tada M, Scambler PJ. Nat Genet. 2007 Jun;39(6):727-9. Epub 2007 Apr 29
- The identification of renal cysts may implicate primary cilia in the aetiology of biliary atresia. Accepted for publication in JPGN July 2010 accepted for publication
- The c.1331C>T / p.V444A *ABCB11* variant in severe intrahepatic cholestasis. American Association for the Study of Liver Disease 2008. Poster presentation