Molecular Genetic Analysis of Inherited Kidney Disease in Saudi Arabia

By

Mohamed Hashem Al-Hamed

A thesis submitted for the degree of Doctor of Philosophy (Ph.D.) in the

Faculty of Medical Sciences of the Newcastle University

Institute of Genetic Medicine

Newcastle University

December 2012

ABSTRACT

Inherited abnormalities of the kidney are frequently observed and represent a significant cause of morbidity and mortality. The globally increasing number of patients with end-stage renal disease (ESRD) urges the identification of molecular pathways involved in renal pathophysiology, to serve as targets for therapeutic intervention.

Data from 2010 estimates the Saudi Arabian population to be 27 million, with one of the highest growth rates in the world. The population is characterized with high consanguinity rate, large family size, and a tribal structure. The consanguinity rate results in a high incidence of autosomal recessive genetic disorders. The population is at high risk of renal failure, with 133 incident cases per million populations per year that require renal replacement therapy. In such a population, characterization of new kidney disease gene loci using homozygosity mapping and positional cloning within consanguineous families is a powerful strategy. This study aimed to adopt this approach in order to search for known and novel molecular causes of inherited kidney diseases in the Saudi population. We studied patients and families with nephrotic syndrome, renal ciliopathies, nephrocalcinosis and renal agenesis. For nephrotic syndrome, we found that the most common genetic cause was a homozygous mutation in the *NPHS2* gene. Novel and reported mutations in known nephrosis genes were detected. In a family with Bardet Biedl Syndrome, we utilized zebrafish and renal epithelial cells to determine the functional significance of a novel BBS5 mutation. In another consanguineous family with an autosomal recessive syndrome of distal renal tubular acidosis, small kidneys, and nephrocalcinosis we identified a novel locus on chromosome 2. We also describe the molecular genetic investigation of families with bilateral renal agenesis. In conclusion, in the highly consanguineous Saudi population we have utilized a variety of genetic approaches to identify and characterize novel genetic variants causing inherited renal disease.

LIST OF PUBLICATIONS

Part of the work presented in this thesis has been published.

Articles

1. Al-Hamed M, Sayer JA, Al-Hassoun I, Aldahmesh M, Meyer B. 2010. **A novel mutation in NPHS2 causing nephrotic syndrome in a Saudi Arabian family**. NDT Plus 3:545-548.

2. Mohamed H Al-Hamed, Essam Al-Sabban, Hamad Al-Mojalli, Naffaa Al-Harbi, Eissa Faqeih, Hammad Al Shaya, Khalid Alhasan, Safaa Al-Hissi, Mohamed Rajab, Noel Edwards, Abbas Al-Abbad, Ibrahim Al-Hassoun, John A Sayer and Brian F Meyer. **A molecular genetic analysis of childhood nephrotic syndrome in a cohort of Saudi Arabian families**. *Journal of Human Genetics*, 18 April 2013; doi:10.1038/jhg.2013.27

3. Mohamed H Al-Hamed, Charles van Lennep, Fatimah Al-Fadhly, Lorraine Eley, Ann Marie Hynes, Roslyn Simms, Brian Meyer and John A Sayer. **Functional modelling of a novel mutation in BBS5.** In review *Journal of Cilia*.

Abstracts

1. The Spectrum of Genes causing Nephrotic Syndrome in Arabia. Presented at ASN Kidney week 2012, American Society of Nephrology Annual Meeting, Oct 30-Nov 4, 2012; San Diego, USA.

   Mohamed H Al-Hamed, John A Sayer, Essam Al-Sabban, Abbas Al-Abbad, Hamad Al-Mojalli, Neefa Al-Harbi, Ibrahim Al-Hassoun, Eissa Faqeih, Khaled Al-Hassan, Safaa Al-Hissi, Mohamed Rajab and Brian F Meyer.

2. Molecular Genetics and Modeling of a Novel BBS5 Mutation. Presented at the American Society of Human Genetics (SHG) 62nd Annual Meeting, November 6-10, 2012. San Francisco, USA.

   Mohamed H Al-Hamed, Charles van Lennep, Fatemah Al-Fadhly, Roslyn J Simms, Brian Meyer and John A Sayer

Abbreviations

Aa	amino acid
BBS	Bardet-Biedl syndrome
Bp	base pair
BRA	Bilateral renal agenesis
BSA	Bovin serum albumin
cDNA	complementary DNA
Chr	Chromosome
CNV	Copy Number Variation
Del	Deletion
DMSO	Dimethylsulphoxide
DNA	Deoxyribonucleic acid
dRTA	distal Renal Tubular Acidosis
DTT	Dithiothreitol
EDTA	Ethylene diamine tetra acetic acid
Hpf	hour post-fertilization
Hr	Hour
IBD	Identical By Descent
Ins	Insertion
Kb	Kilobase
LD	Linkage Disequilibrium
Mb	Megabase pair
Min	Minute
MKS	Meckel syndrome
mRNA	messenger RNA
NC	Nephrocalcinosis
NPHP	Nephronophthisis
NS	Nephrotic syndrome
PBS	Phosphate Buffered Saline
PCR	Polymerase Chain Reaction
PKD	Polycystic Kidney Disease
RNA	Ribonucleic Acid
Rpm	revolutions per minute
RT-PCR	Reverse Transcriptase PCR
SDS	Sodium Dodecyl Sulphate
SNP	Single Nucleotide Polymorphism
Tris	Tris(hydroxymethyl)aminomethane

Acknowledgments

I am extremely grateful to my principal supervisor Dr. John Sayer for his continued support, guidance and concern. I would like to thank him for his endless assistance, ideas and suggestions throughout this Ph.D. He was a supervisor and friend at the same time. I would also like to thank Professor Tim Goodship for his support and guidance.

I am indebted to Dr. Brian Meyer, Chairman of the Genetics Department at King Faisal Specialist Hospital and Research Centre. I would like to thank him and Research Centre administration for facilitating my study and supporting me. Also I would like to thank my colleagues at Genetics department and the clinicians who referred the patients and their families.

I would like to give special and sincere thanks to King Khalid Foundation for supporting my study financially and I hope this work would serve the community.

Finally, I am eternally grateful to my family for their patience, support and love.

Table of Contents

List of Tables

List of Figures

Chapter 1: INTRODUCTION

1.1 Genetics of Kidney

The human kidneys serve important physiological functions to excrete waste products from the body and balance the body's fluids. The kidneys also have endocrine roles and are a key regulator of blood pressure and mineral metabolism.

There is now extensive evidence from a wide variety of approaches that suggests renal disease in the general population has a genetic component. The most numerous studies take the form of evaluating familial aggregation (Bowden, 2003).

At least 40 genes have been shown to be involved in kidney development and many more genes are expressed within the kidney and regulate renal physiology (Drummond, 2008).

Development of the kidney in mammals progresses through three spatially and temporally distinct stages: pronephros, mesonephros and metanephros. The first two stages are transient structures in mammalians and do not contribute to the formation of the mature kidney, but they play major roles in the development of the gonads, adrenal gland and hematopoetic precursors. The third stage, the metanephros persists and gives rise to the functional kidneys. The metanephric or adult kidney is derived from the reciprocal interactions of two primordial mesodermal derivatives, the ureteric bud and the metanephric mesenchyme.(Lechner and Dressler, 1997). Mesenchymal cells near the bud become induced and convert to an epithelium which goes on to generate the functional filtering unit of the kidney, the nephron.

Genetic studies in mice have allowed researchers to begin to unravel the molecular basis of kidney development in human. Transcription factors such as Lim-1, Foxc-1, Wilms' Tumor gene 1 (WT1), Eya1 and Pax2, are all act in the metanephrogenic mesenchyme (MM) to induce the expression of Glial cell-derived neurotrophic factor (Gdnf) and Six2 (Gong et al., 2007). GDNF signaling is the main regulator of ureteric budding and branching morphogenesis, whereas Six2 is required for MM cell renewal (Self et al., 2006). Proteins of the Bone morphogenetic proteins (BMP) family modulate ureteric bud branching and keep bud development. As nephrons form, they express

critical transcription factors such as WT-1, Pax-2, and Hoxa11 and d11, condense, and secrete Wnt-4.Wnt-4 acts in an autocrine loop to stimulate its own synthesis and are required for cells to differentiate into epithelia. As nephrons mature, regions of them differentiate to perform specific physiological functions, a process that requires the proteins WT-1, Lmx-1b, Notch-2, Jagged-1, and Hnf-1. (Davies and Fisher, 2002).

Congenital abnormalities of the kidney are frequently observed in children and represent a significant cause of morbidity and mortality. These conditions are phenotypically variable, leading to a spectrum of renal diseases. Many inherited renal disorders are quite rare, with an incidence of 1:10,000 to 1: 100,000 (National Organization for Rare Disorders., 1992). The following universal principles concerning inherited renal disease can be identified:

1) More than one child in a family may be affected.

2) The child with hereditary renal disease frequently presents with growth failure.

3) The inherited pattern is usually autosomal recessive, but it may also be autosomal dominant or X-linked.

Diseases caused by recessive genes usually manifest prenatally, or in childhood or adolescence (e.g. nephrotic syndrome). Those caused by dominant genes typically manifest in adulthood (e.g. autosomal dominant polycystic kidney disease)(Hildebrandt, 2010). Genotype-phenotype correlation is strong in recessive single-gene renal disorders, this correlation is reduced in those caused by dominant genes because of incomplete penetrance and variable expression in different organs (Hildebrandt, 2010).

Figure 1.1 illustrated classification of genetic renal diseases.

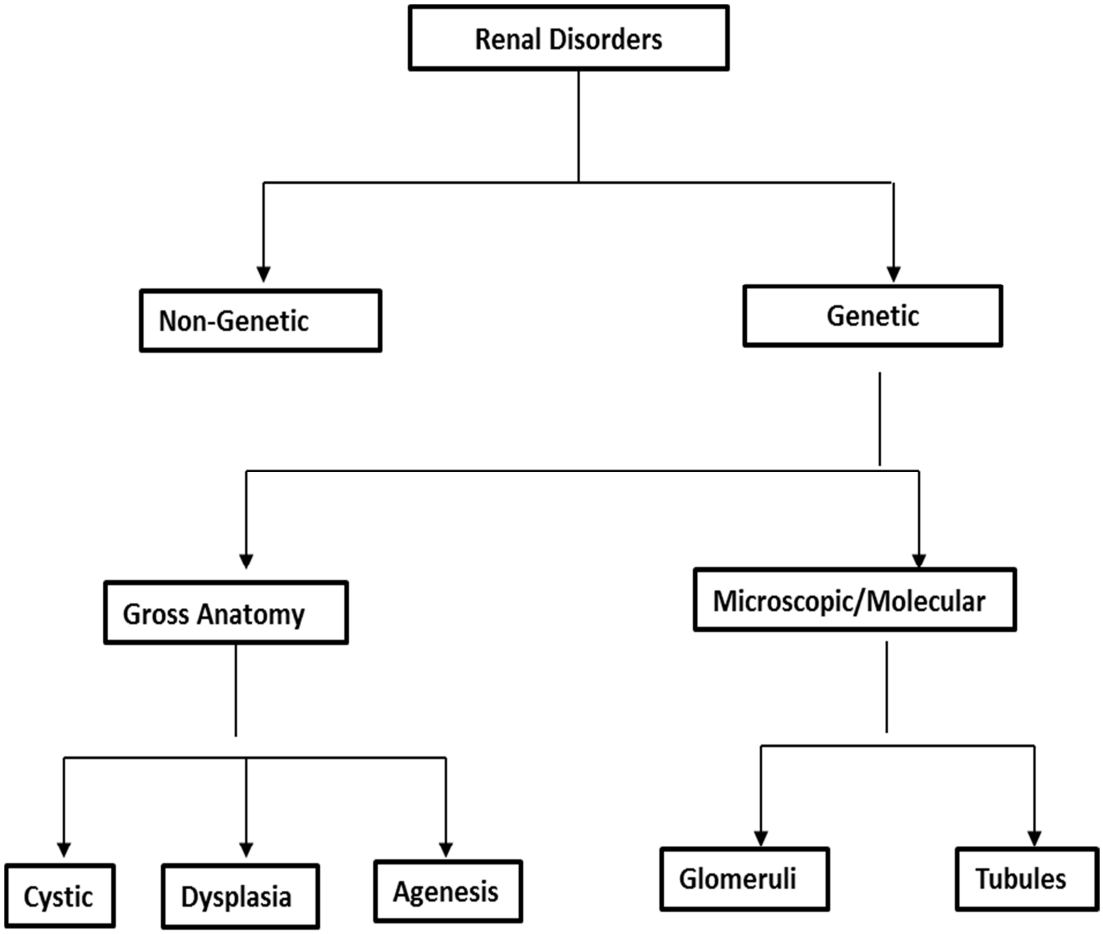

Figure 1.1: Classification of pathological effect of genetics renal disease.

The globally increasing number of patients with end-stage renal disease (ESRD) urges the identification of molecular pathways involved in renal pathophysiology, to serve as targets for intervention. The study of most common genetic cause of renal failure may lead to reduction in the number of kidney diseases progressing to ESRD. Moreover, kidney disease is a growing financial problem. For example, from 1993 to 2005, the total costs of medical care for patients with ESRD increased from US$8 to $20 billion in the US (de Borst et al., 2008). Here in the Kingdom of Saudi Arabia, the estimated cost per annum incurred toward maintenance hemodialysis is US $19,400 and, considering that there are more than 7200 patients on regular dialysis (Shaheen and Al-Khader, 2005).

1.2 Genetics of Saudi population

Saudi Arabia is the largest Arab country occupying most of Arabian Peninsula.

It is bounded on the north by Jordan and Iraq and on the south by Yemen and Oman. It is bounded by Red Sea on the west and Arabian Gulf (Persian Gulf) on the east. The country's total area (2,250,000 km²) is barren desert. The population is cosmopolitan and comprised of large and small minorities. According to a 2010 census, the Saudi population estimated to be 27 million, 70% Saudi native and 30% immigrants, the majority of whom were Arabs. Those considered to be Saudi natives are the early settlers of the urban centres or those originating mostly from neighboring Arab countries in addition to Bedouins (the nomadic Arabs of the desert). Saudi Arabia has one of the highest growth rates in the world (approximately 400000 live births per year). Typically, family size is large, with six children being an average number of offspring per family. Bedouins form tribal communities which are quite isolated. It is estimated that, 8 tribes account for around 10% of the country's population. Islam is the predominant religion among Saudis. Despite the westernization of a significant sector of the population, many still maintain their cultural ties and religious principles. Because they are forbidden by Islam, uncle-niece/aunt-nephew marriages are virtually non-existent. Prenatal diagnosis and pre-implantation genetic diagnosis (PGD) are acceptable for purposes of reassurance or of therapy. Although Islam, according to many religious scholars, discourages consanguineous marriages, such marriages have been traditionally practiced over many generations because of social, economic, and geographical factors. In Saudi Arabia the overall rate of consanguineous marriage is reported to be 57.7% with regional variation from a low of 34.0% to an astounding high of 80.6% (el-Hazmi et al., 1995, al Husain and al Bunyan, 1997, al-Abdulkareem and Ballal, 1998). The most frequent were first cousin marriages (28.4%) followed by distant relative marriages (15.2%) and second cousin marriages (14.6%) (el-Hazmi et al., 1995).

1.3 Inherited Kidney disease in Saudi population

Through years of health services it has become clear that Arabs have a high frequency of genetic disorders, and in particular autosomal recessive traits. The consanguinity rate

together with a cultural preference for endogamous unions results in a high incidence of recessive disorders such as sickle-cell anemia, beta-thalassaemia, inborn errors of metabolism, hereditary hearing impairments (Teebi et al., 2002, Zakzouk, 2002), and a multitude of novel Mendelian diseases. Many novel primarily recessive Mendelian diseases are also very evident in the Saudi population and characterization of new gene loci is facilitated in the population where the average family size is such that pedigrees with multiple affected individuals are readily available.

In 2010, according to Saudi Centre for Organ Transplantation (SCOT Data), the number of Saudi patients with end stage renal failure (ESRF) was 12,633 patients; 11437 patients were treated by haemodialysis and 1196 patients were treated with peritoneal dialysis. Causes of renal failure among haemodialysis patients are presented in Table 1.1. The estimation of new affected cases is 133 for each million yearly (2012).

Cause of Renal Failure	**No.**	**%**
Diabetic Nephropathy	3941	34.5
Hypertensive Nephropathy	3805	33.3
Unknown Aetiology	1272	11.1
Primary Glomerular Disease	697	6.1
Obstructive Uropathy	365	3.2
Hereditary Renal Disease	330	2.9
Congenital Malformation	240	2.1
Primary Tubulo-Interstial Disease	188	1.6
Vasculitis	173	1.5
Pregnancy Related	78	0.7
Others	348	3
Total	11437	100

Table 1.1: Causes of End Stage Renal Failure in Haemodialysis patients - 2010

In contrast to the genetically well studied populations of North America and Europe, Saudi population remained unstudied. Most of studies conducted were describing clinical observation and statistics of inherited kidney diseases. Almost all the studies

published are from referral centres in regions. Preliminary observations indicate that the children in the Saudi Arabia probably have a higher incidence of polycystic kidney disease, familial juvenile nephronophthisis, congenital urological anomalies and familial nephrotic syndrome (Mattoo, 1998, Kari, 2006, Al Harbi, 1997, Al-Ghwery and Al-Asmari, 2004). Studying histological biopsies of patients with glomerular disease was much abundant in many centres in the country. These studies concluded that minimal change disease (MCD) and focal segmental glomerulosclerosis (FSGS) are the most common primary glomerular disorder en-countered in children in Saudi Arabia (Al-Sabban, 1997, al-Rasheed et al., 1996, Mattoo et al., 1990, Kari, 2002, Jalalah and Jamal, 2009, Mitwalli et al., 1996). Molecular studies were conducted lately and covered rare syndromes that have kidney phenotype such as nephrogenic diabetes insipidus (Carroll et al., 2006), cystinosis (Aldahmesh et al., 2009) Bardet Biedl Syndrome (BBS) (Abu Safieh et al., 2010), Joubert syndrome (Alazami et al., 2012) and Meckele Gruber syndrome (Shaheen et al., 2012).

The cohorts of patients studied in this thesis depended upon local referral rates and the use of unique families presenting with possible novel causes. A molecular genetic diagnosis in families with nephrotic syndrome (NS) is useful for the clinician to aid further management. A molecular genetic diagnosis in such families was frequently requested.

NS is characterized namely by heavy proteinuria, hypoproteinaemia, and oedema starting soon after birth, sometimes progressing to ESRD. The incidence of familial cases is from 3 to 5% (Obeidova et al., 2006). Several genes have been implicated in inherited forms of nephrotic syndrome occurring in children. Proteins encoded by these genes influence the function of the podocytes. Mutations in the *NPHS2* gene, encoding protein podocin, are a common cause of childhood steroid-resistant NS. Kidney transplantation is often required as curative therapy for patients with NS. Patients with NS with homozygous or compound heterozygous mutations in the *NPHS2* gene have a reduced risk for recurrence after transplantation (Ruf et al., 2004a), again highlighting the clinical importance of a molecular genetic diagnosis.

Patients with ciliopathies may present with a bizarre and wide spectrum of clinical phenotypes that cause diagnostic uncertainty. A referral for a molecular genetic diagnosis may aid a unifying diagnosis and allow screening of at risk individuals.

Bardet–Biedl Syndrome (BBS; OMIM 209900) is an autosomal recessive disorder clinically characterized primarily by obesity, progressive early-onset retinal degeneration, polydactyly, hypogenitalism, cognitive impairment and renal failure.

BBS is the archetypal ciliopathy, with the multisystem syndrome secondary to defects in ciliary / basal body structure and function. To date, 16 BBS genes have been identified (Forsythe and Beales, 2012). A precise molecular genetic diagnosis also allows genotype phentotype correlations to be made.

The high degree of consanguinity in Saudi Arabia sometimes results in families with a unique phenotype presenting. One such family has recently been identified and reported (Faqeih et al., 2007). The phenotype included a constellation of features including nephrocalcinosis, renal tubular acidosis and learning difficulties. Nephrocalcinosis is an etiologically heterogeneous disorder associated with an increase in the calcium content of the kidneys with underlying complex pathology. It is often seen as a complication of various renal disorders or metabolic disturbances but may be part of inherited syndromes, including rare and exotic syndromes (e.g. Blue Diaper syndrome (Drummond et al., 1964). Distal renal tubular acidosis (dRTA) is caused by a defect in the kidney tubes that causes acid to build up in the bloodstream. It may be caused by a variety of systemic conditions and may be inherited.

Molecular genetic diagnosis has a role in families where stillbirth and early neonatal death has occurred; such cases are not uncommon in the Arabian Peninsula. Such devastating phenotypes need further understanding and investigation to allow insights into disease pathogenesis and to allow screening, if appropriate, of future pregnancies.

Bilateral renal agenesis is an example of neonatal fatal kidney disease. The failure of both kidneys to develop in utero results in oligohydramnios. The lack amniotic fluid may cause compression of the fetus and further fetal malformations. Bilateral renal agenesis is more common in infants with a parent who has a renal anomaly, particularly, unilateral renal agenesis. Studies have shown that unilateral and bilateral renal agenesis

may be genetically related (Kerecuk et al., 2008). For the isolated non-syndromic renal agenesis, only segregation studies have been performed and no loci and/or genes have been mapped so far.

Thus, in this study we used an opportunistic approach to identify disease causing kidney genes in a variety of diseases, using clinical need, diagnostic need and the desire to further our understanding of human renal diseases.

1.4 Approaches to identify inherited kidney Disease

There are many techniques to identify disease causing genes. These include homozygosity mapping, LOD score analysis, association studies, candidate gene screening, Genome-wide knockdown studies, and Whole exome sequencing.

Genetic linkage analysis is the relatively close physical association between a set of genetic markers which lie on the same chromosome. Genome-wide linkage allows finding causal genes without having to know anything on their position on the genome, as necessary in the case of candidate gene studies. It will find such a causal gene, even if there is no assumption for its connection or contribution to the trait.

Homozygosity mapping is a term used to describe in a series of markers (SNPs or microsatellite markers) homozygosity that is identical by descent. Identity by descent (IBD) mapping generally uses marker arrays to survey known polymorphic sites throughout the genome of affected individuals and their parents and/or siblings, both affected and unaffected. People with recessive diseases in consanguineous families are likely to be homozygous for markers linked to the disease locus. If parents are consanguineous by a 2nd degree marriage they would be expected to share 1/32 of all their genes because of their common ancestry, and a child would be homozygous at only 1/64 of all loci (Bittles, 2001, Bittles, 2003). Homozygosity mapping becomes a powerful tool for linkage analysis if families can be found where two or more siblings are affected and they are inbred. It is a model-free analysis, in that no assumptions need to be made about penetrance or gene frequency, although in straightforward applications recessive inheritance is assumed. Homozygosity mapping is an efficient strategy for mapping human genes that cause recessive traits (Lander and Botstein, 1987).

Homozygosity mapping can also be applied in sporadic consanguineous patients (Papic et al., 2011) and in individuals from outbred populations (Hildebrandt et al., 2009).

Homozygosity mapping for a Saudi population would be applicable to identify new gene loci for inherited kidney diseases and would therefore be a useful approach to identify target genes in heterogenetic diseases.

Autozygosity mapping is an extension of homozygosity, in the genomes of affected individuals. It is a form of genetic mapping for autosomal recessive disorders in which affected individuals are expected to have two identical disease alleles by descent. Individuals are frequently observed to have long segments of uninterrupted sequences of homozygous markers. These long homozygous segments arise through numerous mechanisms, including consanguineous marriages, in which parents pass shared chromosomal segments to their child. In a small population, possibly created by the founder effect, there will be a limited gene pool, and thus any inherited disease will probably be a result of two copies of the same mutation segregating on the same haplotype. By taking into consideration the population allele frequencies for all SNPs via autozygosity mapping, the results of homozygosity can be confirmed. The autozygosity mapping can be accomplished by plotting a cumulative LOD score.

LOD score analysis is a statistical test often used for linkage analysis in human, animal, and plant populations. LOD stands for “logarithm of the odds”. It is applied to compare the likelihood that two loci are linked against the likelihood that the two loci are unlinked. Furthermore, if two suspicious regions appear as a result of homozygosity mapping, LOD score analysis may be able to distinguish between the two. A LOD score greater than 3.0 is considered evidence for linkage. A LOD score of +3 indicates 1000 to 1 odds that the linkage being observed did not occur by chance. On the other hand, a LOD score less than -2.0 is considered evidence to exclude linkage.

Linkage by homzygosity mapping is an observation of alleles regardless of whether they are completely identical and inherited from descent or not while linkage by autozygosity mapping searches for identical alleles segregating from the same founder.

Exome sequencing is an efficient strategy to selectively sequence the coding regions of the genome. Exons are short, functionally important sequences of DNA which represent the regions in genes that are translated into protein and the un-

translated region (UTR) flanking them. UTRs are usually not included in exome studies. In the human genome there are about 180,000 exons: these constitute about 1% of the human genome, which translates to about 30 megabases (Mb) in length. It is estimated that the protein coding regions of the human genome constitute about 85% of the disease-causing mutations (Choi et al., 2009). Exome sequencing (Figure 1.2) is only able to identify those variants found in the coding region of genes which affect protein function. It is not able to identify the structural and non-coding variants associated with the disease. This limitation of exome sequencing can be overcome using other methods such as whole genome sequencing or target-enrichment strategies.

Whole genome sequencing (complete genome sequencing) is a process that determines the complete DNA sequence of an organism's genome at a single time. Target-enrichment methods allow one to selectively capture genomic regions of interest from a DNA sample prior to sequencing (Bashiardes et al., 2005).

Using a functional assay is another way to investigate human diseases at a cellular and molecular level. Animal models were used during research and investigation of human disease for the purpose of better understanding the disease without harming actual human being. Whereas a mouse or dog may serve as a mammalian animal model, a baboon or macaque may serve as a less inclusive primate animal model. An animal model for vertebrates is the zebrafish (Kari et al., 2007, Chakraborty et al., 2009). The zebrafish, (Danio rerio) is a tropical freshwater fish that has come to attention recently as a genetically tractable vertebrate model system. It is now a common and useful model organism for studies of vertebrate development and gene function. There are two main reasons to use zebrafish as model system; first, the ability to apply efficient invertebrate style genetics to vertebrate-specific questions, and second the optical clarity of embryos and larvae, which allow easy visualization of developmental processes (Lieschke and Currie, 2007). Zebrafish genome has been fully sequenced, and it has well-understood, easily observable and testable developmental behaviors. Its embryonic development is very rapid, and its embryos are relatively large, robust, and transparent, and able to develop outside their mother (Dahm and Geisler, 2006). Sophisticated mutagenesis and screening strategies on a large scale, and with an economy that is not possible in other vertebrate systems, have generated zebrafish models of a wide variety of human diseases (Lieschke and Currie, 2007).

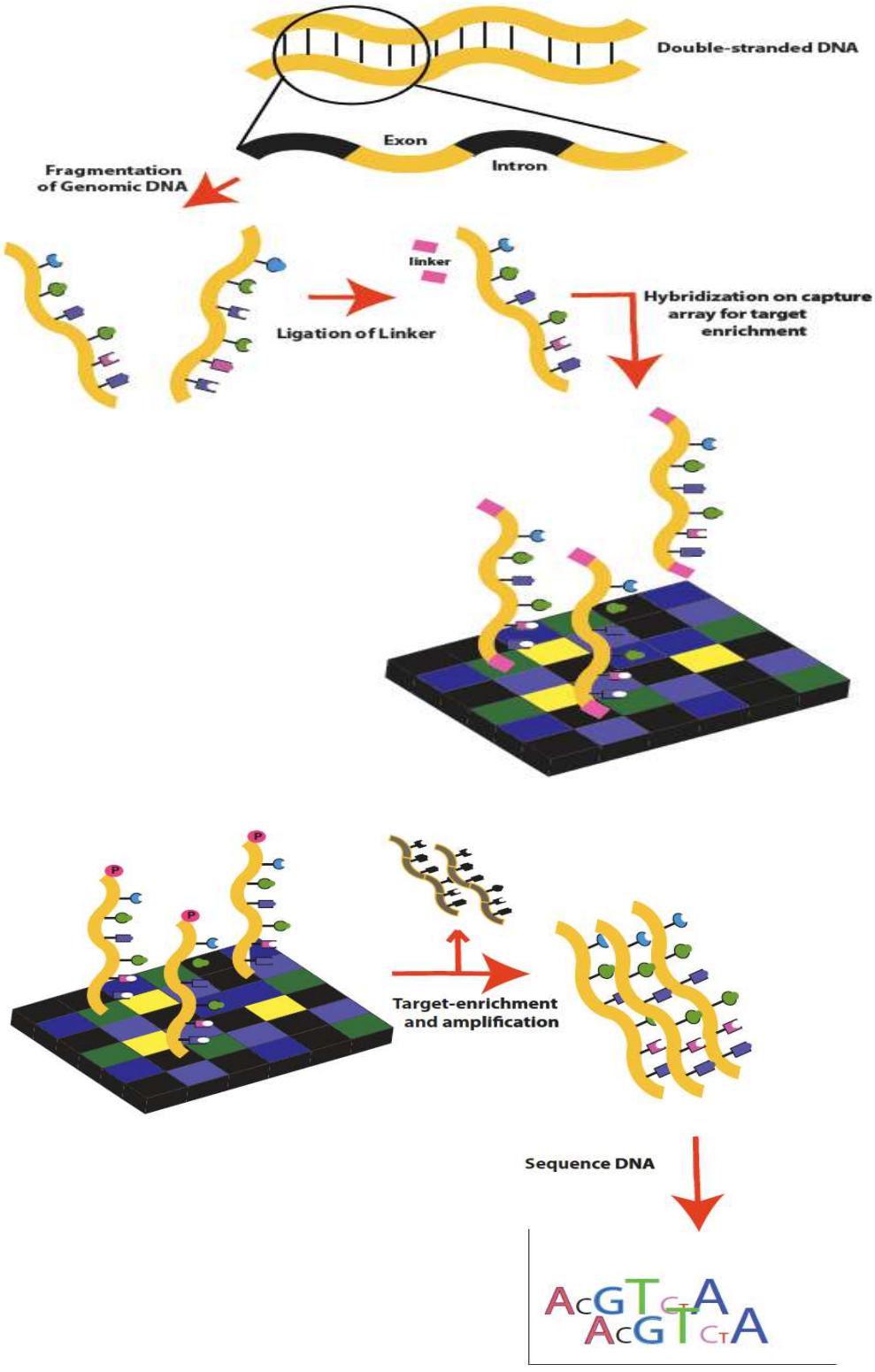

Figure 1.2: Exome sequencing workflow

Studies of kidney development and disease aim to uncover the origins of complexity in the kidney and molecular pathogenesis underlying tissue dysfunction. In zebrafish the functional larval pronephros consists of only two nephrons with glomeruli fused at the embryo midline just ventral to the dorsal aorta as shown in Figure 1.3 (Drummond, 2005).

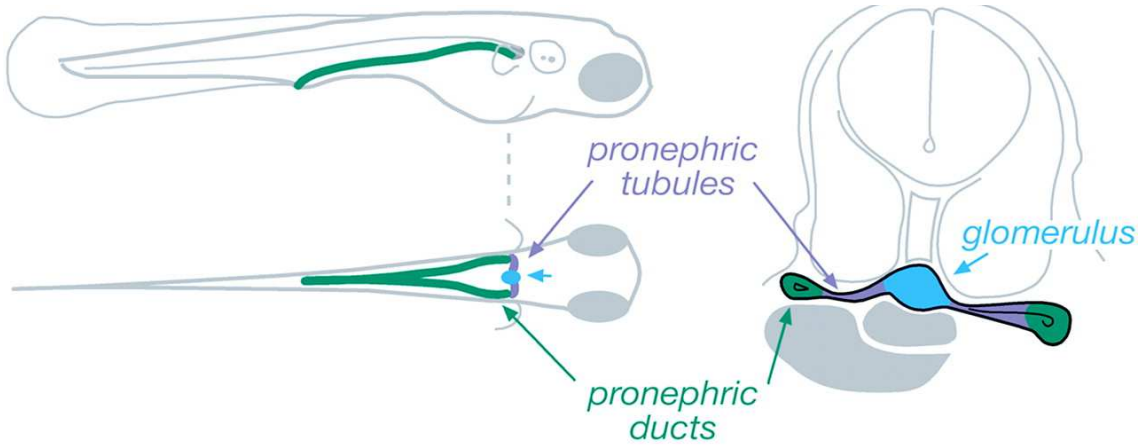

Figure 1.3: Diagram of the mature zebrafish pronephric kidney in 3 days old larva.

The pronephric glomerulus is composed of cell types that are typical of higher vertebrate kidneys, including fenestrated capillary endothelial cells, podocytes, and polarized tubular epithelial cells. Two pronephric tubules connect the glomerulus to the pronephric ducts, which run caudally and fuse just before their contact with the outside world at the cloaca. The zebrafish pronephric nephrons form a closed system of blood filtration, tubular reabsorption, and fluid excretion. The primary function of the fish pronephros is for osmoregulation. Without a functional kidney, larvae die of gross edema because they have salty blood and yet live in a very dilute environment (Drummond, 2005).

Zebrafish could be used as model for defects in glomerulus formation and nephrotic syndromes. Electron microscopy of the zebrafish pronephric glomerulus reveals that like mammalian podocytes, zebrafish podocytes form slit diaphragms between their foot processes. Zebrafish homologs of podocin and nephrin are specifically expressed in podocyte precursor cells as early as 24 hours post fertilization (hpf). These functional similarities between mammalian and zebrafish podocytes, coupled with assays for glomerular filtration, point to applications of fish as a model for study and treatment of

human proteinuria.(Drummond et al., 1998). Several genes that can account for tubule cyst formation have been identified in zebrafish (Ibanez-Tallon et al., 2003). Disruption of other genes resulted in pronephric cyst formation (Sun et al., 2004), kidney cyst (Otto et al., 2003), and glomerular cysts (Sun and Hopkins, 2001). These findings supported usage of zebrafish as model for ciliopathy diseases.

Morpholino oligonucleotides (MOs) are the most common anti-sense “knockdown” technique used in zebrafish (*Danio rerio*) (Bill et al., 2009). MOs are typically employed as oligomers of 25 morpholine bases that are targeted via complementary base pairing to the RNA of interest. Most Morpholinos are used as tools for reverse genetics by knocking down gene function. Gene knockdown is achieved by preventing cells from making a targeted protein. Morpholinos can also modify the splicing of pre-mRNA.(Draper et al., 2001) Knocking down gene expression is a powerful method for learning about the function of a particular protein. In zebrafish embryos, MOs have become a standard knockdown tool that can modify splicing or block translation, depending on the Morpholino's base sequence. Following initial injections into fish embryos at the single-cell or few-cell stages, Morpholino effects can be measured up to five days later (Bill et al., 2009). A cause for concern in the use of Morpholinos is the potential for "off-target" effects. This concern could be solved by coinjecting a rescue mRNA when it feasible. mRNA rescue experiments can often restore the wild-type phenotype to the embryos and provide evidence for the specificity of a Morpholino.

1.5 Aims of the study

The hypothesis being tested was that there are many known and unknown genes causing kidney diseases resulting from consanguineous marriages in the Saudi population. Using patients and families from this population we would identify novel molecular causes of human kidney disease.

Thus, by searching for molecular causes of inherited kidney diseases in Saudi population we aimed to:

- Identify target genes in genetically heterogeneous inherited kidney diseases such as nephrotic syndrome.
- Determine mutations frequency in well characterized genes to provide a molecular genetics diagnosis for known familial nephrotic syndrome.
- Identify new loci for novel familial disorders and kidney diseases in which mutation of well characterized genes have been excluded.
- Scan candidate genes from novel loci (if present) identified in Saudi families for mutations.
- Utilize a functional model for assessing novel mutations using zebrafish as a developmental model.

In this study, Saudi families affected with nephrotic syndrome (NS), Bardet-Biedl Syndrome (BBS), new autosomal recessive distal renal tubular acidosis and nephrocalcinosis (dRTA & NC), and bilateral renal agenesis (BRA) were investigated.

Chapter 2: MATERIALS AND METHODS

This work was performed mainly at the Saudi Diagnostics Laboratory (SDL) at Genetics Department at King Faisal Specialist and Research Centre (KFSH&RC) in Riyadh, Saudi Arabia. Zebrafish modeling and Cell Line Transfection experiments were carried out at the Institute of Genetics Medicine, Newcastle upon Tyne, England.

2.1 Family Materials and Control DNA Samples

In this study we have used DNA and RNA samples obtained, following informed consent, from Saudi Arabian patients affected with inherited kidney diseases. Patients were identified and diagnosed by the Pediatric Nephrology section and/or Obstetrics-Gynecology clinic at King Faisal Specialist Hospital and Research Centre, Riyadh, Saudi Arabia. Where available, in some families we have obtained samples from healthy siblings and parents, in addition to affected cases.

DNA was extracted from blood, cultured cells, amniotic fluid and chorionic villus samples, respectively (Detailed methodologies are given below 2.2). RNA was extracted from cellular material and fresh urine samples where available (See section 2.3). The total number of individuals recruited for this study was 95, comprising 67 individuals with Nephrotic Syndrome; 11 individuals with a “New Autosomal Recessive Distal Renal Tubular Acidosis and Nephrocalcinosis”(Faqeih et al., 2007), 11 individuals with Renal Agenesis, and; 6 individuals with ciliopathy syndromes.

We obtained DNA from 175 Saudi Arabian control DNA samples for estimation of the carrier frequency of novel sequence variants detected in the study. These DNA controls obtained from healthy blood donors and stored at Genetics Department at King Faisal Specialist Hospital and Research Centre for research use. This study has been approved by the research advisory council of King Faisal Specialist Hospital, Riyadh, Saudi Arabia (RAC#2050 045).

Figure 2.1 Outline of the disease gene identification process and some examples of methods that can be used in each step of the process.

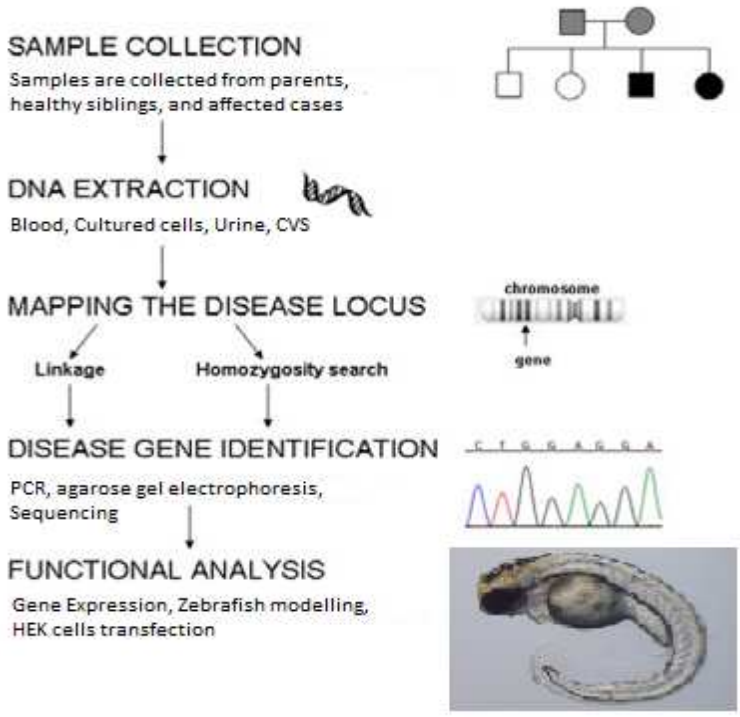

Figure 2.1: Process of sampling, and disease gene identification and functional analysis in the study.

2.2 DNA Extraction

The PUREGENE DNA isolation kit (Gentra) requires whole blood anticoagulated with ethylenediaminetetraacetic acid (EDTA) for extraction of DNA. 3 ml of whole blood was first lysed with 9 ml of RBC lysis buffer [containing Tris (hdroxymethyl) aminomethane, Sodium Chloride, Magnesium Chloride, and Triton X-100] for 15 min to separate red blood cells from white blood cells that contain genomic DNA. After centrifugation (Thermo Scientific Sorvall RC BIOS centrifuge) at 3000 rpm for 15 min, supernatant was removed leaving behind the visible white blood cells. 3 ml of cell lysis buffer [containing Tris, Sodium Chloride, (EDTA), and sodium dodecyl sulfate (SDS)] then was added to the pellet and left overnight at room temperature with gentle mixing. An anionic detergent (SDS) and DNA stabilizer (EDTA) in the cell lysis buffer will isolate DNA by disrupting cell and nuclear membrane. Next day 1 ml of protein precipitation solution (containing 3 M Sodium Citrate) was added to the lysate to precipitate proteins and other contaminants by salt precipitation. After centrifugation for 20 min at 3000 rpm; supernatant that contain the DNA was added to 5 ml of 100% Isopropanol (Fisher A451-4) for DNA recovery. The mix was then centrifuged at 7000

rpm for 5 min; the precipitated DNA is visible as a white pellet. The supernatant was removed and the pellet was washed with 70% ethanol. After centrifugation the pellet was left in the tube to air dry, and then re-suspended in the DNA hydration solution that contains a DNA stabilizer (1 mM EDTA, 10 mM Tris·CI pH 7.5). DNA was allowed to rehydrate overnight at 37 °C in the water bath (Thermo Scientific; Isotemp water bath) The genomic DNA concentration and absorption ratio (260 nm / 280 nm) was measured using a NanoDrop 8000 instrument (Thermo Scientific). Optimal purified DNA has an A260/A280 ratio between 1.7 and 1.9 and up to 200 kb in size. The DNA was stored at 4 °C until further use.

2.3 RNA Extraction

2.3.1 Cell Line RNA Extraction

TRIzol® Reagent (Invitrogen) is a ready-to-use reagent for the isolation of total RNA from cells and tissues. During sample homogenization or lysis, TRIzol® Reagent maintains the integrity of the RNA, while disrupting cells and dissolving cell components. The addition of chloroform followed by centrifugation separates the solution into an aqueous phase and an organic phase. The RNA remains exclusively in the aqueous phase. After transfer of the aqueous phase, the RNA is recovered by precipitation with isopropyl alcohol. The expected yield for 1×10^6 cells of fibroblasts or mammalian cultured cells is 2–7 micrograms (µg).

Cells were pelleted by centrifugation. 1 ml of TRIzol® Reagent was added to the pellet before homogenization. The homogenized samples were incubated at room temperature for 5 min. 200 µl of 1-bromo-3-chloropropane (BCP) was added to the mix and left for 15 min at room temperature. During incubation, the homogenization was mixed thoroughly for 5 times. The mix was then centrifuged at 12000 x g for 15 min at 4 °C. The supernatant was transferred then to a new tube that contains 500 µl of isopropanol. To enhance the precipitation the tube was left at -20 °C for one hour. The tube was then centrifuged at 12000 x g for 15 min at 4 °C. The supernatant was removed and the RNA pellet was washed once by adding 1 ml of 75% ethanol. Samples were mixed and centrifuged at 7500 x g for 5 min at 4 °C. At the end of procedure the RNA was dried briefly for 15 min (air dry). The RNA pellet then was dissolved in 50 µl, RNase free water and incubated in a water bath at 55 °C for 10 min.

RNA concentration and protein ratio (260 nm/280 nm) was measured using NanoDrop 8000 instrument. The optimal extracted RNA has an A260/A280 ratio ≥1.8. Low RNA ratio may results when sample homogenized in an insufficient volume of TRIzol® Reagent or may results when incomplete removal of organic phase. The RNA stored at -80 °C until required for use. Quality of RNA was checked by running 3 µl of purified RNA on a 1.0% agarose gel using standard electrophoresis techniques. Purified RNA may be visualized under UV fluorescence, where there are typically two discrete bands observed in the gel representing 18S and 28S of ribosomal RNA.

2.3.2 Urine RNA Extraction

ZR Urine RNA Isolation kit™ (ZYMO RESEARCH) was used to extract RNA from urine specimens as described by the manufacture. Briefly, to isolate cells from urine, 30 ml of fresh urine sample was taken up in a 30 ml syringe. The urine was pushed completely through the provided ZRC GF™ Filter to isolate the cells in the filter. The remaining urine was removed completely from the filter by pushing through several volumes of air. By using a 1 ml syringe, 700 µl RNA Extraction Buffer Plus™ was pushed through the filter and the flow-through was collected in an RNase-free 1.5 ml tube. By pushing several volumes of air through the filter any residual flow-through would be collected. The content in the tube was mixed by vortexing. For RNA isolation, 700 µl of 100% ethanol was added to the tube containing the flow-through and mixed briefly. The mixture was transferred to a Zymo-Spin™ IC Column in a Collection tube and centrifuged at 12,000 x g for 2 min and the flow through was discarded. 300 µl RNA Wash Buffer was added to the column and centrifuged at 10,000 g for 2 min and again the flow-through was discarded. The column was transferred to an RNase-free tube. 10 µl of Nuclease-free water was added directly to column matrix. After 2 min waiting the column centrifuged at 10,000 x g for 2 min and eluted RNA was collected. The RNA was stored at -80°C until required for use.

2.4 Tissue Culture

Recent developments in genomics have necessitated a constant supply of DNA belonging to specific genotypes. Immortalization by The Epstein–Barr virus (EBV) is an effective procedure for inducing long-term growth of human B-lymphocytes and these EBV-transformed cells can serve as the most reliable source of DNA.

Transformation of peripheral blood lymphocytes by EBV is a laboratory procedure developed involving tissue culture. Immortalization of cells takes 6-8 weeks to generate large amounts of cells that can be indefinitely grown as a continuing source of genetic material. EBV transformation is for immortalization of human B cells only and one limitation of this technique is that there are many genes that not expressed in B lymphocytes.

Whole blood was collected from patients and their relatives in 5 ml Heparin tubes. In a 15 ml centrifuge tube, 1 ml of Ficoll–Paque solution (<0.12 EU/ml) solution of Ficoll™ 400 and sodium diatrizoate with a density of 1.077 + 0.001 g/ml. (Amersham Biosciences) was added and 2 ml of heparinized blood sample was gently layered onto the Ficoll-Paque solution. The tube was centrifuged at 3000 rpm for 20 min. After spinning there should be a very definite band of lymphocytes at the interface. Carefully, the thin band of lymphocytes was transferred into a new 15 ml centrifuge tube and washed with 1X PBS then mixed gently and centrifuged at 3000 rpm for 10 min. Another wash step was performed then lymphocytes re-suspended in tissue media of Roswell Park Memorial Institute (RPMI 1640) containing 20% Fetal Bovine Serum (FBS), 1% antibiotic and 1% L-glutamine, then transferred to a T-25cm² flask (Corning).

0.5 ml of EBV (Epstein Barr Virus) was added into the cell solution. When EBV infects B-lymphocytes *in vitro*, lymphoblastoid cell lines eventually emerge that are capable of indefinite growth. Following this, 5 µl of Phytohaemagglutinin (PHA) was added in order to stimulating cell growth. Finally, 2 µl of cyclosporine (1 µg/ml, Sigma) was added to inhibit proliferation of T cells. The flask was incubated at 37 ˚C in a CO_2 humidified incubator (5% CO_2, 98% humidity). The flasks were placed at a 45˚ angle to concentrate the cells in the corner of the flask and cap remained loose. Media was changed every 3 days to feed the cells. After 7-10 days growth, the cell lines were frozen down using media containing 40% of complete media, 55% FBS and 5% Dimethyl Sulfoxide (DMSO) until further use.

2.5 Polymerase Chain Reaction

Genomic DNA was amplified by Polymerase Chain Reaction (PCR) to amplify coding exons and intron boundaries of genes. Primer 3 software for oligonucleotide primer design (Gudbjartsson et al., 2005) was used to design oligonucleotide primers for exons and intron boundaries of all genes. To facilitate sequencing all genes by the same M13 primers all primers were tagged with following M13 primers:

M13-Forward: GTAAAACGACGGCCAGT,

M13-Reverse: CAGGAAACAGCTATGACC

PCR was performed in a final volume of 25 µl containing approximately 20 ng of genomic DNA, Qiagen master mix kit (Qiagen), including 1X PCR buffer containing 15 mM $MgCl_2$, 100 µmol/L dNTP, 0.5 µmol/L primer pair, and 1 U/reaction HotStar *Taq* polymerase. In addition, 1X Q-solution was used occasionally to facilitate amplification of difficult templates (GC-rich) by modifying the melting behavior of DNA.

For thermocycling we used Touchdown PCR. Briefly, touchdown PCR involves decreasing the annealing temperature by 1 degree ºC every second cycle to a 'touchdown' annealing temperature which is then used for 10 cycles. Thermocycling consisted of an initial denaturation at 95 ºC for 15 min followed by 10 cycles of touchdown PCR. Each cycle consists of denaturation at 95 ºC for 35 sec, annealing at 65 ºC down to 56 ºC for 35 sec and extension at 72 ºC for 1 min. After these 10 cycles another 25 cycles was carried out consisting of denaturation at 95ºC for 30 sec, annealing at 58 ºC for 30 sec, and extension at 72 ºC for 50 sec. A final extension step of 10 min at 72 ºC is then carried out before holding at 4 ºC.

In all PCR reactions, a negative control was added to check for DNA contamination, also normal control DNA was added to be used as reference in sequencing. All PCR products were visualized in a 2% agarose gel (Sigma) to check for PCR contamination and PCR amplification. 5 µl of PCR product mixed with 1 µl loading dye was loaded in 2% agarose gel. The gel was electrophoresed for 30 min at 120 V. Also a 100 bp size marker (Qiagen) was loaded in the gel to allow the PCR product size to be confirmed. The gel then was checked under UV-transilluminator (BioRAD).

2.6 Sequencing of Amplicons

Direct sequencing of PCR amplicons was used for mutation screening. PCR products were treated with the Agencourt® AMPure® PCR purification system that utilizes Agenocourt's solid-phase paramagnetic bead technology for high-throughput purification of PCR amplicons. Agencourt AMPure utilizes an optimized buffer to selectively bind PCR amplicons 100 bp and larger to paramagnetic beads. Excess oligos, nucleotides, salts, and enzymes then removed using a simple washing procedure. Following magnetic beads treatment; products were sequenced using BigDye™ Terminator Cycle Sequencing kit (PE Applied Biosystems) as described by the manufacturer. Following thermocycling the products were precipitated, washed and re-suspended in Hi Di formamide. Advantages of using Hi Di formamide instead water are stability and no need for heat/cool step. Products electrophoresed on an ABI 3730xl capillary sequencer (PE Applied Biosystems). Using Applied Biosystems Genetic Analyzers; raw data converted to ABI and EditSeq files formats that exported for visualization and analysis of sequence.

DNA sequencing was performed mainly on one strand or both strands for positive mutations for confirmation. For all detected variants, samples and all related family members if available were re-PCR and re-sequenced in both directions. Most of sequencing was performed by Sequencing Core Facility at Genetics Department at King Faisal Specialist Hospital and Research Centre in Riyadh.

Sequencing data were analyzed using Mutation Surveyor® software Version 3.24 (SoftGenetics LLC State College) and SeqMan II software 6.1 (DNAStar). Reference sequences of genes were downloaded from NCBI website for comparison with sequencing results. Database search for novel variants detected was carried out. The following websites were searched for reported and novel variants:

- Ensembl Genome Browser (www.ensembl.org) / 1000 Genomes
- The Human Gene Mutation Database [HGMD] (www.hgmd.org)
- Human Genome Variation Society [HGVS] (www.hgvs.org)
- Exome Variant Server (http://evs.gs.washington.edu/EVS/)

2.7 Gene Expression

2.7.1 Reverse Transcriptase Polymerase Chain Reaction (RT-PCR)

SuperScript III First-Strand Synthesis System for RT-PCR (Invitrogen) was used to synthesize first strand cDNA from total RNA.

The following RNA/primer mixture was combined in 0.2 ml tube:

3 µg total RNA	3.0 µl
50 µM oligo $(dT)_{20}$	1.0 µl
10 mM dNTP	1.0 µl
DEPC-treated water	5.0 µl
Total	10.0 µl

The tube then incubated at 65 °C for 5min then placed on ice for 2 min.

The cDNA Synthesis mix was prepared in the following order:

10X RT buffer	2.0 µl
25 mM $MgCl_2$	4.0 µl
0.1 M DTT	2.0 µl
RNaseOUT™ (40 U/ µl)	1.0 µl
SuperScript™ III RT (200 U/ µl)	1.0 µl
Total	10.0 µl

On ice the prepared volume (10.0 µl) of cDNA Synthesis mix then added to RNA/primer mixture and incubated at 50 °C for 1 hour. The reaction terminated at 85 °C for 5 min. After brief centrifugation, 1.0 µl RNAse H was added to each tube and incubated at 37 °C for 20 min. Addition of RNAse H to increase the sensitivity of the

PCR by removing the RNA template from cDNA:RNA hybrid by digestion with RNAse H. The cDNA synthesis reaction was stored at -20 °C for further analysis.

2.7.2 Quantitative Polymerase Chain Reaction (qPCR)

EXPRESS SYPER®GreenER™ qPCR SuperMixes with premixed ROX Kit (Invitrogen) provide components for real time quantitative PCR (qPCR). A 384-well plate was used with a maximum reaction volume of 10 µl per well. On ice; a master mix of common components for multiple reactions was prepared as follow:

EXPRESS SYPER®GreenER™ qPCR SuperMixes with premixed ROX	10.0 µl
10 µM forward primer	0.4 µl
10 µM reverse primer	0.4 µl
cDNA (20 folded diluted)	2.0 µl
DEPC-treated water	7.2 µl
Total	20.0 µl

Each sample reaction was performed in triplicate. In addition no-template control (NTC) reactions were included to test DNA contamination of the enzyme / primer mixes. After brief centrifugation, reactions placed in the real-time instrument AB7900HT (Applied Biosystems) programmed as follow:

50 °C for 2 min (UDG incubation)

95 °C for 2 min

40 cycles of:

95 °C for 15 seconds

60 °C for 1 minute

2.7.3 Data Analysis of qPCR

One of the most important steps in relative quantitation experimental design is the selection of an appropriate endogenous control. Normalization to an endogenous control

(often referred to as a housekeeping gene) allows us to correct results that can be skewed by differing amounts of input nucleic acid template.

The Comparative C_T Method ($\Delta\Delta C_T$ Method)

By measuring the amount of cellular RNA, one is able to determine to what extent that particular gene is being expressed. Two different methods of presenting quantitative gene expression exist: absolute and relative quantification. Absolute quantification calculates the copy number of the gene usually by relating the PCR signal to a standard curve. Relative gene expression presents the data of the gene of interest relative to some calibrator or internal control gene. A widely used method to present relative gene expression is the Comparative C_T method, also referred to as the $\Delta\Delta C_T$ Method. It is possible to eliminate the use of standard curves and to use the $\Delta\Delta C_T$ Method for relative quantitation as long as the PCR efficiencies between the target(s) and endogenous control(s) are relatively equivalent (Schmittgen and Livak, 2008).

Arithmetic Formula:

The amount of target, normalized to an endogenous reference and relative to a calibrator, is given by: $2^{-\Delta\Delta C_T}$. An excel formula table was used to calculate relative quantification of gene expression ΔΔCT.

2.8 Genotyping

2.8.1 Homozygosity Mapping

The markers of choice that have emerged for whole-genome linkage scans and association studies are single nucleotide polymorphisms (SNPs). Although there are multiple sources of genetic variation that occur among individuals, SNPs are the most common type of sequence variation and are powerful markers due to their abundance, stability, and relative ease of scoring.

To identify disease loci we carried out a genome wide linkage search using Affymetrix GeneChip® Human Mapping 250K Sty Arrays, or Affymetrix Axiom® 2.0 Assay Arrays, or Affymetrix SNP 6.0 Arrays, or Affymetrix CytoScan assay platforms according to the manufacturer's protocol (http://www.affymetrix.com). 250 ng genomic DNA from each family member was hybridized on the array.

All these platforms almost have similar workflow as described in the original article (Matsuzaki et al., 2004) :

Genomic DNA digestion with restriction enzyme → Adaptor Ligation to Digested DNA by T4 → Polymerase Chain Reaction (PCR) for Ligated Samples → Fragmentation of Purified PCR Products by DNAseI → The 3′ ends of the fragmented amplicons were labeled with biotinlyated by using Terminal Deoxynucleotidyl Transferase → Hybridization → Washing, Staining and Scanning → Data Analysis.

The main difference between these arrays is the SNP marker density in the genome. SNP 6.0 and CytoScan arrays could be used for molecular karyotype analysis as they have copy number probes designed to interrogate Copy Number Variations (CNVs) in the genome. For CytoScan assay; digestion reaction was set as follow:

Ingredients	Concentration	X 1.0
Nuclease Free Water	-	11.55 µl
NE Buffer 2	10X	2.0 µl
BSA	100X	0.2 µl
Nsp I restriction enzyme	10 u/µl	1.0 µl
Total of Master Mix		14.75 µl
DNA sample	50 ng/µl	5 µl
Total volume per tube		19.75 µl

The reaction tubes were mixed and spun down for 1 min then incubated in the Thermal Cycler for 2 hours at 37 °C then inactivated the enzyme at 65 °C for 20 min. Reaction was held at 4 °C.

The ligation reaction was set as follow:

Ingredients	Concentration	X 1
Adapter *Nsp 1*	50 µM	0.75 µl
T4 DNA Ligase Buffer	10 X	2.50 µl
T4 DNA Ligase	400U/µl	2.0 µl
Total	-	5.25 µl
Digested DNA	-	19.75 µl
Ligation mix		5.25 µl
Total volume per tube		25.0 µl

The reaction tubes were mixed and spin down for 1 min and incubated in the Thermal Cycler at 16 °C for 3 hrs. The T4 Ligase was denatured by heating to 70 °C for 20 min. The reaction was hold at 4 °C. Reaction tubes were spin down at 2000 rpm for 1 min and diluted by adding 75 µl nuclease free water. Four 10 µl aliquots of each sample were transferred to the PCR plate and PCR reaction was set as follow:

Ingredients	Concentration	X 1
Nuclease Free H_2O	-	39.5 µl
TITANIUM *Taq* PCR Buffer	10 X	10.0 µl
GC-MELT	5mM	20.0 µl
dNTPs	2.5mM	14.0 µl
PCR primer 002	100uM	4.5 µl
TITANIUM *Taq* DNA Polymerase	50X	2.0 µl
Total of Master Mix		90.0 µl
DNA sample (ligated product)		10.0 µl
Total volume per tube		100.0 µl

Thermocycling consisted of an initial denaturation at 94ºC for 3 min followed by 30 cycles. Each cycle consists of denaturation at 94 ºC for 30 sec, annealing at 60 ºC for 45 sec and extension at 68 ºC for 15 sec. A final extension step of 7 min at 68 ºC is then carried out before holding at 4 ºC. PCR products were checked in 2% agarose gel. Majority of product should be between 150 and 2000 bp (Figure 2.2).

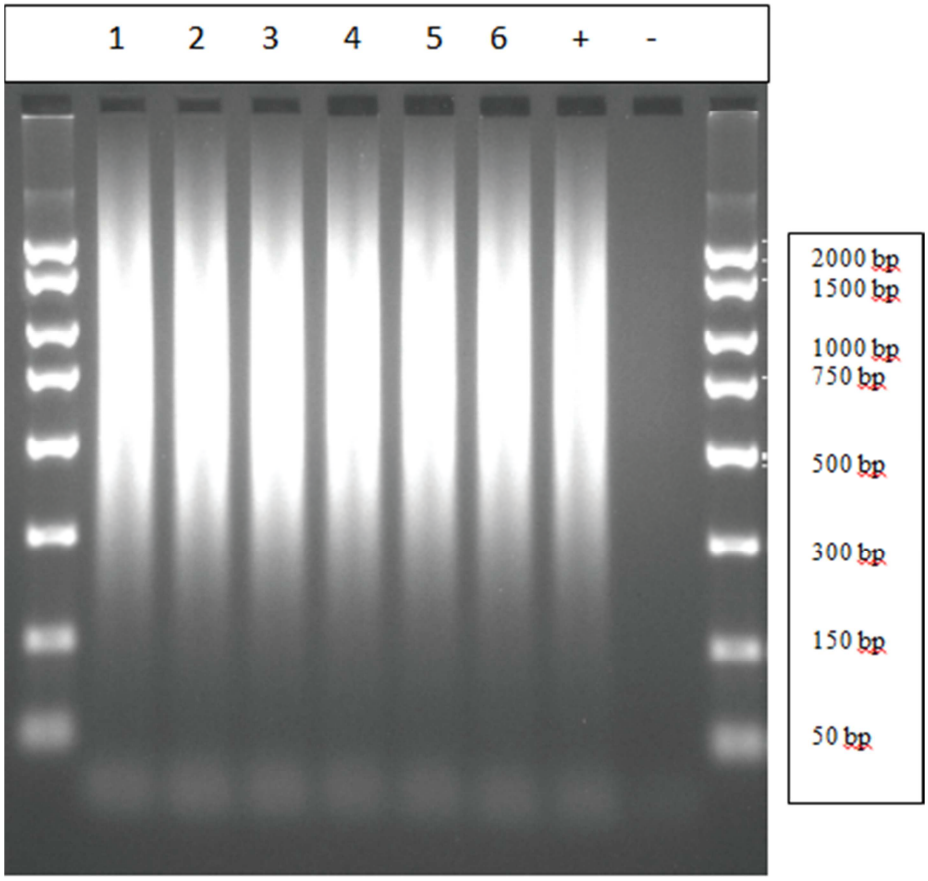

Figure 2.2: PCR Check: PCR products run on a 2% TBE gel at 5 V/cm for 45 min. Majority of product should be between 150 and 2000 bp. (+) indicates positive control DNA, while (-) indicates negative control H_2O.

All 4 PCR products for each sample were pooled by transferring all PCR reactions to the appropriately marked 1.5 mL Eppendorf Safe-Lock tube. To purify PCR products, 720 µl of purification beads was added to each tube and mixed well by inverting several times and then incubated at room temperature for 10 min. Tubes then centrifuged for 3 min at speed (16100 rcf). Tubes were then placed in a magnetic stand and the supernatant was pipetted off. 1 ml of Purification Wash Buffer was added to each tube. Tubes were mixed vigorously and centrifuged for 3 min at high speed (16100 rcf). By placing tubes back to magnetic stand supernatant was pipetted off and 52 µl of elution buffer was added to the tubes. Tubes loaded into the foam adapter, and vortexed at maximum power for 10 min to resuspend the beads. Tubes were then centrifuged again at speed 16100 rcf for 3 min. 47 µl of eluted samples were transferred to an appropriate new tube. Purified PCR products were quantified by measuring OD at 260 nm in NanoDrop instrument.

For fragmentation step, the following reaction was set up:

Ingredients	Concentration	X 1
Nuclease Free H_2O	-	122.4 µl
Fragmentation Buffer	10 X	158.4 µl
Fragmentation Reagent	2U/µl	7.2 µl
Fragmentation Master Mix		10.0 µl
Purified PCR Product		45.0 µl
Total volume per tube		55.0 µl

Reaction tubes were incubated at the Thermal Cycler at 37 ºC for 35 min then heated to 95 ºC for 15 min and hold at 4 ºC forever. 4.0 µl of each fragmented sample was run in 4% agarose gel to check for fragmentation quality. Average fragment distribution is between 25 to 125 bp as in Figure 2.3.

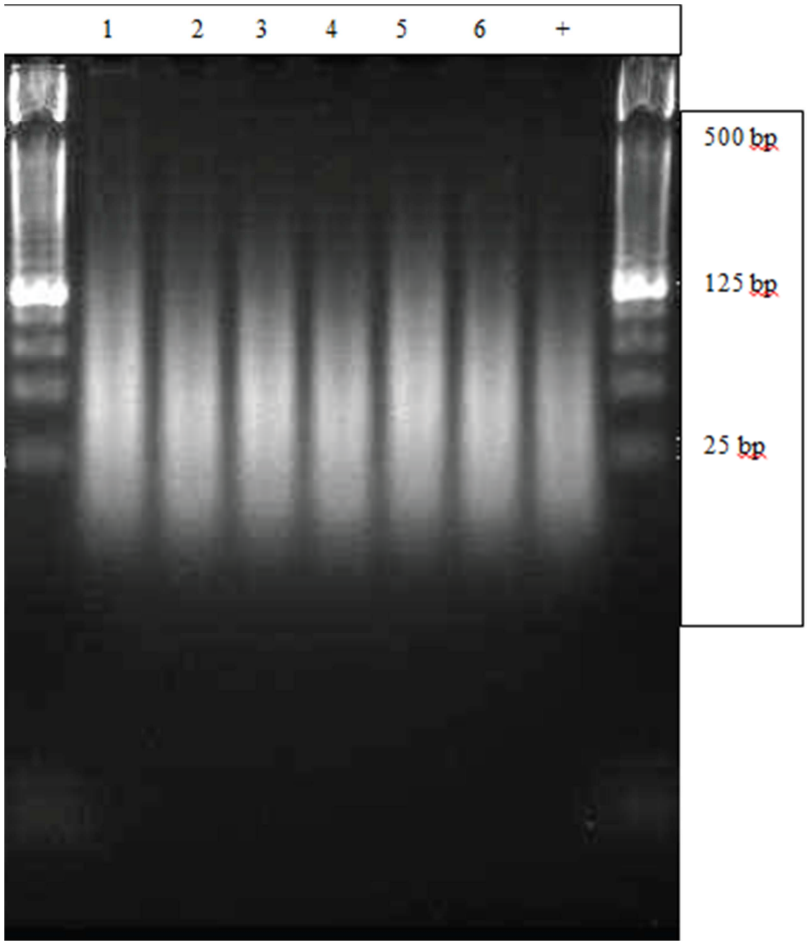

Figure 2.3: fragmented samples run on a 4% TBE gel at 5 V/cm for 45 min. Average fragment distribution is between 25 to 125 bp.

After passing Fragmentation Quality Control, samples were set for labeling reaction as follow:

Ingredients	Concentration	X 1
TdT Buffer	5X	14.0 µl
DNA Labeling Reagent	30 mM	2.0 µl
TdT		3.5 µl
Labeling Mix		19.5 µl
Fragmented DNA		51.0 µl
Total volume per tube		70.5 µl

Reaction tubes were loaded in the Thermal Cycler and incubated at 37 ºC for 4 hrs. The reaction inactivated by heating to 95 ºC for 15 min. The reaction was held at 4 ºC. The reaction then proceeded to the hybridization step.

A hybridization master mix was prepared as followed:

Ingredients	X 1
Hyb Buffer Part 1	165.0 µl
Hyb Buffer Part 2	15.0 µl
Hyb Buffer Part 3	7.0 µl
Hyb Buffer Part 4	1.0 µl
Oligo Control Reagent 0100	2.0 µl
Hybridization Mix	190.0 µl
Labeled DNA	70.5 µl
Total volume per tube	160.5 µl

Tubes were loaded onto the thermal cycler and Hybridization program was run as follow: 95 ºC for 10 min then hold at 49 ºC. While the reaction tubes incubated at 49 ºC, 200 µl of each sample was loaded onto each array. Arrays were immediately loaded into the hybridization oven. The arrays hybridized for 18 hrs at 50 °C and 60 rpm. Following the hybridization, the arrays were washed with 6X SSPE (Phosphate buffer, pH 7.4, sodium chloride, EDTA) and 0.01% Tween-20 at 25°C, then more stringently washed with 0.6X SSPE and 0.01% Tween-20 at 45 °C. Hybridization signals were generated in a three-step signal amplification process: 10 µg/ml streptavidin (Pierce) was added to the biotinylated targets hybridized to the oligonucleotide probes, and washed with 6X SSPE and 0.01% Tween-20 at 25 °C, followed by the addition of 5 µg/mL biotinylated goat anti-streptavidin (Vector Laboratories) to increase the effective number of biotin molecules on the target; finally, streptavidin R-phycoerythrin (SAPE) conjugate (Molecular Probes) was added and washed extensively with 6X SSPE and 0.01% Tween-20 at 30 °C. The Streptavidin, Antibody, and SAPE were added to arrays in 6X SSPE, 1X Denhardt's solution, and 0.01% Tween-20 at 25 °C for 10 min. The washing and staining procedures were performed using Affymetrix fluidics stations. The arrays

were scanned using the Affymetrix Gene Chip® Scanner 3000 7G with the GeneChip® Genotyping Analysis Software (GTYPE). The analyzed data was classified by chromosome and sorted based on the physical position of the SNPs. We looked for regions with informative SNPs that were homozygous for the same allele for the affected children and heterozygous for the parents, and heterozygous or homozygous for the opposite allele for the unaffected. Homozygosity regions were identified using the following software: CNAG (Version 2.0), Genotyping Console (Version 3.0.1), and Chromosome Analysis Suite (ChAS). CNAG2.0 was developed to enable high-quality analysis of copy number alterations and allelic imbalances in cancer genomes, congenital disorders, and normal individuals using Affymetrix GeneChip® platforms.

2.8.2 Linkage and LOD Score Analysis

Raw data received from Genotyping Core facility was formatted to create manageable files that could be used for an initial linkage analysis. The linkage analysis would be performed using easyLINKAGE-Plus, therefore the pedigree and genotype data files that we had created from the raw data were appropriately formatted.

LOD scores were calculated with the easyLINKAGE-Plus software (http://www.uniwuerzburg.de/nephrologie/molecular_genetics/molecular_genetics.htm) ALLEGRO and GENEHUNTER programs were used to calculate a multipoint parametric linkage analysis LOD score in autosomal recessive pedigrees.

2.9 Zebrafish Model

2.9.1 Zebrafish husbandry

Adult zebrafish were kept and maintained in a zebrafish facility according to standard procedures. Transgenic fluorescent reporter fish were used for studying cardiac morphology (*cmlc2*:GFP, green fluorescent protein gene under the control of the cardiac myosin light chain 2 gene promoter) (Lange et al., 2008) and renal morphology (*cldnb*:lynGFP, a membrane-localised variant of the green fluorescent protein gene under the control of the *claudin b* gene promoter) (Anzenberger et al., 2006). Other zebrafish lines used were AB and Golden. Zygotes were collected from natural spawnings and placed in petri dishes of E3 medium (5 mM NaCl, 0.17 mM KCl, 0.33 mM CaCl2, 0.33 mM MgSO4) (Detrich et al., 2011), one dish per clutch. Embryos for injection were loaded onto the edge of a glass slide in single file, after discarding dead embryos (which appear grey or black under light microscopy). MOs were microinjected

under light microscopy into the yolk of 1-8 cell embryos, using a glass micropipette (made by a Sutter Instrument Co. P-97 flaming micropipette puller) and an Eppendorf FemtoJet pneumatic microinjector. The micropipettes were calibrated by measuring droplet size on a reticle so that 2 nl of MO solution was injected into each embryo. Successful injection was confirmed by the presence of red solution in the middle of the yolk, which diffused slowly outwards into the rest of the embryo. The number of microinjected embryos was recorded from the counter on the microinjector display, and the number of uninjected embryos retained for comparison was counted manually.

2.9.2 Preparation of Moepholino (MO) solution for ZF injection:

Morpholino antisense oligonucleotides are widely used to modify gene expression by blocking translation of a targeted protein or by modifying pre-mRNA splicing. Morpholinos in the zebrafish serve as a powerful reverse genetics tool by knocking down gene function.

We used microinjection of a morpholino oligo targeted to the translational initiation site of *BBS5* gene (5-GATCACTGTCTGCGTATATTGTCGA-3); BBS5_ATG oligo (300nmol, mw=8465), obtained from Gene Tools.

150 µL, sterile water was added to lyophilized oligo to make a 2 mM stock solution. The solution was aliquoted and stored at -20 °C until ready for use. On the day of the injection, the morpholino solution heated at 65 °C for 10 min then snaps cooled on ice immediately and spun briefly. This step denatures any secondary structures in the oligo and ensures that the solution is completely solubilized. An injection buffer was prepared: 45 ul Danieau solution + 5 ul of 0.5% phenol red dye (Sigma-Aldrich Co., P0290). The phenol red serves as a visible marker for the injection of the solution into the embryo.

Finally a working solution prepared by diluting the morpholino solution in the injection buffer; A typical working solutions of 0.4 mM to 0.1 mM *bbs5* morpholino was prepared. The delivered amount of morpholino was estimated from volume of morpholino injected. By injecting 1.8 nl of 0.4 mM of morpholino the dose given to the embryo is approximately equal to 6 ng.

2.9.3 Preparation of mRNA solution for ZF injection:

For gene function studies, we prepared capped RNA for microinjection into zebrafish embryos. Capped RNA behaves similarly to eukaryotic mRNAs found in vivo due to the presence of the CAP analog.

Human *BBS5* EST clone (ref. 5272889) inserted into vector pBluescriptR was obtained (GeneService) and the clone was sequenced to verify clone length and fidelity. To obtain mutant BBS5 (c.966insT) site directed mutagenesis was employed (using a commercial Mutagenesis service (Mutagenex Inc)). Clones (wild type and mutant BBS5) contained a T7 site upstream of ATG. Using a forward primer AE37 (T7 primer) and reverse primer AE91 (in 3'UTR of BBS5) PCR products of size 1100bp from wild type and mutant BBS5 was obtained. These products were cleaned using Qiagen PCR clean up kit. The resulted PCR products were used to perform an *in vitro* synthesis reaction of large amounts of capped RNA using the mMessage mMachine T7kit (Ambion, Inc., AM1340). cRNA samples purified by phenol:chloroform extraction and isopropanol precipitation. Resulted cRNA was suspended in RNAse free H_2O and quantified then stored at -80°C until ready for use. Working solutions with sterile water and a final concentration of 0.05% phenol red (Sigma-Aldrich Co., P0290) was prepared. Phenol red serves as a visible marker for the injection of the solution into the embryo. Injecting 1.78nl into embryos to deliver 90pg and 180pg of mRNA requires concentration of 50 and 100 ng/µL cRNA respectively. Clones were designed and prepared by Dr. John Sayer (Figure 2.9.1 A, B, and C)

MO injected alone or together with rescue cRNA (Mutant and wild type) as described above.

After microinjection, the embryos were placed in fresh E3 medium, approximately 50 embryos per petri dish and incubated at 28.5 °C. Mortality counts were performed at 3 and 24hpf on both injected and uninjected control embryos. The mortality rate in the 0-3 hour timeframe was used to assess death due to needle trauma; the 3-24 hour timeframe was used to assess death due to MO effects. Clutches where the mortality rate in uninjected embryos was above 50% at 24 h post-ferilization (hpf) were discarded.

2.9.4 Phenotyping zebrtafish

Light and fluorescent (GFP filter) microscopy was performed to monitor the embryos at 24, 48 and 72 hpf using a Leica MZ16FA fluorescent microscope on embryos immobilised in 0.1% tricaine in E3. The appearance of each fish at 72 hpf was recorded, noting the presence of pericardial effusion, tail length and appearance, body oedema and venous stasis in the cardiac venous sinus. Pronephric duct morphology assessment in *cldnb*:lynGFP fish was performed at 48hpf for better visualization. To assess cardiac

morphology (D-loop, L-loop or failure of looping and cardiac chamber appearance), 72hpf cmlc2:GFP embryos were placed on their backs against the edge of a glass slide.

Fish were phenotyped under light microscopy and classed as following:

Normal: if they were indistinguishable from un-injected fish and

Mild: if they had small eyes, mild pericardial effusion and mild or no tail defects.

Moderate: if they had a large pericardial effusion and moderate tail defects, (obvious shortening and large kinks / curls in their tails).

Severe: if they had no or very short malformed tail, widespread oedema, malformed eyes and minimal cardiac muscle contraction.

These classifications have been suggested by (Bedell et al., 2011).

Images were taken with a Leica DF425C camera and the Leica Application Suite V3 program. The data was then analysed by counting the number of fish in each group exhibiting certain phenotypes. A chi-squared test was performed to discover if a difference in the proportion of fish exhibiting certain phenotypes between two groups was statistically significant using the Prism GraphPad software.

2.10 Transfection:

2.10.1 Cell culture

Routine tissue culture techniques were used to facilitate the growth of the Human Embryonic Kidney 293 cells (HEK293) mammalian cell line (Graham et al., 1977). These were performed in class II laminar flow hoods with sterile handling techniques. Cell lines were cultured in roux of 25 cm^2, 75 cm^2 or 162cm^2 with Dulbecco's Modified Eagle Media (DMEM).

2.10.2 Methods of Transfection and Factors influencing Transfection

Transfection is the process of introducing nucleic acids into eukaryotic cells by non-viral methods. Using various chemical, lipid or physical methods, this gene transfer technology is a powerful tool to study gene function and protein expression in the context of a cell. Transcription of the transfected gene can be analyzed 24-48 hours after transfection. Transfection is dependent upon multiple parameters which affect

efficiency of the reaction. It is important to have a healthy cell culture, preferably of low passage number. Cells should be routinely split 24 hours prior to transfection, which allows normal cell metabolism to be re-established and increases the take up rate of DNA. Contamination with bacteria, mycoplasma and fungi must be avoided. The type of vector construct also influences the efficiency of transfection. Linearised plasmid DNA allows improved integration into the host genome and may be used for stable transfection, whereas transient transfection is most efficient with supercoiled plasmid DNA. Best results are obtained with high purity plasmid DNA, and removal of endotoxins may be required in sensitive cell lines. Classical transfection methods include the calcium–phosphate method and electroporation. In this study transfection was carried out using the proprietary reagent Lipofectamine 2000.

Liposomes are positively charged lipid reagents that interact with the negatively charged DNA molecules to form complexes, which are then endocytosed. Compared to calcium phosphate this method allows higher efficiency transfection. Cells must often be transfected in the absence of serum to achieve high efficiency. Lipofectamine 2000 is a new highly efficient liposomal transfection reagent.

Mutant *BBS5* (c.966insT) and wild type plasmids were engineered into pcDNA3.1-NT-GFP vector (Invitrogen) to test the effect of the mutation on protein expression. A green fluorescent protein (GFP) reporter system was introduced in both constructs. Following subcloning into pcDNA3.1-NT-GFP clones were fully sequenced to confirm sequence fidelity, reading frame and orientation (data not shown). Following successful cloning, large scale preparations of plasmid cDNA were obtained using overnight culture in LB broth and isolation of DNA using a MaxiPrep kit (Qiagen). Plasmids were engineered and prepared by Dr. John Sayer (Figure 2.9.1 D, E, F, and G).

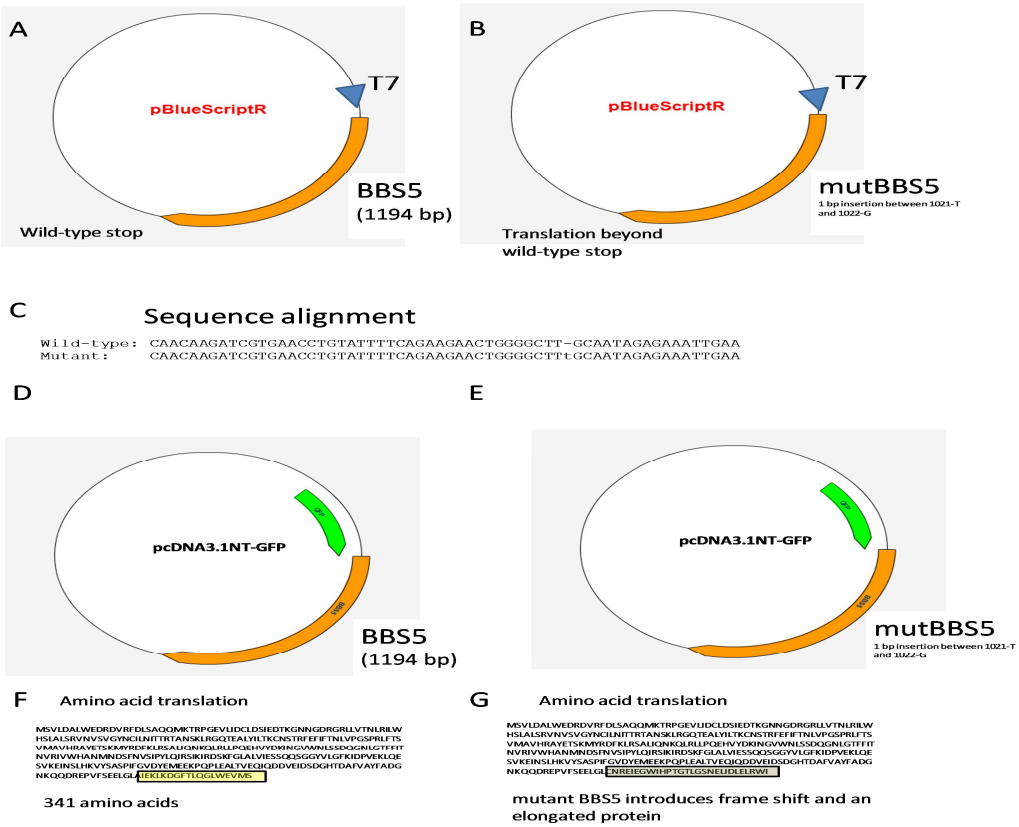

Figure 2.4: Constructs used for BBS5 in vitro studies

A. pBlueScriptR vector is shown containing full length human BBS5 nucleotide sequence. This vector DNA was used as a template for the production of cRNA, using the T7 promoter region as shown (Blue triangle).

B. pBlueScriptR vector is shown containing full length human BBS5 with a single nucleotide insertion exactly mimicking the novel mutation. This mutant cDNA was used as a template for the production of cRNA, using the T7 promoter region as shown (Blue triangle). The frame shift causes translation beyond the wild type stop codon, producing a predicted elongated mutant protein.

C. Nucleotide sequence showing single t insertion in the mutant construct (verified by Sanger sequencing (data not shown)).

D. Subcloning ofBBS5 from pBlueScriptR vector into pcDNA3.1NT-GFP was performed. Shown is the vector containing full length human BBS5 nucleotide sequence with a N-terminal GFP tag (Green). This vector DNA was used for transfection studies.

E. Subcloning of mutant BBS5 from pBlueScriptR vector into pcDNA3.1NT-GFP was performed. Shown is the vector containing mutant BBS5 nucleotide sequence with a N-terminal GFP tag (Green). This vector DNA was used for transfection studies.

F. Amino acid translation of wild-type BBS5. Box identifies residues that would be affected by the novel mutation

G. Amino acid translation of mutant BBS5, with alternate reading frame and new translated amino acids shown in boxed region.

2.10.3 Transfecting Human Embryonic Kidney 293 cells (HEK 293 Cells):

Human Embryonic Kidney 293 cells (HEK 293) are a specific cell line originally derived from human embryonic kidney cells grown in tissue culture (Graham et al., 1977). HEK 293 cells grow easily and transfect very readily and have been widely-used in cell biology research for many years. They are also used by the biotechnology industry to produce therapeutic proteins and viruses for gene therapy.

Lipofectamine™ 2000 Transfection Reagent (Invitrogen Life Technologies) is a commonly used reagent for achieving transient transfection of HEK293 cells with high efficiency. Lipofectamine 2000 reagent was used empirically as per the manufacturer's protocol. The day before transfection, HEK 293 cells were trypsinized by removing the media and adding 4 ml of 0.05% Trypsin-EDTA (Gibco). After 5 min incubation at 37 C°, 10 ml of complete growth medium (Dulbecco's Modified Eagle Medium (DMEM), Gibco) was added to 0.5 ml of trypsinized cells. In 6-well format plate, 0.5 ml of trypsinized cells (approximately 6.25 x 10^5 cells) and 1.5 ml of DMEM media were added to each well on top of a sterilized cover slip. The plate was incubated at 37 °C for 24 hrs. Cell density was 50-80% confluent on the day of transfection, to allow efficient transfection. On the day of transfection, the plate was checked and the media was replaced with 1.5 ml of complete growth media. Typically the 6 well plate of HEK293 cells was divided into two groups, three wells for transfection with wild type plasmid and remaining three for mutant plasmid.

For each well of cells to be transfected, 2.5 µg of plasmid DNA was diluted in 500 µl of Opti-MEM® Reduced Serum Media without serum. For each well of cells, 5 µl of Lipofectamine™ 2000 was added to the above diluted Opti-MEM®: DNA solution and mixed gently. The mix was left at room temperature for 30 min to form DNA-Lipofectamine™ 2000 Reagent complexes. After incubation, 500 µl of the mix was added directly to each well containing cells and mixed gently by rocking the plate back and forth. The plate then was incubated at 37 °C for 24 hrs in a CO_2 incubator.

2.10.4 Immunofluorescence

For fixation, 40 grams of para-formaldehyde (Sigma) was dissolved in 1X PBS (Sigma) up to 1 Liter to make 4% of para-formaldehyde (4% PFA). Media was removed from

the plate and cells were fixed with 2 ml of 4% PFA at room temperature for 10 min. Cells then were washed twice with 500 µl 1X PBS for 5 min each. Cells permeabilization was achieved with 0.5% Ttriton-X100 in PBS solution for 10 min at room temperature. Cells were then washed twice with 1X PBS for 5 min each. Prior to immunofluorescence staining, cells were blocked with 5% BSA in PBS for 30 min to block non-specific binding of immunoglobulin. 1.0 ml of monoclonal anti mouse acetylated tubulin (Sigma) diluted in block solution at 1 in 1000 was added to each well for 1 hour at room temperature. Cells were washed with PBS three times for 5 min each. Secondary antibody Alexa Fluor 549 donkey anti mouse (Invitrogen) diluted 1 in 200 in block solution was added to cells at room temperature for 1 hour. Cells were washed with PBS three times for 5 min each. Coverslips were mounted on microscopic slides with vectashield (Vector Laboratories H-1000 with DAPI (DNA stain)) and sealed with nail polish. Staining protocol was also carried out using a pericentrin primary antibody (Abcam) and secondarily detected using an anti-rabbit Cy3 antibody.

For fluorescence visualization, digital images of the green, red, blue and overlay fluorescent pattern were captured using confocal laser scanning microscopy (Nikon and Olympus).

2.11 Bioinformatics:

Bioinformatics is the application of computer science and information technology to the field of biology and medicine.

The following software packages were used in the study:

2.11.1 easyLINKAGE-Plus for Linkage Analysis

easyLINKAGE was designed to make the use of linkage programs user-friendly and to enable those analyses on Microsoft Windows based operating systems. The software was written in a PERL/Tk-based program. The programe was generated in 2005 (Lindner and Hoffmann, 2005, Hoffmann and Lindner, 2005) to combine automated setup and performance of linkage analyses and simulation under an easy to handle graphical user interface for Microsoft Windows 2000/XP and standard UNIX systems.

An Improved easyLinkage program (easyLINKAGE Plus) is enabling linkage analyses for large-scale SNP data in addition to those of microsatellites. A major benefit of

easyLINKAGE is the generation of structured text outputs and graphical plots of LOD scores, P values, and many other parameters. The implementation of SNP projects derived from the Affymetrix 10k/50k/100k/250k/500k/axiom and Illumina chips is another major step ahead. Those projects can be analyzed with Allegro (single-/multipoint analyses), GeneHunter/-Plus, Merlin, FastLink, SimWalk and SuperLink. For SNP projects only two files are necessary: one file that contains all genotypes, other file with the pedigree information in linkage format. Only files that start with “p” and end with “pro” will be recognized as pedigree information files.

Information in our pedigree files were arranged according to the Table 2.1 below in a Notepad txt files. The name of pedigree file was saved as (pESRF_ADfamily.pro) to be recognized as a pedigree information file.

Family ID	Individual #	Father	Mother	Gender 1=Male 2=Female	Affection status, 1=unaffected 2=affected	DNA availability 0= No 2= Yes
1	1	7	8	1	2	2
1	2	7	8	2	2	2
1	3	7	8	2	2	2
1	4	7	8	2	2	2
1	5	7	8	2	2	2
1	6	9	2	2	2	2
1	7	0	0	1	2	0
1	8	0	0	2	1	0
1	9	0	0	1	1	0
1	10	7	8	1	1	2
1	11	7	8	1	1	2

Table 2.1: Arrangment of pedigree data in Notepad txt files for run in easyLINKAGE software.

The file that contains genotypes was first exported to an Excel file to assign each individual an identifier code for the pedigree. If there was no call for a SNP, this was coded as 00. The data for each individual were then combined to produce a single genotype data file. The file then paste in a Notepad txt file and saved as Snapshot files (ESRF_ADfamily.snp). Both files (pedigree and genotype) were then uploaded in easyLINKAGE-Plus software and Allegro program was used to perform parametric linkage analysis in autosomal recessive and dominant models as shown in Figure 2.4.

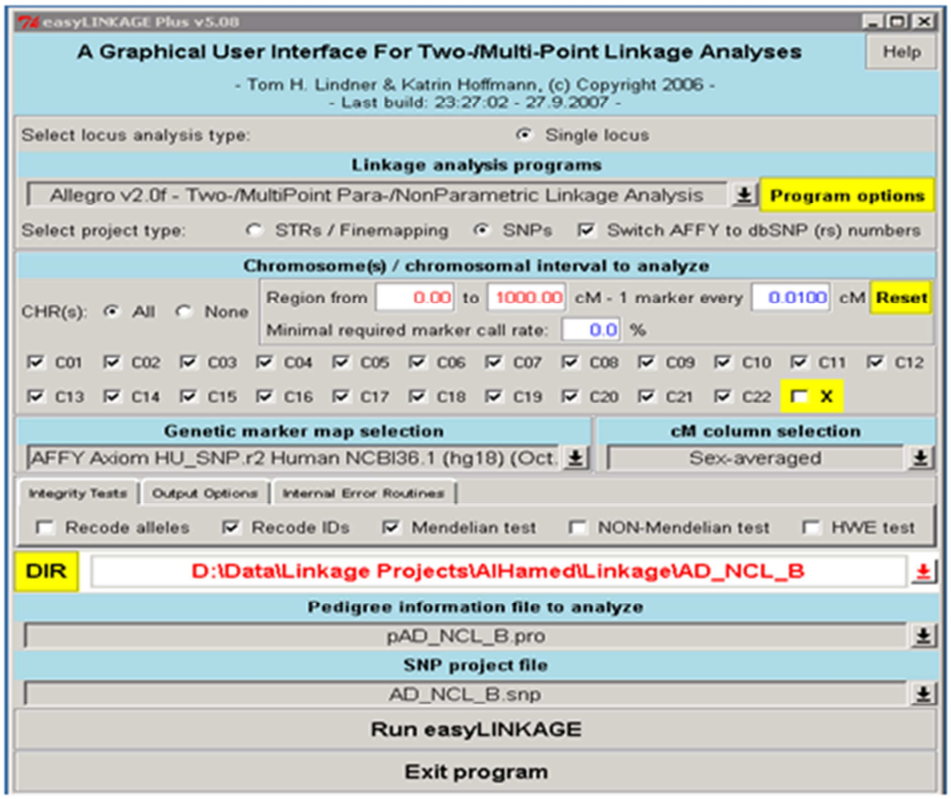

Figure 2.5: Parameters used to run Allegro program to perform linkage analysis in easyLINKAGE-Plus software.

Allegro is a computer program for multipoint genetic linkage analysis and related calculations (Gudbjartsson et al., 2005). Allegro multipoint linkage analysis is based on the Lander-Green Hidden Markov model (HMM).

It is the program that used to calculate parametric LOD score in inherited kidney disease families in our study. The biggest advantages of Allegro over other programs are the allele sharing models that it provides and a much shorter execution time. In addition to parametric LOD scores, Allegro calculates allele sharing LOD scores (Kong and Cox, 1997) and non-parametric (NPL) scores. The linkage analysis of Allegro involves three steps. Firstly, determination of single point probabilities of individual inheritance vectors, secondly multipoint calculation, where genotype information on neighboring markers is taken into account, and thirdly score calculation. The computational engine of Allegro is based on the Fourier transforms to calculate convolutions (Kruglyak and Lander, 1998). Although Allegro is not limited in terms of the number of markers, using large quantities of SNPs can lead to a wrong haplotype assignment.

GeneHunter is another program implemented in easyLINKAGE. In 1996, Kruglyak et al. introduced GeneHunter (Kruglyak et al., 1996) and since then it has undergone many improvements. It does perform multipoint analysis of pedigree data that include: LOD-

score computation, non-parametric linkage analysis, information-content mapping, and haplotype reconstruction. It is similar to Allegro in using the Lander-Green Hidden Markov model of extracting complete multipoint linkage data from incomplete marker information. The program was recompiled for the use in Microsoft Windows. GeneHunter program is limited to run in nuclear pedigrees or small multigenerational pedigrees (less than 16 individuals). In addition Allegro and other programs result in faster analysis speeds than GeneHunter. In our study we used GeneHunter to calculate LOD score and to confirm results obtained by Allegro.

2.11.2 Homozygosity Mapping Software

The following software programs were used to analyze homozygosity in autosomal recessive families. CNAG (Version 2.0): Copy Number Analyzer for Affymetrix GeneChip Mapping 100K arrays (CNAG) version 2.0 was developed (Nannya et al., 2005) to enable high-quality analysis of copy number alterations and allelic imbalances in cancer genomes, congenital disorders, and normal individuals using Affymetrix GeneChip® platforms. It employs a robust algorithm to correct between chip variations introduced by the differences in experimental conditions, enabling high quality copy number analysis. When the constitutive DNA is available, it also allows for allele-based copy number estimation to sensitively detect allelic imbalances. Copy number calls are automated according to the copy number calls based on the hidden Markov model analysis and the LOH inference is also enabled even when constitutional DNA is not available. This algorithm improves signal-to-noise (S/N) ratios by reducing variation in the raw signal ratios and optimizes the selection of the reference. In addition, availability of accurate genotyping information further enables LOH inference and allele-based copy.

The CNAG software was used initially as a tool to detect loss of heterozygosity (LOH) regions in the families studied. The software is user friendly but was not designed mainly for homozygosity mapping and therefore many homozygosity regions were missed. The software is free and can be downloaded from the following website: http://www.genome.umin.jp/CNAGtop2.html

Genotyping Console (Version 3.0.1): was the second software to be used for detecting homozygosity regions in families. The software provided by Affymetrix® is the first commercially available software that integrates SNP genotyping, copy number

polymorphism (CNP) genotyping, and rare copy number variation (CNV) identification in one data analysis application. It generates genotyping calls, copy number calls for CNV regions and individual probe sets, loss of heterozygosity (LOH) data, cluster graphs, and quality control metrics. The software was used for analysis of 500K Array Sets and SNP 6.0 Array.

The Affymetrix® Chromosome Analysis Suite 1.2 (ChAS) Software was designed specifically for cytogenetic researchers with an easy-to-use graphical interface. The software is the newest and most used software in this study. It has many advantages over other software packages. In addition to its ease of use, the software can provide full view of Chromosomal aberrations that may include copy number gain or loss, mosaicism, in addition to loss of heterozygosity (LOH). ChAS provides several options for viewing and studying the loaded CYCHP or CNCHP data as graphic displays or tables. Unfortunately, the computer algorithm for ChAS is not provided by the company.

2.11.3 Sequence Alignment Software

Mutation Surveyor® software Version 3.24 from softgentics and Lasergene's SeqMan II version 6.0 software were used for sequence alignment and analysis. Lasergene's SeqMan II offers quick and accurate sequence assembly and analysis of Sanger sequencing data. Sequencing results can be copied and blast in other browser for sequence search. Mutation Surveyor® is DNA Sequencing analysis software capable of performing variant analysis of up to 2000 Sanger sequencing files (.ab1) generated by Applied Biosystems Genetic Analyzers electrophoresis system in 15 minutes. Utilizing patented anti-correlation technology Mutation Surveyor delivers excellent accuracy, sensitivity, low false positive and negative rates in the analysis of DNA variants.

2.11.4 Web-based Genetics Tools

The following websites were used for displaying genes, gene mapping, gene findings, gene predictions, reference sequences, mRNA and expressed sequence tag alignments, simple nucleotide polymorphisms, expression and regulatory data, and pairwise and multiple-species comparative genomics data in addition to clinical information.

- The University of California Santa Cruz (UCSC) Genome Browser (genome.ucsc.edu): The UCSC Genome Browser presents a diverse collection of annotation tracks including mRNA alignments, mappings of DNA repeat

elements, gene predictions, gene-expression data, and disease-association data. The Genome Browser was used for gene predictions in a region, gene expression in tissues, designing primers through ExonPrimer, to detect change conservations through vertebrate. BLAT (the BLAST-Like Alignment Tool) was used to map sequence to the genome. It was developed to identify similarities between DNA sequences and protein sequences and to assist in the annotation of the human genome sequence.

- The National Center for Biotechnology Information (www.ncbi.nlm.nih.gov): Through this site, the NCBI provides access to Genbank and to the Entrez databases, as well as to many tools for sequence analysis and data mining. Locuslink is also at NCBI, and for non-redundant and curated sequence data covering the fly, human, mouse, rat and zebrafish genomes. The NCBI website was used for downloading reference sequences (Gene), to determine candidate genes in locus interval (Map Viewer), References citation (PubMed), to have information of human genes and genetic disorders (Online Mendelian Inheritance in Man (OMIM)), and Database of Short Genetic Variations (dbSNP) that contains population-specific frequency and genotype data, and mapping information for both neutral variations and clinical mutations.
- Ensembl Genome Browser (www.ensembl.org): The Ensembl project produces genome databases for vertebrates and other eukaryotic. It is a joint project between EMBL - EBI and the Wellcome Trust Sanger Institute to develop a software system which produces and maintains automatic annotation on selected eukaryotic genomes. Ensembl was used to display gene sequence, transcripts (cDNA), and exons. Also used to detect variations at nucleotide and structural levels.
- The GeneCards Human Gene Database (www.genecards.org)
- Genetics Home Reference (ghr.nlm.nih.gov)
- UniProt (www.uniprot.org)
- Tissue-specific Gene Expression and Regulation [TiGER] (http://bioinfo.wilmer.jhu.edu/tiger)
- GeneDistiller 2 (www.genedistiller.org)
- GenAtlas (www.genatlas.org)
- GeneTests (www.ncbi.nlm.nih.gov/sites/GeneTests/?db=GeneTests)

For searching and assessing reported and novel mutations detected in the study the following websites were visited:

- The Human Gene Mutation Database [HGMD] (www.hgmd.org)
- Human Genome Variation Society [HGVS] (www.hgvs.org)
- PolyPhen-2 prediction of functional effects of human nsSNPs (http://genetics.bwh.harvard.edu/pph2/): The score range in PolyPhen-2 is fall between 0 and 1. Where 0 is benign change and 1 is probably damaging.
- SIFT (http://sift.jcvi.org): SIFT score, ranges from 0 to 1. The amino acid substitution is predicted damaging if the score is ≤ 0.05, and tolerated if the score is > 0.05.
- SNPs3D (www.snps3d.org): SNPs3D scoring follows the higher the entropy is, the more tolerant a position is to a mutation. i.e. scores less than zero are more damaging than above zero.

2.11.5 Statistics (GraphPad Prism) software

Chi-square is a statistical test commonly used to compare observed data with data we would expect to obtain according to a specific hypothesis. The chi-square test is always testing what we call the null hypothesis, which states that there is no significant difference between the expected and observed result.

A chi-squared test was performed to discover if a difference in the proportion of fish exhibiting certain phenotypes between two groups was statistically significant using the GraphPad Prism software.

All data were stored at server's computers of KFSH&RC and IGM and stored in personal laptop. Copies of all data were stored in external hard drive. Data was shared between KFSH&RC and IGM through RAS logon at IGM and through Dropbox website (www.dropbox.com).

Chapter 3: RESULT

3.1 Nephrotic Syndrome

3.1.1 Introduction

Nephrotic Syndrome (NS) is a renal disease characterized by heavy proteinuria, hypoalbuminemia, edema and hyperlipidemia (Franceschini et al., 2006). Urinary losses of macromolecules such as albumin reflect a dysfunction of the highly permselective glomerular filtration barrier (GFB) (Mundel and Shankland, 2002). The GFB structure consists of podocyte foot-process, glomerular basement membrane (GBM), endothelial fenestration and the slit-diaphragm (SD). The SD, which represents the main size selective filter barrier in the kidney, is thought to be a modified adherens junction and a growing number of proteins have been identified to have a crucial role in this structure (Mundel and Shankland, 2002).

The identification of mutations leading to defects in proteins highly expressed in the podocyte and slit diaphragm (Figure 3.1.1) has helped to unravel the basis of GFB physiology and pathophysiology (Machuca et al., 2009). Kidney biopsies in NS patients may show non-specific changes such as minimal change, as well as focal segmental glomerulosclerosis (FSGS) and diffuse mesangial sclerosis (DMS) (Godefroid and Dahan, 2010). A molecular genetic diagnosis is important for making treatment decisions including suitability for renal transplantation and to enable screening other family members at risk of disease.

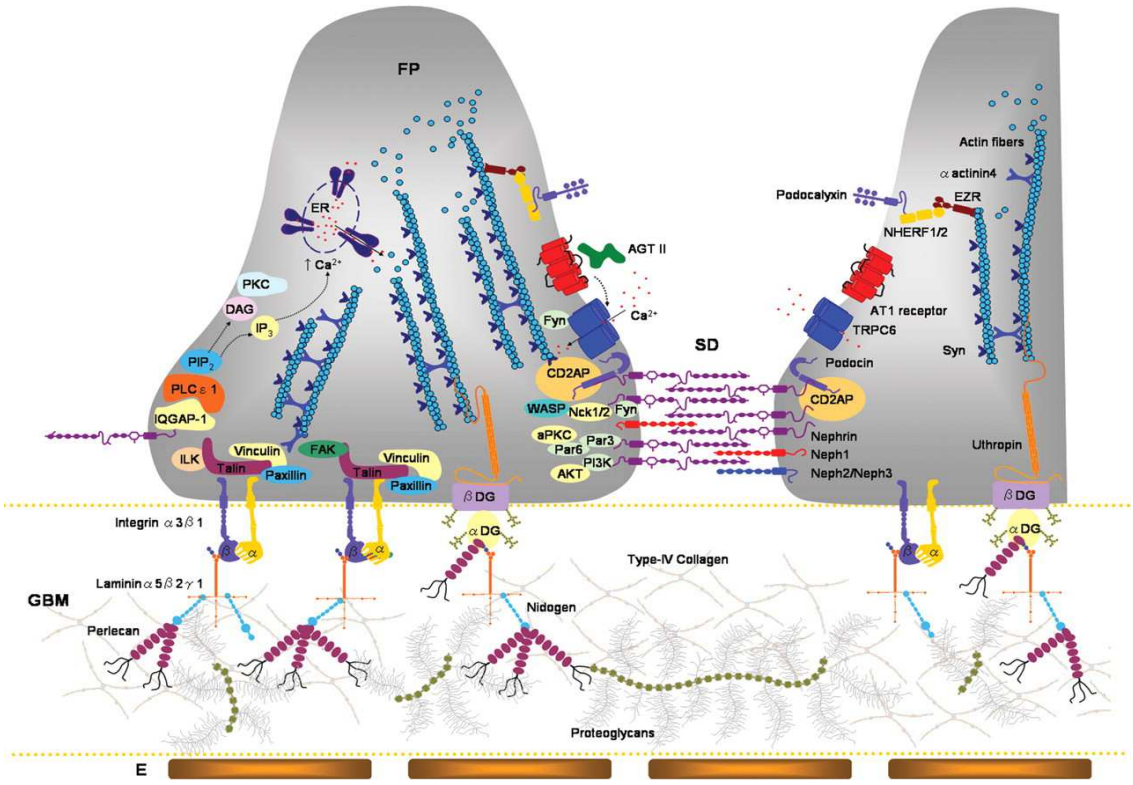

Figure 3.1.1: Molecular overview of the slit-diaphragm and podocyte cell–matrix interactions (Machuca et al., 2009).

Nephrotic syndromes may be given a series of descriptive labels dependent upon their age of presentation. Congenital NS (CNS), manifests in utero or during the first 3 months of life (Hinkes et al., 2007); infantile NS, has an onset between 3 months and 1 year of age (Hinkes et al., 2007) and childhood steroid resistant nephrotic syndrome (SRNS) may be defined as no urinary remission within 4 weeks of prednisone therapy 60 mg/m2/day (Brodehl et al., 1982). In recent years, the identification of genes coding for podocyte and slit diaphragm proteins by positional cloning in NS patients helped our understanding of molecular basis of inherited NS. To date defects in over ten genes have been recognized; *NPHS1* (Kestila et al., 1998), *NPHS2* (Boute et al., 2000), *PLCE1* (Hinkes et al., 2006), *WT1* (Jeanpierre et al., 1998), *LAMB2* (Zenker et al., 2004), *PTPRO* (Ozaltin et al., 2012), *ACTN4* (Kaplan et al., 2000), *TRPC6* (Winn et al., 2005), *CD2AP* (Kim et al., 2003), *INF2* (Brown et al., 2010), and *MYO1E* (Mele et al., 2011).

***NPHS1 gene* (Nephrin):**

Nephrin is the major component of SD which is a 1241-residue transmembrane adhesion protein of the immunoglobulin superfamily. *NPHS1* is the coding gene for Nephrin. Kestila et al. (Kestila et al., 1998) mapped the CNF gene to chromosome 19q13.1 and subsequently cloned the *NPHS1* gene, which contains 29 exons and has a size of 26 kb. Mutations in *NPHS1* cause the congenital nephrotic syndrome (CNS) of Finnish type that characterized by autosomal recessive inheritance. Two mutations, Fin-major and Fin-minor account for more than 90% of mutations in Finland. To date there are 173 mutations in the *NPHS1* gene reported to cause NS [HGMD]. Compared with many other genetic disorders, *NPHS1* shows relatively little phenotypic variation.

***NPHS2* gene (Podocin):**

Podocin is a hairpin-like protein at slit diaphragm which lines the podocytes and assists in maintaining the barrier at the GBM. Podocin is a 383-residu and a member of the stomatin family, which is one of the most important membrane proteins and is exclusively expressed in the podocytes at the foot processes in the place of anchorage of the slit diaphragm. Podocin interacts with nephrin and *CD2AP*, and is part of a protein complex within podocytes that includes Neph1 and Neph2 proteins.

NPHS2 gene on chromosome 1q25 consists of 8 exons and codes for podocin. Mutations in *NPHS2* cause autosomal recessive NS. Mutations in *NPHS2* accounts for 42% of familial and 10% of sporadic cases of childhood-onset Steroid Resistant Nephrotic Syndrome (SRNS) and have also been found in 39% of patients with CNS (2). There are 116 mutations reported so far in the *NPHS2* gene.

***PLCE1* gene (phospholipase C epsilon)**

Phosphatidylinositol-4,5-bisphosphate phosphodiesterase epsilon-1 is an enzyme that in humans is encoded by the *PLCE1* gene. *PLCE1* belongs to the phospholipase family that catalyzes the hydrolysis of polyphosphoinositides such as phosphatidylinositol-4,5-bisphosphate (PtdIns(4,5)P2) to generate the second messengers Ins(1,4,5)P3 and diacylglycerol. These products initiate a cascade of intracellular responses that result in cell growth and differentiation and gene expression. The *PLCE1* gene located on Chromosome 10q23 and consists of 32 exons coding for 2302 amino acids.

PLCE1 gene mutations have been recently described in patients with early onset nephrotic syndrome (NS) and diffuse mesangial sclerosis (DMS) (Boyer et al., 2010). In addition, two cases of *PLCE1* mutations associated with focal segmental glomerulosclerosis (FSGS). To date there are 28 mutations reported in the gene according to HGMD database.

***CD2AP* gene (CD2-Associated Protein):**

CD2 associated protein seems to act as an adapter protein between membrane proteins and the actin cytoskeleton. It is may anchor the podocyte slit diaphragm to the actin cytoskeleton in renal glomerolus. Also it may play a role in receptor clustering and cytoskeletal polarity in the junction between T-cell and antigen-presenting cell.

CD2AP gene located in chromosome 6p12 and consists of 18 exons coding for 639 amino acids. Defect in *CD2AP* gene modify the interaction with CD2 in lymphocytes and alter the composition of the renal slit diaphragm. Mutations in the *CD2AP* results in presence of segmental sclerosis in glomeruli and resulting in proteinuria, reduced glomerular filtration rate and edema. Only six mutations reported on this gene.

***LAMB2 gene (*Laminin subunit beta-2):**

LAMB2 encodes the basement membrane protein laminin beta2. Laminin is thought to mediate the attachment, migration and organization of cells into tissues during embryonic development by interacting with other extracellular matrix components. *LAMB2* gene located in chromosome 3p21 consisting of 32 exons coding for 1798 amino acids. Mutations of *LAMB2* are the cause of Pierson syndrome that is characterized by nephrotic syndrome with neonatal onset, diffuse mesangial sclerosis and eye abnormalities. Mutations in *LAMB2* also cause congenital nephrotic syndrome that may be associated with ocular abnormalities. Non truncating *LAMB2* mutations may display variable phenotypes ranging from a milder variant of Pierson syndrome to isolated congenital nephrotic syndrome. Fifty five mutations were reported to cause diseases in *LAMB2* gene so far.

***MYO1E* gene (myosin IE):**

This gene encodes a member of the nonmuscle class I myosins which are a subgroup of the unconventional myosin protein family. The unconventional myosin proteins

function as actin-based molecular motors and serve in intracellular movements. *MYO1E* gene is located on Chromosome 15 with 28 coding exons. Mutations in this gene are the cause of focal segmental glomerulosclerosis-6 (Mele et al., 2011). To date there are only two mutations reported in the *MYO1E* gene (A159P and Y695X).

***PTPRO* gene (protein tyrosine phosphatase, receptor type, O):**

PTPRO gene encodes a member of the R3 subtype family of receptor-type protein tyrosine phosphatases. The function of PTPRO is to possess tyrosine phosphatase activity. It is localized mainly in the glomerulus of kidney also detected in brain, lung and placenta. The *PTPRO* gene located on Chromosome 12 and consisted of 26 coding exons. Recently, (Ozaltin et al., 2012) reported two splicing mutations that cause nephrotic syndrome, childhood-onset. No other mutations reported in the gene up to date.

Recently, *NEIL1* (Nei Endonuclease VIII-like 1) which encodes a base-excision DNA repair enzyme was postulated as a candidate gene for NS in a single consanguineous family (Sanna-Cherchi et al., 2011). Its role in NS remains uncertain.

Despite increasing genetic evidence supporting the fact that underlying mutations lead to NS, the majority of NS occurs sporadically; The incidence of familial cases is estimated to be just 3 to 5% (Obeidova et al., 2006). In this study we screened for mutations in 9 genes implicated in inherited NS in a Saudi Arabian population with either CNS, infantile NS or childhood SRNS. Such a study has never been conducted for this part of the world. The Saudi Arabian population has a tribal structure and the overall rate of consanguineous marriage is reported to be around 58%, with regional variations (el-Hazmi et al., 1995). In such a population, the identification of mutations in known recessive disease genes is an important consideration and we were interested in our ability to detect a molecular genetic cause of NS. Identification of a molecular genetic cause of NS allows both a definitive diagnosis and improved clinical management of the patient and at risk relatives. It is noteworthy that the Saudi population is at high risk of renal failure, with 133 incident cases per million population per year that require renal replacement therapy (2011).

We identified 67 cases, representing 53 families with CNS, infantile NS or childhood SRNS and undertook mutational analysis in known / candidate NS genes. We found a

high rate of mutations in this cohort, solving 47% of cases. Direct sequencing identified novel and likely pathogenic genetic variants in *NPHS1, NPHS2, MYO1E* and *PLCE1*, increasing the known spectrum of mutations in these genes.

3.1.2 Methods

3.1.2.1 Study Cohort

Following informed consent, DNA was extracted from peripheral blood cells using the Gentra Systems PUREGENE DNA Isolation kit. A total of 67 samples from 53 different families were obtained. Altogether, 29 samples were obtained from 15 families with evidence of familial nephrotic syndrome and 38 samples were obtained from families with a single affected individual with NS. Clinical phenotypes included patients with CNS, infantile NS and childhood SRNS. Consanguineous marriages were noted; where 36 families were consanguineous and five families were not. Consanguinity status was unknown in 12 families.

3.1.2.2 Mutation Analysis

Mutational screening was undertaken of known genes implicated in nephrotic syndrome; *NPHS1, NPHS2, LAMB2, PLCE1, CD2AP, MYO1E, WT1, PTPRO* and *NEIL1*. Direct sequencing of all coding exons and exon-intron boundaries was performed. Oligonucleotide primers for PCR amplification of genomic DNA were designed using Primer3 software (http://frodo.wi.mit.edu/) and synthesized by Metabion International AG (Germany). Primer sequences are listed in Appendix A. PCR was performed in a final volume of 25µl. PCR products were treated with the Agencourt® AMPure® PCR purification system. PCR products were sequenced using BigDye™ Terminator Cycle Sequencing kit (PE Applied Biosystems) as described by the manufacturer. Sequences were analyzed using Mutation Surveyor® software Version 3.24 (SoftGenetics) and SeqMan II software 6.1 (DNAStar).

Computational analyses of novel missense mutations were performed with PolyPhen-2 (http://genetics.bwh.harvard.edu/pph2/), SIFT (http://sift.jcvi.org/), SNPS3D (http://www.snps3d.org/) and mutationtaster (http://www.mutationtaster.org/). To assess splicing effects we used the GeneSplicer software (http://www.cbcb.umd.edu/software/GeneSplicer/gene_spl.shtml). In addition to database searches, a control DNA panel from 175 individuals from a Saudi Arabian population was used to screen for all novel sequence variants.

3.1.3 Results

Sixty seven individual cases affected with congenital NS, infantile NS, childhood SRNS representing 53 families from the Arabian Peninsula were screened for mutations in the following genes: *NPHS1, NPHS2, LAMB2, PLCE1, MYO1E, WT1, PTPRO, NEIL1* and *CD2AP*. We identified mutations known NS genes in 11 out of 15 (73%) families where there was reported to be more than 1 affected member with NS, suggesting an inherited cause. A molecular genetic diagnosis was obtained in 37% (14 of 38) of families with a single affected member with NS. All families in whom we detected mutations had a history of consanguinity, most commonly first cousin marriages. Overall, in this population, by screening 9 genes implicated in inherited NS we established a likely molecular genetic cause in 47% of families (Figure 3.1.2 and Table 3.1.1).

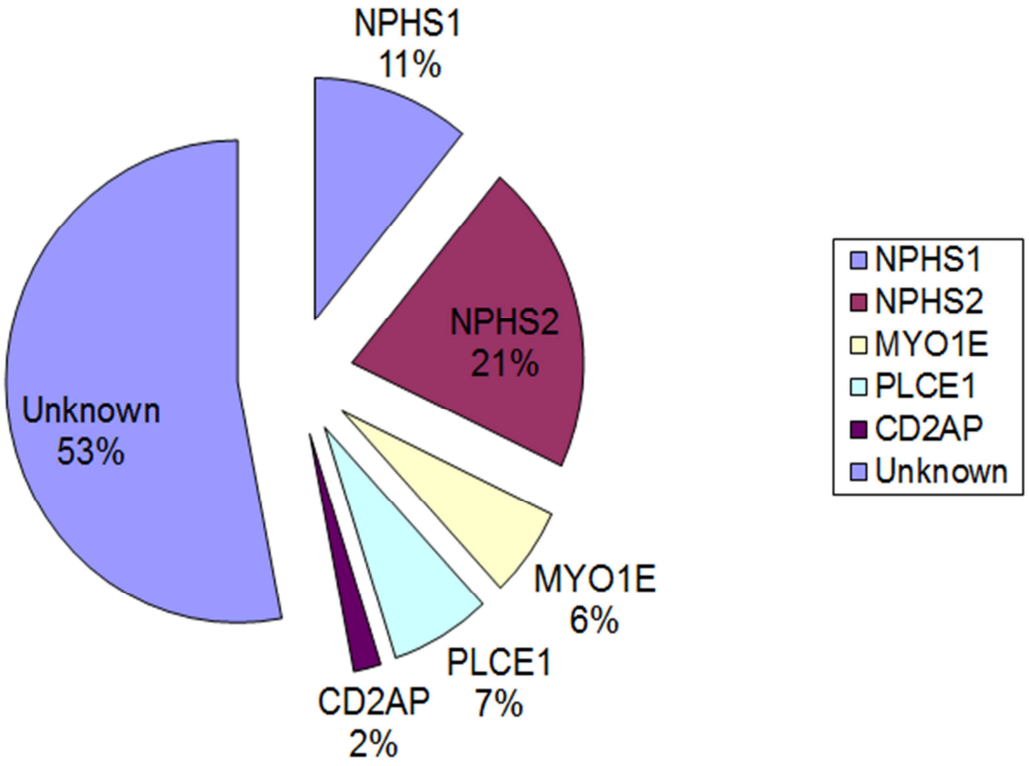

Figure 3.1.2: Percentage of cases with mutations in genes causing NS in Saudi population

Family ID	Number of affected	Familial (F)/Sporadic (S)	Consanguinity	Phenotype & Pathology	*Gene*	DNA variant	Homozygous or Heterozygous	Predict effect on Protein	Previously reported?
S3#	1	S	yes	CNS	*NPHS1*	c.515delCCA	Homozygous	Frame-shift	Lenkkeri, et al., 1999
S25	1	?	yes	SRNS	*NPHS1*	c.515delCCA	Homozygous	Frame-shift	Lenkkeri, et al., 1999
S22	1	F	yes	CNS	*NPHS1*	c.1103C>T / c.1379G>A	Compound Heterozygous	P368L/R460Q	Beltcheva, et al., 2001
S37	1	S	yes	CNS	*NPHS1*	c.1627+2T>G	Homozygous	**Splicing defect**	novel
F16	1	S	yes	CNS	*NPHS1*	c.2404C>T	Homozygous	R802W	Lenkkeri, et al., 1999
F4	1	S	yes	SRNS	*NPHS1*	c.3250insG	Homozygous	Frame-shift	Kestila, et al., 1998
F19	1	F	yes	SRNS	*NPHS2*	c.115C>T	Homozygous	**Q39X**	novel
F1	3	F	yes	SRNS	*NPHS2*	c.385C>T	Homozygous	Q129X	Al-Hamed, et al., 2010
S36	1	S	yes	SRNS	*NPHS2*	c.413G>C	Homozygous	**R138P**	novel
F3	2	F	yes	CNS	*NPHS2*	c.503G>A	Homozygous	R168H	Weber, et al., 2004
F12	2	F	yes	SRNS+FSGS	*NPHS2*	c.538G>A	Homozygous	V180M	Boute, et al., 2000
F18	1	F	yes	SRNS+FSGS	*NPHS2*	c.538G>A	Homozygous	V180M	Boute, et al., 2000
S24	1	S	yes	SRNS+FSGS	*NPHS2*	c.538G>A	Homozygous	V180M	Boute, et al., 2000
F7	1	S	yes	SRNS	*NPHS2*	c.538G>A	Homozygous	V180M	Boute, et al., 2000
F2	2	F	yes	SRNS+FSGS	*NPHS2*	c.779T>A	Homozygous	V260E	Weber, et al., 2004
F6	2	F	yes	SRNS+FSGS	*NPHS2*	c.779T>A	Homozygous	V260E	Weber, et al., 2004
F8	2	F	yes	SRNS	*NPHS2*	c.779T>A	Homozygous	V260E	Weber, et al., 2004
S20*^	1	?	yes	SRNS+FSGS	*PLCE1*	c.3058C>T	Homozygous	**Q1020X**	novel
F5^	1	S	yes	SRNS	*PLCE1*	c.3058C>T	Homozygous	**Q1020X**	novel
S26^	1	S	yes	SRNS	*PLCE1*	c.3058C>T	Homozygous	**Q1020X**	novel
F15^	1	S	yes	CNS	*PLCE1*	c.3058C>T	Homozygous	**Q1020X**	novel
F13+	2	F	yes	SRNS+MC	*MYO1E*	c.141C>G	Homozygous	**Y47X**	novel
F14	1	S	yes	SRNS+MC	*MYO1E*	c.141C>G	Homozygous	**Y47X**	novel
S17	1	S	yes	SRNS+FSGS	*MYO1E*	c.356C>T	Homozygous	**T119I**	novel
F23	1	F	yes	CNS	*CD2AP*	c.600T>G / c.1120A>G	Compound Heterozygous	**F220L/T374A**	rs139926926 / Gigante, et al., 2009

Table 3.1.1: Saudi Arabian nephrotic syndrome cases identified with mutations in disease associated genes

Gene	Homozygous Nucleotide Change	Segregation from parents	Amino Acid change	PolyPhen-2	SIFT	SNPS3D	Allele frequency in Saudi population	Expected pathogenicity
NPHS1	c.1627+2T>G	Yes	**Predicted Splicing defect**	N/A	N/A	N/A	0	Pathogenic
NPHS2	c.115C>T	N/A	**Q39X**	N/A	N/A	N/A	0	Pathogenic
NPHS2	c.413G>C	N/A	**R138P**	Probably damaging (1.00)	Damaging (0)	Deleterious (-1.28)	0	Pathogenic
PLCE1	c.3058C>T	Yes	**Q1020X**	N/A	N/A	N/A	0	Pathogenic
MYO1E	c.141C>G	Yes	**Y47X**	N/A	N/A	N/A	0	Pathogenic
MYO1E	c.356C>T	N/A	**T119I**	Probably damaging (1.00)	Damaging (0)	Deleterious (-2.65)	0	Pathogenic

Table 3.1.2: Novel genetic variants detected in *NPHS1*, *NPHS2*, *PLCE1* and *MYO1E* and in silico analysis of pathogenicity

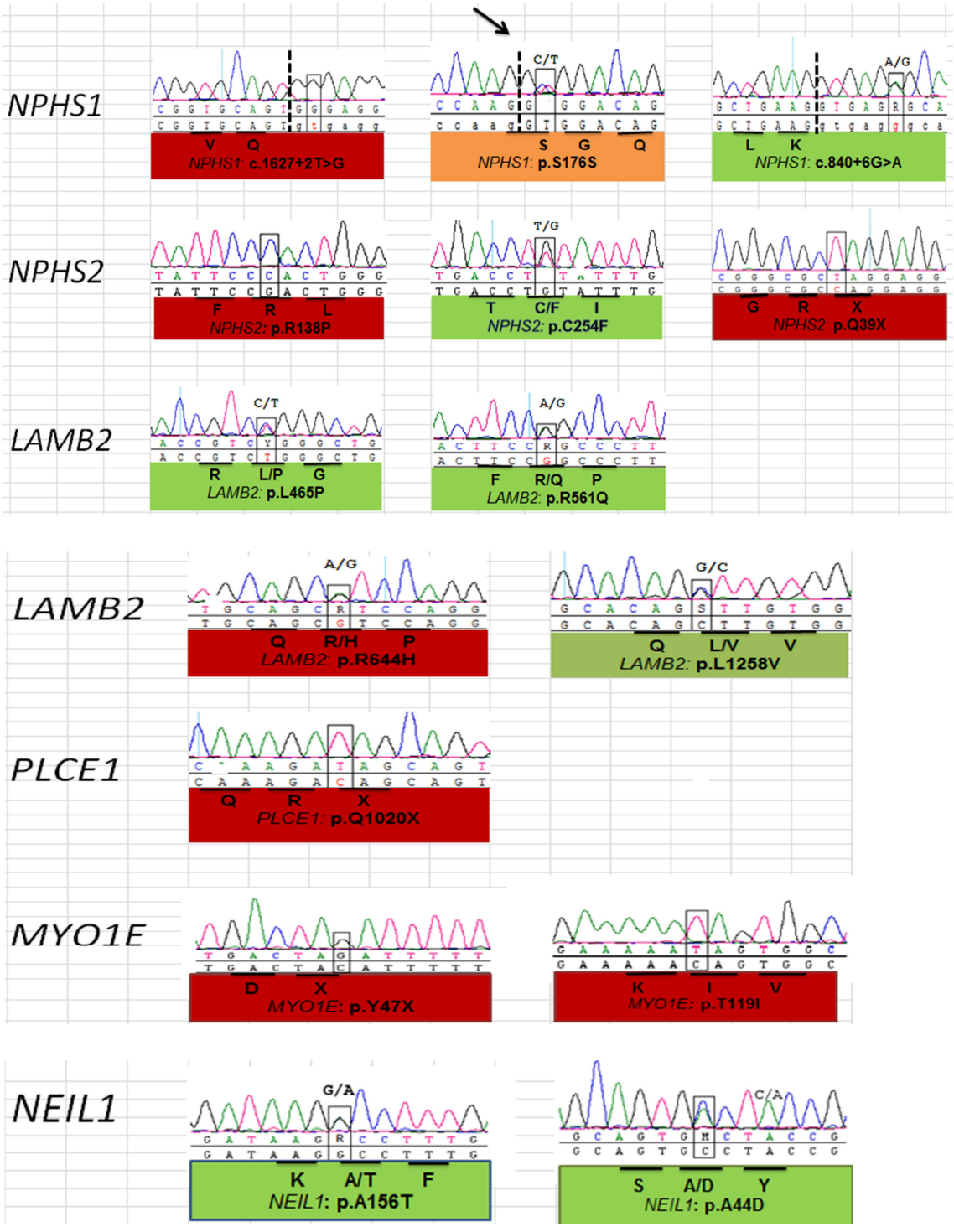

Figure 3.1.3: Sequencing chromatograms of pathogenic (in red) and benign (in green) variants detected in Saudi Arabian NS cases. Sequence chromatogram is shown with nucleotides shown, with reference sequence shown below. Reading frame is indicated by black bars and amino acid translations are given. Exons are marked with capital letters while introns marked with small letters.

No mutations were found in *WT1* following screening of exons 8 and 9, *NEIL1* or *PTPRO* by screening all coding exons. In this Saudi Arabian cohort, mutations in the NS genes *NPHS2, NPHS1* and *PLCE1* account for 21%, 11% and 8% of cases, respectively of NS (Figure 3.1.2). Mutations in *NPHS2* represent over two-fifths of genetically proven NS in this population.

Previously published data from a world-wide cohort suggested that ~50% of CNS is caused by mutations in *NPHS1* (Schoeb et al., 2010). In our cohort, mutations in *NPHS1* accounted for 4 out of 7 cases (57%) of CNS which had a confirmed molecular genetic diagnosis.

Of the novel homozygous sequence variants in coding regions (Table 3.1.2 and Figure 3.1.3), we identified 3 novel nonsense mutations, a donor splice site mutation with a high likelihood of aberrant splicing and 2 missense mutations. In silico predictions suggested that these missense changes are pathogenic (Table 3.1.2).

A detailed analysis of the results for each gene is given below.

***NPHS1*:** In this gene we found 6 families carrying known or predicted pathogenic mutations. 5 of the mutations were found in the homozygous state, and one (S22) was a compound heterozygous mutation involving two previously reported variants P368L and R460Q (Beltcheva et al., 2001) (Table 3.1.1). Here, segregation of the mutation from each parent was confirmed (data not shown). None of our cohort had Finmajor (nt121delCT, L41fsX91) or Finminor (c.3325C>T, R1109X) mutations (Kestila et al., 1998). The affected patient from family S3 had a homozygous frame shift mutation in *NPHS1* combined with a novel heterozygous *NPHS2* (c.761G>T; p.C254F) allele (Table 3.1.1 and 3.1.3). A p.P264R missense change found in *NPHS1* in 2 families (Table 3.1.4). This allele has previously been identified as contributing to a compound heterozygous mutation in *NPHS1* in cases of severe CNS (Koziell et al., 2002, Caridi et al., 2009). It has also been noted in cases of tri-alellic inheritance of steroid resistant FSGS (Lowik et al., 2008). Tri-allelism (sometimes called oligogenicity) refers to the finding of 2 recessively inherited mutations at one loci plus a third mutated loci in a different gene. It has been frequently described in patients with Bardet Biedl Syndrome (BBS) (see Chapter 4). The P264R variant has a minor allele frequency of 1% (rs34982899, dbSNP 1000 Genomes) and its pathogenicity is not proven. Indeed it has recently been classified as a polymorphism (Machuca et al., 2010). One novel

homozygous sequence variants was detected in the *NPHS1* gene (S37 Table 3.1.1 and Table 3.1.2). This sequence variant is predicted to be a splice site mutation affecting the 5'-splice donor site (c.1627+2T>G) predicting exon skipping of exon 12. A novel heterozygous *NPHS1* variant c.528T>C (p.S176S) is in a highly conserved region, and may affect splicing at acceptor site, given that the 528T nucleotide is the second nucleotide within exon 5 (Table 3.1.3). This variant was found in a heterozygous state in two patients from the same family (F22) in combination with a heterozygous missense change in *LAMB2* (c.1982G>A; p.R561Q) (Table 3.1.3).

Another novel heterozygous *NPHS1* variant c.840+6G>A was found in 7 families, 2 of whom also had a homozygous mutation p.Q1020X in *PLCE1* gene (Table 3.1.1 and 3.1.3). This *NPHS1* c.840+6G>A variant was also found in 1.15% of our normal controls, suggesting it is unlikely to be pathogenic in isolation (Table 3.1.3).

Of the six families with homozygous or compound heterozygous *NPHS1* mutations, four presented with CNS, whilst the others had milder phenotypes, consistent with previous reports (Heeringa et al., 2008).

***NPHS2*:** Among our NS cohort there were 11 families with *NPHS2* mutations (Table 3.1.1). All mutations in affected individuals were in a homozygous state consistent with known parental consanguinity. Mutations detected in *NPHS2* accounted for 44 % of all the mutations detected in our cohort.

Two novel homozygous *NPHS2* mutations were detected: a nonsense mutation c.115C>T (p.Q39X) which is the most premature truncating mutation reported so far in *NPHS2* and a missense mutation c.413G>C (p.R138P) in patients with SRNS. Of note, this arginine residue at position 138 (R138) of podocin has been the residue mutated in other cohorts, specifically (R138Q) and (R138X) (Boute et al., 2000). Both mutations R138Q and R138X were reported in SRNS patients. Indeed R138Q has been described as a founder mutation in European populations (Boute et al., 2000). Huber and et al. demonstrated that this mutation causes a failure to recruit nephrin into lipid rafts (Huber et al., 2003), providing important insights into the pathogenesis of this *NPHS2* mutation.

A novel heterozygous *NPHS2* variant C254F was detected in a patient homozygous for a c.515delCCA mutation in *NPHS1* gene (Table 3.1.1 and 3.1.3). The variant C254F

was not predicted to be definitely pathogenic by in silico models (Table 3.1.3) but neither was it found in our normal controls. This suggests that C254F could be a modifier allele.

Whilst screening our control population for novel *NPHS2* variants, we observed a *NPHS2* variant P20L (data not shown) in 2.32% of Saudi Arabian controls. This variant is of unknown functional significance (Caridi et al., 2009) and is rare in control European populations (Ruf et al., 2004a) and was not associated with a NS disease phenotype, suggesting that this change should be considered as a rare polymorphism.

***PLCE1*:** A novel homozygous mutation in *PLCE1* (Q1020X) was detected in 4 families in our cohort, with a broad range of phenotypes (Table 3.1.1). All affected individuals had the identical mutation, and segregation was confirmed in 2 of the families (F5 and F15, data not shown). This truncating mutation is predicted to have a severe effect on the PLCE1 protein with loss of the PLC catalytic domain, the protein kinase C conserved region and the Ras association domain (Hinkes et al., 2006). Interestingly, each of the Q1020X mutations in exon 7 of *PLCE1* gene was associated with the variant (c.4665+52G>C) in intron 18 (Table 3.1.3).

Affected and unaffected daughters in a family of African ancestry with SRNS (F17) were found to be homozygous for a rare variant (K2173R) in *PLCE1* (Table 3.1.4). We did not found this variant in our Saudi Arabian control cohort. The affected F17 member was also heterozygous for novel variants c.840+6G>A in *NPHS1* gene and R644H in *LAMB2* gene (Table 3.1.3).

Gene	Family ID	Nucleotide Change	Amino Acid change	Homozygous or Heterozygous	PolyPhen-2	SIFT	SNPS3D	Allele frequency in Saudi population
NPHS1	F22	c.528T>C	S176S, possible splicing defect	Heterozygous	N/A	N/A	N/A	0
NPHS1	S20^, F9, F17, S18, S26^, S30, S31	c.840+6G>A	Intronic, possible splicing defect	Heterozygous	N/A	N/A	N/A	1.15%
NPHS2	S3#	c.761G>T	C254F	Heterozygous	Benign (0.302)	Tolerated (0.16)	Deleterious (-0.39)	0
LAMB2	S18	c.1393T>C	L465P	Heterozygous	Benign (0.137)	Tolerated (0.35)	Non-deleterious (0.44)	0
LAMB2	F22	c.1982G>A	R561Q	Heterozygous	Probably damaging (0.994)	Tolerated (0.3)	Non-deleterious (0.91)	0.27%
LAMB2	F17	c.1931G>A	R644H	Heterozygous	Probably damaging (1.00)	Tolerated (0.44)	Deleterious (-1.16)	0.82%
LAMB2	S11, S17	c.3772C>G	L1258V	Heterozygous	Possibly damaging (0.539)	Tolerated (0.38)	Non-deleterious (0.18)	0.55%
NEIL1	S9	c.131C>A	A44D	Heterozygous	Probably damaging (0.999)	Tolerated (0.37)	Non-deleterious (0.19)	0
NEIL1	S21	c.466G>A	A156T	Heterozygous	Benign (0.217)	Tolerated (0.33)	Non-deleterious (0.91)	0
PLCE1	F5^, F15^, S20^, S26^	c.4665+52G>C	Intronic	Homozygous	N/A	N/A	N/A	0

Table 3.1.3: Novel genetic variants of unknown significance and polymorphisms detected in *NPHS1*, *NPHS2*, *LAMB2*, *NEIL1* and *PLCE1*.

Gene	*Family ID*	**Nucleotide Change**	**Amino Acid change**	**Homozygous or Heterozygous**	**SNP ID**	**Allele frequency in Saudi population**
NPHS1	S39	c.791C>G	P264R	Homozygous	rs34982899	N/A
NPHS1	S34, F22	c.791C>G	P264R	Heterozygous	rs34982899	N/A
NPHS2	F13	c.709G>C	E237Q	Heterozygous	rs146906190	N/A
LAMB2	S18, S29	c.5293G>A	A1765T	Heterozygous	rs74951356	N/A
LAMB2	S20, S26	c.2099G>A	G700E	Heterozygous	rs142860588	N/A
LAMB2	S39	c.3443G>A	R1148H	Heterozygous	rs138774635	0.82%
PLCE1	F17	c.6518A>G	K2173R	Homozygous	rs111929795	0%
CD2AP	F5, S14, S16	c.902A>T	K301M	Heterozygous	rs141778404	2.79%

Table 3.1.4: Nonpathogenic sequence variants / polymorphisms identified in known nephrotic syndrome genes

***MYO1E*:** We detected novel homozygous *MYO1E* mutations in 3 families (Table 3.1.1). Two families presented with SRNS and minimal change glomerulonephritis on renal biopsy. The affected patients both had a Y47X nonsense mutation, segregating from each parent (data not shown). The affected patient in family S17 presented with SRNS and FSGS on renal biopsy and was found to have a T119I homozygous mutation (Table 3.1.1 and Figure 3.1.3). The Threonine residue at position 119 is conserved throughout vertebrates (including zebrafish) and in the myosin-1d homolog of the amoeba Dictyostelium discoideum. The Y47X is a severe truncating mutation, with a predicted loss of all functional domains of the *MYO1E* protein, including the motor-head domain, the calmodulin binding IQ domain and the tail domain (Mele et al., 2011).

***CD2AP*:** A single patient from family F23 in our NS cohort was found to have 2 known heterozygous variants in *CD2AP* (Table 3.1.1). The first F220L is a novel variant of unknown significance (rs139926926), in a residue conserved to Danio rerio and was not detected in 350 healthy Saudi Arabian alleles. This variant was combined with T374A, a likely pathogenic mutation (Gigante et al., 2009). The heterozygous T374A mutation in *CD2AP* was previously reported in a 2 year old child presenting with SRNS and histological features of FSGS, and the missense change disrupts a proline rich domain important for protein-protein interactions (Gigante et al., 2009).

In addition, the *CD2AP* variant K301M was found as a heterozygous allele in 3 of our patients; one was associated with homozygous Q1020X mutation in *PLCE1* gene (F5), other two (S14 & S16) were detected alone (Table 3.1.4). Although K301M has been reported previously as a pathogenic mutation (Gigante et al., 2009), we also found this variant in 2.79% of our normal controls suggesting that this variant is a polymorphism.

***LAMB2*:** We found novel heterozygous *LAMB2* sequence variants in 5 families from our cohort (Table 3.1.3). The affected patient in family S18 had a L465P heterozygous variant in *LAMB2* gene together with a A1765T polymorphism in *LAMB2* (Schoeb et al., 2010) (Table 3.1.3 & 3.1.4). The L465P missense change is novel, and was absent from control samples, but predicted to be benign using in silico testing (Table 3.1.3). In another family (F22), two affected patients (brothers) with FSGS had a single heterozygous R561Q variant in *LAMB2*. This novel variation of unknown severity was present in 0.27% of healthy controls and was in combination with a novel heterozygous

NPHS1 S176S variant, predicted to have some deleterious impact upon splicing (Table 3.1.3).

Additional novel (and likely polymorphic) variants were detected in the *LAMB2* gene (Table 3.1.3, Figure 3.1.3). A heterozygous L1258V missense variant was detected in two families and also detected in 0.55% of our normal controls. The variant R644H was detected in one allele of our patients and in 0.82% of normal controls.

Family S39 had a known heterozygous variant in *LAMB2* (c.3443G>A; p.R1148H) in association with a heterozygous P264R NPHS1 polymorphism (Table 3.1.4). This R1148H *LAMB2* variant was found also in 0.82% of our normal controls (Table 3.1.4).

3.1.4 Discussion

This study is the first to describe the molecular basis of NS in a Saudi Arabian population. Despite finding a high rate of mutations in known NS genes within our cohort (47%), these findings suggest there are other novel genetic causes of NS yet to be discovered. The genetic heterogeneity underlying NS, even in this highly consanguineous population, is evident. It is noteworthy that all patients with identified mutations in NS associated genes were from consanguineous marriages and most of the pathogenic mutations identified were in the homozygous state. There were two families (S22 and F23) where we identified compound heterozygous mutations in *NPHS1* and *CD2AP*, respectively. Thus, reliance on screening genes in homozygous regions alone in known consanguineous families may miss compound heterozygous changes in relevant genes. Our experience is not unique. Using a homozygosity mapping approach in 12 families from different backgrounds with CNS, Schoeb et al. solved just 5 families by detecting homozygous mutations in *NPHS1* (Schoeb et al., 2010). A more systematic search, including all known NS genes, is therefore important for achieving a high mutation detection rate. A strategy of targeted gene sequencing for patients manifesting NS in the first year of life in a world-wide study has previously been reported, and noted that two-thirds of patients could be explained by mutations in 1 of 4 genes (*NPHS1*, *NPHS2*, *WT1* or *LAMB2*) (Hinkes et al., 2007). In the modern era of whole exome sequencing and targeted gene capture and sequencing, these approaches will allow mutations in known NS genes and novel NS genes to be detected with great efficiency.

In our study, which included families with CNS, infantile NS and childhood SRNS, following molecular analysis of 9 known NS genes, mutations were detected in around 47% of cases, which is comparable to the mutation detection rate in SRNS by other groups (Santin et al., 2011). Cohorts of more restricted phenotypes have identified mutations ~80% of cases (non-Finnish CNS) by a systematic screen of implicated genes (Machuca et al., 2010). Our mutation detection rate was, as one might predict, higher in patients where there was evidence of familial disease (more than 1 affected member) (73%) than in single individuals with NS (37%). Sporadic SRNS is likely to account for a greater proportion of NS, where just one individual is affected and where there is no family history of NS. Out of 36 consanguineous families we were able to detect likely causative mutations in 24 families and no mutations were found in other 12 families.

In five non-consanguineous families, causative mutation was found only in one family. No mutations were detected in remaining 12 families with unknown consanguinity status.

Mutations in *NPHS2* gene are the most frequent identified genetic cause of NS in our Saudi Arabian cohort, accounting for 44% of all pathogenic mutations detected, in comparison to ~40% of a European cohort (Hinkes et al., 2007), and ~30% in a large Turkish study (Berdeli et al., 2007). Common *NPHS2* mutations in our cohort included missense mutations V260E and V180M. All cases with mutations in *NPHS2* presented with childhood SRNS except one family (F3) that had CNS in association with a *NPHS2* R168H mutation. All mutations detected in *NPHS2* were homozygous, consistent with parental consanguinity in these cases.

NPHS1 mutations were detected, as expected, mainly in patients presenting with CNS patients. Three cases presented with childhood SRNS.

In *PLCE1* gene we found one mutation common to four families with NS patients. Each of these families were descendants from large Saudi Arabian tribes. In Saudi Arabia, 8 tribes account for around 10% of the country's population. The association of the Q1020X in *PLCE1* with the intronic variant (c.4665+52G>C) in *PLCE1* in each of these cases may indicate a founder effect (Table 3.1.3).

The original description of *MYO1E* mutations identified just 2 consanguineous families with missense mutations A159P and nonsense mutation Y695X, originating from Italy and Turkey respectively (Mele et al., 2011). There have been no additional reports of *MYO1E* mutations to date. Here, by identifying two novel *MYO1E* mutations in an Arabic consanguineous population, we confirm the pathogenicity of *MYO1E* in NS and expand the spectrum of mutations.

Most of the mutations previously reported in the *CD2AP* gene have been heterozygous changes, except in one case with biopsy proven FSGS where a homozygous mutation was found (Lowik et al., 2008). In our cohort we identified a single patient from a consanguineous family harboring a compound heterozygous mutation (F220L/T374A) in the *CD2AP* gene. In a previously published cohort of 35 families with SRNS in whom *NPHS1*, *NPHS2*, and *PLCE1* mutations had been previously excluded, no *CD2AP* mutations were identified, confirming their rarity (Benoit et al., 2010).

We found one case (S3), presenting as CNS, where a previously reported homozygous mutation (c.515delCCA) in *NPHS1* (Lenkkeri et al., 1999) was associated with a heterozygous missense change (c.761G>T; C254F) in *NPHS2*. This emphasizes the need to screen multiple NS associated genes, even within a consanguineous pedigree, to determine modifier gene effects. The presence of triallelism in NS has been noted before (Koziell et al., 2002, Caridi et al., 2003) (Schultheiss et al., 2004) and additional alleles in NS genes may modify the renal phenotype and the clinical presentation and course. Weber et al. described a child with CNS in whom a combination of a heterozygous de novo splice mutation in *NPHS1* and a homozygous *NPHS2* R138Q mutation was detected (Weber et al., 2004). In contrast, it is noteworthy that in Bardet Biedl syndrome (BBS), an inherited ciliopathy where oligogenicity has frequently been previously reported (Badano et al., 2003, Pereiro et al., 2011), a Saudi cohort failed to exhibit these phenomena (Abu-Safieh et al., 2012).

We were extremely careful not to define all the variants we have detected as pathogenic mutations, although many of them were found in less than 1% of our normal controls. With each new variant, we carefully checked for the presence of additional mutations (in the same and other NS associated genes) and confirmed the variant co-segregated, where parental samples were available, with clinical status within the family. As an example, the P264R variant we observed in *NPHS1* (Table 3.1.4) has previously been implicated as pathogenic in cases of severe CNS (Koziell et al., 2002) however in a more recent paper it has been classified as a polymorphism (Machuca et al., 2010). Often novel (heterozygous) sequence variants were also detected in our ethnically matched control population (Table 3.1.3). We also note that the high rate of consanguineous marriage may make it difficult to estimate pathogenicity of novel mutations without segregation and extended family analysis.

Carrier frequency of NS in Saudi population could be established using Hardy–Weinberg equilibrium (HWE) providing the prevalence of NS in the population is known. Determination of allele frequency and heterozygous carrier frequency in a population for which the prevalence of disease is known is the most important medical application of the HWE. The HWE holds only for large population in which there is random mating. Consanguinity in Saudi population is around 55% and this is may alter allele frequency in the population by increasing proportion of homozygous in the next generation.

Based on data obtained, allele frequency of NS in Saudi population is a challenge due to consanguinity and unknown prevalence of the disease in the population.

The molecular genetic diagnosis of NS remains a vital aid to the clinical management of families with NS. It allows for the appropriate long term management to be undertaken, genetic counseling to be undertaken and where necessary, screening of siblings and other at risk family members. Decisions regarding immunosuppression and transplantation are aided by a molecular genetic diagnosis. It is well recognized that children with causative mutations in NS genes do not respond well to treatment with corticosteroids and other immunosuppressants (Hinkes et al., 2007).

In conclusion, in a Saudi Arabian cohort, 47% of families with NS were explained by mutations on 5 known nephrotic syndrome genes *(NPHS1, NPHS2, PLCE1, MYO1E* and *CD2AP*). *NPHS2* gene mutations were the most common molecular genetic cause of NS in this cohort. Unsolved patients from consanguineous families suggest additional novel genetic causes of NS are likely. The genetic heterogeneity of NS, suggests that screening strategies should continue to include multiple NS genes, including rare and recently discovered genetic causes, to allow a high yield of molecular genetic diagnoses. This will then lead to improvements in both precise diagnosis and clinical management.

Chapter 3: RESULTS

3.2 Bardet-Biedle Syndrome and Modelling of a Novel *BBS5* Mutation

3.2.1 Introduction

Bardet–Biedl syndrome (BBS; OMIM 209900) is a ciliopathic disorder that affects multiple organs and body systems. It is characterized by polycystic kidney, hypercholesterolemia, hypogonadism, mental retardation, obesity, polydactyly, renal failure, retinitis pigmentosa, rod-con dystrophy, and syndactyly (Moore et al., 2005). Beales and et al proposed a scheme for diagnosing BBS (Beales et al., 1999); They suggested four primary features or three primary features and two secondary features are required for a clinical diagnosis of Bardet–Biedl syndrome as shown in Table 3.2.1.

BBS is inherited in an autosomal recessive fashion. To date; at least 16 BBS genes have been identified (*BBS1, BBS2, ARL6, BBS4, BBS5, MKKS, BBS7, TTC8, BBS9, BBS10, TRIM32, BBS12, MKS1, CEP290, WDPCP* and *SDCCAG8*). The frequency, function and locus of each gene are shown in Table 3.2.2.

Primary features	**Secondary features**
Four features are required to be present:	*Or three primary plus two secondary features are required:*
Rod-cone dystrophy	Speech disorder/delay
Polydactyly	Strabismus/cataracts/astigmatism
Obesity	Brachydactyly/syndactyly
Learning disabilities	Developmental delay
Hypogonadism in males	Polyuria/polydipsia (nephrogenic diabetes insipidus)
Renal anomalies	Ataxia/poor coordination/imbalance
	Mild spasticity (especially lower limbs)
	Diabetes mellitus
	Dental crowding/ hypodontia/small roots/high arched palate
	Left ventricular hypertrophy/congenital heart disease
	Hepatic fibrosis

Table 3.2.1: Diagnostics criteria of BBS:

Gene	***Frequency***	***Locus***	***Function***	***Refernces***
BBS1	23%	11q13	BBSome protein	(Mykytyn et al., 2002)
BBS2	8%	16q21	BBSome protein	(Nishimura et al., 2001)
BBS3/ARL6	0.40%	3p12-p13	GTPase	(Chiang et al., 2004)
BBS4	2%	15q22.3-q23	BBSome protein	*(Mykytyn et al., 2002)*
BBS5	0.40%	2q31	BBSome protein	(Li et al., 2004)
BBS6/MKKS	6%	20p12	Part of chaperonin complex	(Kim et al., 2005)
BBS7	2%	4q27	BBSome protein	(Badano et al., 2003)
BBS8/TTC8	1%	14q32.1	BBSome protein	(Ansley et al., 2003)
BBS9/B1	6%	7p14	BBSome protein	(Nishimura et al., 2005)
BBS10	20%	12q21.2	Part of chaperonin complex	(Stoetzel et al., 2006)
BBS11/TRIM32	0.10%	9q31-q34.1	E3 ubiquitin ligase	(Chiang et al., 2006)
BBS12	5%	4q27	Part of chaperonin complex	(Stoetzel et al., 2007)
BBS13/ MKS1	4.50%	17q23	Centriole migration	(Leitch et al., 2008)

BBS14/CEP290 /NPHP6	1%	12q21.3	Basal body: RPGR interaction	(Leitch et al., 2008)
BBS15/ WDPCP	1%	2p15	Basal body: localisation of septins and ciliogenesis	(Kim et al., 2010)
BBS16/SDCCA G8	1%	1q43	Basal body: interacts with OFD1	(Otto et al., 2010)

Table 3.2.2: The frequency, function and locus of BBS genes

Although BBS inherited in autosomal recessive pattern, there is evidence of non Mendelian mode of inheritance named oligogenic inheritance or triallelic inheritance (Katsanis et al., 2001). Oligogenic inheritance is a term that describes 2 recessively inherited mutations at one loci plus a third mutated loci in a different gene. Trialleleic inheritance has been observed. This phenomenon was observed in two ciliopathy diseases, BBS (Katsanis et al., 2001) and nephronophthisis (Hoefele et al., 2007). However, in a number of families, *trans* mutant alleles interact with the primary causal locus to modulate the penetrance and/or the expressivity of the phenotype.

The proteins encoded by these genes are involved in the maintenance and function of cilia. Identification of *BBS8* led to linking BBS to ciliary dysfunction where situs inversus (a defect of left-right axis determination) was diagnosed. It was found that all BBS proteins studied to date localize primarily to centrosomes, basal bodies, or cilia (Zaghloul and Katsanis, 2009). BBS proteins are required for the localization of G protein-coupled receptors to primary cilia on central neurons. Recent studies suggest that the BBS complex can associate with the RAB8 GDP/GTP exchange factor to promote trafficking of vesicles to the cilium (Nachury et al., 2007). In addition studies using C. elegans have shown that *bbs8* may be involved in the regulation of the RFX transcription factor, which regulates expression of ciliary protein–encoding genes (Ansley et al., 2003). These findings suggest an important role for BBS proteins in

ciliary function, and these proteins have been associated with ciliogenesis and intraflagellar transport (IFT), the motor-dependent trafficking of cargo along the cilium (Zaghloul and Katsanis, 2009).

In Europe and North America, BBS has a prevalence of 1 in 140,000 to 1 in 160,000 newborns (Beales et al., 1997)

The prevalence of BBS in Arabs varies from one country to another. It is estimated to be 1 in 13,500 newborn among Bedouins of Kuwait (Farag and Teebi, 1989), 1 in 30,000 in Oman (Rajab et al., 2005) and 1 in 156,000 individuals in Tunisia (M'Hamdi et al., 2011) . The incidence of BBS is estimated by the Centre for Arab Genomic Studies to be 6-10 per 100 000 live birth in Arab population (CTGA Database). Recently there have been a few reports describing clinical and molecular BBS in Saudi Arabia (Abu Safieh et al., 2010, Cherian and Al-Sanna'a, 2009).

The *BBS5* gene was identified and mapped by Li (Li et al., 2004) to chromosome 2q31. Nachury et al found that BBS5 is 1 of 7 BBS proteins (BBSome) that form the stable core of a protein complex required for ciliogenesis (Nachury et al., 2007). It was suggested that BBS5 mediated binding to phospholipids, predominantly phosphatidylinositol 3-phosphate.

Pathogenic mutations in *BBS5* gene were identified in several patients with BBS. There are 17 reported mutations to date (Table 3.2.3) according to HGMD (http://www.hgmd.org). These mutations distributed among many ethnic groups.

Nucleotide change	Amino acid change	Mutation type	Phenotype	Reference
c.1A>T	p.M1L	Missense	Bardet-Biedl syndrome	Muller (2010) Hum Genet 127, 583
c.2T>A	p.M1K	Missense	Bardet-Biedl syndrome	Harville (2010) J Med Genet 47, 262
c.149T>G	p.L50R	Missense	Bardet-Biedl syndrome	Chen (2011) Invest Ophthalmol Vis Sci 52, 5317
c.158C>T	p.T53I	Missense	Bardet-Biedl syndrome	Abu-Safieh (2012) Eur J Hum Genet 20, 420
c.166A>G	p.R56G	Missense	Bardet-Biedl syndrome	Muller (2010) Hum Genet 127, 583
c.214G>A	p.G72S	Missense	Bardet-Biedl syndrome	Hjortshoj (2008) Am J Med Genet A 146A, 517
c.547A>G	p.T183A	Missense	Bardet-Biedl syndrome	Hjortshoj (2008) Am J Med Genet A 146A, 517
c.551A>G	p.N184S	Missense	Phenotype modifier	Li (2004) Cell 117, 541 Zaghloul (2010) Proc Natl Acad Sci U S A 107: 10602 [Functional characterisation]

				Song (2011) Invest Ophthalmol Vis Sci 52: 9053 [Additional phenotype]
c.620G>A	p.R207H	Missense	Phenotype modifier	Li (2004) Cell 117, 541 Zaghloul (2010) Proc Natl Acad Sci U S A 107: 10602 [Functional characterisation]
c.889G>A	p.D297N	Missense	Bardet-Biedl syndrome	Feuillan (2011) J Clin Endocrinol Metab 96, E528
c.522+3A>G	IVS6 ds A-G +3	Splice site	Bardet-Biedl syndrome	Li (2004) Cell 117, 541
c.619-1G>C	IVS7 as G-C -1	Splice site	Bardet-Biedl syndrome	Feuillan (2011) J Clin Endocrinol Metab 96, E528
c.123delA		Deletion		Smaoui (2006) Invest Ophthalmol Vis Sci 47, 3487
c.54dupC		Insertion	Bardet-Biedl syndrome	Feuillan (2011) J Clin Endocrinol Metab 96, E528
263_271delTACGAGGCCinsGCTCTTA		Insertion/Deletion	Bardet-Biedl syndrome	Li (2004) Cell 117, 541

Table 3.2.3: HGMD reported mutations of *BBS5* gene.

Many investigators have shown that the zebrafish provides experimentally tractable and physiologically relevant models of important aspects of ciliary dysfunction(Zaghloul et al., 2010, Badano et al., 2006, Chiang et al., 2006, Leitch et al., 2008, Ross et al., 2005). It is proven that human mRNA for ciliopathy genes may be utilized to rescue both morphant and mutant zebrafish phenotypes efficiently, providing a robust platform for interpretation of the pathological relevance of identified mutated alleles (Zaghloul et al., 2010, Chiang et al., 2006, Leitch et al., 2008, Parfitt et al., 2009). Rescue experiments in zebrafish for the *AHI1* ciliopathy gene was demonstrated recently by our group(Simms et al., 2012).

Suppression of some BBS proteins in zebrafish causes gastrulation defects that include shortened and curvature body axes, longer somites, broad and kinked notochords, cardiac abnormalities (including *situs inversus*) and pronephric duct dilatation (Stoetzel et al., 2006, Stoetzel et al., 2007, Leitch et al., 2008, Ross et al., 2005, Gerdes et al., 2007).

Here, we report a consanguineous Saudi Arabian family with clinical features of BBS, together with a novel mutation in *BBS5* gene. For first time, we have developed a zebrafish model of *BBS5* knockdown and mimicked the human mutation in this organ system. Using cell culture systems we also verify the pathogencity of the mutation by demonstrating mislocalisation of mutated BBS5 protein in renal epithelial cells.

3.2.2 Results

3.2.2.1 Clinical Phenotype:

A first degree consanguineous family from Saudi Arabia was given a diagnosis of BBS. The family consisted of 3 affected children, unaffected child and parents (Figure 3.2.1).

The affected children all had post axial polydactly of the hands and feet (primary feature 1). Clinical examination of their eyes suggested a mild eye phenotype, with some decreased visual acuity with occasional rotator eye movements but with preserved retina. The affected children were all of short stature and obese (primary feature 2) with hyperphagia. They had variable degrees of developmental delay with learning disabilities (primary feature 3) and speech delay (secondary feature). They have variable degree of kidney phenotype (primary feature 4).

They all had subclinical hypothyroidism. The affected male and one of the affected females had a small Ventricular Septal defect (VSD) (secondary feature) (Table 3.2.4). Thus, each of the affected individuals had enough primary and secondary features to establish a clinical diagnosis of BBS.

	Gender	Age	polydactyly	obesity	cognitive impairment	Renal insufficiency	retinal dystrophy	CHD	Others
Affected-1	Female	9	+	+	+	-	-	-	Hypothyroidism / squint / decreased visual acuity
Affected-2	Female	7	+	+	+	+	-	+	Hypothyroidism
Affected-3	Male	3	+	+	-	+	-	+	Hypothyroidism

Table 3.2.4: Clinical characteristics of the patients of BBS family

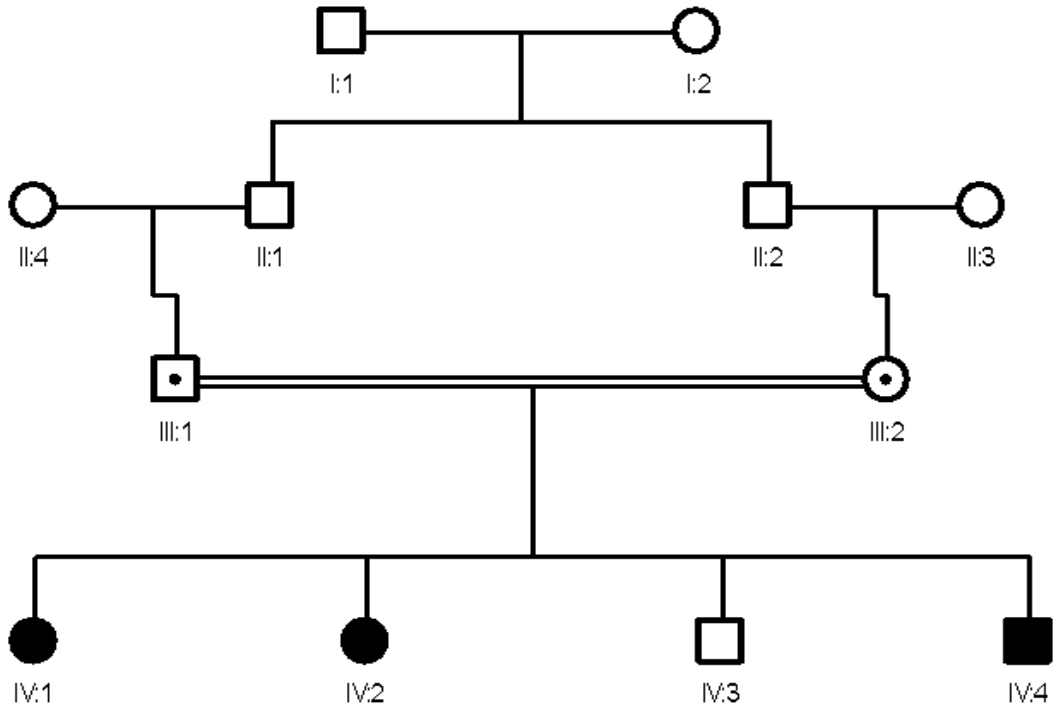

Figure 3.2.1: Family pedigree of the family with *BBS5* mutation. Squares indicate male family members, circles female family members; solid symbols indicate affected persons, circles and squares with dots indicate likely carriers, and double horizontal bars consanguinity. Roman numbers underneath circles and squares indicate the sibling number.

3.2.2.2 Genome wide SNP Genotyping:

To search for homozygous regions, all members of the BBS family were genotyped with an Affymetrix GeneChip probe array containing around 238000 biallelic SNPs. The SNP genotype call rate was ≥ 96% (Data not shown).

All three affected children shared regions of homozygosity in the genome. The chromosome 2 region of homozygosity contained the *BBS5* gene as shown in Figure 3.2.2.

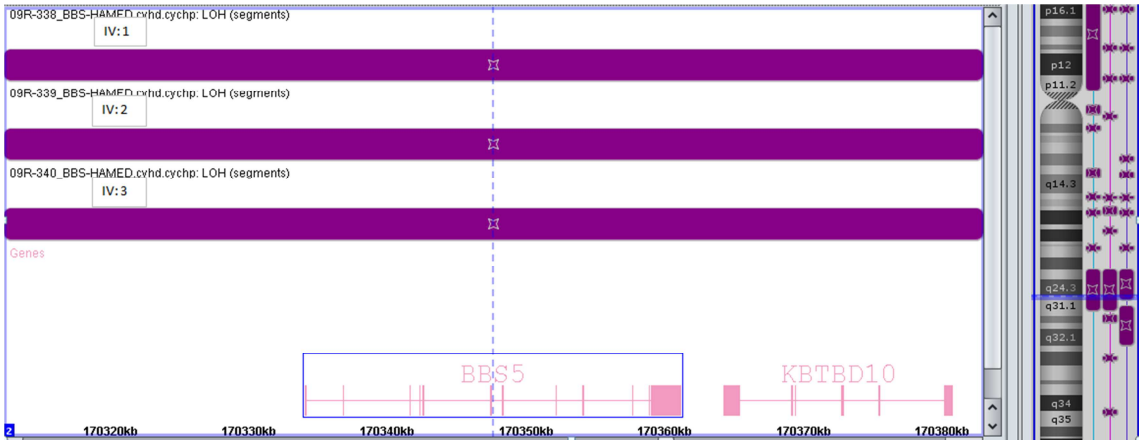

Figure 3.2.2: A screen grab of ChAS2 software for chromosome 2 showing homozygosity in *BBS5* gene in all three affected with BBS.

3.2.2.3 BBS5 gene analysis and Mutation Screening:

As candidate gene, *BBS5* was screened by direct sequencing of the entire coding sequence and consensus splice sites of the gene. Oligonucleotide primers are listed in the appendix. In exon 12 of the *BBS5* gene we found a novel mutation c.966insT (p.A323CfsX55) predicted to cause a frameshift that cause truncation 55 amino acids downstream (Figure 3.2.3A). The mutation resulted in a predicted elongated peptide that has 379 amino acids instead of 342 in the wild type (Figure 3.2.3B). The three affected siblings were (as expected) homozygous for the mutation, whilst both parents and unaffected son were heterozygous for the mutation. The mutation was not found following screening of 96 DNA samples obtained from normal controls of an ethnically matched population. Mutation Taster software predicted a disease causing mutation with a high score for this alteration.

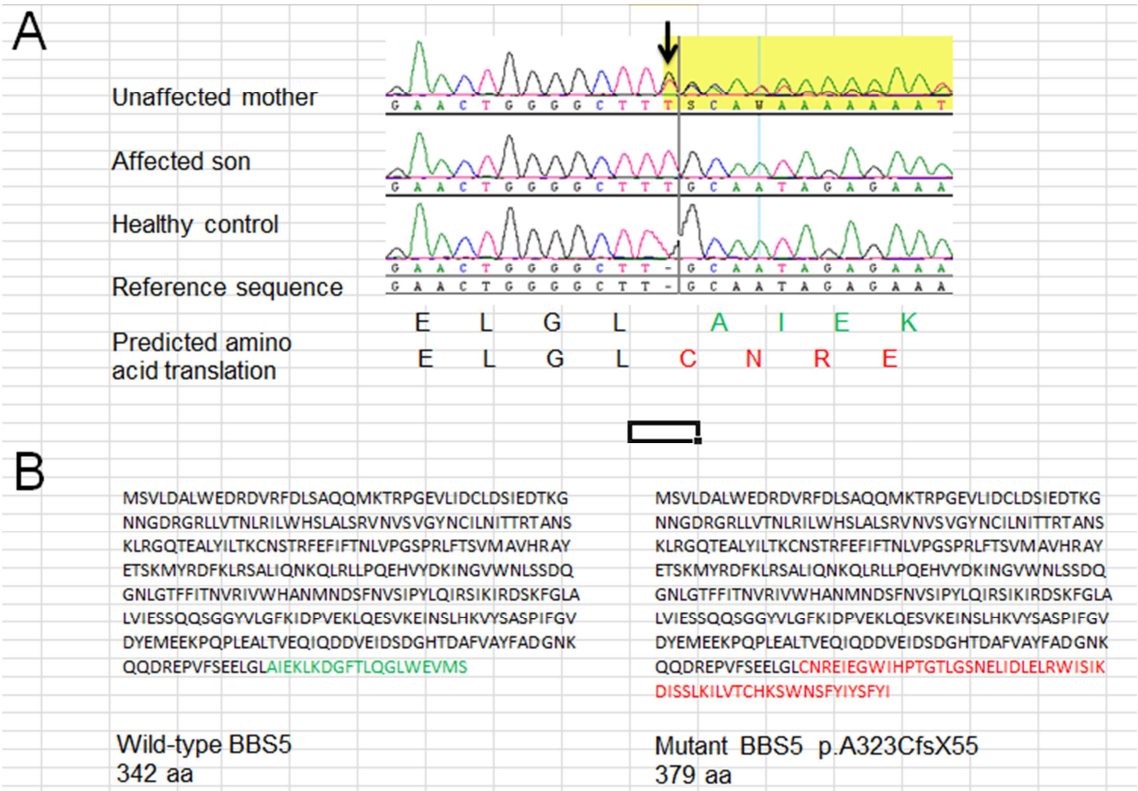

Figure 3.2.3: (A): Sequencing chromatogram of exon 12 of *BBS5* gene showing insertion of T nucleotide in both alleles of affected son and in one allele of the mother. (B): Showing differences between wild type BBS5 protein (green) and mutant BBS5 (red).

3.2.2.4 Knockdown of BBS5 in Zebrafish:

In order to establish a developmental model for *BBS5*, we explored genetic knockdown in an animal model. Using zebrafish (ZF) embryos, morpholino antisense oligonucleotides (MO) were designed to modify gene expression of *BBS5* (see materials and methods chapter, section 2.9.2). In general, MOs were used to block translation of a targeted protein or to target a predicted intron-exon boundary, thus modifying pre-mRNA splicing.

Using microinjection techniques in single cell stage embryos, 6ng of the morpholino oligo (MO) targeted to the translational initiation site of *BBS5* gene was injected into zebrafish embryos. This amount of morpholino injected was chosen after performing titration experiments, because high dose of morpholino (12ng) lead to death of all embryos and low dose (3ng) showed no effect (data not shown). Six hours post morpholino injection; embryos were inspected under a light microscope to remove death embryos (as a results of poor egg quality). 10 hours embryos were counted and

dead embryos were removed. By 48 hour post fertilization (hpf), morphant embryos displayed a typical ciliopathy phenotype with body axis curvature, cardiac abnormalities (including *situs inversus*) and pronephric duct dilatation as shown in Figure 3.2.4. Phenotypes were classified as mild, moderate and severe using criteria (described in the materials and methods chapter; section 2.9.4) based on morphology and body axis changes seen at set time point during development (Leitch et al., 2008, Zaghloul et al., 2010, Huang et al., 2011).

Injection of *BBS5* MO (ATG blocking) alone showed phenotype in approximately 20% of fish (n=194). The main phenotypes seen were severe and moderate boxy axis curvature and cardiac anomalies (Figure 3.2.4).

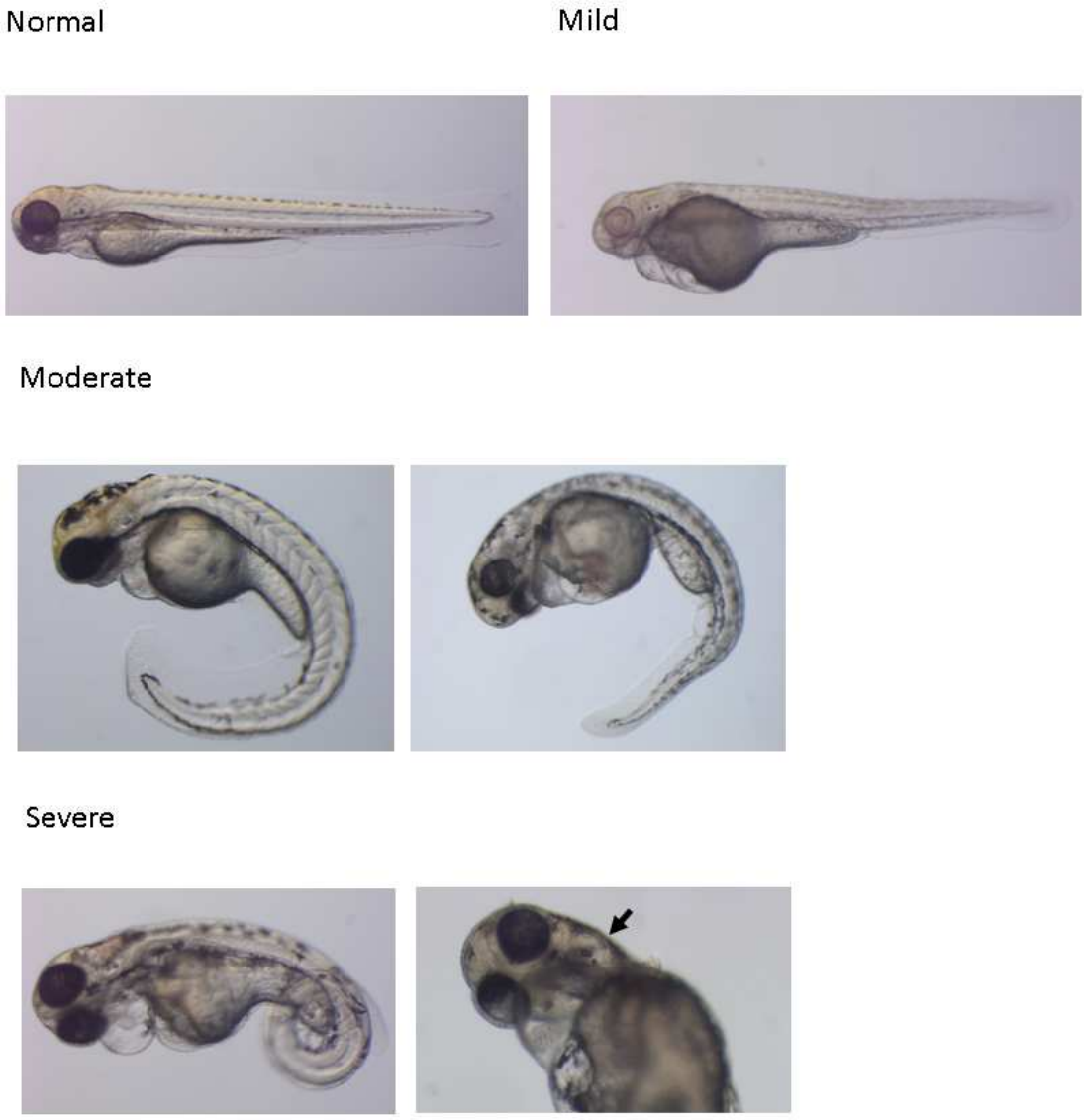

Figure 3.2.4: Zebrafish displayed a ciliopathy phenotype at 72 hpf following injection with BBS5 morpholino. Arrow shows 3 ears instead of 2 in the zebrafish as an indication of BBS phenotype in zebrafish.

In order to mimic the novel human *BBS5* mutation co-injection of MO with mutated *BBS5* mRNA (c.966insT) (see Figure 2.4 B) showed dysmorphic phenotype in 33% of fish embryos (n=78) while the co-injection of MO with wild type *BBS5* mRNA to rescue the phenotype, showed the phenotype in approximately 5% of fish (n=103) (Figure 3.2.5). These results confirm the pathogenicity of the (c.966insT) *BBS5* mutation that led to BBS phenotype in the patients. Using the GraphPad Prism software, Chi-square analysis showed significant differences in the proportion of fish exhibiting phenotypes between groups.

Chi-square analysis was used to compare the effects of treatment upon phenotype of the fish as shown in Table (3.2.5).

In order to compare the groups; three contingency tables were designed. In the first table a comparison of the treatment of fish with BBS5 ATG MO alone and MO + wild type rescue mRNA showed a significant difference in the phenotype of fish ($p = 0.009$). In the second table, a comparison of the treatment of fish with MO alone and MO + mutant mRNA was performed. There is no significant difference between these two categories ($p = 0.2283$), indicating pathogenicity of the mutant mRNA. The third table compares the treatment of fish with MO + wild type rescue mRNA and MO + mutant mRNA. This group showed significance difference between these two categories ($p = 0.0002$) again implicating that the mutant BBS5 mRNA produced a disease phenotype in this model system.

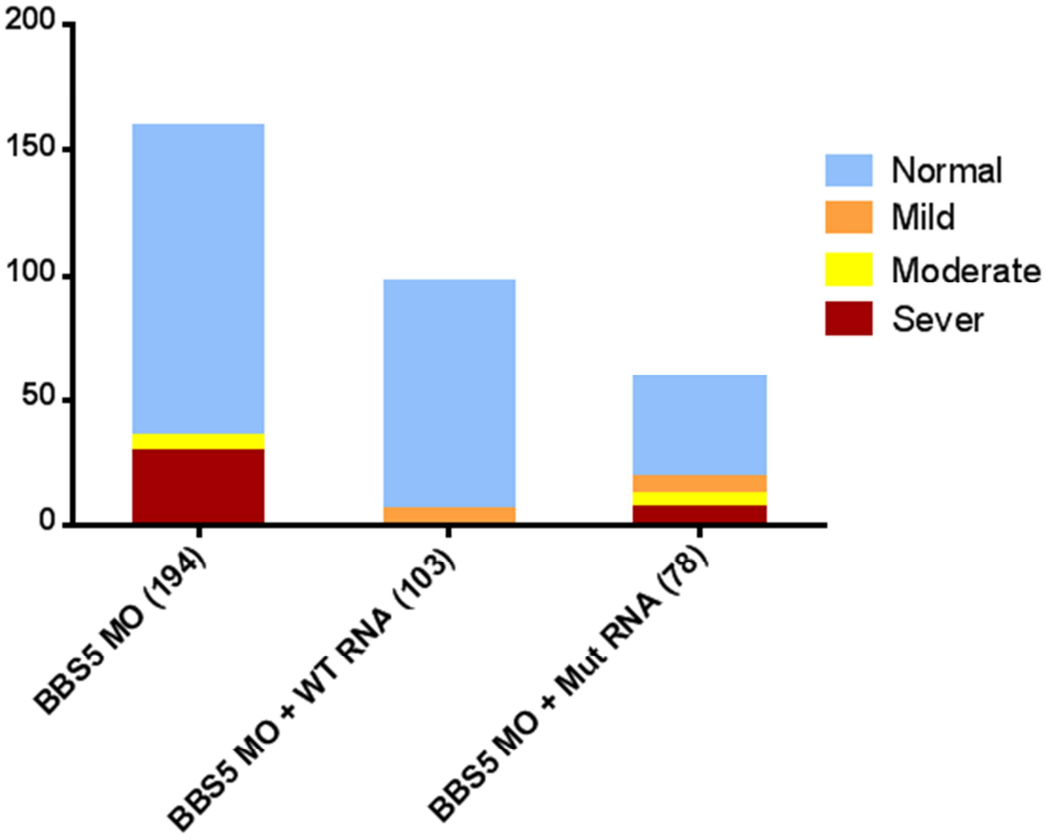

Figure 3.2.5: Analysis of morphant phenotypes and rescue with co-injection with wild-type or mutant BBS5 RNA.

Treatment	Wild Type Phenotype (n)	Affected Phenotype (n)	P value	Significance
BBS5 MO vs. BBS5 MO + WT rescue mRNA	82 94	18 6	0.009	Yes
BBS5 MO vs. BBS5 MO + MUT mRNA	82 75	18 25	0.2283	No
BBS5 MO + WT rescue mRNA vs. BBS5 MO + MUT mRNA	94 75	6 25	0.0002	Yes

Table 3.2.5: Chi-square analysis of comparing the effects of treatment upon phenotype of the fish

3.2.2.5 Transfection of HEK293 with BBS5:

Wild type and mutant BBS5 were transfected in HEK293 to assess the pathogenicity of the mutation detected (c.966insT) at cellular level (see Figure 2.4 D & E) . It is reported that BBS proteins localize to the basal body of ciliated cells (Fan et al., 2004) and BBS is caused by defect at the basal body of ciliated cells (Ansley et al., 2003).

To detect the localization of BBS5 protein and to determine the effect of mutation, we transfected a construct encoding a green fluorescent protein(GFP)-tagged version of the protein into HEK293 cell line. Localization of GFP-tagged WT BBS5 protein (green) in HEK293 cells shows presence at the basal body, as indicated by colocalization with acetylated tubulin and pericentrin in red as shown in Figure 3.2.6. Transfection of the mutant *BBS5* mRNA (containing the c.966insT mutation) in *BBS5* had an observable effect on the localisation of *BBS5* (Figure 3.2.6). The cellualar localisation of the mutated BBS5 appeared to be more intracellular and difuse and the colocalation of BBS5 with pericentrin was lost (Figure 3.2.6).

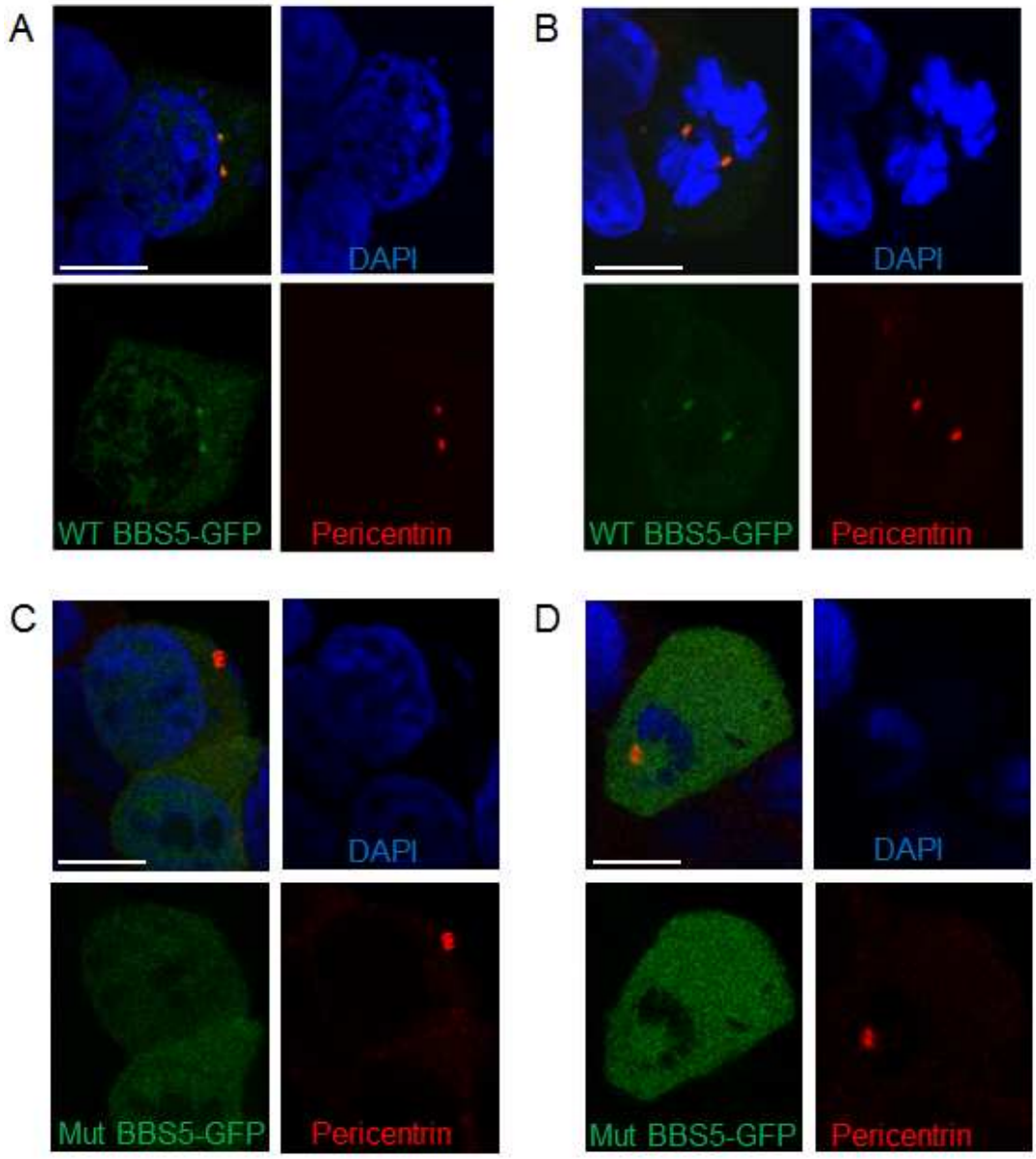

Figure 3.2.6: Comparison of wild type and mutant BBS5 localization in HEK293 cells. (A) Wild type (WT) BBS5-GFP colocalizes with pericentrin at centrosomes. (B) WT BBS5-GFP colocalizes with pericentrin at the centrosomes in a cell undergoing mitosis. (C & D) Mutant (Mut) BBS5-GFP is distributed diffusely throughout the cell and fails to colocalize with the centrosomal marker pericentrin.

3.2.3 Discussion

The identification of mutations causing BBS is challenging because of genetic heterogenity of BBS. There are mutliple genes implicated and evidence of oligogencitiy is strong. The oligogencity in reported with BBS however was not observed in a cohort of BBS pateints from Saudi Arabia (Abu-Safieh et al., 2012).

Homozygosity mapping remains a robust approach that is highly suited for genetically heterogeneous autosomal recessive disorders in populations in which consanguinity is prevalent. Using genchip platforms of single nucleotide polymorphisms (SNP) helped in advancing homozygosity mapping and made it easier by saving time and resources. By homozygosity mapping we targetted a region containing *BBS5,* rather than screening all 16 genes causing BBS. Phenotype-genotype correlation is very poor in BBS and this is evident in our case. In the original *BBS5* gene discovery paper (Nachury et al., 2007) there was no clinical phenotype properly reported. Although this study identified a mutation in *BBS5*, the phenotype resemble reported phenotypes in others from same population who had mutations in *BBS1*, *BBS3*, and *BBS4* (Abu Safieh et al., 2010). This is expected because all BBS proteins localized primarily to the basal body of mammalian cells. BBS proteins regulate proteasome function in addition to the transport of inversin from the basal body into the cytoplasm(Zaghloul and Katsanis, 2009).

All three of the affected patients had hypothyroidism and it is not clear if this is related directly to the *BBS5* mutation. Congenital hypothyroidism is a condition that affects infants from birth and results from a partial or complete loss of thyroid function. Most cases of congenital hypothyroidism are sporadic. It is estimated that 15 to 20 percent of cases are inherited in an autosomal recessive fashion. Mutations in the *DUOX2*, *SLC5A5*, *TG*, *TPO*, and *TSHB* genes prevent or reduce the production of thyroid hormones, even though the thyroid gland is present (http://ghr.nlm.nih.gov). Performing homozygosity mapping by decent in all three affected patients showed no homozygosity in the region of these genes, which may indicate other causes of hypothyroidism in this family. It would be worth sequencing all congenital hypothyroidism genes to rule out a compound heterozygous mutation in the family.

BBS5 mutations are a minor contributor to BBS as only 2% of families from various ethnic backgrounds harbor BBS5 mutations (Li et al., 2004). Of 17 reported pathogenic

mutations in *BBS5* gene, four of them reported in Arabs (Hjortshoj et al., 2008, Chen et al., 2011). The type of mutation identified (c.966insT) is unique because this frameshift resulted in peptide elongation and such mutation never been reported in BBS genes to date. Unfortunately, due to only genomic DNA being available for this family, our options to investigate this elongated BBS5 protein by using mRNA and Western blot studies are limited.

A molecular diagnostic test covering all 16 known BBS genes is important to confirm clinical diagnosis although it will not affect treatment modalities. The predictive power of each genotype will remain limited until the genetic basis of the phenotypic variability is understood (Zaghloul and Katsanis, 2009). At present, molecular diagnosis using whole exome sequencing or ciliopathy genes array is amenable at low cost rate in comparison to Sanger sequencing of all 16 BBS genes. These technologies offer extensive information that may complicate diagnosis of the disease at time being.

Although this is not the first time of modeling *BBS5* mutation in zebrafish, others (Yen et al., 2006) have not described the phenotype of morphant in detail. By using zebrafish to investigate the function of the mutated *BBS5* gene in BBS syndrome, we have helped to characterize the function of this protein in embryogenesis. The advantages of using zebrafish for modeling human disease include the use of a vertebrate that shares many developmental processes and anatomical features with humans and the assaying of a physiologically relevant phenotype of known significance to the syndrome in humans (Zaghloul et al., 2010). In order to determine whether the *BBS5* mutation had a dominant negative affect, further experiments would be required. It would be interesting, for example, to see the effect of injecting mutant *BBS5* mRNA alone, to see whether this overrides (in a dominant negative manner) the wild type *BBS5* mRNA. It would also be interesting to determine the expression pattern of *BBS5* in zebrafish embryos to try and explain the multi system effects that were seen in the *BBS5* MO injected fish.

At cellular level genes that are mutated in ciliopathies, can affect ciliary signalling in different ways: through changes in cilia structure; in targeting signaling molecules appropriately; or at the level of the sensory and/or signaling molecules (Bettencourt-Dias et al., 2011) . It is known that *BBS5* localized to basal bodies just beneath the cilia (Li et al., 2004). The BBSome, associates with the GDP/GTP exchange factor Rab8,

which allows for vesicle trafficking to the base of the cilium (Nachury et al., 2007). We were limited in not having cell line of affected individuals to use BBS5 antibodies to help in localizing wild type and mutant BBS5 proteins in the cell by immunohistochemistry.

Mislocalization of mutated BBS5 protein was confirmed by transfecting HEK293 cell line. Mislocalization also observed in the the dominant-negative mutation L327P in *BBS4* gene in post mitotic IMCD3 cells (Zaghloul et al., 2010). There are no known ciliary or centrsomal targetting motifs within BBS5, however our data suggests that such a motif may be within the C-terminus of the protein beyond amino acid 322. Further experiments mutating each residue in turn (322-341 amino acid residues of BBS5) at the C-terminus would help identify such a motif. An alternative strategy would be to engineer a series of 3'-terminal tagged mutants, with a varying severity of truncation to identify the precise region of any putative localisation signal.

In conclusion we describe a family with clinical features of BBS, together with a novel mutation in BBS5. Modeling this disease in zebrafish mimics the human disease, and pathogencity of the novel *BBS5* mutation is demonstrated by mislocalisation in renal epithelial cells.

Chapter 3: RESULTS

3.3 A New Syndrome of Autosomal Recessive Distal Renal Acidosis associated with Nephrocalcinosis

3.3.1 Introduction

The kidney maintains systemic acid-base homeostasis through proximal tubular reabsorption of filtered bicarbonate (HCO3−), and excretion of the protons (H+) by the alpha intercalated cells of the collecting duct. Impairment of either process produces a phenotype known as renal tubular acidosis (RTA)(Alper, 2010). RTA is a clinical syndrome secondary to this accumulation of acid in the body. A clue to diagnosis is a failure of the kidneys to appropriately acidify the urine. There are four types of RTA, and these may have different clinical syndromes and different inherited and acquired causes (Katzir et al., 2008). These are outlined briefly below.

Type 1 distal RTA (dRTA) is characterized by a failure of acid secretion by the alpha intercalated cells of the cortical collecting duct of the distal nephron.

Type 2 proximal RTA (pRTA) is caused by a failure of the proximal tubular cells to reabsorb filtered bicarbonate from the urine, leading to urinary bicarbonate wasting and subsequent acidemia.

Type 3 RTA is a combined proximal and distal RTA in which a reduction in proximal tubular reclamation of filtered bicarbonate is combined with a disturbance in distal tubular mechanism of maximally acidifying the urine.

Type 4 RTA is associated with a mild (normal anion gap) metabolic acidosis due to either to a deficiency of aldosterone, or to a tubular resistance to its effects. Table 3.3.1 summarizes the clinical features of type 1, 2, and 4 (type 3 is usually excluded from modern classifications).

Type	**Location**	**Metabolic Acidosis**	**Potassium disturbance**	**Nephrocalcinosis**	**Pathophysiology**
Type 1	Distal tubule	Yes (severe)	Hypokalemia	Common	Failure of H+ secretion by the α intercalated cells
Type 2	Proximal tubule	yes	Hypokalemia	uncommon	Failed HCO_3^- reabsorption from the urine by the proximal tubular cells
Type 4	Adrenal	Mild when present	Hyperkalemia	uncommon	Deficiency of aldosterone, or a resistance to its effects, i.e. hypoaldosteronism/pseudohypoaldosteronism (commonly found in diabetes mellitus patients)

Table 3.3.1: Renal Tubular Acidosis types and clinical features

Causes of RTA generally can be classified as inherited or acquired RTA. Acquired forms of RTA can develop quite commonly and may be secondary to autoimmune diseases (e.g. Sjogren's syndrome, systemic lupus erythematosus (SLE), and thyroiditis), and drugs or toxins (e.g. amphotericin B, toluene inhalation).

Inherited causes of both proximal and distal RTA have been reported. Proximal RTA is often seemed with disorders that result in a generalized proximal tubular dysfunction, such as cystinosis. dRTA is rarer and there are three main types: (a) autosomal dominant dRTA (the commonest form), (b) autosomal recessive dRTA with sensorineural deafness, (c) and autosomal recessive dRTA without or with mild sensorineural deafness (Laing et al., 2005). Genetic causes of RTA include the following genes: *SCL4A1*(Devonald et al., 2003), *ATP6V1B1*(Karet et al., 1999b), *ATP6V0A4* (Karet et al., 1999a), *SLC4A4* (Igarashi et al., 1999), and *CA2* (Sly et al., 1983). Mutations in these genes may cause a spectrum of proximal and distal RTA. Table 3.3.2 represents a summary of the inherited renal tubular acidoses.

Type of RTA	Inheritance	Age at Presentation	Gene(s)	Protein
Distal (Type 1)	Dominant	Older/adult	*SCL4A1*	AE1
-	Recessive	Childhood	*SCL4A1*	AE1
-	Recessive with early onset hearing loss	Infancy/childhood	*ATP6V1B1*	B1 subunit of H+-ATPase
-	Recessive with later onset hearing loss	Infancy/childhood	*ATP6V0A4*	a4 subunit of H+-ATPase
Proximal (Type 2)	Recessive with ocular abnormalities	Infancy	*SLC4A4*	NBC1
Combined proximal and distal (Type 3)	Recessive with osteopetrosis	Infancy/childhood	*CA2*	CA II

Table 3.3.2: Summary of the Inherited Renal Tubular Acidoses

Both dominant and recessive dRTA commonly result in bone diseases and nephrocalcinosis (NC).

Nephrocalcinosis (deposition of calcium in the kidney) is an aetiologically heterogeneous disorder associated with an increase in the calcium content of the kidneys with underlying complex pathology. It is often seen in premature infants as a complication of various renal disorders or metabolic disturbances or as a consequence of loop diuretics. Nephrocalcinosis is related to, but not the same as, kidney stones (nephrolithiasis). In general any disorder that leads to high levels of calcium in the blood or urine may lead to nephrocalcinosis. Inherited conditions that can cause nephrocalcinosis include: Bartter's syndrome (Cumming and Ohlsson, 1984), other genetic causes of hypercalciria including mutations of *CLDN16* (Simon et al., 1999) and *CLDN19* (Konrad et al., 2006) genes, familial hypomagnesemia (Praga et al., 1995), primary hyperoxaluria (Morris et al., 1982), and inherited renal tubular acidosis.

A new autosomal recessive syndrome with the constellation of distal renal tubular acidosis, small kidneys, nephrocalcinosis, neurobehavioral impairment, short stature, and distinctive facial features was recently described in a first degree cousin marriage in Saudi family(Faqeih et al., 2007). It was suggested that the syndrome is likely to be of autosomal recessive inheritance based on similarly affected siblings from both sexes associated with parental consanguinity. We initiated a molecular study to identify the gene variants that may be underlying the syndrome.

3.3.2 Results

3.3.2.1 Family Data

This large family consisted of parents who are phenotypically normal. The parents were first cousins from Saudi Arabia with four affected and seven unaffected siblings (2 of whom are not included in this study) (Figure 3.3.1).

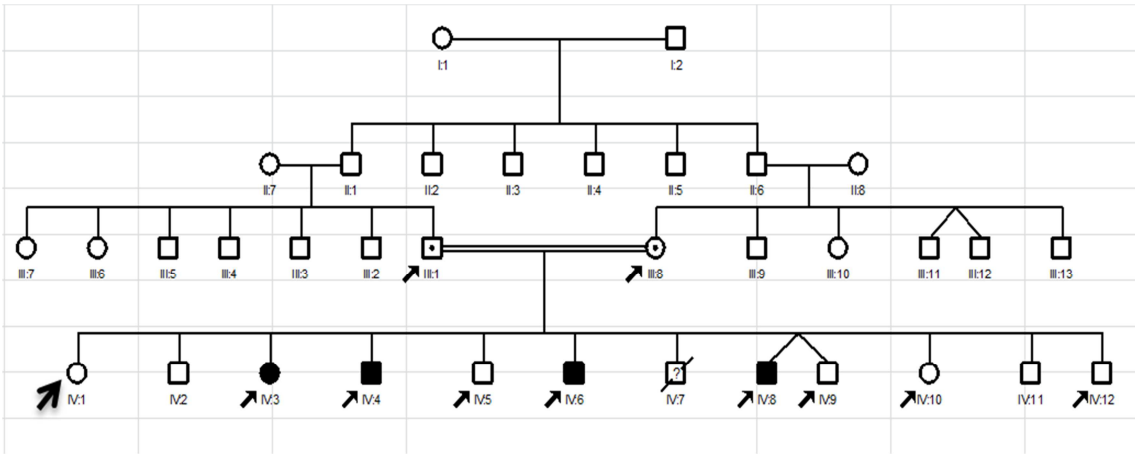

Figure 3.3.1: Family pedigree of the family (F2) with new autosomal recessive dRTA and NC. Squares indicate male family members, circles female family members; solid symbols indicate affected persons, circles and squares with dots indicate likely carriers, and double horizontal bars indicate consanguinity. Roman numbers underneath circles and squares indicate the sibling number. Arrows indicate individuals with available DNA.

In addition, one pregnancy ended at term with a male baby (IV: 7) with severe hydrops and renal failure who died at 1 week of life (further data is not available). There were no similar clinical cases reported on either side of the family (following questioning of known close relatives). Clinical presentations were described fully in the original paper (Faqeih et al., 2007). The major central nervous system (CNS) and renal abnormalities found in the affected patients are summarized in Table 3.3.3.

Clinical/laboratory finding	***Patient 1***	***Patient 2***	***Patient 3***	***Patient 4***
Short stature	yes	yes	yes	yes
CNS abnormalities				
Ventriculomegally	yes	yes	yes	yes
Abnormal myelination	yes	yes	yes	yes
Renal abnormalities				
Small kidneys	yes	yes	yes	yes
Proteinuria	yes	yes	yes	yes
Cortical hyper-echogenicity	yes	yes	yes	yes
Medullary nephrocalcinosis	yes	yes	yes	no
Poor cortico-medullary differentiation	yes	yes	yes	yes
Metabolic acidosis and RTA	yes	yes	yes	yes

Table 3.3.3: Summary of major CNS and renal abnormalities in the family F2

The family was screened first by direct sequencing for *CA2* and *ATP6V1B1* genes, both genes represent candidate genetic causes of renal tubular acidosis and *CA2* mutation (c.232+1G>A) is common in Saudi population. Sequencing revealed no mutations in either of these genes. Other genes that cause RTA including *SLC4A1* and *ATP6V0A4* were screened lately and no mutations were detected.

3.3.2.2 Homozygosity Mapping

Following a hypothesis of inheritance by descent and to identify a novel gene locus as the cause of this new syndrome, we carried out a genome wide linkage search using Affymetrix GeneChip® Human Mapping 250K Sty Arrays according to the manufacturer's protocol (http://www.affymetrix.com). We looked for regions with informative SNPs that were homozygous for the same allele for the affected children and heterozygous for the parents, and heterozygous or homozygous for the opposite allele for the unaffected individuals. A region of homozygosity in chromosome 2 was identified in four affected using CNAG (Version 2.0) software. To rule out large chromosomal abnormalities and to confirm our findings suggesting a region of homozygosity in chromosome 2 we used CytoScan arrays on all family members as described in the methods and materials chapter. From this scan the primary data was analyzed using Chromosome Analysis Suite (ChAS) software. These results did not detect any chromosomal abnormalities and the region of homozygosity was confirmed at the same region of chromosome 2 between genes *TSGA10* and *ST6GAL2* at physical map position Chr2: 99688169-107503563. Figure 3.3.2 (A&B) represents the region of homozygosity using two different analysis algorithms, CNAG2 versus ChAS2.

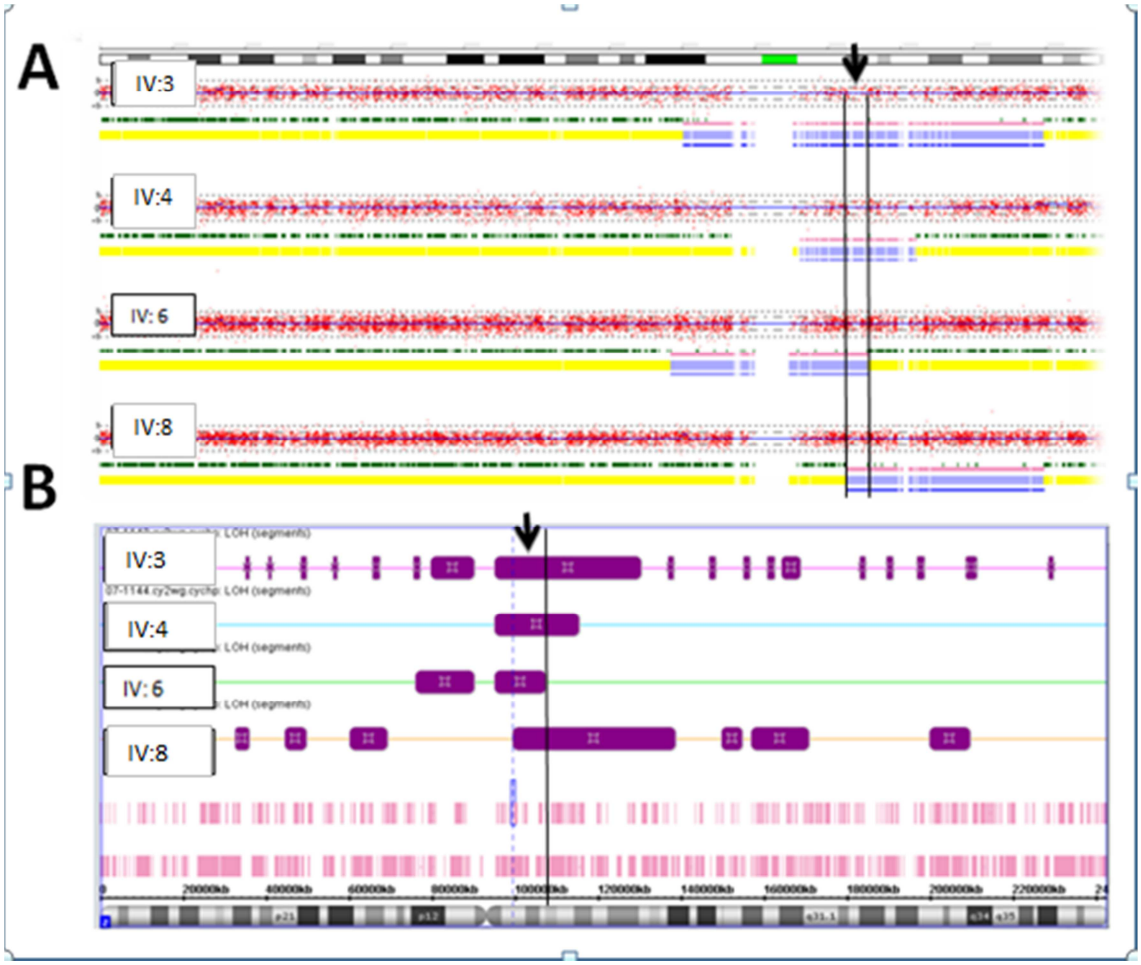

Figure 3.3.2: Homozygosity mapping using two comparable approaches

(A): This is a screen grab of CNAG2 software for chromosome 2. The arrow indicates shared homozygosity regions (shown in violet) on chromosome 2 between all 4 affected patients of the family using 250K chip array. Affected family members are marked IV: 3, 4, 8, 9 respectively.

(B): A screen grab of ChAS2 software for chromosome 2. Family members are annotated as in A. The arrow between 2 lines indicates shared homozygosity region in chromosome 2 between all 4 affected patients of the family using CytoScan chip array. Pink color indicates genes. Cytogenetic map is shown graphically map below.

3.3.2.3 Linkage Analysis and LOD Score Calculation

To confirm the region of homozygosity identified on chromosome 2 in parametric values, a LOD score was calculated with the easyLINKAGE-Plus software using the raw Axiom SNP data in all four affected, parents and 5 unaffected individuals (n=11). In order to perform this, the numbers of SNP calls for each patient were first reduced by

the software program based on (i) informative SNPs (heterozygous SNPs are informative for linkage analysis therefore SNPs that are always homozygous for one allele in all typed samples were removed (equal to 285114 SNPs)), (ii) Mendelian error rates (e.g. genotyping errors due to non -specificity of experimental assay, inappropriate allele calling, or simply random assay instability; these errors were excluded (46 SNPs)), and (iii) genotype call rate (e.g. any SNP with a genotyping call rate of less than 80% was excluded (0 SNPs)). The net number of SNPs used to calculate the LOD score were 268293 SNPs per sample. Using a disease allele frequency of 0.001 the GENEHUNTER program was used to calculate a multipoint parametric linkage analysis LOD score in an autosomal recessive pedigree model. The program was run at 1 marker every 0.1000 cM. The highest LOD score obtained was **3.37** at the same region of homozygosity, located within the chromosome 2 locus (Figure 3.3.3). LOD score results showed peaks at other loci (chromosomes: 6 and 11). The highest LOD score at chromosome 6 was **2.9** and **2.77** at chromosome 11. Haplotypes results of all three chromosomes 2, 6, and 11 are presented in Table (3.3.4).

SNP_ID	dbSNP RS ID	Chromosome	Chromosomal Position	IV:3	IV:4	IV:6	IV:8	III:1 Father	III:8 Mother	IV:5	IV:9	IV:10	IV:11	IV:1
AX-11404704	rs2679876	2	105938107	AA	AA	AA	AA	AB	AB	BB	BB	AB	AB	BB
AX-11437319	rs3217439	2	105953938	AA	AA	AA	AA	AB	AB	BB	BB	AB	AB	BB
AX-11425315	rs28930676	2	105959532	BB	BB	BB	BB	BB	AB	AB	AB	AB	BB	AB
AX-11399282	rs2576768	2	105960627	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB
AX-11581814	rs6736939	2	105962370	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA
AX-11475121	rs36099361	2	105975463	BB	BB	BB	BB	AB	BB	AB	AB	BB	AB	AB
AX-11380949	rs2278501	2	105979506	AA	AA	AA	AA	AB	AA	AB	AB	AA	AB	AB
AX-11380950	rs2278502	2	105979730	BB	BB	BB	BB	AB	BB	AB	AB	BB	AB	AB
AX-11671739	rs880427	2	105985228	BB	BB	BB	BB	BB	AB	AB	AB	AB	BB	AB
AX-11368408	rs2118291	2	105985671	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB
AX-11570730	rs6543286	2	105991702	AA	AA	AA	AA	AB	AB	BB	BB	AB	AB	BB
AX-11036108	rs10182476	2	105994276	BB	BB	BB	BB	AB	AB	AA	AA	AB	AB	AA
AX-11251971	rs13386004	2	105994498	AA	AA	AA	AA	AB	AB	BB	BB	AB	AB	BB
AX-11538044	rs4946858	6	108087830	AB	AB	BB	AB	AB	AB	AB	AA	AB	AB	AA
AX-11386650	rs2354095	6	108090237	AB	AB	AA	AB	AB	AB	AB	BB	AB	AB	BB
AX-11327786	rs17528372	6	108093657	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA
AX-11643520	rs7757683	6	108101564	AB	AB	BB	AB	AB	AB	AB	AA	AB	AB	AA
AX-11243708	rs13191632	6	108108756	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB
AX-11689588	rs9486696	6	108109472	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB
AX-11537972	rs4945790	6	108109985	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA
AX-11689589	rs9486702	6	108112832	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA
AX-11595967	rs6941259	6	108122113	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB
AX-11684879	rs9386672	6	108134840	AB	AB	BB	AB	AB	AB	AB	AA	AB	AB	AA
AX-11303058	rs17068819	6	108139570	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB
AX-11196472	rs12214852	6	108141213	AB	AB	AA	AB	AB	AB	AB	BB	AB	AB	BB
AX-11593804	rs6910909	6	108143093	AB	AB	BB	AB	AB	AB	AB	AA	AB	AB	AA
AX-11130125	rs10894669	11	133206940	AA	AA	AA	AA	AB	AB	AB	AA	BB	AB	AB
AX-11503595	rs4387356	11	133209187	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB
AX-11152000	rs11223441	11	133212242	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB
AX-11654156	rs7925902	11	133213019	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA
AX-11654609	rs7933617	11	133220663	BB	BB	BB	BB	AB	BB	AB	BB	AB	AB	BB
AX-11227646	rs12804596	11	133220898	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA
AX-11227200	rs12791663	11	133222400	BB	BB	BB	BB	AB	AB	AB	BB	AA	AB	AB
AX-11139008	rs12282747	11	133227132	BB	BB	BB	BB	BB	AB	BB	BB	AB	BB	AB
AX-11465025	rs35501691	11	133228369	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB
AX-11654976	rs7939409	11	133233949	BB	BB	BB	BB	AB	BB	AB	BB	AB	AB	BB

Table 3.3.4: Haplotypes of regions of chromosomes 2, 6 and 11 with high LOD score in the F2 family. Red color indicates genotypes in all 4 affected, yellow in parents, and green in 5 unaffected siblings.

				Marker	CHR	cM	pLOD	NPLnpLOD	P val	INFO	Alp
Project:	NS2-AXIOM	Inheritance:	Recessive								
Family name:	TOTALS	Common allele:	99.90 %	1.rs1866040	2	118.61	2.4236	3.6845	2.9479	0.0156	1.0000	1.00
Used map:	Axiom GW Hu (sex-averaged)	Disease allele:	0.10 %	2.rs6543286	2	118.71	2.4236	3.6845	2.9479	0.0156	1.0000	1.00
Marker positions:	5346 ok / 28912 ? / 0 outside	LC1 PCOPY rate:	0.00 %	3.rs4851056 - rs11885188	2	118.63	2.4236	3.6845	2.9479	0.0156	1.0000	1.00
Allele frequencies:	Codominant	LC1 PENET wt/mt:	0.00 %	4.rs13429686 - rs7600139	2	118.65	2.4236	3.6845	2.9479	0.0156	1.0000	1.00
HAP algo / SCFct:	Off / Hom	LC1 PENET mt/mt:	99.00 %	5.rs2679876 - rs28930676	2	118.67	2.4236	3.6845	2.9479	0.0156	1.0000	1.00

Pair setting: All pairs of affected / phenotyped sibs

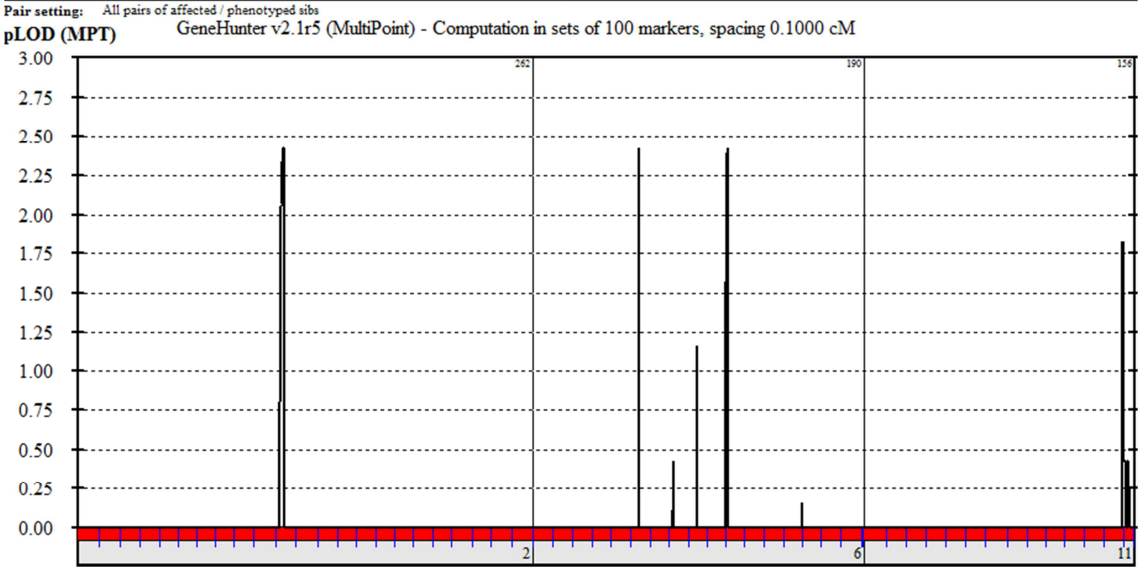

Figure 3.3.3: Graphical plots of LOD scores result from easyLINKAGE Plus software of the family F2 with new autosomal dRTA and NC. The y axis indicates the LOD score value, whilst the x axis shows chromosomes 2, 6 and 11.

Homozygosity was not detected in chromosome 11 using the homozygosity mapping approaches and this is may be explained by SNP density and small homozygosity region. Chromosome 6 haplotype may indicate compound heterozygous mutation in the family.

Because both homozygosity mapping and LOD score analysis approaches determined same region on chromosome 2 that had highest LOD score obtained **3.37** we focused initially on chromosome 2 locus.

The Chromosome 2 region falls between the SNPs (rs4851019) and (rs920217) respectively. The physical map for the region is Chr2: 102636083-105363180 using UCSC genome browser hg18 assembly. The region contains 20 known coding and hypothetical genes (Figure 3.3.4).

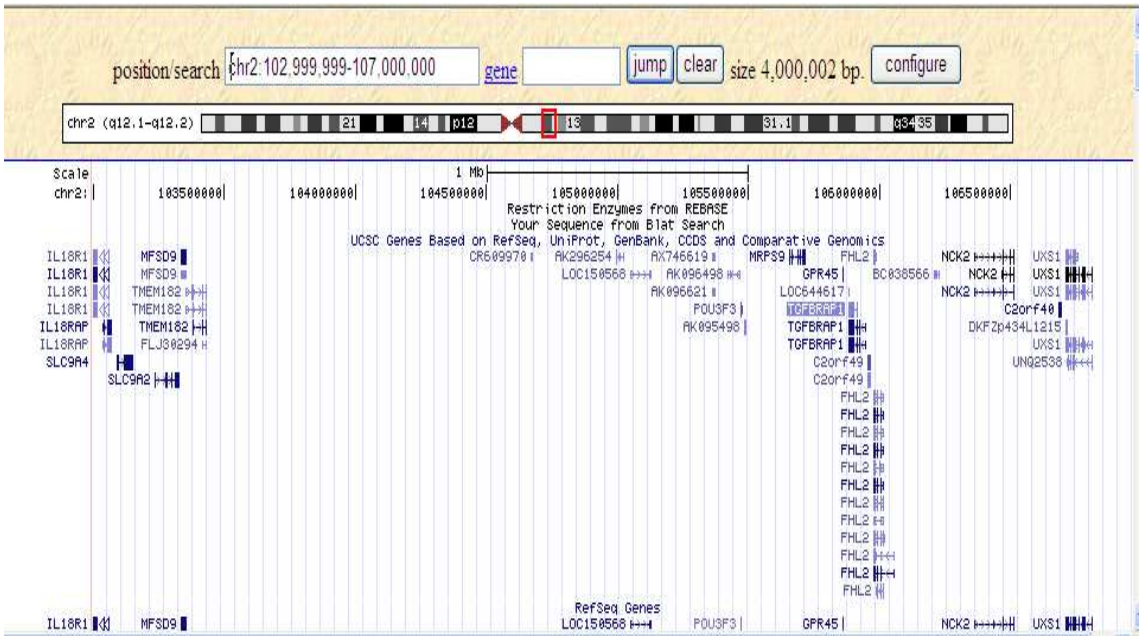

Figure 3.3.4: Screenshot of UCSC genome browser hg18 of chromosome 2 showing the region of homozygosity mapping and linkage analysis results. All coding and hypothetical genes in the region are shown.

3.3.2.4 Candidate Gene Sequencing

Using these annotated genes within the Chromosome 2 putative locus as "candidate genes" we attempted to detect the mutated gene by exon PCR, followed by direct sequencing.

All genes in the region, known coding and hypothetical genes (Table 3.3.5) according to UCSC genome browser hg18 assembly were screened by direct sequencing of PCR products. All four affected patients, their parents, and one unaffected member of the family together with a normal DNA control (from an ethnically matched control) were sequenced for all 20 genes (a total of 148 amplicons per sample).

Primer-3 software for oligonucleotide primer design was used to design exons and intron boundaries of genes. Oligonucleotide primers are listed Appendix A. These oligonucleotide primers were amplified by PCR. DNA sequencing was performed on one strand. Sequencing data was analyzed using SeqMan II software 6.1 (DNAStar).

Screening all of the annotated genes in the region introduced a new interval between the *SLC9A4* gene and the *BC038566* hypothetical gene, where one of the affected individuals was heterozygous for SNPs in both genes. The new physical map position Chr2: 103,236,166 - 106,015,68 is shown in Table 3.3.5. However, no mutations were detected within the coding sequence of the 20 genes (known and hypothetical) screened (data not shown). In addition, promoters regions of *SLC9A2* and *POU3F3* genes were sequenced given these were strong candidates for a renal phenotype. However these regions were negative for mutations (data not shown).

Gene	Physical position	Status of Gene	Sequencing results on affected	Tissues of expression	MGI knockout mouse results
SLC9A4	103,089,762-103,150,431	Coding Gene	Heterozygous		
SLC9A2	103,236,166-103,327,809	Coding Gene	Homozygous	Colon, Testis, Kidney, Brain	Gastric acid secretion is impaired in homozygous mutant mice
MFSD9	103,333,666-103,353,337	Coding Gene	Homozygous		
TMEM182	103,378,490-103,434,138	Coding Gene	Homozygous	Heart, Bone Marrow	
AK054856	103,353,394-103,432,398	mRNA Sequence	Homozygous		
AK296254	104,995,308-105,024,790	mRNA Sequence	Homozygous		
LOC150568	105,050,805-105,129,215	non-coding RNA	Homozygous		
AX746619	105,363,095-105,372,213	non-coding RNA	Homozygous		
AK096621	105,363,098-105,374,177	mRNA Sequence	Homozygous		
AK096498	105,421,977-105,467,916	mRNA Sequence	Homozygous		
POU3F3	105,471,969-105,473,471	Coding Gene	Homozygous	Kidney, Brain	Homozygous mutation of this gene results in death within 36 hours after birth
AK095498	105,481,955-105,488,840	mRNA Sequence	Homozygous		
MRPS9	105,654,483-105,716,418	Coding Gene	Homozygous	Kidney, Brain	

Gene	Physical position	Status of Gene	Sequencing results on affected	Tissues of expression	MGI knockout mouse results
GPR45	105,858,200-105,859,924	Coding Gene	Homozygous		
LOC644617	105,880,847-105,882,690	mRNA Sequence	Homozygous		
TGFBRAP1	105,883,540-105,946,148	Coding Gene	Homozygous	Thymus, Kidney, Brain	
C2ORF49	105,954,013-105,961,984	Coding Gene	Homozygous	Bone, Kidney, Brain	
FHL2	105,977,283-106,015,681	Coding Gene	Homozygous	Heart, Kidney, Brain	Mice homozygous for a targeted null mutation exhibit no discernable phenotype; mice are viable and fertile with normal cardiac development and physiology
BC038566	106,209,554-106,227,016	mRNA Sequence	Heterozygous		
NCK2B	106,471,502-106,509,632	Coding Gene	Heterozygous		

Table 3.3.5: Summary of genes sequenced, results of sequencing, tissue of gene expression, and knockout mouse model.

Higher density mapping Affymetrix GeneChip® Human Mapping 6.0K was employed to detect other smaller regions of homozygosity, in part because of the spurious peaks obtained from the LOD score results. Running this raw data through Genotyping Console (Version 3.0.1) software confirmed that, the region on chromosome 2 was the only region to share homozygosity between the 4 affected siblings. Figure 3.3.5 shows homozygosity results of SNP 6.0 arrays for all affected siblings. These results confirmed that extra LOD scores peaks in the genome were likely to be false positives and that the region of chromosome 2 remained the putative locus.

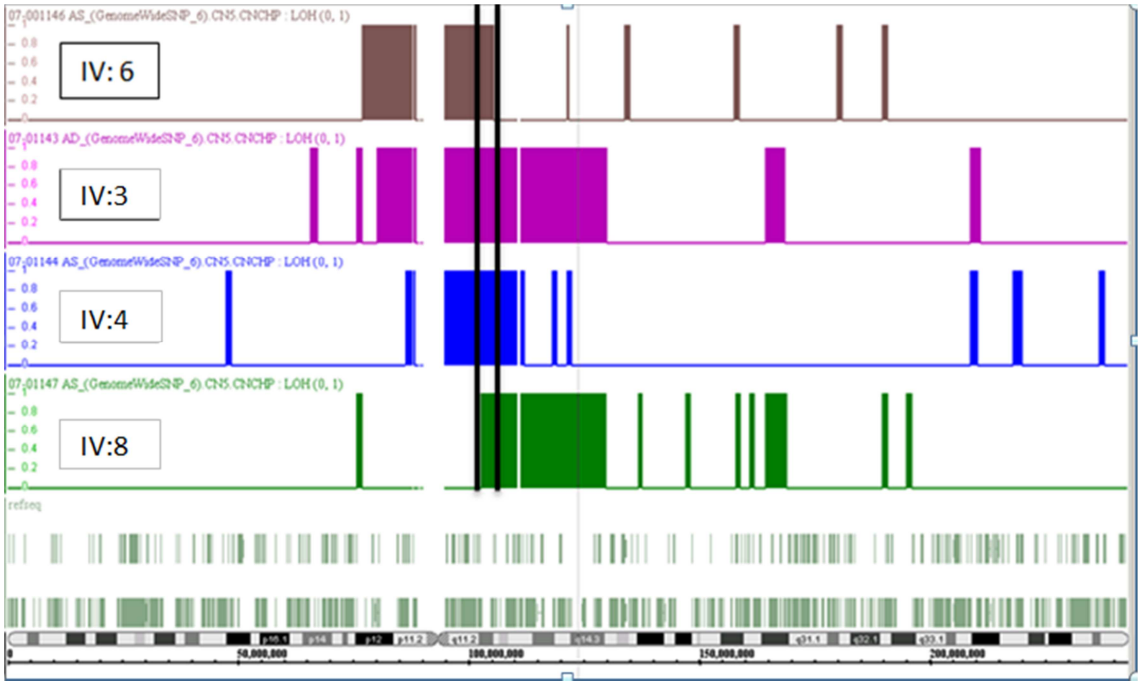

Figure 3.3.5: A Screenshot of Genotyping Console (Version 3.0.1) software for chromosome 2 using SNP 6.0 arrays. Homozygosity region of all 4 affected family members falls between two vertical black lines. No other homozygosity regions were detected in the genome.

3.3.2.5 miRNA Sequencing

Because no pathogenic mutations were detected in coding exons and exon-introns boundaries in the genes screened within the Chromosome 2 locus, we decided to sequence all miRNA in the new interval of homozygosity (Chr-2:103,236,166 - 106,015,681). MicroRNA (miRNA) is a short ribonucleic acid (RNA) molecule found in eukaryotic cells. A microRNA molecule has very few nucleotides (an average of 22) compared with other RNAs.

miRNAs are post-transcriptional regulators that bind to complementary sequences on target messenger RNA transcripts (mRNAs), usually resulting in translational repression or target degradation and gene silencing. Using UCSC Genome Browser on Human Feb. 2009 (GRCh37/hg19) Assembly we found 16 miRNAs regulatory sites (Figure 3.3.6).

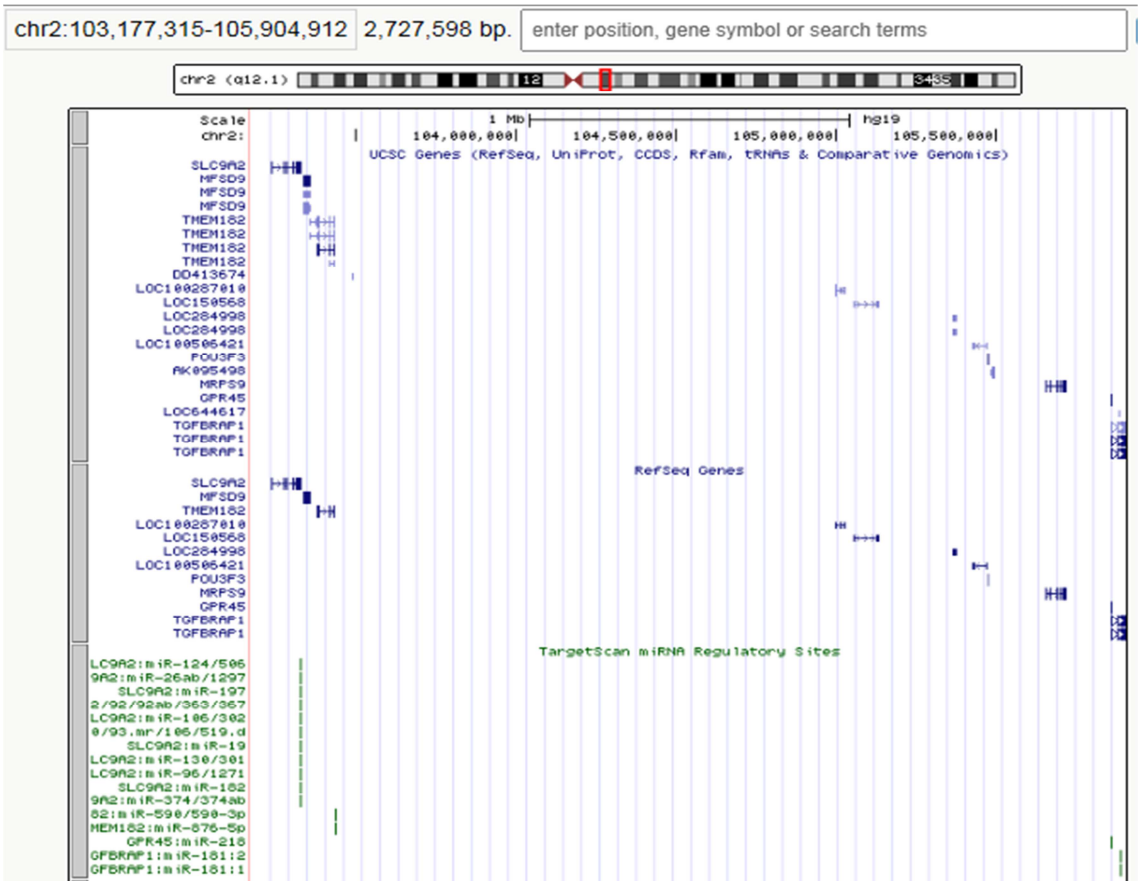

Figure 3.3.6: Screenshot of UCSC genome browser hg19 of chromosome 2 showing the miRNA (green) in the region of homozygosity. UCSC genes (top) and RefSeq genes (bottom) are in blue.

Primers were designed using Primer 3 software. Oligonucleotide primers are listed in Appendix A. These primers were amplified by PCR and direct sequencing was carried out. By sequencing all 16 miRNA no sequence variants were detected (Data not shown). Sequencing TGFBRAP1: miRNA-181:1; resulted in all 4 affected to be negative for rs70953504 (28bp deletion) while parents and other siblings vary between

homozygous and heterozygous for the SNP (Figure 3.3.7). The SNP was found in 1 out of six ethnically matched controls.

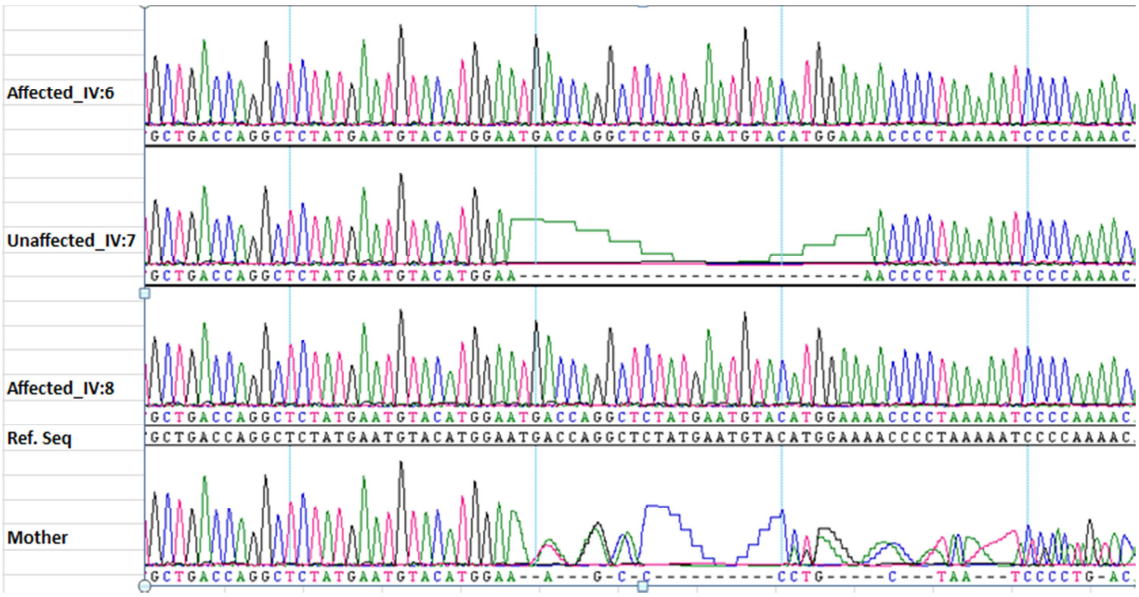

Figure 3.3.7**:** Sequencing chromatogram of rs70953504 of TGFBRAP1: miRNA-181:1 showing deletion of 28 nucleotides in both alleles of unaffected son. Both affected were negative for the deletion and mother had deletion in one allele (heterozygous).

3.3.2.6 Gene Expression

Assuming the mutation may be located in promoters regions or non-coding regions of genes, gene expression profiling for all known genes in the homozygous region of chromosome 2 was carried out. We performed real time quantitative PCR in order to attempt to measure gene expression.

Whole blood samples from two affected individuals (IV-6 and IV-8), the father (III-2), and one unaffected (IV-5) were obtained following informed consent. Transformation of peripheral blood lymphocytes by EBV was performed using standard tissue culture techniques. After tissue growth, RNA was extracted from lymphoblast cultures then reverse transcribed to cDNA. The cDNA was then used in real time quantitative PCR (qPCR). Expression oligonucleotide primers for the following genes were designed using Primer3 software: *SLC9A2*, *MFSD9*, *TMEM182*, *POU3F3*, *MRPS9* and *TGFBRAP1*. *GAPDH1* oligonucleotide primers were used as an endogenous control. Oligonucleotide primers are listed in Appendix A. Relative gene expression ΔΔCT identified no differences between gene expressions of affected versus others.

Because gene expression is tissue specific we considered preferentially expression of these genes and some of them may not expressed lymphocytes. The phenotype of disease within this family involves kidneys and brain. In order to assess renal expression, we obtained urine samples from the family; as some cells (proximal tubular cells and others) pass into the urine under normal physiological conditions(Sutherland and Bain, 1972).

Urine samples from three affected individuals (IV-4, IV-6, and IV-8), Father (III-2), two unaffected individuals (IV-1 and IV-5) and urine sample from an ethnically matched normal control were obtained. The yield of RNA was little (around 0.3 µg) therefore we were able to test for three genes only: *SLC9A2*, *POU3F3*, and *MRPS9*. *GAPDH* expression was used as an internal control. *SLC9A2*, *POU3F3*, and *MRPS9* are known to be expressed in kidney and brain tissues (http://bioinfo.wilmer.jhu.edu/tiger). We were limited to replicate the experiment results only once to confirm our findings. Results for genes expression are shown in Table 3.3.6. There was no difference in expression of candidate gene *SLC9A2* and *POU3F3* in urine from father and affected, using *GAPDH* as an internal reference control. *MRPS9* results indicated increased expression but these results could not be replicated in the same affected sample nor be confirmed in other affected siblings.

	MRPS9				***GAPDH***					
Sample ID	Sample1	Sample2	Sample3	Average	Sample1	Sample2	Sample3	Average	delta CT	induct
Nill	31.75909	31.29385	32.92431	31.99	20.44519	21.7517	22.61472	21.6		**1**
III-2	26.4636	26.542	28.5378	27.18	15.844	16.5905	17.4255	16.62	-0.1725	**1.13**
IV-4_A	32.59318	31.23417	31.47773	31.77	18.58563	18.44696	18.618	18.55	-2.8296	**7.11**
IV-4_B	31.6233	30.8716	31.6323	31.38	18.3553	18.5794	18.276	18.4	-2.5836	**5.99**

Table 3.3.6: Primary results of gene expression of *MRPS9* gene in urine sample. The III-2 specimen (Father) and two specimens A and B from the affected (IV-4) were analyzed. Each specimen was amplified in triplicate (Labeled Sample1, 2, 3 respectively) for *MRPS9* and *GAPDH* genes. Induct results indicated that the expression of the gene in the affected is higher by 7.11 and 5.99 folds than normal.

Because no mutations were detected in chromosome 2 region, we looked at other regions with high LOD score (chromosomes 6 and 11). The region in chromosome 11 contained one gene (*OPCML* gene) where no mutations were detected in the family. The region in chromosome 6 contained five candidate genes and sequencing these genes still in progress.

3.3.3 Discussion

We have had the opportunity to investigate at a molecular genetic level a possible new autosomal recessive disease described in the Saudi population. Our molecular genetics work confirmed a likely autosomal recessive inheritance of the disease by homozygosity mapping, and subsequently we explored both linkage analysis, and candidate gene screening. Homozygosity mapping and linkage analysis methods confirmed a putative novel locus of the disease in the family on Chromosome 2 by virtue of the family size and the number of affected individuals. LOD score results showed peaks at other loci (Chromosomes: 6 and 11). There are many factors can affect linkage analysis. These include disease allele penetrance, marker allele frequency, genotyping errors, and linkage heterogeneity. However, higher positive LOD scores and a larger number of them are expected around true rather than around false peaks (Hoh and Ott, 2000). In addition, using high density SNP arrays for homozygosity mapping, none of the additional regions (outside of Chromosome 2) were revealed by using detailed homozygosity mapping. The physical map of the putative novel locus is Chr-2:103,236,166 - 106,015,681, and the cytogenetic map reference 2q12.1. No renal phenotype has been associated with any gene mutations in this region, according to OMIM. Knockout mice of three genes in the region (*SLC9A2*, *POU3F3,* and *FHL2*) have shown variable phenotypes (Table 3.3.5) as described by the Mouse Genome Informatics (MGI). None of these phenotypes however, are similar to affected individuals of the family F2.

Despite screening all known and hypothetical genes in the region by exon PCR, we were unable to determine mutated gene. On the other hand, exon PCR confirmed the region of homozygosity where all four affected shared the same genotypes at the region. Candidate gene screening covered only entire cDNA regions, and introns and promoter regions were not included. Development of Next Generation Sequencing (NGS) has

revealed the importance of deep intron mutations. There are a few reports describing lethality of intron mutations (Webb et al., 2012, Abu Safieh et al., 2010).

The promoter regions of *SLC9A2* and *POU3F3* genes were fully sequenced and they were also negative for mutations or disease causing variants. The promoter regions of other candidate genes have not been studied yet.

In addition sequencing of all the 16 miRNAs in the region revealed no pathogenic mutations. Results for rs70953504 in all affected confirmed once again the region of homozygosity. Profiling miRNA expression in the affected siblings could be achieved using commercially available miRNA RT-PCR kits. This is may help in detecting reduced miRNA expression as a mechanism of disease in this family.

The *POU3F3* promoter region, a strong candidate gene based on murine model data (Castro et al., 2006), is a GC rich region (Sumiyama et al., 1996), and methylation studies of the region have not been performed. Methylation of DNA is a biochemical process that is important for normal development in eukaryotic species. It occurs at cytosines residues within the sequence context of CpG islands (Bird, 1986). It is known that genes are not active at all times and DNA methylation is one of several mechanisms that cell use to control gene expression. Given the critical role of DNA methylation in gene expression and cell differentiation, it seems that errors in methylation could give rise to a number of lethal diseases such as Rett syndrome, which is caused by mutations in the gene encoding methyl-CpG-binding protein-2 (*MECP2*) (Amir et al., 1999) and Fragile-X syndrome where CGG repeat expansion at the 5′-untranslated region and aberrant de novo methylation of the CpG island upstream of the gene can cause the disease (Kremer et al., 1991, Oberle et al., 1991). Testing for gene expression of some of the candidate genes through real time qPCR was performed, but no clear results of abnormal gene expression were obtained. Gene expression is tissue specific and we were uncertain of expression of genes in B lymphocytes. The RNA was extracted from lymphoblast and the effect of EBV in genes tested unknown. From a technical point of view, real time qPCR has to be reproducible and a result has to be consistent. This was a difficult task with the limited amount of RNA extracted from the urine samples obtained in this study. The primary purpose of the gene expression experiment was to lead to a possibly defective gene, rather than identify the molecular cause. Gene expression usually used to prove consequences of molecular defect or in a

reverse manner to highlight the effect of the gene to help in identifying the molecular cause.

Sequencing whole mRNAs of all genes in the region may help in identifying causative gene that lead to the disease in the family. Next generation sequencing (NGS) is another approach that could help in detecting mutated gene(s).

In order to proceed with NGS, there are three options, Firstly sequencing the whole genome, secondly performing exome sequencing, which sequences the coding regions of the genome selectively, or thirdly targeted region sequencing (target region capture) by enriching specific region by microarray or solution hybridization based on probes designed according to the bases sequence of interested genomic regions. The region of homozygosity in Chromosome 2 is around 3 Mb and the third option would be the preferred option to try and solve genetically this family, because this is a focused and in-depth investigation on genetic variants in the region. Exome sequencing would not cover completely all coding exons in the genome and we already sequenced all known coding exons in the region but still valid options for other loci.

Given that NGS typically result in many changes of unknown significance, sequence variants will need confirmation. In such circumstances we can reduce such effort by sequencing a pooled sample of all four affected, and a pooled sample of both parents. Sequence variants can also be limited to the region of homozygosity already established in this study.

In conclusion, using homozygosity mapping, we have been able to localize a small region on Chromosome 2 that represents a putative novel locus for a syndrome that results in dRTA and nephrocalcinosis in a large family of Saudi Arabian origin. Efforts to unravel the molecular cause of the disease in chromosome 2, 6, and 11 are in progress.

Chapter 3: RESULTS

3.4 Bilateral Renal Agenesis

3.4.1 Introduction

Renal Agenesis (RA) or Potter's syndrome is a congenital absence of one (unilateral) or both (bilateral) kidneys. The kidneys develop between the 5th and 14th week of fetal development, and by the 14th week they are normally producing urine (Harewood et al., 2010). Bilateral Renal Agenesis (BRA) belongs to a group of prenatally lethal renal diseases, including severe bilateral renal dysplasia, and severe obstructive uropathy. The incidence of BRA is approximately 1/10000 births (Sanna-Cherchi et al., 2007). Unilateral renal agenesis (URA) is more common 1/1000-2000 (Yalavarthy and Parikh, 2003, Pohl et al., 2002) and is usually clinically silent. BRA is often detected on fetal ultrasound because there will be a lack of amniotic fluid (oligohydramnios). The common association between oligohydramnios and BRA often leads babies to have additional birth defects, including abnormalities of the urinary system, genitals, limbs, heart, and lungs. Approximately 40% of fetuses with bilateral renal agenesis will be stillborn, and if born alive, the baby will usually live only a few hours (Kelalis et al., 1985). Currently, the precise causes of renal agenesis are not known. Studies have shown that unilateral and bilateral renal agenesis may be genetically related (Kerecuk et al., 2008). Several lines of evidence suggest a genetic contribution to the development of BRA and renal adysplasia. First, the incidence of BRA is increased in fetuses who have first degree relatives with congenital renal anomalies (Roodhooft et al., 1984, Bankier et al., 1985), Second, renal anomalies are associated with some autosomal recessive and dominantly inherited disorders. According to Winter-Barraitser Dysmorphology Database (London Medical Databases, London) there are 148 different syndrome entities associated with renal agenesis (Harewood et al., 2010). Third, animal genetic models have implicated many genes important for renal development (Bates, 2000). Furthermore, human renal agenesis has been characterized as a multifactorial disorder with an apparent genetic contribution (Yalavarthy and Parikh, 2003). Table 3.4.1 (Modified from (Sanna-Cherchi et al., 2007)) lists principal genes targeted in mice leading to renal agenesis, hypoplasia, and dysplasia.

Gene	Human homolog	Kidney phenotype	Reference
Nuclear proteins/Transcription factors			
Foxd1	*FOXD1*	Small, fused, undifferentiated kidneys	Hatini et al., 1996
Eya1	*EYA1*	Absent kidneys	Johnson et al., 1999
Emx2	*EMX2*	Absent kidneys	Miyamoto et al., 1997
Hoxa11/Hoxd11	*HOXA11/HOXD11*	Small or absent kidneys	Davis et al.,1995
Lhx1	*LHX1*	Absent kidneys	Shawlot and Behringer, 1995
Pax2	*PAX2*	Small or absent kidneys	Torres et al., 195
Sall1	*SALL1*	Absent kidneys	Nishinakamura et al.,2001
Pbx1	*PBX1*	Small or absent kidneys	Schnabel et al., 2003
Wt1	*WT1*	Absent kidneys	Kreidberg et al., 1993
Secreted factors/Other			
Bmp4	*BMP4*	Altered ureteric bud (UB) branching	Miyazaki et al.2000
Bmp7	*BMP7*	Disrupted nephrogenesis	Dudley et al., 1995
Wnt4	*WNT4*	Undifferentiated kidneys	Stark et al., 1994
Gdnf	*GDNF*	Absent kidneys, severe dysgenesis	Sanchez et al., 1996
Receptors			
Ret	*RET*	Absent kidneys, severe dysgenesis	Schuchardt et al., 1994
Fgfr1/Fgfr2	*FGFR1/FGFR2*	Absent kidneys	Poladia et al., 2006
Agtr2	*AGTR2*	Multiple urinary tract malformations	Nishimura et al.,1999

Table 3.4.1: Genes targeted in mice leading to renal agenesis, hypoplasia, dysplasia

For the isolated, non-syndromic renal agenesis/hypoplasia and dysplasia, segregation studies have been performed (Sanna-Cherchi et al., 2007). Mutations in genes such as *TCF2* (Bingham et al., 2000), *PAX2* (Sanyanusin et al., 1995), *EYA1* (Abdelhak et al., 1997), *SIX1* (Ruf et al., 2004b), *SALL1*(Kohlhase et al., 1998), *RET* (Skinner et al., 2008), *BMP4* and *SIX2* (Weber et al., 2008) have been shown to be associated with renal agenesis or pediatric congenital anomalies of the kidney and urinary tract (CAKUT), which is a larger category of renal diseases including renal agenesis. Due to its complexity and association with other phenotypes, many genes are likely to play a role in renal agenesis.

Medications and therapy for RA in addition to prevention by molecular testing could be improved by understanding its molecular genetics basis.

Here we report identification of a putative novel gene causing BRA in a Saudi Arabian family.

3.4.2 Results

3.4.2.1 Family Data

Two consanguineous families were recruited for the study. The first family (RA-FAM1) belongs to a large tribe in Saudi Arabia and consisted of two branches, labeled A and B (Figure 3.4.1). Branch A of the RA-FAM1 had visited Obstetrics and Gynecology (OB/GYN) clinic at KFSH&RC since 1991 presenting with a series of stillborn fetuses. Following ethical approval we were involved with this family from 2008. At that time we were able, following informed consent, to obtain DNA from family members RA-FAM1_III:5 (unaffected mother) and RA-FAM1_IV:8 (a stillbirth fetus). In 2009 the DNA was obtained RA-FAM1_IV:9 (stillborn fetus). The DNA from RA-FAM1_III:4 (unaffected father) and RA-FAM1_IV:6 and RA-FAM1_IV:7 (two healthy siblings) was also obtained, following informed consent.

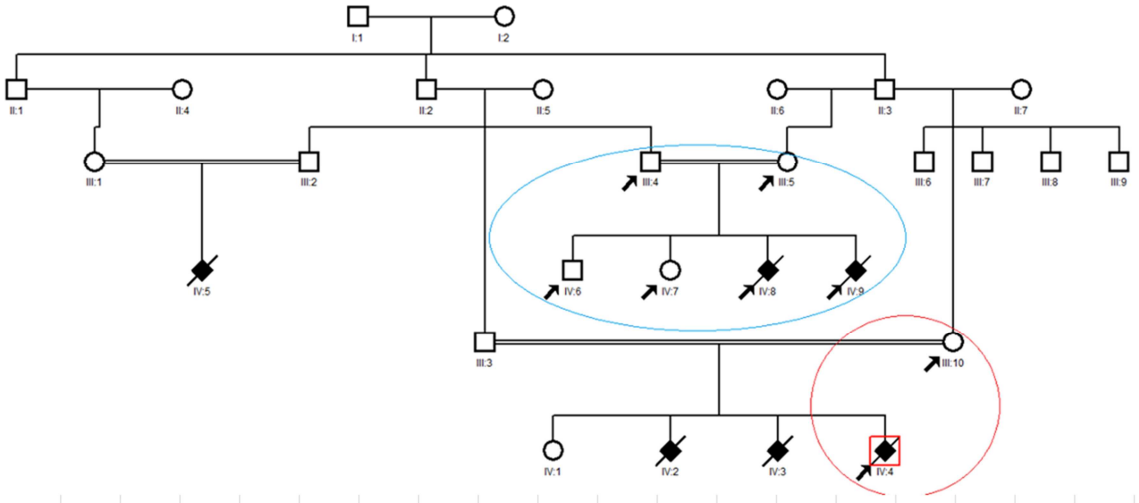

Figure 3.4.1: Family pedigree of the family (RA-FAM1) with bilateral renal agenesis. Squares indicate male family members, circles female family members; solid symbols indicate affected persons, slashes deceased fetuses and double horizontal bars consanguinity. Roman numbers underneath circles and squares indicate the sibling number. Arrows indicate individuals where DNA is available. Individuals (III:4, III:5, IV:6, IV:7, IV:8 and IV:9) represent branch A of the family (circle in blue); whilst (III:10 and IV:4) represent branch B (circle in red).

Branch B (the mother RA-FAM1_III:10 and stillbirth fetus RA-FAM1_IV:4, Figure 3.4.1) of the family where referred to OB/GYN clinic at KFSH&RC in 2003. In 2010 we obtained permission to use DNA from the mother (III:10) and a fibroblast culture from dead fetus (IV:4).

A second family (RA-FAM2) was referred to OB/GYN clinic at KFSH&RC in December 2011 after initiation the study with the RA-FAM1. We were able to obtain DNA only from the affected dead fetus with bilateral renal agenesis and both parents. No healthy siblings available (Figure 3.4.2).

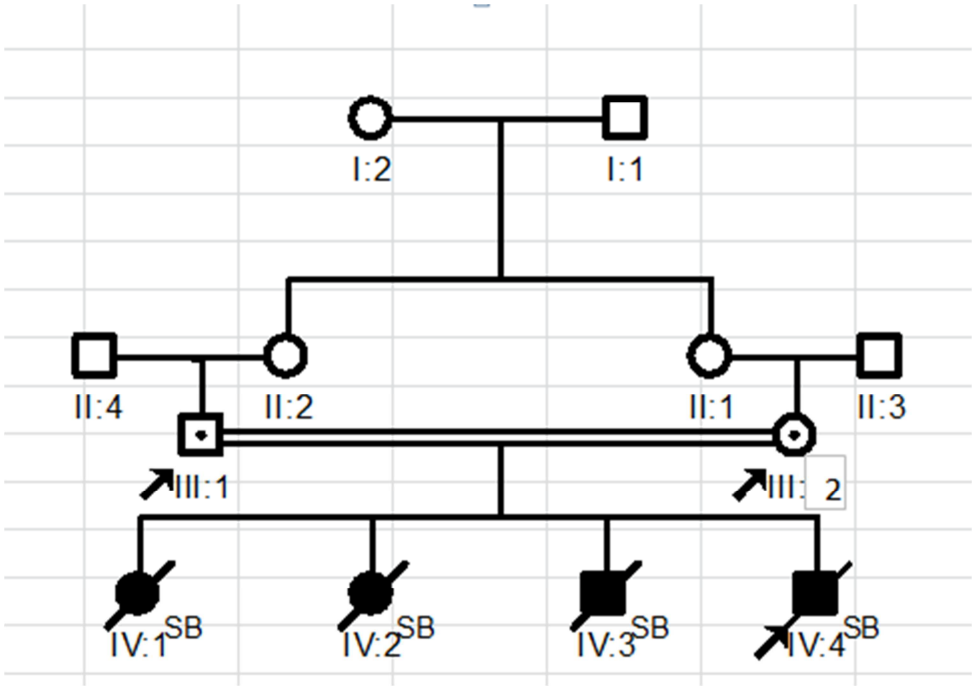

Figure 3.4.2: Pedigree of RA-FAM2. Arrows indicate individuals where DNA is available. SB indicates still birth fetuses.

3.4.2.2 Clinical Presentation

Branch A of the family RA-FAM1 lost a total of seven pregnancies with an antenatal ultrasound diagnosis of BRA together with absence of the bladder in association with oligohydramnios. Branch B of RA-FAM1 lost three pregnancies due to oligohydramnios with different features in each pregnancy. The stillborn fetuses had additional abnormalities including cleft palate, abnormal limbs, and small chest.

Parents in both branches were healthy, and healthy siblings were present in both branches of the RA-FAM1 (Figure 3.4.1). The pedigree and clinical presentation are suggestive of autosomal recessive inheritance.

RA-FAM2 lost three pregnancies in another regional hospital. RA-FAM2 was referred to our hospital when they presented with severe oligohydramnios in a fourth pregnancy. The antenatal ultrasound diagnosis revealed BRA with the absence of a bladder. Fetal blood sample together with the parent's blood was collected for extraction. Both parents are healthy. Again, the presentation of disease and the pattern of inheritance suggest an autosomal recessive disorder (Figure 3.4.2).

3.4.2.3 Homozygosity mapping using 250K Arrays

In order to try and define a novel disease locus for bilateral renal agenesis in these two families (RA-FAM1 and RA-FAM2) we initially performed homozygosity mapping, using DNA from branch A of RA-FAM1. We used Affymetrix GeneChip® Human Mapping 250K Sty Arrays to identify regions of homozygosity between the two affected family members (IV:8 and IV:9). Using CNAG2 software we were able to identify five regions of homozygosity; two regions on Chromosome 1((q21.3-q24.2) and (q32.2-q41) respectively), one region on Chromosomes 5 (q31.1-q33.1), and two regions on Chromosome 17 ((p13.2-p12) and (q11.2-q12) respectively (Figure 3.4.3).

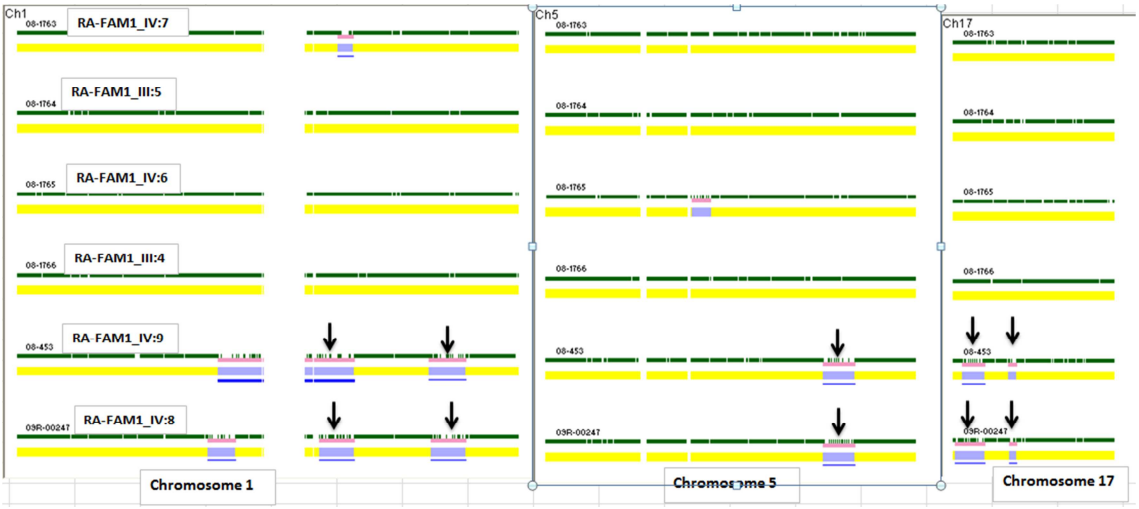

Figure 3.4.3: CNAG2 software analysis showing regions of homozygosity detected in branch A of RA-FAM1 using 250K chip. Arrows indicate regions of homozygosity in both affected.

These regions failed to contain genes that were known to be involved in renal agenesis in humans or mice.

Three possible candidate genes for renal agenesis were identified within these regions. These include fibroblast growth factor 1 (*FGF1*) (Jaye et al., 1986) and GDNF family receptor alpha 3 (*GFRA3*) (Baloh et al., 1998) on Chromosome 5 and transforming growth factor beta 2 (*TGFB2*) (Schlunegger and Grutter, 1992) on Chromosome 1. Direct sequencing (using exon PCR and Sanger sequencing, oligonucleotide primers sequences given in Appendix A) of these genes however, revealed no pathogenic mutations (Data not shown).

3.4.2.4 Homozygosity Mapping and Linkage Analysis using SNP 6.0K Arrays

After obtaining DNA samples from branch B of family RA-FAM-1, we carried out additional homozygosity mapping using a higher density strategy. We employed Affymetrix GeneChip® Human Mapping 6.0K chip arrays in family members (affected members IV:8, IV:9, and IV:4 and unaffected members III:4, III:5, III:10, IV:6, and IV:7) in order to detect shared regions of homozygosity in all affected members. We also used data from 6.0K SNPs genotyping for the calculation of LOD score by linkage analysis.

Homozygosity mapping in family RA-FAM1 revealed just one distinctive region shared by all three affected and abolished most of the homozygosity regions detected by 250K arrays. The region of homozygosity on Chromosome 1 had previously been detected using the more limited 250K arrays for two of the affected (IV:8 and IV:9) see Figure 3.4.3. The confirmed region of homozygosity was located on the long arm of Chromosome 1 (q32.2-q41). The physical map position of the region is chr1: 204,586,303-223,963,718 (Figure 3.4.4).

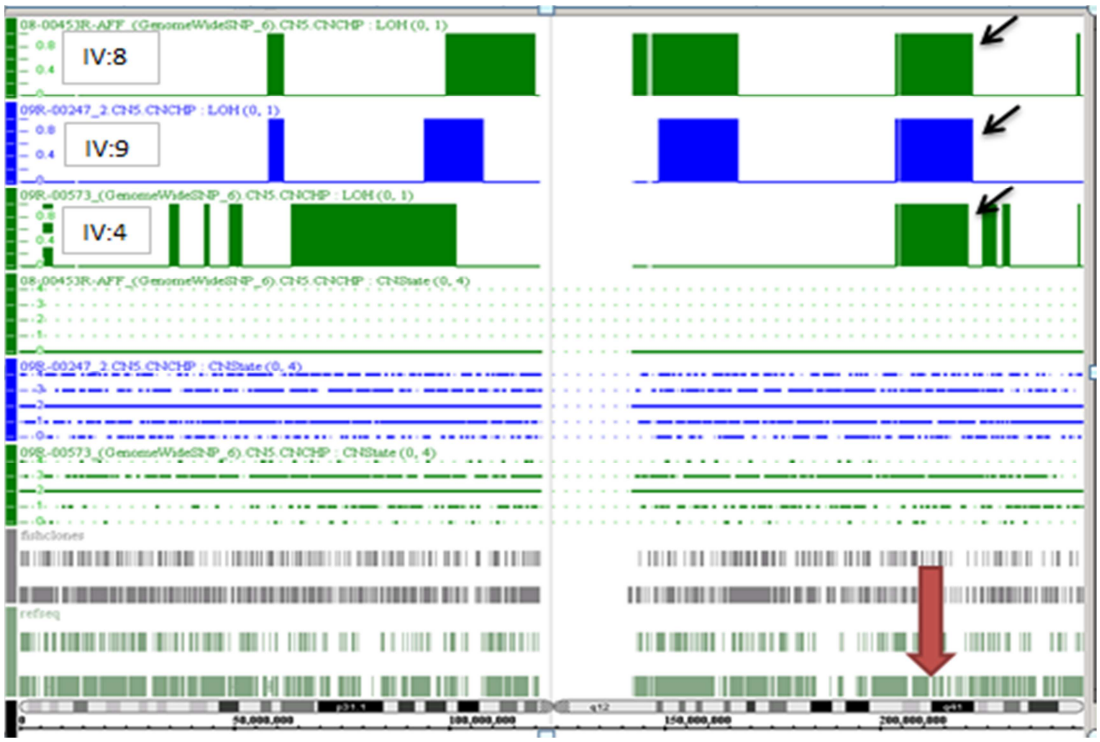

Figure 3.4.4: A screenshot of Genotyping Console (version 3.0.1) software for chromosome 1 using SNP 6.0 arrays. Black arrows indicate shared homozygosity blocks in all three affected. Red block arrow indicates the cytogenetic locus.

Linkage analysis was performed in order to confirm the locus of homozygosity by parametric analysis. Data from the 6.0K SNP genotyping arrays was handled and run using the easyLINKAGE-Plus software (see Materials and methods). Data from 7 members of family RA-FAM1 (IV:4, IV:8, IV:9, IV:6, IV:7, III:4, III:5) was used for the analysis. The ALLEGRO program was run using a disease allele frequency of 0.001 to calculate a multipoint parametric linkage analysis LOD score in an autosomal recessive pedigree model. The highest LOD score obtained equal to **2.65** on Chromosome 1 (Figure 3.4.5). The net SNPs number used for LOD score calculation was 10527 SNPs after removal of non-informative and Mendelian errors SNPs. The linkage interval falls between rs7542334 and rs12025220 that is equal to physical map position Chr1: 193,239,154 - 219,634,630.

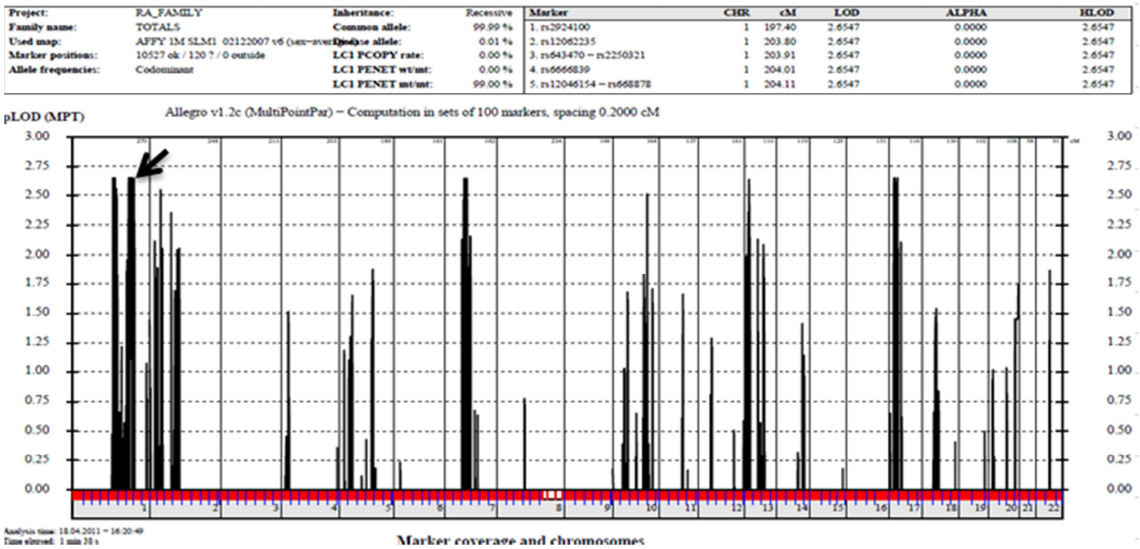

Figure 3.4.5: Graphical plots of LOD scores result from easyLINKAGE Plus software for the family with renal agenesis. The y axis indicates the LOD score value, whilst the x axis shows Chromosomes 1 through 22. The arrow indicates the highest LOD score on Chromosome 1.

Although the LOD score is not equal to 3 the value which indicates linkage, the score is high enough in the context of the number of affected (3) and the pedigree where the genetic background of one affected is not equal to that of the other two. Moreover, the large block of homozygosity in the same region support the linkage between the locus and disease causing gene.

3.4.2.5 Candidate gene screening

The region of homozygosity and linkage analysis identified in family RA-FAM1 on Chromosome 1 was large and contained numerous genes. Screening all genes in the region (using exon PCR and Sanger sequencing) would be very costly and time consuming. In such a situation, there are several approaches that could be helpful. Firstly, using a next generation sequencing (NGS) approach could be employed to screen all genes in the region. Secondly a more selective screening of certain genes in the region could be carried out, based on their known or predicted function, transcription, and phenotype, aided by software such as GeneDistiller2 software (www.genedistiller.org). Another approach would be to combine data from families with a similar phenotype (e.g. family RA-FAM2).

Initially, we adopted a selective screening of candidate genes. Using GeneDistiller2, searching for genes with a known involvement of renal/urinary system phenotype; ten genes in the Chromosome 1 region were proposed to be candidates. These genes were *ADORA1, REN, KISS1, ESRRG, RASSF5, CD55, CR2, CR1L, FLVCR1,* and *TGFB2* gene. The later (*TGFB2*) had already been sequenced, given its known role in renal development (see section 3.4.2.3). Of the remaining 9 genes, we directly sequenced the following genes *REN, CR2, FLVCR1* and *ESRRG*. These genes looked like appropriate candidates, based on their known function and expression data. Defects in *REN* gene are a cause of renal tubular dysgenesis (RTD) [OMIM: 267430]. RTD is an autosomal recessive severe disorder of renal tubular development characterized by persistent fetal anuria and perinatal death, probably due to pulmonary hypoplasia from early-onset oligohydramnios (the Potter phenotype) (Gribouval et al., 2005). Genetic variations in *CR2* gene are associated with susceptibility to systemic lupus erythematosus type 9 (SLEB9) (Wu et al., 2007). The *FLVCR1* gene encodes a member of the major facilitator superfamily of transporter proteins. Defects in *FLVCR1* are the cause of posterior column ataxia with retinitis pigmentosa (PCARP) (Rajadhyaksha et al., 2010). Estrogen-related receptor gamma (*ESRRG*) is an orphan receptor that acts as transcription activator in the absence of bound ligand. It is expressed at high level in fetal kidney and brain (Greschik et al., 2002). Direct sequencing of *REN, CR2, FLVCR1* and *ESRRG* genes did not identify any pathogenic mutations (data not shown). Because of emerging data on the ciliopathy gene *CENPF*, we also sequenced this gene in its entirety, as it falls in the Chromosome 1 region of homozygosity in family RA-FAM1 but no mutation was detected.

In summary a total of 93 amplicons per sample were amplified and sequenced by screening *REN, CR2, FLVCR1, ESRRG* and *CENPF* genes. The region of homozygosity was confirmed in all the affected patients when screened for these genes.

3.4.2.6 Homozygosity mapping using CytoScan arrays

Given the lack of success in the candidate gene approach, another approach would be to utilize NGS (exome sequencing) to screen for all genes in the putative Chr1 locus. However, at this time we received DNA samples from family RA-FAM2, also with bilateral agenesis. Using DNA samples from family RA-FAM2 plus all three affected from family RA_FAM1, we carried out homozygosity mapping using CytoScan arrays. The CytoScan arrays were used to compare regions of homozygosity in all four affected (IV:8, IV:9 and IV:4 from family RA-FAM1 and IV:4 from family RA-FAM2) and to check for any molecular karyotype abnormalities. Data analysis by ChAS software reported no chromosomal aneuploidy in affected fetuses. The single affected in family RA-FAM2 (IV:4, Figure 3.4.2) shared with other three affected in family RA-FAM1 (IV:8, IV:9, IV:4) the region of homozygosity previously identified on Chromosome 1 (Figure 3.4.6). The shared region of homozygosity from family RA-FAM1 was reduced by combining data from the family RA-FAM2 affected member. This new region of homozygosity on Chromosome 1 contained only eight genes. These are: *MIR205, CAMK1G, LAMB3, G0S2, HSD11B1, TRAF3IP3, C1orf74,* and *IRF6.*

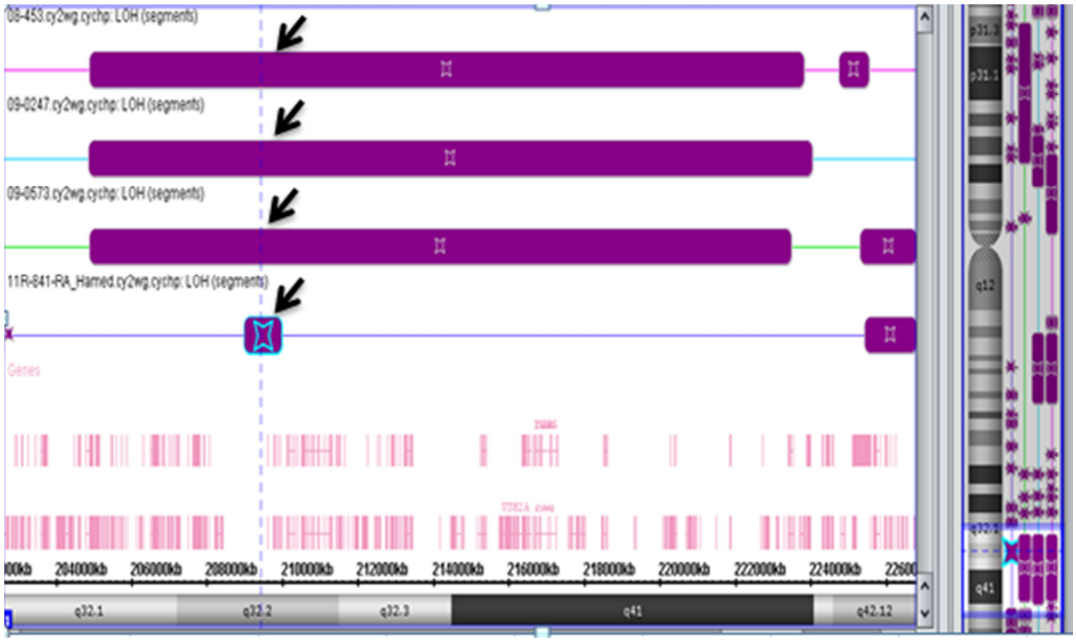

Figure 3.4.6: A screen grab of ChAS2 software for chromosome 1 for four affected in both families. Arrows indicate shared homozygosity region in chromosome 1 between all 4 affected patients of the two families using CytoScan chip array. Pink color indicates genes. Cytogenetic map is shown graphically map on right side.

3.4.2.7 Exon PCR Sequencing

Using these annotated genes (*MIR205, CAMK1G, LAMB3, G0S2, HSD11B1, TRAF3IP3, C1orf74,* and *IRF6*) within homozygosity region on Chr1 as "candidate genes" we attempted to detect the mutated gene by exon PCR, followed by direct Sanger sequencing.

Primers for all eight genes were designed using Primer-3 software for oligonucleotide primer design (Primer sequences are given in Appendix A). The entire cDNA of all genes was amplified using DNA from III:4 of family RA-FAM1 (father), together with one affected and unaffected individual from family RA-FAM1 branch A (IV:8 and IV:6), and mother and affected from branch B of family RA-FAM1 (III:10 and IV:4) in addition to mother and affected (III:2 and IV:4) in family RA-FAM2. One ethnically matched normal control was included in addition. A total of 72 amplicons per sample were sequenced. Analysis of sequencing chromatograms from all genes using SeqMan and Mutation Surveyor® software resulted in the detection of a novel homozygous missense mutation in calcium/calmodulin-dependent protein kinase IG (*CAMK1G*) gene in affected members of family RA-FAM1. At nucleotide 619 (NM_020439.2) of the

cDNA of *CAMK1G* gene the wild type Guanine replaced by mutant Adenine (c.619G>A) predicting an amino acid change from Valine to Isoleucine at position 207 (NP_065172.1) of the protein (p.V207I) as shown in Figure 3.4.7. The mutation was detected in two branches of family 1 and segregated in autosomal receive fashion where parents were heterozygous and all three affected were homozygous for the missense change. Unaffected siblings were heterozygous for the sequence variant. The mutation was confirmed by sequencing both directions in all family members. Unfortunately, no mutations were detected in family RA-FAM2. Although all affected in both families showed homozygosity, checking SNP haplotype results in chromosome 1 region where *CAMK1G* gene located revealed no allele sharing between both families (Table 3.4.2). We determined the frequency of the change c.619G>A in 190 ethnically matched normal controls from our population. This variant was not detected, suggesting that this gene may be a novel cause of renal agenesis.

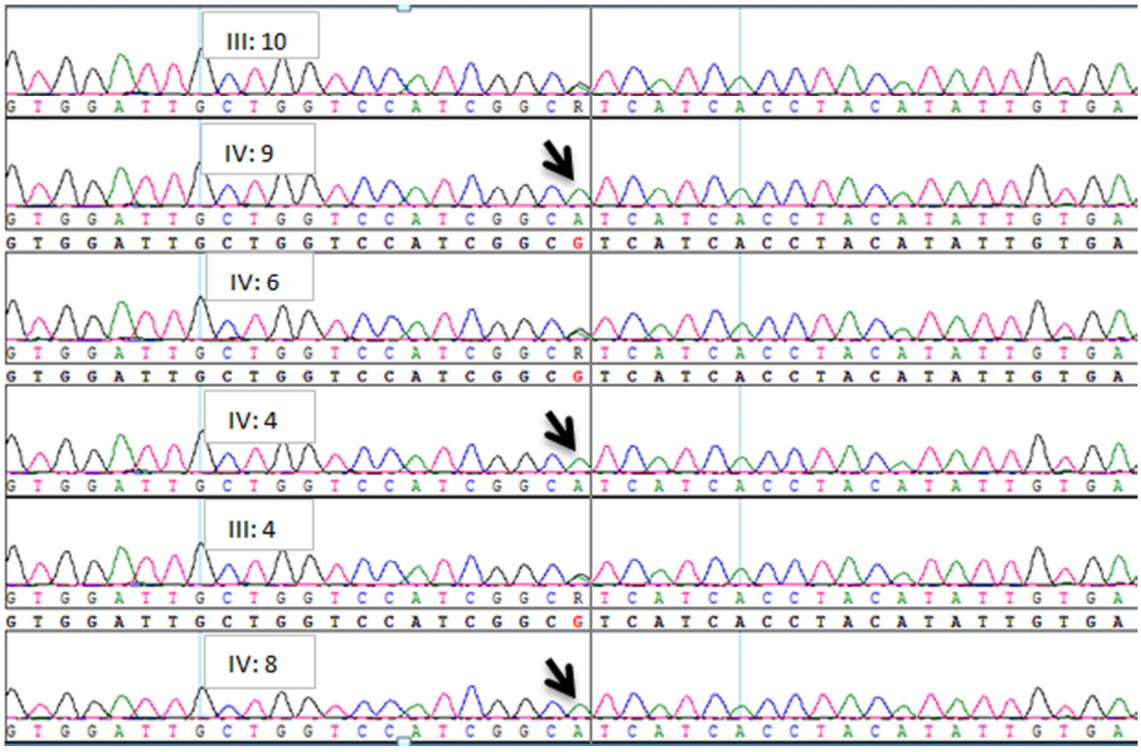

Figure 3.4.7: Sequencing chromatogram showing part of exon-7 of *CAMK1G* gene. Vertical line indicates position of nucleotide change. Arrows indicate homozygous change in 3 affected. Heterozygous parents and unaffected sibling denote R [G/A].

Probe Set ID	dbSNP RS ID	Chr	Chr position	RA-FAM1_III:10 mother	RA-FAM1_IV:7	RA-FAM1_III:5 mother	RA-FAM1_IV:6	RA-FAM1_III:4 father	RA-FAM1_IV:8 affc	RA-FAM1_IV:9 affc	RA-FAM1_IV:4 affc	RA-FAM2_IV:4 affc	RA-FAM2_III:2 mother	RA-FAM2_III:1 father
AX-11578969	rs6695896	1	209745395	BB	AB	AB	BB	BB	BB	BB	BB	AB	AB	BB
AX-11322285	rs17388121	1	209748418	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB
AX-11203994	rs12408760	1	209749159	AB	AB	AB	AB	AB	AA	AA	AA	BB	BB	AB
AX-11192837	rs12143936	1	209751951	BB	BB	BB	BB	BB	BB	BB	BB	AB	AB	BB
AX-11624336	rs742253	1	209759045	AB	BB	BB	AB	AB	BB	BB	BB	AB	AB	AB
AX-11458351	rs35104617	1	209762237	AB	AB	AB	AB	AB	AA	AA	AA	BB	BB	AB
AX-11608774	rs713074	1	209763973	BB	AB	AB	BB	BB	BB	BB	BB	BB	BB	BB
AX-11386825	rs2356934	1	209764211	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB
AX-11627758	rs7528698	1	209765087	AA	AA	AA	AA	AA	AA	AA	AA	AB	AB	AA
AX-11687074	rs9429820	1	209769630	BB	BB	BB	BB	BB	BB	BB	BB	AB	AB	BB
AX-11381262	rs2281139	1	209774889	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB
AX-11687082	rs9430004	1	209775800	AB	AB	AB	AB	AB	BB	BB	BB	AB	AB	NoCall
AX-11377147	rs2235451	1	209778052	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB
AX-11671778	rs880782	1	209779953	AB	BB	BB	AB	AB	BB	BB	BB	AB	AB	AB
AX-11319145	rs17316402	1	209785148	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB
AX-11144990	rs11119315	1	209785206	BB	BB	BB	BB	BB	BB	BB	BB	AB	AB	BB
AX-11380309	rs2272880	1	209786078	BB	BB	BB	AB	AB	BB	BB	BB	AB	AA	BB
AX-11398759	rs2566	1	209788514	BB	BB	BB	AB	AB	BB	BB	BB	BB	BB	BB
AX-11365957	rs2076227	1	209789086	BB	BB	BB	BB	BB	BB	BB	BB	AB	AA	BB
AX-11436571	rs3179860	1	209795909	AA	AA	AA	AA	AA	AA	AA	AA	AA	AB	AA
AX-11376731	rs2229466	1	209797253	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA

Table 3.4.2: SNP haplotype results for both families (RA-FAM1 and RA-FAM2) for the region in chromosome 1 that contained *CAMK1G* gene. All four affected from both families were highlighted by yellow. Genotypes in red indicate differences between two families.

3.4.2.8 in silico analysis

Because the change identified in family RA-FAM1 was novel and the fact that no pathogenic mutations have been reported before in the *CAMK1G* gene (according to HGMD (www.hgmd.org)) we performed bioinformatics analysis to verify the pathogenicity of the change.

There are many software packages that can help in assessing the pathogenicity of new sequence variants. We used several programs including PolyPhen-2, SIFT, SNPs3D and Mutation Taster. The results summary of these in silico analyses are given in Table 3.4.3.

Using in silico analyses, all four programs used for the mutation analysis concurred that the amino acid change would be predicted to be a disease causing mutation. The amino

acid change from Valine to Isoleucine at position 207 was assessed for its possible effect on the structure and function of the CAMK1G protein. Conservation of wild type amino acid Valine among different species was tested using Mutation Taster program. The results showed the conservation of the amino acid Valine among 10 species as shown in Figure 3.4.8. Mutant protein features also assessed by Mutation Taster program where indicated that the protein has had protein kinase domain between amino acids 23 and 227 which is lost by the mutation. Also a helix structure between amino acid 198 and 214 is predicted to be lost because the changed amino acid falls in this region (Figure 3.4.8).

Gene	UniProt #	Nucleotide Change	Amino Acid change	PolyPhen2 (Score)	SIFT	SNPS3D	Mutation Taster (prediction value)	Allele frequency in Saudi population	Pathogenicity
CAMK1G	Q96NX5	c.619G>A	p.V207I	Probably damaging (1.00)[1]	Damaging (0.04)[2]	(-0.7)[3]	Disease causing (0.999)[4]	0	Yes

Table 3.4.3**:** Summarized results of in silico analysis of *CAMK1G* variant c.619G>A.

[1] PolyPhen2 results range from 0-1, with highest score indicating the probability of damage of the mutation.
[2] SIFT scores range from 0 to 1, where amino acid substitutions are scored and are deemed to be damaging if the score is ≤ 0.05, and tolerated if the score is > 0.05.
[3] SNPs3D scoring follows the higher the score entropy is, the more tolerant a position is to a mutation.
[4] Mutation Taster scores mutations from 0 to 1, where a value close to 1 indicates a high 'security' of the prediction.

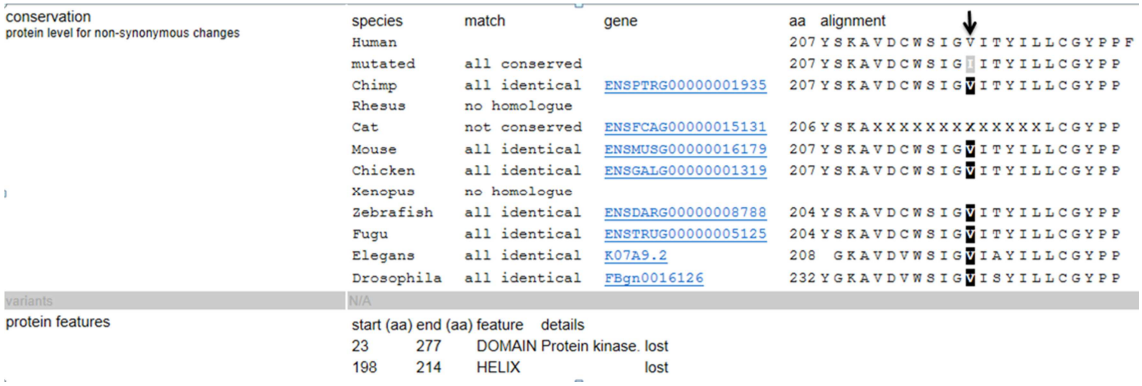

Figure 3.4.8: A screen grab from Mutation Taster program showing conservation of the amino acid Valine at position 207 in *CAMK1G* gene among different species.

At the protein structure level the protein homology/analogy recognition engine (Phyre2) program was used to check a comparison in structure and three dimensions configuration between wild type *CAMK1G* protein and mutant *CAMK1G* (p.V207I). Wild type and mutant proteins sequence of *CAMK1G* gene were uploaded. At three dimensions (3D) model of the two sequences resulted in a different configuration between wild type and mutant protein (p.V207I) as illustrated in Figure 3.4.9. In addition a helix structure between amino acids 198 and 214 was disrupted as predicted by Phyre2 program (Figure 3.4.10) in consistent with Mutation Taster prediction (Figure 3.4.8).

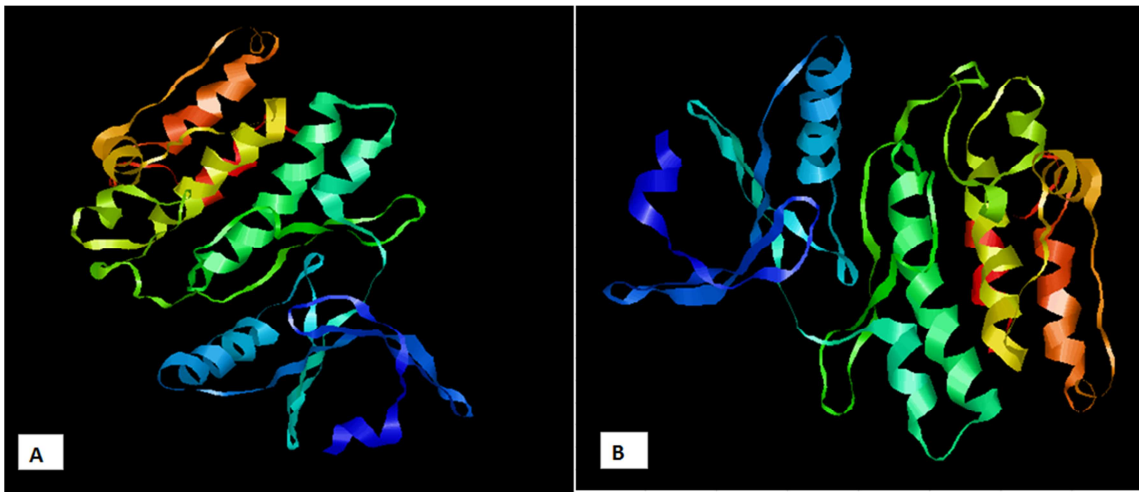

Figure 3.4.9: 3D protein configuration of wild type CAMK1G (A) and mutant CAMK1G (B).

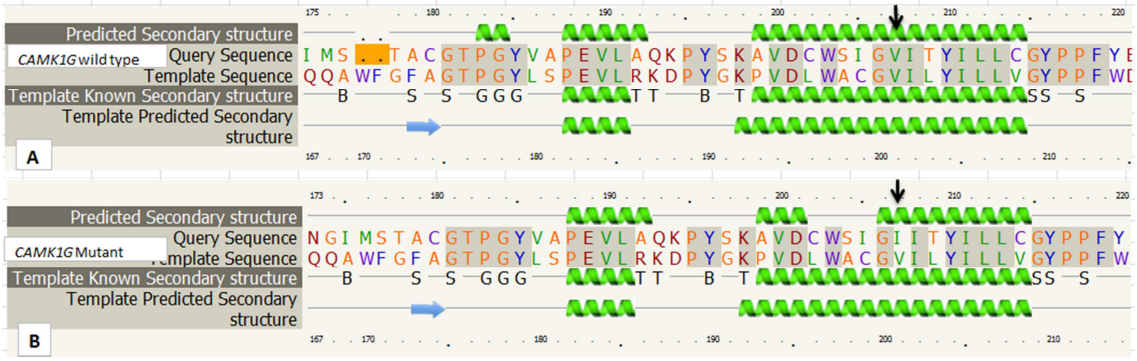

Figure 3.4.10: Part of protein structure of CAMK1G gene AA175-AA220. Arrow indicates amino acid change from Valine to Isoleucine in wild type (A) and in mutant (B). Green indicates helix.

Information about *CAMK1G* gene interactions is scarce. *CAMK1G* interactions were searched for using the SNPsD3 website (http://www.snps3d.org). This identified groups of genes where interactions may be relevant. The first group consisted of known disease genes that included *HMGCR* (HMG-CoA reductase is the rate-limiting enzyme for cholesterol synthesis), *LAMB3* (epidermolysis bullosa junctional Herlitz type disease), *HSD11B1* (cortisone reductase deficiency [CRD]), and *IRF6* (cause van der Woude syndrome and popliteal pterygium syndrome). The second group of genes predicted or known to interact with *CAMK1G* are those one with an unknown phenotype and included *SLC16A1*, *HDAC4*, *CAMK1*, *CAMKK2*, *HDAC5*, *NEK2*, and *G0S2* as illustrated in Figure 3.4.11.

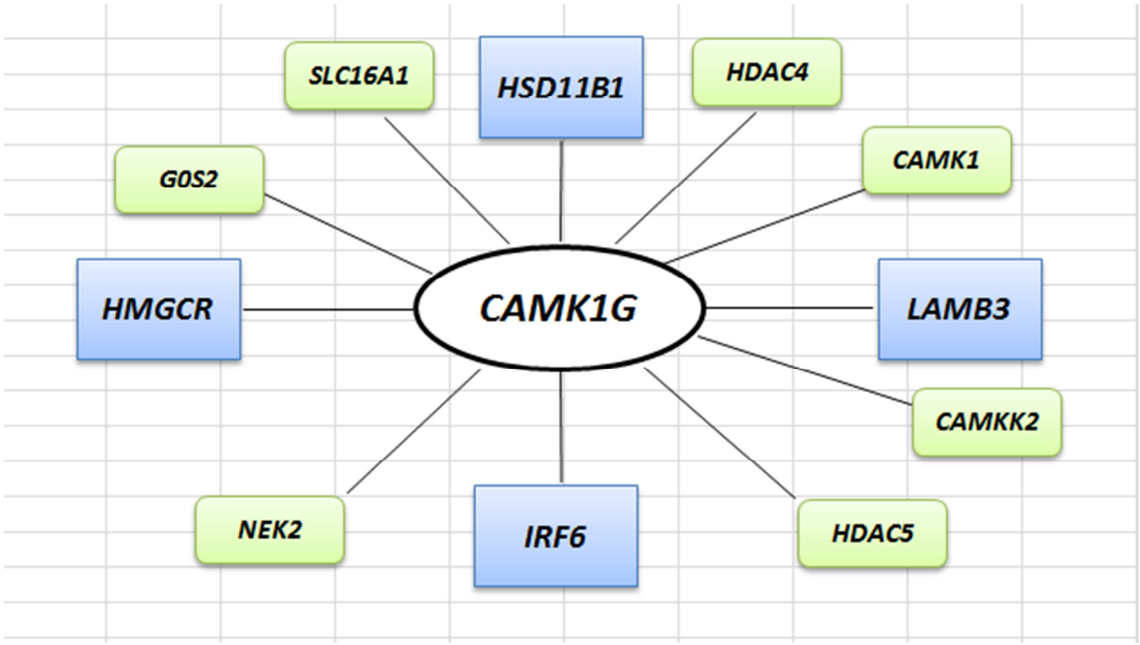

Figure 3.4.11: Illustration of *CAMK1G* gene interactions. Blue rectangles indicate known disease genes. Green rounded rectangles indicate genes of unknown phenotype.

3.4.3 Discussion

Using a large consanguineous family with renal agenesis and renal dysplasia we have identified a putative novel gene. A search for the disease causing gene used three integrated techniques, homozygosity mapping, linkage analysis and candidate gene screening. This led to the identification of *CAMK1G* gene as putative causative gene of bilateral renal agenesis in a single family. This gene encodes a protein similar to calcium/calmodulin dependent protein kinase, however, its exact function is not known. A novel missense mutation c.619G>A (p.V207I) in the *CAMK1G* gene was found in a consanguineous family with three stillborn affected by BRA. The distinctive homozygosity region between all affected, a high LOD score, and confirmed homozygosity in all genes screened in the region in addition to finding a lethal mutation segregating in all family members supported our hypothesis of a novel disease locus on Chromosome 1 and highly rule out other loci. The region of homozygosity on Chromosome 1 was apparently reduced by the addition of data from an unrelated affected fetus from family RA-FAM2. However, this may have only been coincidental, as we were unable to detect the causative mutation in family RA-FAM2. In family RA-FAM1 linkage analysis resulted in a LOD score less than 3 (LOD score $\geq$ 3 indicate linkage). This is expected because of the number of affected and genomic background of the three affected is not similar (grandmother of IV: 4 different than grandmother of

III: 13 and III: 14). We did not find a mutation in *CAMK1G* in family RA-FAM2. This may be because this family does not map to the disease locus, and contributed to the homozygosity mapping by chance especially no linkage analysis was performed for this family. Renal agenesis is likely to be heterogeneous disease and may be caused by a defect in many genes that code for cell structure or signaling pathways.

Because of the frequent stillbirth rate of BRA infants, genes causing this severe disease have not been carefully studied in humans. The putative gene we have detected belongs to CAMK Ser/Thr protein kinase family and subfamily of CaMK by similarity (http://genome.ucsc.edu). The exact function of the gene is unknown and further studies will hopefully reveal the importance of the gene. In mice, renal agenesis is commonly found in animals lacking the genes for the tyrosine kinase receptor Ret, its co-receptor Gfra1, or the Gfra1 ligand Gdnf (Schuchardt et al., 1994, Pichel et al., 1996, Cacalano et al., 1998).

However, in humans, the genes responsible for kidney development have remained elusive. Genes interacting with *CAMK1G* gene include *IRF6* gene which can lead to the autosomal dominant van der Woude syndrome (VWS) (Kondo et al., 2002). The *IRF6* and *CAMK1G* genes are located nearby in terms of physical position at same homozygous region of Chromosome 1. De Medeiros and et al reported a novel mutation in *IRF6* resulting in VWS-PPS spectrum disorder with renal aplasia (de Medeiros et al., 2008). The main phenotype of VWS is cleft palate and this is seen in some of stillborn fetuses in family RA-FAM1. These data could support the hypothesis of interaction between these two genes in addition to other genes in the region such as *HSD11B1* and *LAMB3*.

Bioinformatics analyses are widely used for new sequence variants. Many molecular genetics approaches are enhance by bioinformatics analysis. Here, we used homozygosity mapping and linkage analysis software packages to aid us to identify a putative novel locus. Following the identification of sequence variant, mutation prediction software programs helped in identifying the change as pathogenic. Protein fold recognition software (Phyre2) also predicted a defect in mutant protein.

The use of a gGene interactions predcition program helped in giving a very preliminary idea about the possible functions of the gene. It is important to note that all these in silico models are modest support for pathogenicity of any sequence variant and experimental confirmation will be required to determine in vivo effects.

Further studies are required to determine *CAMK1G* gene function and the localization of the encoded protein.

Gene expression may help in identifying the effect of the mutation on mRNA expression in developing renal tissues. Western blot analysis would be required to determine the effect of the missense mutation on peptide translation although this is not a truncating mutation. Localization of the protein in renal epithelial cells (for example) could be studied to better understanding of the function of the protein. Cells transfection with wild type and mutant CAMK1G may give some clues as to protein mislocalisation. Animal models are another approach that would highlight the effect of knockout gene and its phenotype. Interestingly, mice homozygous for a knock-out allele of *CAMK1G* exhibited impaired dendritogenesis (http://www.informatics.jax.org/).

In conclusion, we described a putative new gene underlying bilateral renal agenesis in a consanguineous Saudi Arabian family. Further detailed analyses on the gene and its encoded protein will need to be conducted to confirm our findings at experimental level

Chapter 4
CONCLUDING DISCUSSION

The human genome has around 22,000 genes and around 180,000 coding exons. There are at present 3500 known simple disease causing genes. There are 1800 genes with an unknown molecular basis, and 1900 further genes suspected to be Mendelian inheritance genes. This is points to the fact that there are more genes yet to be discovered and mechanisms of genes function yet to be investigated. It seems that our understanding of the genome is still at the beginning of reading letters and making basic sense of words.

Our study demonstrates recent challenges in the field of genetics. Each chapter had its own challenge, from the mutation detection rate in monogenic genes causing nephrotic syndrome to variants of unknown significance in novel genes as in our families with renal agenesis.

Because autosomal recessive (AR) diseases have strong genotype-phenotype correlations, the diagnostic approach of AR disease is a very powerful tool to manage and detect monogenic kidney disease. Our study investigating genes causing NS in the Saudi population was aimed to detect the frequency of mutations in known genes causing the disease. Seventeen mutations in 5 genes covered 47% of patients with NS. On the other hand, the 53% of patients who were negative for mutations in known NS genes are a potentially valuable cohort for the identification of new genes causing NS.

A functional assay for investigating the novel *BBS5* mutation employed an animal model, the zebrafish. The morphant phenotype that was noted in the animal model as consequence of *BBS5* knockdown resembled the human disease. Animal models are useful to generate a picture of the disease phenotype but still limited in explaining interaction mechanisms within the cell. On the other hand, *in vitro* cell line experiments can reveal molecular mechanisms but do not show the whole picture compared to that of a living organism. Functional assays should be an integral part of the investigation in all

genetics studies for better understanding of human genome interactions and mechanisms of disease.

For the potentially novel cause of autosomal recessive distal tubular acidosis and nephrocalcinosis, we identified a putative locus. Finding the causative gene in this family remains a challenge. It is now clear that mutations within coding exons are not the only cause of human genetic diseases. Variants within regulatory elements such as promoters and enhancers, in addition to variants in introns may play a role in their pathogenicity. New generation sequencing (NGS) may overcome Sanger sequencing's limitations and help to reveal the importance of variants outside coding regions. Other landmarks in the genome such as repetitive segments and methylated regions should be investigated as well, as they may contribute to our understanding of the genome.

Proving the functional importance of a potentially novel missense mutation in a family with bilateral renal agenesis remains challenging. This is now the case with many variants of unknown significance resulting from exome sequencing technology. The technology has added much genomic information that we are still unable to put it in context of genome function. Not only that, this new technology has diverted our attention of understanding gene function *in vivo* to bioinformatics software in order to handle huge data and predict the effect of defective genes. Trends of using *in silico* gene prediction software instead of *in vivo* experimentation is now expanding and this may affect our imagination to develop new *in vivo* assays that would help in our understanding of the genome.

In spite of all the challenges mentioned before, this study establishes a solid basis for understanding some of the inherited kidney diseases in Saudi Arabia. The new findings include novel mutations in nephrotic syndrome genes and BBS genes. The study also identified a putative novel locus in a new autosomal recessive distal tubular acidosis and nephrocalcinosis syndrome, and a potential novel gene causing bilateral renal agenesis. These studies add valuable information to the scientific community and support other research to investigate more in renal disease in Saudi Arabia.

Appendix A

List of oligonucleotide primer sequences used in the thesis

Chapter 3.1 Nephrotic Syndrome

NPHS1			
NPHS1-Fw_ex1&2	Ggacccagagaaagccagac	*NPHS1-Rv_ex1&2*	GGGGAGCCGGAAGTCAG
NPHS1-Fw_ex3&4	GATCAGAGCCTGAGACCCTC	*NPHS1-Rv_ex3&4*	TCAGGAACACACACCCTTCC
NPHS1-Fw_ex5	TCCCCAGAATCTATCTTGCG	*NPHS1-Rv_ex5*	GGTTCCCATGGGGAAAATTAG
NPHS1-Fw_ex6-7	GACTCCCCAAATTTCAGATG	*NPHS1-Rv_ex6-7*	AGCTCAGGACTGGCTCCC
NPHS1-Fw_ex8	ACAGTGGGGTCTGGGAGC	*NPHS1-Rv_ex8*	CCAGTAGGCATAATTTGGGG
NPHS1-Fw_ex9	GGGCTGGTCTGTGAGAAATC	*NPHS1-Rv_ex9*	CCCTTCCCTATCCACGAGTC
NPHS1-Fw_ex10-11	TAGCACGATGGATAGGGGTG	*NPHS1-Rv_ex10-11*	TGCCCAAGATTCTGAAGGAG
NPHS1-Fw_ex12	CCAGTGGGCAGGGTAGG	*NPHS1-Rv_ex12*	acctggctctgtccctcc
NPHS1-Fw_ex13-14	ggtggACGCAGACCAGG	*NPHS1-Rv_ex13-14*	AGACCCAAGGAGTAGTTTAGGG
NPHS1-Fw_ex15-16	Cctgtgcctgatctccaatc	*NPHS1-Rv_ex15-16*	GAGACTCCACAATGGGCAAG
NPHS1-Fw_ex17	AAGACATCCCTCCCACCTG	*NPHS1-Rv_ex17*	caagAAACCACAACCCCTTG
NPHS1-Fw_ex18-19	GGGCAGTGATGGATCTGG	*NPHS1-Rv_ex18-19*	GACTCAGGGAGGGGAAGTG
NPHS1-Fw_ex20	CAAGCATAAGTGGTTGGGTG	*NPHS1-Rv_ex20*	ATCAGGGATGTGGGAATGG
NPHS1-Fw_ex21-22	GGGAAAACCTGGACAGAATC	*NPHS1-Rv_ex21-22*	ttcaccatactaccctacaCATCC
NPHS1-Fw_ex23	CATGAATCTAATAGGCTTAAGAAGAGG	*NPHS1-Rv_ex23*	CAAGCAGAGGAGGTAGGGTC
NPHS1-Fw_ex24-26	TTCTCGGGGAGACCCAC	*NPHS1-Rv_ex24-26*	CTAACGGCAGGGCTTCAG
NPHS1-Fw_ex27-28	AGGTTGATCATTGCCCTTCC	*NPHS1-Rv_ex27-28*	ATAGGAGGTAGGCTCCCAGC
NPHS1-Fw_ex29	TTCTGGGAAGTTAAGCAGGG	*NPHS1-Rv_ex29*	TGGGTTTTATGGAGCTCACC
NPHS2			
NPHS2-Fw_ex1	ACACAGCTGTCGGGGTTC	*NPHS2-rv_ex1*	CTCTGGCTTCAGTGGGTCTC
NPHS2-Fw_ex2	GGACCAACAGATGCTAGTCAGTG	*NPHS2-rv_ex2*	ggtgtgGCCAGTGAGAGG
NPHS2-Fw_ex3	TATGCCAAGGCCTTTTGAAG	*NPHS2-rv_ex3*	CATGGGTTGAAGAAATTGGC
NPHS2-Fw_ex4	AAAGGTGAAACCCAAACAGC	*NPHS2-rv_ex4*	tcattttgtcCACGGTAGGTAG
NPHS2-Fw_ex5	ACCCCACATAGGAAAGGAGC	*NPHS2-rv_ex5*	GGGATGGAACTGGCCATAG
NPHS2-Fw_ex6	AGGGTTTAGGCATGCTCTCC	*NPHS2-rv_ex6*	TGGCTGTAAGATATTAGGTGATTTG
NPHS2-Fw_ex7	TTTGGCCCCTAAATCATGG	*NPHS2-rv_ex7*	GGACAGTAAGGAAGCAAAGGG

NPHS2-Fw_ex8-1	Gatgctcagtgcttgtctgc	*NPHS2-rv_ex8-1*	TGTTGACCAAAGCTGTTTCC
NPHS2-Fw_ex8-2	CATGGAATGCACGAAGTCAC	*NPHS2-rv_ex8-2*	TCTCCCTCAGATTTTAAGCCAC
PLCE1			
PLCE1-Fw_ex1	AAGTTGCAGAGTGCAATCCC	*PLCE1-rv_ex1*	AATCTGGCATGATTTTCTCCC
PLCE1-Fw_ex1-2	TGGAAGCAACTTGCCAAAG	*PLCE1-rv_ex1-2*	CAAAAGTATTTATCATTCAAGGTGTC
PLCE1-Fw_ex1-3	AATTGTGTACCACTACCTGGGG	*PLCE1-rv_ex1-3*	AATTGTGTACCACTACCTGGGG
PLCE1-Fw_ex2	TTGCACTTGGAGCATCTGAG	*PLCE1-rv_ex2*	GTGAACTTAATTTTCCATCAGGAG
PLCE1-Fw_ex3	TATCTGTTGCAGGGGACACC	*PLCE1-rv_ex3*	GGGGAGGTATGCTTGAAAATC
PLCE1-Fw_ex4	cccagccAGGACCTACAG	*PLCE1-rv_ex4*	CTGGCTTGGGTAAAGGTGAG
PLCE1-Fw_ex5	AATTTAGGCTCCTTGCTGTAAAC	*PLCE1-rv_ex5*	GACAGGGCCTGATGAGACAG
PLCE1-Fw_ex6	GCCAAGTATGTTATCCAGGC	*PLCE1-rv_ex6*	AAGTAGGAAACAACTCATCAACAGG
PLCE1-Fw_ex7-1	CAGTGAAAATGTCTTTGGGTTG	*PLCE1-rv_ex7-1*	AGATCCAAGACCCCTTCCG
PLCE1-Fw_ex7-2	CTCCCCAGCCAGCAGTAAAG	*PLCE1-rv_ex7-2*	tcacccagtgtgtgctcttg
PLCE1-Fw_ex8	CATTAATATATTTACTTCCCCATTGC	*PLCE1-rv_ex8*	TCACCATACAGCGGTCCTG
PLCE1-Fw_ex9	GCAGTATTGAAGGTGGGTAGG	*PLCE1-rv_ex9*	AATGTTTGCAATGCTTAAATCAC
PLCE1-Fw_ex10	GGTTTGTATTTGGTTCTGAGGG	*PLCE1-rv_ex10*	ATATGCCCTTCCACACAACTG
PLCE1-Fw_ex11-12	AATTGGAGCAAGTCTGGTGG	*PLCE1-rv_ex11-12*	AGCACACTATGTGGGACAGG
PLCE1-Fw_ex13	ACTTTGCAAAAGATGCTGGC	*PLCE1-rv_ex13*	TCCCAGATTTAAAGGCTTTGG
PLCE1-Fw_ex14-15	GATGTGGTGGTTTCTTTCCC	*PLCE1-rv_ex14-15*	GCTCTTCAAAGGAGTCTGGG
PLCE1-Fw_ex16	TTCGAGGCTTTATCTCCAGG	*PLCE1-rv_ex16*	aacagagcgagaacccg
PLCE1-Fw_ex17	CTACCCTGCCTTCTGACCAC	*PLCE1-rv_ex17*	GACAATTGCAAAACAAGGGC
PLCE1-Fw_ex18	GCAGGGGACAGCTTCTTTC	*PLCE1-rv_ex18*	GGTCTATGGAATGAAAAGCCC
PLCE1-Fw_ex19	ACGATTGTGTTAAACATCAGGG	*PLCE1-rv_ex19*	ATTACTGGTCTTTGGCGCTC
PLCE1-Fw_ex20	GCTCTTAAGTGAAGCATTGGG	*PLCE1-rv_ex20*	ACAGGGAGCAAGTGGAATC
PLCE1-Fw_ex21	Aaaagagctttgggaatccag	*PLCE1-rv_ex21*	TGAGAGTGTTCACAATGCCC
PLCE1-Fw_ex22	ATGTTGAGTTGCCTTGCTTG	*PLCE1-rv_ex22*	AGGGAAGTGCTTAGACAGTAAAATATC
PLCE1-Fw_ex23	TGCTATGACTGTTTACTGGGATG	*PLCE1-rv_ex23*	CAGTCTGGAATCTTGCCCC
PLCE1-Fw_ex24	TAAGGTGATACTTAAGCGACAAAG	*PLCE1-rv_ex24*	CCAGTTCGTTTTCAGGAAGG
PLCE1-Fw_ex25	AAGCAGTGAGGTGCAGAGG	*PLCE1-rv_ex25*	aaattctggcatccatcacc
PLCE1-Fw_ex26	TTCCAGGAGCATCTTCTTTTC	*PLCE1-rv_ex26*	AAGCCCACCTTCAGTGTAGC
PLCE1-Fw_ex27	GCACCTCTGTATCAAATAGAGCTTAG	*PLCE1-rv_ex27*	ACACACAGGCATTTATACACACAC
PLCE1-Fw_ex28	TGAACACCATGAAAGTTGATTTG	*PLCE1-rv_ex28*	TTCAGGGAGGTTGTGAGTGG
PLCE1-Fw_ex29	CATCACCAAGATACAAGCTCAG	*PLCE1-rv_ex29*	GAACTTCACTCATAAAATGAGCC
PLCE1-Fw_ex30	GAATGAATGCAAATGTTGGAG	*PLCE1-rv_ex30*	TCCTAAATTTCACCAGCTTCC
PLCE1-Fw_ex31	GTCATGTGACAGAGGAACCTG	*PLCE1-rv_ex31*	AAAAGACATTGTACTAATTATGCCTTC
PLCE1-Fw_ex32	TGGCTGGAAAACTCTGAACC	*PLCE1-rv_ex32*	atggcccctgaattaagaatag

PTPRO			
PTPRO-Fw_ex1	AGTTCGCCATTGTGAGCC	*PTPRO-rv_ex1*	ACACGCGAGGACAACAAAG
PTPRO-Fw_ex2	TTAGCATAACTTTATTGCCAATTATC	*PTPRO-rv_ex2*	AGCCCTACGGGAAGAAACAG
PTPRO-Fw_ex3	ATCGGCCTTTCTCTCTCCTC	*PTPRO-rv_ex3*	GAGAACATCATTTGGGAAAAGC
PTPRO-Fw_ex4	AAAGCCCTTCAGGATTGAGTC	*PTPRO-rv_ex4*	GGGTAAGTCAGTTAATTATAACCTTTG
PTPRO-Fw_ex5	TCTGCTTTTAGAAGTGGTTGAATTAC	*PTPRO-rv_ex5*	TCATTTGGGTTTTCTGCTCTC
PTPRO-Fw_ex6	TCTATACCCAGGGCCTTTCC	*PTPRO-rv_ex6*	TCCCACCAGTTCCATTGTTC
PTPRO-Fw_ex7	GAATCTTTTATCTCAATCAGTGTAACG	*PTPRO-rv_ex7*	ATTGCCCTGTAAGGTTGTGC
PTPRO-Fw_ex8	TCTGAATATTATACCAGCCTCTGTG	*PTPRO-rv_ex8*	ACCCCTCCGTCAATACCC
PTPRO-Fw_ex9	GGCCATTGCTAAGACATTCC	*PTPRO-rv_ex9*	CATAAAGGTTTGCCTGTTCCC
PTPRO-Fw_ex10	ATTGGAAATCTTTGGCTCTATATC	*PTPRO-rv_ex10*	TGAAATTGATGCTCACCCAG
PTPRO-Fw_ex11	TATGGGATTTCATGCTCTGC	*PTPRO-rv_ex11*	gacgtgtcaataaatgattgcttag
PTPRO-Fw_ex12	GGTGGTGCACTGATTCATTC	*PTPRO-rv_ex12*	GAGGAGCTCTTGCTGAATCG
PTPRO-Fw_ex13	GGGGATGCTTCACCTGC	*PTPRO-rv_ex13*	TTTTGCCCAGATAATTAGCCTTAC
PTPRO-Fw_ex14	GTCAGATGGGGAATGGGG	*PTPRO-rv_ex14*	aagcttgggatttctTGGTG
PTPRO-Fw_ex15	TCATCGTAAGCTCACTGCCC	*PTPRO-rv_ex15*	TAAAATCCAATATACCGCTATACAAC
PTPRO-Fw_ex16	GGTCAACGTAAGAATGCTTTGTC	*PTPRO-rv_ex16*	AATCGACTTTGGAGCACACAC
PTPRO-Fw_ex17	CGCACATTTATATTTCACAGAATG	*PTPRO-rv_ex17*	ACCGGTAACTGCTTTAATTCAC
PTPRO-Fw_ex18	CAAATGGGAAGTATAAAGTGCAG	*PTPRO-rv_ex18*	GCCTCTTTCTTAAGCGTCCC
PTPRO-Fw_ex19	TCTATTTTCTCCAAGAAGATTAAGAAG	*PTPRO-rv_ex19*	GACATGACCGCAACACAAAC
PTPRO-Fw_ex20	aaagggcttgTGAGCAATAAAC	*PTPRO-rv_ex20*	GGTATGGCCAGAGATTGGTG
PTPRO-Fw_ex21	TTTAACTTAATATTGGATAGAAAGGGC	*PTPRO-rv_ex21*	TTGGGAAAGGCTCTCTTTTG
PTPRO-Fw_ex22	TGAAGCAGCATTTGGTGTTG	*PTPRO-rv_ex22*	CAGCAAGCATTTCATCAACG
PTPRO-Fw_ex24	TACTGACCAGGAGTGGAACG	*PTPRO-rv_ex24*	TGGTGAGCTGTTCCCCTTAC
PTPRO-Fw_ex25	Acttgaagccgagaggcag	*PTPRO-rv_ex25*	GCCTCAAAGCAGTTTGAACATATAC
PTPRO-Fw_ex26	TGTGTCTCTGTAAATAATGAAAATGG	*PTPRO-rv_ex26*	TTTGTCCCATCCATCACATC
WT1			
WT1-Fw_ex8	CCTTTAATGAGATCCCCTTTTCC	*WT1-Rv_ex8*	GGGGAAATGTGGGGTGTTTCC
WT1-Fw_ex9	CCTCACTGTGCCCACATTGT	*WT1-Rv_ex9*	GCACTATTCCTTCTCTCAACTGAG
CD2AP			
CD2APex1-Fw	GATGGAGGCGACTCTTCG	*CD2APex1-rv*	AGAAAGGGCACAGCCTCC
CD2APex2-Fw	CAGCCTTGCTGATGTCTTTG	*CD2APex2-rv*	TATCCAGCAGGACACAGAGC
CD2APex3-Fw	GCAGAGCAAAGGTTGGAAAC	*CD2APex3-rv*	AAATTAAGATTGACTACACAGGGC
CD2APex4-Fw	TTATATAAGTCCTGTGACTAAACATGG	*CD2APex4-rv*	TTTCTAAAATCCATAGCATCAATTACC

CD2APex5-Fw	CCCATAGGGTGACGATTTTC	*CD2APex5-rv*	GACTAATTGAAGCTGACTCAAAG TTC
CD2APex6-Fw	AGGATGTCAAGCGTTTTGTG	*CD2APex6-rv*	CAGCCTCCTCATTCATAACC
CD2APex7-Fw	TTAGGGAAAGCTGGGTAACTG	*CD2APex7-rv*	TGTAACAAATGCTTCCAACAAG
CD2APex8-Fw	TAATGAAATTGTTACTTAACATA CGG	*CD2APex8-rv*	acgtgccaagttttcaggac
CD2APex9-Fw	Ctggccaacagagcaacac	*CD2APex9-rv*	CAAAAGAATAGTGGCAAAATGT TC
CD2APex10-Fw	TGTCAAGGATTGAATGAAGAGG	*CD2APex10-rv*	TTTGCATTAGATCCAATTTTACC
CD2APex11-Fw	Aattccggtacattgttctttc	*CD2APex11-rv*	CTCCCAGAAAACTCCAAGC
CD2APex12-Fw	CTCTGTCACACTTTTCCAGCC	*CD2APex12-rv*	GAGCCATTAAGGTCAAAGGTG
CD2APex13-Fw	AATCGGTATTAGGCCATACAATC	*CD2APex13-rv*	TCAACTGTGTAATTTCAACTTTAT CC
CD2APex14-Fw	GCACAGGTCAATTTGTGAGTTC	*CD2APex14-rv*	AATACCAGCATTTGCCACTTG
CD2APex15-Fw	TGCTTTATTATCCAAAGACTTACT ACC	*CD2APex15-rv*	TTCCCTATGTTTGATGTCCAC
CD2APex16-Fw	TCCAAATTTGCCACTTATCTTG	*CD2APex16-rv*	gaactaccacacccagccAC
CD2APex17-Fw	TCCCATGAACTGGTAGGTTAGG	*CD2APex17-rv*	AGGAACCGATTACCCAGTTG
CD2APex18-Fw	GCATGGGAAACTGGATTTTAG	*CD2APex18-rv*	GCTGAAGTTCATAGCGTTGC
LAMB2			
LAMB2-ex1Fw	CCCTCGTATAGGTGTGGGC	*LAMB2-ex1rv*	GGGCAACTACCAGGACCTAC
LAMB2-ex2-3Fw	GAGGTGGAGCCACTGGATAG	*LAMB2-ex2-3rv*	ACAGGAATCGGAAGTCAAGG
LAMB2-ex4Fw	TGGGGCAAAAGTGAGCAG	*LAMB2-ex4rv*	cTCTGTCCAGGGGCAATG
LAMB2-ex5-6Fw	TCTTGATGGCTCTGGGTTG	*LAMB2-ex5-6rv*	ACTGAGGCCCCAAATAGTCC
LAMB2-ex7-8Fw	GGACATGGGGCAGCAAG	*LAMB2-ex7-8rv*	AGGTTCTACCCAGGGCACAG
LAMB2-ex9Fw	ACCCTTGTGACCTGGTTCC	*LAMB2-ex9rv*	CTCTCTAGTTCCTGCCCCAG
LAMB2-ex10-11Fw	GGACTGGCAGTGAGCTAGTTG	*LAMB2-ex10-11rv*	ATAGATGTGACACCCCAGCC
LAMB2-ex12-13Fw	AGACGTGTGGTTGGGATTTC	*LAMB2-ex12-13rv*	CCAGTCATAGACGTGTCCACC
LAMB2-ex14-15Fw	CTGTGGCAGAAGAGGCAAAG	*LAMB2-ex14-15rv*	CCTGCAGGGAGCTTAGGG
LAMB2-ex16Fw	CCAGATGATTTGTACAGGGC	*LAMB2-ex16rv*	TATCCCTCAAGTCCCACACC
LAMB2-ex17-18Fw	CTCTGGCAATGCCTACGTG	*LAMB2-ex17-18rv*	AGGTCCAGAAGGAGGAGAGAG
LAMB2-ex19-20Fw	CCACCCACACTCCCCTC	*LAMB2-ex19-20rv*	CCCACGGTTAAGAGGAAGC
LAMB2-ex21Fw	GTGGGATCAGGGTGAGTTG	*LAMB2-ex21rv*	GCACAGTAGTCAAGAGGAGGC
LAMB2-ex22-23Fw	AGTGTGGGAGGGCATGG	*LAMB2-ex22-23rv*	GGTCACTGTGGGGCATTC
LAMB2-ex24Fw	GGGTAGATGGAGGACTTCCC	*LAMB2-ex24rv*	TTCAGCCCTGGTCCACTG
LAMB2-ex25Fw	GAGGCCCAGTGGACCAG	*LAMB2-ex25rv*	GGCACCTAAATTGGGCAGAG
LAMB2-ex26Fw	ACTAGGCAGGTGGACCAGG	*LAMB2-ex26rv*	CAGGCACAGAACAAGTCAGG
LAMB2-ex27Fw	CAGGTGGGTGTAGATGTTCC	*LAMB2-ex27rv*	GGAGATACAGACACCGGGAG
LAMB2-ex28-29Fw	TGAGTCCATACCCCAAAACTG	*LAMB2-ex28-29rv*	CTGCCATCCTCTCCTGTACC
LAMB2-ex30-31Fw	CAGTGGGACAAGGTTATGGG	*LAMB2-ex30-31rv*	CTAGGAAGGGCAGGTGTCAG
LAMB2-ex32Fw	TACAGGGTAAGAAATGGGGC	*LAMB2-ex32rv*	CCTCGAGGTTCCCCTGG

MYO1E			
MYO1-ex1FW	AACTTTTGAAGTTCGCCCAG	*MYO1-ex1rv*	CGTCCACCTTCTCCACCC
MYO1-ex2FW	TTTGGGAAATTAACAGCATTTTG	*MYO1-ex2rv*	acacacgtgagccactgc
MYO1-ex3FW	TGGACAGTCATTGTGAATTCTTG	*MYO1-ex3rv*	ATGTCTGCATCATAAAAGCCC
MYO1-ex4FW	GGAAAAGAAATTGAGTGTATTGCC	*MYO1-ex4rv*	AAACCACTGCTTATCGACCG
MYO1-ex5FW	AGACAAGGAAGGGAGCTGG	*MYO1-ex5rv*	TGGAGCCTCTTACATTTCCC
MYO1-ex6FW	TGCATTCAGAAACCAGGTCC	*MYO1-ex6rv*	AAGGCTCCCATTTCCTGTG
MYO1-ex7FW	ATTCAAAGCCTGTTTGGGAC	*MYO1-ex7rv*	TCCTCCCTGGCTTCCTAGAC
MYO1-ex8FW	AGGATGCAGGAGTGACTTCG	*MYO1-ex8rv*	TTTTGGCCACAGGAAATGG
MYO1-ex9FW	CAGCTGCTGTGCCTTGG	*MYO1-ex9rv*	TCCCAATATTCCCATTTGAG
MYO1-ex10FW	TGTGCACATAAGTGCCTCC	*MYO1-ex10rv*	cgcacccagccTACTAGTTTTAC
MYO1-ex11-12FW	CCGTGTAGAATCCAGCTTCC	*MYO1-ex11-12rv*	GGAACCGAAATCTATTAACAAAAGAC
MYO1-ex13FW	TTTGTCCATTCAGGGTCTGTC	*MYO1-ex13rv*	TCAGAGTTGTCACTTTGCCTG
MYO1-ex14FW	CCTTGCCTAAGAAAGGACCTG	*MYO1-ex14rv*	TGAATCAATAAAATACGGCGAG
MYO1-ex15FW	GGGTAGCCATGACAGCTTTG	*MYO1-ex15rv*	taggaacacaccaccacacc
MYO1-ex16FW	AGGCTTTCAGCCCCAGAG	*MYO1-ex16rv*	AGGGGTGCAGTTCCTTACTC
MYO1-ex17FW	CAAGTCTCCTTTGAAATTTGGTC	*MYO1-ex17rv*	GGCTCATCTTGAGCAGAGG
MYO1-ex18FW	TTGGCTGGTAAGAACACGAG	*MYO1-ex18rv*	actcgtacgcttgaagtggg
MYO1-ex19FW	GTAGAAAACCAGTGCGGTCC	*MYO1-ex19rv*	aagccagctgaggggtg
MYO1-ex20-21FW	GTCTAAACTGAGGGTCCCTGG	*MYO1-ex20-21rv*	acaacgaacacattctgatttg
MYO1-ex22FW	Aggtgtgagccaccatgc	*MYO1-ex22rv*	CTTCCTCCTGGCTGTTTGG
MYO1-ex23FW	TGTAAAATCTTAGCACTATCTTGCAG	*MYO1-ex23rv*	AGCCAGACCCCACTCTTCTC
MYO1-ex24FW	TGAGCACCGTTTTCTCACTG	*MYO1-ex24rv*	TTTGCTCGGGAGACTCGG
MYO1-ex25FW	AGGATCCATAAAACCAACCAG	*MYO1-ex25rv*	TGCATTGATTTGCTAATGGG
MYO1-ex26FW	GAAATGTGGAGAGACTTTTGGC	*MYO1-ex26rv*	TTGAATGATGGAGATGGAGC
MYO1-ex27FW	ATGCTCCAAACCACCGTTC	*MYO1-ex27rv*	GATCTTACAGCAATGTGACTGC
MYO1-ex28FW	CATGGATGGAGCCAAATCAC	*MYO1-ex28rv*	CATTGTGGATTGTAAGGGGAG

Appendix A

List of oligonucleotide primer sequences used in the thesis

Chapter 3.2 Bardet–Biedl Syndrome and Modeling of a Novel BBS5 Mutation

BBS5			
BBS5-Fw_ex1	GCCTTGGAGCCAGAGAGAC	*BBS5-RV_ex1*	GGGTCCCTAACGCGCAC
BBS5-Fw_ex2	GAAGACCTTATTTCTAAATGCATGAAC	*BBS5-RV_ex2*	ACACTGACAAATGATGCTGTTTC
BBS5-Fw_ex3	AAGTTCCATATCCATCTCATTCAG	*BBS5-RV_ex3*	TGCAAATTCACCTGACAAAG
BBS5-Fw_ex6	ggcttgggcaaggtattgag	*BBS5-RV_ex6*	AAAAGTAATACGCTTTCCTATAATGC
BBS5-Fw_ex7	TCAATTTTGGTATTATTTGCTGC	*BBS5-RV_ex7*	GGGAGAAAGAACTCATCTAAAAGTC
BBS5-Fw_ex8	GCAGTTTAATGAAGGAAGGAGG	*BBS5-RV_ex8*	TCTGGGGACAATTTCTGTATGAC
BBS5-Fw_ex9	tgaagctggatgatggctc	*BBS5-RV_ex9*	ATCCAGGAGCTATGCCCAC
BBS5-Fw_ex10	TCTCTTTCTGGTTCTTTTCGTG	*BBS5-RV_ex10*	cacctggccTCAAAGTATTTC
BBS5-Fw_ex11&12	TCTGCATATTAAGTTCCTTAGCCC	*BBS5-RV_ex11&12*	TCCATGACTTATGGCAGGTG

Appendix A

List of oligonucleotide primer sequences used in the thesis

Chapter 3.3 A New Syndrome of Autosomal Recessive Distal Renal Yubular Acidosis associated with Nephrocalcinosis

CA2			
CA2-Fw_ex1	GGGAGCCTATAAAAGCTGGTG	*CA2-RV_ex1*	CGGTAAACAGCATGTGCG
CA2-Fw_ex2	GGTCTGGGTGTACACTTTCCC	*CA2-RV_ex2*	TGGCTCCTTCCATTAAGCTC
CA2-Fw_ex3	AGGCATGTGTTTCatgtgtg	*CA2-RV_ex3*	TTCACTGGGCGTGAGTTG
CA2-Fw_ex4	GACCATTGAATAAAATCTGTCAGC	*CA2-RV_ex4*	TCCTGTATGTTACAAATACTC TGTCTG
CA2-Fw_ex5	AGTGGAGAATTTGGGCTCAC	*CA2-RV_ex5*	AACTGCTCATCCAAACACCAC
CA2-Fw_ex6	TGTGTCTGCTGCTCTCCTACC	*CA2-RV_ex6*	ccaaggtcttCTGCCTTGAC
CA2-Fw_ex7	Atgagccactgcgcctg	*CA2-RV_ex7*	GACACAAAGCAACCAGGGTC
ATP6V1B1			
ATP6V1B1-Fw_ex1	CCAGCTGGACCTGAAGTCTC	*ATP6V1B1-RV_ex1*	GAACCTGGCAGCCCACC
ATP6V1B1-Fw_ex2	GTGTGTGAGCAGGGTGTTTG	*ATP6V1B1-RV_ex2*	CATGGAGGGAAAGGGTAAT TTAG
ATP6V1B1-Fw_ex3&4	AGGAGGAGAAGGGACTTTGC	*ATP6V1B1-RV_ex3&4*	Gtttggtcagagaagctggg
ATP6V1B1-Fw_ex5	AGTAGTGAGGGACACAGGGC	*ATP6V1B1-RV_ex5*	GCAGAACTTCCAAGAGTGGG
ATP6V1B1-Fw_ex6	GAGAGCAGGGAAGGGTTTG	*ATP6V1B1-RV_ex6*	CAGGAGCTGTCCTCACTGG
ATP6V1B1-Fw_ex7	CAGTGGTGCTCAGTGGGAC	*ATP6V1B1-RV_ex7*	CTGAGGGAGACAGGCTGC
ATP6V1B1-Fw_ex8	CTGAGGGAGACAGGCTGC	*ATP6V1B1-RV_ex8*	TAGGTCTTGGAAGGGGACG
ATP6V1B1-Fw_ex9	TATGTGGTGCATTAGCCCAG	*ATP6V1B1-RV_ex9*	CTCTATTTGCAGGGTGACCG
ATP6V1B1-Fw_ex10&11	CCCTTCCTAGCTTCAGCCTC	*ATP6V1B1-RV_ex10&11*	GGGGCAGTGAAACATCTG
ATP6V1B1-Fw_ex12&13	ACAATTTGGGGACAGGGG	*ATP6V1B1-RV_ex12&13*	TGTTAGGAATGTGTGTGGGG
ATP6V1B1-Fw_ex14	CTGGTGTGGAGCCAGTAACC	*ATP6V1B1-RV_ex14*	CACCACCTCGGAGCCAG
SLC9A4			
SLC9A4-Fw_ex1	CCGACTTCAGATGTGTGGC	*SLC9A4-RV_ex1*	CCCGTTATCTATAAGGCATTT TC
SLC9A4-Fw_ex2	AGGTACACCCAGACCGTTTC	*SLC9A4-RV_ex2*	CACCCCGAACCCTCATC
SLC9A4-Fw_ex3	ATGTGCATGCAAAGTGTTCC	*SLC9A4-RV_ex3*	TCATGTGATGTGGTTCTAAA CTTTC
SLC9A4-Fw_ex4	TCATGCATCTGTGGCATTG	*SLC9A4-RV_ex4*	GTTTTAACATTGCGGAAGGG
SLC9A4-Fw_ex5	GAATTAGGCTGCATTGTTTGC	*SLC9A4-RV_ex5*	TCTCCGGAAATTTAGTTTGG G

SLC9A4-Fw_ex6	AAAGGAATGGAAAACTGAAATG	*SLC9A4-RV_ex6*	GTGGGGAAGGAAAGAGGAAG
SLC9A4-Fw_ex7	TGTGTGTCTTTCAGAGTTCGC	*SLC9A4-RV_ex7*	ttTATGCCCAGCTTCCAATG
SLC9A4-Fw_ex8	ATGAATTCCAGGCACTGTCC	*SLC9A4-RV_ex8*	AAATGAGGCAATGAGTCAgc
SLC9A4-Fw_ex9	CACCTCTTGCCAAATGAAAG	*SLC9A4-RV_ex9*	TGAACATCAACAGACTCCCC
SLC9A4-Fw_ex10	GCTTCCTTGGCTGAAAGTTG	*SLC9A4-RV_ex10*	CTCAGCAGATGACAAACACTTG
SLC9A4-Fw_ex11	TGCCATGTGAAGCTCTTAATTG	*SLC9A4-RV_ex11*	Ctcttggcccacacccc
SLC9A4-Fw_ex12	TCCTGCCATGTGGATATTAAGTC	*SLC9A4-RV_ex12*	TTTCACAGTTATAGGAAGACAGACTC
SLC9A2			
SLC9A2-Fw_ex1	CCTCTGGTTGCAGAGACCC	*SLC9A2-RV_ex1*	AGAAGCTGCTCCATCTGTCG
SLC9A2-Fw_ex2	CCTGTCTCCCAAATTCCTCC	*SLC9A2-RV_ex2*	GTTGAGGTGCAGAGAGCCAC
SLC9A2-Fw_ex3	GTGTTTGTGCCCAGAGTGAC	*SLC9A2-RV_ex3*	Ctgaccttgtgatccacgc
SLC9A2-Fw_ex4	GAGATTGCAGGATTTCCGAG	*SLC9A2-RV_ex4*	ACAAAAGGAAAGCCCCATTC
SLC9A2-Fw_ex5	TTTTCATATTTTGGCCTCACC	*SLC9A2-RV_ex5*	TTTGCTACTTTGCATCCTGTAAAC
SLC9A2-Fw_ex6	TCAAAGCTTAGGAGAAGTTACCC	*SLC9A2-RV_ex6*	TTCAGGAAATGGACAACGTG
SLC9A2-Fw_ex7	AGTAGAGTGAAAATTATTGATTTCAGG	*SLC9A2-RV_ex7*	TGCTTTCAGAATGACCCTGG
SLC9A2-Fw_ex8	CCAGATTAGCTAAATGCTGCTAC	*SLC9A2-RV_ex8*	tcagtatttgcAgctgggc
SLC9A2-Fw_ex9	TGAAATAAGAGTCACAGAATTTGG	*SLC9A2-RV_ex9*	CCAGAAGGCATCCTGTTAGC
SLC9A2-Fw_ex10	AAGTCTTGGAATTTCTGGGG	*SLC9A2-RV_ex10*	AATGGAATAGCTGTCGGCTC
SLC9A2-Fw_ex12	TTACTTGCAATGCCTTCTCTC	*SLC9A2-RV_ex12*	CCACTGTTGCATCAGGTTTC
MFSD9			
MFSD9-Fw_ex1	GGCTCCCGGGGTAGGTC	*MFSD9-Rv_ex1*	CCGGAAGCTCCAAGAACC
MFSD9-Fw_ex2	GGCAGCGGATTTGTATGG	*MFSD9-Rv_ex2*	TCATTTCCCACTTACCTTTGG
MFSD9-Fw_ex3	TCTGTTGGTTTCTGTGTAATTGTTATC	*MFSD9-Rv_ex3*	AAAAGCAGGGAATTGTAGGAAG
MFSD9-Fw_ex4	CATTTGCATGGTGGTAGTCG	*MFSD9-Rv_ex4*	AGCAAGACCACAGAGGTGGC
MFSD9-Fw_ex5	agctgccttccaaAGAAACC	*MFSD9-Rv_ex5*	Tttggacgatgagtctgtctattc
MFSD9-Fw_ex6A	CCTTCCCCAAGGAGGTAGTG	*MFSD9-Rv_ex6A*	TGTGCAGGTGAGTATGCTGG
MFSD9-Fw_ex6B	CCAAGGTGACAGGCTACCTC	*MFSD9-Rv_ex6B*	TCCAAAGTGCCCTTTTCATC
TMEM182			
TMEM182-Fw_ex1&2	AGCTAGGACAGCCTTCTCAAG	*TMEM182-Rv_ex1&2*	CAATCCAGTCCGCTAGAAGTG
TMEM182-Fw_ex3	TGCTTGTGTAACCCTTGTGC	*TMEM182-Rv_ex3*	TTTTGATCTGGGCATCCTTC
TMEM182-Fw_ex4	CTGACTGTGACAATCTGCTGTG	*TMEM182-Rv_ex4*	TTACTAGAAGGCTGACGCCTG
TMEM182-Fw_ex5	GCTTGTCATCCGGTCAATG	*TMEM182-Rv_ex5*	TGGGATCAATGAAACAAAATG
POU3F3			

POU3F3-Fw_ex1-1	tgctgctgGTGACCAAGG	*POU3F3-Rv_ex1-1*	GCACCAAGCTCTGGGCTc
POU3F3-Fw_ex1-2	GCACCAGTGGGTCACAGC	*POU3F3-Rv_ex1-2*	GCCAGCTCTGGCGTGTC
POU3F3-Fw_ex1-3	CACCAGTGGGTCACAGCC	*POU3F3-Rv_ex1-3*	CTTCACTGAACGCTCGTCTG
POU3F3-Fw_ex1-4	GGGGCCTGGACTCAACAG	*POU3F3-Rv_ex1-4*	ATGCCTGTCTCTGGGTGG
MRPS9			
MRPS9-Fw_ex1	CTGCGCCGTCTCTAGGC	*MRPS9-Rv_ex1*	GCTGGCTCCACATCCTC
MRPS9-Fw_ex2	AAGCAAGGCTTTGAGTAGGG	*MRPS9-Rv_ex2*	AAACACAGAGCTAAGGGGCTC
MRPS9-Fw_ex3&4	AAACACAGAGCTAAGGGGCTC	*MRPS9-Rv_ex3&4*	TACGATATGTGAGCAACCGC
MRPS9-Fw_ex5	TGGTTAAATGCAGCTATTTGG	*MRPS9-Rv_ex5*	CATCAACACACTGTCCTACCTC
MRPS9-Fw_ex6	TCAGAAGTTATAATGTACCATTGCTC	*MRPS9-Rv_ex6*	TGAACAAATTTTAGGCTTCAGG
MRPS9-Fw_ex7	AGAGTAGATTGAGGAAATGGAATG	*MRPS9-Rv_ex7*	CGTAGGAACACCTGCCTCTG
MRPS9-Fw_ex8	AGCGTGAAGATGCAGAAGTG	*MRPS9-Rv_ex8*	GGAAATCCAAAGTAACGGTAGC
MRPS9-Fw_ex9	GTCTCAAAGCGCAGGTCTTC	*MRPS9-Rv_ex9*	CAGAATTAGAAGATCACAAGCCTTC
MRPS9-Fw_ex10	GAAAAGTATAACAGTGGCATGCAG	*MRPS9-Rv_ex10*	CGAACTTCAGGCACTAGGAC
MRPS9-Fw_ex11	TTACTGGTCTGGTTGCTTTCC	*MRPS9-Rv_ex11*	TGACAGTACTGCCCTCATGC
GPR45			
GPR45-Fw_ex1-1	GCAAATCCCAGCTAGTGCTC	*GPR45-Rv_ex1-1*	TGATGATGAGCAGGATGGC
GPR45-Fw_ex1-2	TCTGCTGCATGCCCTTC	*GPR45-Rv_ex1-2*	GAAGGCCTTGGTCTTGAAGC
GPR45-Fw_ex1-3	ACAACCAGTCGGACAGCC	*GPR45-Rv_ex1-3*	TGTGCTTAAATTGCCACCAC
GPR45-Fw_ex1-4	CAGTCTGCGGTTTAGGGG	*GPR45-Rv_ex1-4*	GGCATGGGAAGAAATGCAC
TGFBRAP1			
TGFBRAP1-Fw_ex1	CGTCTCGGCCTGTGCTC	*TGFBRAP1-Rv_ex1*	Ccactacacctgccccag
TGFBRAP1-Fw_ex2-1	GCCTCTGTTTCTGCTTCTTCC	*TGFBRAP1-Rv_ex2-1*	CACAGGGTTCTCGTTCAGTG
TGFBRAP1-Fw_ex2-2	CTGTGTGACAACTCCATCAGC	*TGFBRAP1-Rv_ex2-2*	ACACACGCCCAATGTCAC
TGFBRAP1-Fw_ex3	ACAGGAAAATGACACCATGC	*TGFBRAP1-Rv_ex3*	caAGGCCCAGTTCCTAAATG
TGFBRAP1-Fw_ex4	AGTTTGGGGAAAGCAGTGTG	*TGFBRAP1-Rv_ex4*	TCTGAATGTCCTCTAACTGGTGAC
TGFBRAP1-Fw_ex5	AATTTAACCCTGCCGTGATG	*TGFBRAP1-Rv_ex5*	GTTCTGGCCACAGAGGGAC
TGFBRAP1-Fw_ex6	GCTGTTTTGTGTGCCTATGTG	*TGFBRAP1-Rv_ex6*	GCAAAGAGCCTCAGTAAGGG
TGFBRAP1-Fw_ex7	AATATGGGGAGCTTGGACTG	*TGFBRAP1-Rv_ex7*	GGCAAAAGTGGTCTCTATCAAC

TGFBRAP1-Fw_ex8	TTCTCAACTAATTCTGAGATATGTGC	*TGFBRAP1-Rv_ex8*	AAAGCAAAGCTTTTGGGGTC
TGFBRAP1-Fw_ex9	AGCTTGAATGTGCAATGACC	*TGFBRAP1-Rv_ex9*	GTTCAATTCCGTCTCTCCAG
TGFBRAP1-Fw_ex10	GGGGAGAGCTTTGACTTTCC	*TGFBRAP1-Rv_ex10*	CCCACCCGCTTGATATGAG
TGFBRAP1-Fw_ex11	GAAGCCCAGAGCAGCAAG	*TGFBRAP1-Rv_ex11*	TTTCCATCAGTAAAATCTACCTTCAG
TGFBRAP1-Fw_ex12	CCATGGAGTAATGCATGTCAAC	*TGFBRAP1-Rv_ex12*	GAGTCTAAAGGTTGGAGTGTGG
C2ORF49			
C2ORF49-Fw_ex1	TCAGAGTGACGACACACGC	*C2ORF49-Rv_ex1*	CAAGCTCGGGTGCAAGC
C2ORF49-Fw_ex2	CGCGTGTTATTGTGAAAATGG	*C2ORF49-Rv_ex2*	Cagtgcctggtgctcaatag
C2ORF49-Fw_ex3	ATCCCAGTGAACTGGTTTGC	*C2ORF49-Rv_ex3*	GCAGACTGCTCAAGTTTCATTC
C2ORF49-Fw_ex4	ATGGCAGCAGTGATCTTCC	*C2ORF49-Rv_ex4*	TTCCTAATCAGGGCAAGACAG
FHL2			
FHL2-Fw_ex1	CGCTCCCGATAACTACTGTG	*FHL2-Rv_ex1*	GTTCGGGTCCCCTCTCC
FHL2-Fw_ex2	ACAGTTAATGCCTTTGCGTG	*FHL2-Rv_ex2*	AAACCACAGAGAAACCCGTG
FHL2-Fw_ex3	GTCCAAGCCCTCCTGTCC	*FHL2-Rv_ex3*	CAGAGGAACGTGCACAAGG
FHL2-Fw_ex4	GCTCTGAAGGCAGATTGCTG	*FHL2-Rv_ex4*	TGGACCATGGAGGCAAAC
FHL2-Fw_ex5	GTCCAGTGGCAGTAGCTGAG	*FHL2-Rv_ex5*	TAGGCCTTGTTCAAGCTTCC
FHL2-Fw_ex6-1	CAGTCCCAGTTGCCTGATTC	*FHL2-Rv_ex6-1*	ACCATCTTCCCACCTAGGGC
FHL2-Fw_ex6-2	TTCTGGTTTCTTCCAGCCAC	*FHL2-Rv_ex6-2*	AGACTCAGTGCCTCGACCTG
NCK2B			
NCK2B-Fw_ex1	GCTGTCCCTCTGAAGCAGTC	*NCK2B-Rv_ex1*	AATGGCTCAATGACAGCCC
NCK2B-Fw_ex2	TTGGTGCAAAGGATGAAGTG	*NCK2B-Rv_ex2*	CGAACAGTGTCCCTTGAAATG
NCK2B-Fw_ex3-1	TAGGGAGTGTGGTGGTGCC	*NCK2B-Rv_ex3-1*	TCTCGAAGTTGAGCTCCTCC
NCK2B-Fw_ex3-2	AGCTTCCTGAGCCTGCG	*NCK2B-Rv_ex3-2*	GCGCAAGGCGCATTTAC
NCK2B-Fw_ex-4-1	GATTTAGACACAGCTCCCCG	*NCK2B-Rv_ex-4-1*	TAAAGAGGGGTTCCATGAGC
NCK2B-Fw_ex-4-2	CTCCCATTTGCCATCCAG	*NCK2B-Rv_ex-4-2*	AGTGTCAACAAACGGCTACAG
NCK2B-Fw_ex-4-3	GGAACCGGTGACTCAGAAAG	*NCK2B-Rv_ex-4-3*	GATTTCCACTCCGCAATGAG
Hypothetical genes			
LOC644617x1AF	TGAAACTCCATGACGAACAC	*LOC644617x1AF*	TGAAACTCCATGACGAACAC
LOC644617x1BF	CTCCCGAGGACATCCAGAC	*LOC644617x1BF*	CTCCCGAGGACATCCAGAC
LOC644617x1CF	TTTCTCATTCCTAACACACagacag	*LOC644617x1CF*	TTTCTCATTCCTAACACACagacag
LOC644617x1DF	GAGCAATTAGATCCATCAGACC	*LOC644617x1DF*	GAGCAATTAGATCCATCAGACC

AK95498x1F	GCAGTTGGAAGCAGGATTTC	*AK95498x1R*	CTGGAGAACTTCCCCTGC
AK95498x2AF	TCAAGAAAGGGGACACCAAG	*AK95498x2AR*	ACTCACAGCGGCTAGTTGG
AK95498x2BF	CCAGGGTTACCCCAGAGTTC	*AK95498x2BR*	CAAAGCTATAACAAAGGCTCCC
AK95498x2CF	ACAGTCAGCGGGCTCTGC	*AK95498x2CR*	CAAGACTACTGGGACCTGGC
AK95498x2DF	CCCAGGCTAAGGATTTGAGG	*AK95498x2DR*	GAGAGATCAGTTACCTTTCCCC
AK95498x2EF	GATGCAAATGGAGGTGGAAG	*AK95498x2ER*	TTGAGAATCTTTAGGTGACAATGG
AK95498x2F-F	GACTTGCCTGAGCGATTTTG	*AK95498x2F-R*	AAGAATGATCACTGCTGCCC
AK96498x1Fw	CAAGGGGAGAACAGTTAGTGC	*AK96498x1Rv*	TGAAAATAAGCTCCCAAACCTAC
AK96498x2CFw	GCTAGAAATATATGAGCCTCCAGC	*AK96498x2CRv*	Actgatacaaggatacaagggaag
AK96498x4AFw	GAGCATGGCACAAAGACCTC	*AK96498x4ARv*	AAATTAACTCACAGCCAAAGGC
AK96498x4BFw	tgccaattcctgATGTTCTC	*AK96498x4BRv*	ctggcttactatatgatcttggg
AK96498x4CFw	GTGGTTGGGAGGATCAAGTG	*AK96498x4CRv*	CAAGCCTGGATCATTTTCATC
AK96498x4DFw	CCAAGGAGTTGGCTTTTATTTG	*AK96498x4DRv*	Gtctccccagccatgcag
AK96498x4EFw	TGCCCTTTTAGTAGGGGAAG	*AK96498x4ERv*	CTCTCAGAGGTCTTTGCAGG
AK96621x1AFw	TTGCTATTGTTTTGGCATGG	*AK96621x1ARv*	GATCATAACTAATGTTGCAAAGCC
AK96621x1BFw	TCTTGTTCCCGTGTATGTGC	*AK96621x1BRv*	GGAATGTTGGACATAGGTATACTGG
AK96621x1CFw	TTCAAGCATGGTGACATTCC	*AK96621x1CRv*	CTACAGGCACTGGCCCAC
AK96621x1DFw	GGGATTGAGAAACTCCTTGC	*AK96621x1DRv*	AGCTCCAGGACACCCTCAC
AK96621x2Fw	ATGCCCTCCTCCTGAGTTC	*AK96621x2Rv*	TTCCTAAAGCTTGATGACTGC
AK96621x3Fw	AGTCTCTCTCACCTGCCCC	*AK96621x3Rv*	AAAATGATGAAGAACTGGCTAGG
AK96621x4AFw	TGGTGCTATCAATTTGGCTG	*AK96621x4ARv*	CAATGCCTCTTCCGTGAGAC
AK96621x4BFw	AGAGCTGTGGATCTGCGTC	*AK96621x4BRv*	CACAAGACAGAGATCTGGGG
CR607634-Fw_ex1	AACATACACGGCAGCAAAATC	*CR607634-Rv_ex1*	GCAAATGGAAGGGGAAAGAG
CR607634-Fw_ex2	CCTATATTGTGGGAATTGAAATG	*CR607634-Rv_ex2*	ATGAGGGGTTTGCAACAATC
CR607634-Fw_ex3	TCAAGCTGCCTCAAAATATCC	*CR607634-Rv_ex3*	CGGTATAAAGGTCACTGGAACC
CR607634-Fw_ex4-1	CAGCCACTCGACCTCCTTAG	*CR607634-Rv_ex4-1*	TGTGCACGAAGGCATTG
CR607634-Fw_ex4-2	agaagcaaggaatgcagaag	*CR607634-Rv_ex4-2*	Cactgaactgcatgagaacaag
AX746619-Fw_ex1-1	GGAGTGGCTTCCCTCAGAAG	*AX746619-Rv_ex1-1*	GGACAGAGCTTGGCACTAGG
AX746619-Fw_ex1-2	CTTTCTGTGGCTCTCCCATC	*AX746619-Rv_ex1-2*	TGCTCAGTGCAGGTTGATTG
AX746619-	AGCGAGGATCTTGCCATCTC	*AX746619-Rv_ex1-*	TGATGACTGCATATTGCTTAT

Fw_ex1-3		*3*	CAC
AX746619-Fw_ex2	AGTCTCTCTCACCTGCCCC	*AX746619-Rv_ex2*	AAAATGATGAAGAACTGGCTAGG
AX746619-Fw_ex3-1	TGGTGCTATCAATTTGGCTG	*AX746619-Rv_ex3-1*	CAATGCCTCTTCCGTGAGAC
AX746619-Fw_ex3-2	AGAGCTGTGGATCTGCGTC	*AX746619-Rv_ex3-2*	CACAAGACAGAGATCTGGGG
CR609970-Fw_ex1	TCAACAGGTTGAAACGTAAGAGTAG	*CR609970-Rv_ex1*	CAGTGTGAAAAGCTGAATCGG
CR609970-Fw_ex2-1	ggaaggcctagaaaatcatgaag	*CR609970-Rv_ex2-1*	Aacttcagttgctgtacttttcag
CR609970-Fw_ex2-2	ccctgagaaagcccagatac	*CR609970-Rv_ex2-2*	Tctgatgagaaattggctgatac
CR609970-Fw_ex2-3	gcaataccaaattggataaaccc	*CR609970-Rv_ex2-3*	cactatttgcaacttctcccttg
Gene expression primers			
SLC9A2Exp-Fw_A	Ttggagagtccctgctgaat	*SLC9A2Exp-Rv_A*	ttcatagtcattgcacaagcag
SLC9A2Exp-Fw_B	Ccactggtggagtttcttga	*SLC9A2Exp-Rv_B*	Cactggacgtgaccttcctt
POU3F3Exp-Fw_A	agttcgccaagcagttcaag	*POU3F3Exp-Rv_A*	Gcacttgaggaagtggctct
POU3F3Exp-Fw_B	Gttctcgcagaccaccatct	*POU3F3Exp-Rv_B*	Cttgaggaagtggctctcca
TGFBRAPExp-Fw_A	cgatcgtcaagaggataggg	*TGFBRAPExp-Rv_A*	Tgtggcaacgatcactcttc
TGFBRAPExp-Fw_B	Gtccgggagctgatctctct	*TGFBRAPExp-Rv_B*	Ctgcagcagcatcttggtta
TMEM182Exp-Fw	Tcgccagccattttctctac	*TMEM182Exp-Rv*	tgctgggatcaatgaaacaa
MFSD9Exp-Fw	Tgtttgtcctggctagagtcc	*MFSD9Exp-Rv*	Caggcccttctctgtactgc
MRPSExp-Fw	cccctgagttgcaaacaaat	*MRPSExp-Rv*	acaaaccacttgggaaaagg

Appendix A

List of oligonucleotide primer sequences used in the thesis

Chapter 3.4 Bilateral Renal Agenesis

FGF1			
FGF1-Fw_ex1	CAGCTCTGTGGCATGGTATC	*FGF1-Rv_ex1*	TCCAAGTAGCATTACATTTGCAC
FGF1-Fw_ex2	AAAGTGCATTTTGTGCCTTTG	*FGF1-Rv_ex2*	TAAATGTGCACCTTCCTCCC
FGF1-Fw_ex3	TCCTATCGGCAAAAGCGTAG	*FGF1-Rv_ex3*	AGTTTGCCTCACTCCATGTAAC
FGF1-Fw_ex4-1	tgaggcatggagagctttag	*FGF1-Rv_ex4-1*	GGGGAATCACAGATAGGGTTTAC
FGF1-Fw_ex4-2	AGTCTGGGTCTGGGTTTGG	*FGF1-Rv_ex4-2*	CCCTTTCTTTCTCTGCACTCC
FGF1-Fw_ex4-3	GGTGACTTCAGGTGGAGAGC	*FGF1-Rv_ex4-3*	GCATTTTATAGGCATTGGGC
FGF1-Fw_ex4-4	CCGAACCTCACCTGACCTC	*FGF1-Rv_ex4-4*	TCATATGAAACAGGAGCTAACACTC
FGF1-Fw_ex4-5	TGATACTGATGCCATGTAAAGG	*FGF1-Rv_ex4-5*	TGCCCTGGGAGAAGATACAG
FGF1-Fw_ex4-6	CAAACTCAAGCATAACATTGGC	*FGF1-Rv_ex4-6*	AATGGGAATGTCCCCAGG
FGF1-Fw_ex4-7	AATCAGCCAAAGAGCCTGTC	*FGF1-Rv_ex4-7*	ACACCTCAGGGATGCTGG
GFRA3			
GFRA3-Fw_ex1	CCTCGGAAAGAGCAACCTC	*GFRA3-Rv_ex1*	GCTCCAGTTCCCCTTGTTC
GFRA3-Fw_ex2	GGAATATCCCTCTCCCAACC	*GFRA3-Rv_ex2*	ATGTGGCCCATATCCTGC
GFRA3-Fw_ex3	GAGGTCTGAGCCAGCAGC	*GFRA3-Rv_ex3*	CCTTCCTCCAGGGTCCAG
GFRA3-Fw_ex4	GTTCCTGTGGAACCTGGG	*GFRA3-Rv_ex4*	AGGAAGGAGTGGCGTGTG
GFRA3-Fw_ex5&6	cctgggcaagagagagtgag	*GFRA3-Rv_ex5&6*	TCCTTCCCTTGAGTCCAGC
GFRA3-Fw_ex7	CCTTGGGAGCCTGAGAAATC	*GFRA3-Rv_ex7*	CTGCCTAGTTTGGGTTTTCC
GFRA3-Fw_ex8-1	GGCTGAGGTAGTGAAAGGGC	*GFRA3-Rv_ex8-1*	GGTGATCCTGGTAGTCAAGAGG
GFRA3-Fw_ex8-2	CATGCTGCCCCTCCTTG	*GFRA3-Rv_ex8-2*	tgtgttctttcatctataacaccc
TGFB2			
TGFB2-Fw_ex1-1	CTTTGTTGAAGGCAGACACG	*TGFB2-Rv_ex1-1*	TCCTTCTCTTGCTCCAAACG
TGFB2-Fw_ex1-2	AGCCAGGGTGTCGCAAG	*TGFB2-Rv_ex1-2*	GAACTGACGGGAGGGGC
TGFB2-Fw_ex1-3	CTGCTGCTCCTGCTCTCAG	*TGFB2-Rv_ex1-3*	AAGCGCTCAGCACACAGTAG
TGFB2-Fw_ex1-4	TTTCCACTTTTGGAACTACTGG	*TGFB2-Rv_ex1-4*	AACTGAAAAGGAGGTGGTGG

TGFB2-Fw_ex2	GTGCCAGGCATCTCTTCAG	*TGFB2-Rv_ex2*	TCCCCACAGGAGACAAACAC
TGFB2-Fw_ex3	TTCCCTTATGGTTTCTTGGG	*TGFB2-Rv_ex3*	Tgtactctagcctgtgacagagag
TGFB2-Fw_ex4&5	GCAAATAGCCTGGTGTTGTG	*TGFB2-Rv_ex4&5*	CCAAATGAATGAAACTCCCC
TGFB2-Fw_ex6	TGGCTGACTATATTTGATGAAGG	*TGFB2-Rv_ex6*	GGGCAAAGGCAGTTTAAGAG
TGFB2-Fw_ex7	TGGTGGTAGAGTGAGGGTGG	*TGFB2-Rv_ex7*	CAGCAGGTGCACATGGTAAC
TGFB2-Fw_ex8-1	ACGAATTGCGTTCATTTTCC	*TGFB2-Rv_ex8-1*	CGGTACGTACAGCAACTCCAC
TGFB2-Fw_ex8-2	CCTACTTTGTAAGTGAGAGAGACAAG	*TGFB2-Rv_ex8-2*	TTTTAAGGTCTGAAGTTTGACCAG
TGFB2-Fw_ex8-3	AAAAGAAACTTTCAGTCAGAATAAGTC	*TGFB2-Rv_ex8-3*	CCACTGAACTCGAACCCATC
TGFB2-Fw_ex8-4	TTTGCCCTACTTGTGCTTTG	*TGFB2-Rv_ex8-4*	CTCTTCTGCAATGATGTGGC
TGFB2-Fw_ex8-5	CCAGGTCAGCATAAGTCATTTTG	*TGFB2-Rv_ex8-5*	GGGAAATAAAGACCTTACAGAACC
TGFB2-Fw_ex8-6	TCTGCAGACATTTTCCTCTCAG	*TGFB2-Rv_ex8-6*	TCCTTGATGGATATGGCCTG
TGFB2-Fw_ex8-7	ACCTTTCCGATTGCCCTC	*TGFB2-Rv_ex8-7*	TGATGGTTTGTTTCTAGATGGTG
REN			
REN-Fw_ex1	ATTGCAATCACCCCATGC	*REN-Fw_ex1*	GTGTTGGGAAGGTGGGAAC
REN-Fw_ex2	CACTGTGGTGGAGGTACAGC	*REN-Fw_ex2*	AAAGCCCTAGGTCCATGAGG
REN-Fw_ex3	AGAGTGAAGCCAGGCAAGAC	*REN-Fw_ex3*	CAGGATTGCTCATGCACAG
REN-Fw_ex4	CCCACAATGCACCCTGC	*REN-Fw_ex4*	GGTCAGGAGAGGCCTGG
REN-Fw_ex5	AGACCAGACTCCAACCAAGG	*REN-Fw_ex5*	TCCACTCCCCATCTCTTCTC
REN-Fw_ex6	acagcagggctgatgagg	*REN-Fw_ex6*	ATTCTAGGTCAGGTGCTCGG
REN-Fw_ex7	GAGTTGCTGGGTCCTGGAG	*REN-Fw_ex7*	GTAATGCAGTCCTTCCCCAC
REN-Fw_ex8	CCCATCAGCCTtctgtctg	*REN-Fw_ex8*	TTTGTGAGCCGATACCAGG
REN-Fw_ex9	AGCGCAGGACTCCTTGTC	*REN-Fw_ex9*	TTGAGGGTCTCTGTCCAGC
REN-Fw_ex10	ACGATGGGGCAAAATGG	*REN-Fw_ex10*	GCTAGGCGAATGTGTAGCTTG
CR2			
CR2-Fw_ex1	CTTGTCCCACCCTCACCG	*CR2-Rv_ex1*	AGAAACTTTCCCTGCCCG
CR2-Fw_ex2	GAGTTTGGGAGTAAGCAGGG	*CR2-Rv_ex2*	CAGGAGGGTACAAAACGTGG
CR2-Fw_ex3&4	ccccttgATTCTAGATTGTGAAG	*CR2-Rv_ex3&4*	CAGAGATCCTCACAGAGGAGAC
CR2-Fw_ex5	TGCTGTTCTTCAGCACAAAC	*CR2-Rv_ex5*	GAGTCTGTAAGGTAATGACTGAACC
CR2-Fw_ex6	AGCATCTGGGGCATTCTTTG	*CR2-Rv_ex6*	AGCCCACAACCCCTTCTTC
CR2-Fw_ex7&8	TGGCTCAGTTTCTTTCTGTGG	*CR2-Rv_ex7&8*	ATAGAGCCTGCAGTGTGTGC
CR2-Fw_ex9	GAAGTTGGTGCTGATGTT	*CR2-Rv_ex9*	TGCAACCATAACTATTGG

	GG		CAG
CR2-Fw_ex10	TGGTGGTTCTTTGTTCTTTGG	*CR2-Rv_ex10*	CTTGCCTCTTTCCATGATGC
CR2-Fw_ex11	TCCAGAGTTGTCCTTCTCTTTG	*CR2-Rv_ex11*	TGAAGGTCAAGTTAATCCACAAAAC
CR2-Fw_ex12	TTTTGCATGCAGTTTGAAGAG	*CR2-Rv_ex12*	CCTGCCCTCTCTAATCCTCC
CR2-Fw_ex13	AATGAAGACCGCTCACCTTG	*CR2-Rv_ex13*	TGTCAAGTGCAGATTCAAAACC
CR2-Fw_ex14	GCAGGCTATGTGTTCCTCTG	*CR2-Rv_ex14*	ATTTTAAGTTGAAAGTACCCAACTG
CR2-Fw_ex15	GGGAAATTTCTGCTTTCTTGG	*CR2-Rv_ex15*	TGAGGGAAACTCTCACTCTCTTC
CR2-Fw_ex16	AAATCGACTTGGAAGGGAGC	*CR2-Rv_ex16*	TCCAACTCTATGCCAAAGTGAC
CR2-Fw_ex17	TCAGAACAATGTAGGTGATCGTC	*CR2-Rv_ex17*	GCTTCTGTGTTTAGTTTCAGACAAG
CR2-Fw_ex18	CAACTGCTACATTGAAACATGGTC	*CR2-Rv_ex18*	CGCTTATTGATAAATACAAGTTGAAGG
CR2-Fw_ex19-1	CTAATAGACTTGATGCTTTGTTGG	*CR2-Rv_ex19-1*	GATTAACACCACCAGAGGAGG
CR2-Fw_ex19-2	GCCAGGCCATGGCTATAAAC	*CR2-Rv_ex19-2*	GACCAAGTGATGGGATCCTG
FLVCR1			
FLVCR1-Fw_ex1-1	AGGGCCGTAGTCATGCAG	*FLVCR1-Rv_ex1-1*	GAAGATCAGGAGCACCACG
FLVCR1-Fw_ex1-2	AGAGGAGACCCAGGCCC	*FLVCR1-Rv_ex1-2*	ACAGGCCATAGTTCTCACGG
FLVCR1-Fw_ex2	AATTTTCAAAGCACCTTTATAAACAC	*FLVCR1-Rv_ex2*	GGGCATGGAAGAATGAAAC
FLVCR1-Fw_ex3	GAAATCAGTGAAAATGAAAACCC	*FLVCR1-Rv_ex3*	AATCCACAATGCTGAAAGGC
FLVCR1-Fw_ex4	TCCACAATTGTCACATAATATTCC	*FLVCR1-Rv_ex4*	CCCTTAATCTCATGACCCTGC
FLVCR1-Fw_ex5	TGAAAAGGTGTGGGAATGTG	*FLVCR1-Rv_ex5*	AAAAGCAAATTGTCCTTTAAAATC
FLVCR1-Fw_ex6	ATGGACAAGAAGGGGTGAAC	*FLVCR1-Rv_ex6*	TTGAAGAAACTTTGGATGAACTG
FLVCR1-Fw_ex9	CCAAGAAAGACTTTGAGCTGC	*FLVCR1-Rv_ex9*	TGGTTCTTGGTCTGTGGGAC
FLVCR1-Fw_ex10-1	GCATTAATCAAGTCTGATCTGC	*FLVCR1-Rv_ex10-1*	TCATCTGAAAGCCTGATCCC
FLVCR1-Fw_ex10-2	AAGGCTGGGTTTGTATGTGG	*FLVCR1-Rv_ex10-2*	AGAGCATGTACCTGGGCATAG
FLVCR1-Fw_ex10-3	TCATGGATAAATGCTAACGCTG	*FLVCR1-Rv_ex10-3*	TCAAAATTGCTATTCACTTCAAGG
FLVCR1-Fw_ex10-4	ATTTCCACAGTGCCTACCAC	*FLVCR1-Rv_ex10-4*	tgccattcgcattcaagc
FLVCR1-Fw_ex10-5	gttaggccatccttgcattg	*FLVCR1-Rv_ex10-5*	ggatacttgtccctccaatctc
FLVCR1-Fw_ex10-6	ccagcctgggagataaagtg	*FLVCR1-Rv_ex10-6*	TCTGGCTGGATTCAAGTTCC
FLVCR1-Fw_ex10-7	caagtatcccaactatatcaGGCAG	*FLVCR1-Rv_ex10-7*	TGTTCCAAACCTCAGGAACTC
FLVCR1-Fw_ex10-8	TCCAATCATTAGCTTCCTTTG	*FLVCR1-Rv_ex10-8*	tcacctagtgatgccacagc
FLVCR1-Fw_ex10	TGAAAGTTGTCTGTAATATGTCAAGC	*FLVCR1-Rv_ex10*	CCATGCACTACCTCATTCCC

ESRRG			
ESRRG-Fw_ex1	CCAAGTAAGAGGCATCCAGC	*ESRRG-Rv_ex1*	ATTGTCACATTTCCCTCTGC
ESRRG-Fw_ex2	tttggtctctgtgcttgacac	*ESRRG-Rv_ex2*	CACCATTAAAGCTTAACAGTAAGAGG
ESRRG-Fw_ex3	TTGGGAAGCTAATGTTTCCTG	*ESRRG-Rv_ex3*	TTTTGCCTCCCATCCCC
ESRRG-Fw_ex4	GGTATTTGATTGGTCTCAAACATAG	*ESRRG-Rv_ex4*	ACATAGAGCCAACTGGGAGG
ESRRG-Fw_ex5	GCCATAATTGGAATATTTAGCG	*ESRRG-Rv_ex5*	GGATTCTATTAGGATATGCAAAACTC
ESRRG-Fw_ex6	TGTTGCATTATGAAATTATGGC	*ESRRG-Rv_ex6*	CCTAAAGAAGTTAAGGAGAGGGG
ESRRG-Fw_ex7-1	GAAATGCAAGACTCTGAACAAATC	*ESRRG-Rv_ex7-1*	aaaagaaagaaggcaggcag
ESRRG-Fw_ex7-2	TCttctttcctctttcttcccac	*ESRRG-Rv_ex7-2*	AATGATAAGAAGCAGCCGATG
ESRRG-Fw_ex7-3	GAGAGAACCAACGTTCCTAAAAG	*ESRRG-Rv_ex7-3*	TGCACTCCTGACAATTCTACG
ESRRG-Fw_ex7-4	TGCAATTTATGTTGTGTGCC	*ESRRG-Rv_ex7-4*	GGGAGCAGCTACTAACACTGG
ESRRG-Fw_ex7-5	TGAAAGTCAAGTCAGGTTCCAG	*ESRRG-Rv_ex7-5*	GAGCGAGTTAATAACAACATCTGAG
ESRRG-Fw_ex7-6	GGATGTACAATTCCTCAGGAGAC	*ESRRG-Rv_ex7-6*	CCATCTCAGGTTTTCAGTTGG
ESRRG-Fw_ex7-7	AACAGAAAATACAACTCCTGGG	*ESRRG-Rv_ex7-7*	CAATCACCAACACCCATTTTC
ESRRG-Fw_ex7-8	CCTTGTAGTTACACTTGTGTTTCAATC	*ESRRG-Rv_ex7-8*	GGGCAGTGTTTATCAATTTGG
CENPF			
CENPF-Fw_ex1	TAAAATCCATCCGAAAGGGC	*CENPF-Rv_ex1*	GGAGAGTGATCCCAGAGGC
CENPF-Fw_ex2	TCATTAATTTCTGAGACTTTGTTTCC	*CENPF-Rv_ex2*	CAGTTTTAAATACCAGCACTTCTCTG
CENPF-Fw_ex3	TGTTCATATGGCTTATTGCAGC	*CENPF-Rv_ex3*	TCTTCCACTCTTTCAAGTTACGG
CENPF-Fw_ex4	TCTGGGAATGTAAGGCATTG	*CENPF-Rv_ex4*	TGCAGAATTAAAATCTCAAGCTAAAC
CENPF-Fw_ex5	TGATCTGTGAATTCCTTCATGG	*CENPF-Rv_ex5*	TGAGCCCAAAACCTTTTCTC
CENPF-Fw_ex6	TGTTAACTTCTTGGGATTATGGC	*CENPF-Rv_ex6*	CACCTGTGAAATTACCTCAAGC
CENPF-Fw_ex7	TTTGTTTCGTCAAAGATTTCAG	*CENPF-Rv_ex7*	tgcccaacttaaaacagctttc
CENPF-Fw_ex8	ATAGTTCCTAAACATAAATTGTGTGC	*CENPF-Rv_ex8*	ACAACTCCCCTTTCCAAGTG
CENPF-Fw_ex9	TAGGATGCTCATGCCTCACC	*CENPF-Rv_ex9*	AACAAGTACAAGGAACAAGAGGG
CENPF-Fw_ex10	TTGATATTCCATGACCATTTTATTG	*CENPF-Rv_ex10*	CTTTTCCCCTTGCACTTGG
CENPF-Fw_ex11	GGGACCTTTATTAAGGCAGAAG	*CENPF-Rv_ex11*	ATCCCCTCTCACGGGAATAG
CENPF-Fw_ex12-1	TTTGTTTGGATTTTGCCAGG	*CENPF-Rv_ex12-1*	CTGAGAACTCAGCTTTCTGCTG
CENPF-Fw_ex12-2	GAGTAAGAACGCTGGAGATGG	*CENPF-Rv_ex12-2*	TGACTTGTTTCAGCCACAAAAC
CENPF-Fw_ex12-3	GCCATCCTACAAAATAGAGTTGATTC	*CENPF-Rv_ex12-3*	TCTTTTCCTGTGCTGCTTTG

CENPF-Fw_ex12-4	TCCAGAAAAGTGAGAGTTTTGC	*CENPF-Rv_ex12-4*	CAAGATCCATCTGAGGTTTAAAATTAC
CENPF-Fw_ex12-5	TGAACAGCTGATGAAGGTAATG	*CENPF-Rv_ex12-5*	CACCATGGAGAAGACCACTG
CENPF-Fw_ex12-6	GAAAATATGTGAAATACTGCAGGC	*CENPF-Rv_ex12-6*	CTACACTGCACTGGTTTGCTG
CENPF-Fw_ex12-7	TCTTGTGTGCCTGACAGCTC	*CENPF-Rv_ex12-7*	CATTTTCCTCAAGTTCTGGTTTG
CENPF-Fw_ex13-1	GTGTTTAGCAGAGGCCACG	*CENPF-Rv_ex13-1*	TGAGGCCTTCAGAAGTTTCTAC
CENPF-Fw_ex13-2	GCTAAATGAAATGAAAGAATTAGACTC	*CENPF-Rv_ex13-2*	ACTCACTTTGATGAGACTCAAGC
CENPF-Fw_ex13-3	GAGAGAAACCAGCTTCGTGG	*CENPF-Rv_ex13-3*	AGTGAATCATTCTCCCGCTC
CENPF-Fw_ex13-4	TCGAGAAACTGAGAGTTCGC	*CENPF-Rv_ex13-4*	TTTCATCAGCTTCTAGGCGG
CENPF-Fw_ex13-5	ATGAGGCAGTAGCAGCCTTG	*CENPF-Rv_ex13-5*	GCCTTCTCAAGTTCAAGACACTC
CENPF-Fw_ex13-6	GAGGTCTGGAATTAGATGTTGTTAC	*CENPF-Rv_ex13-6*	CATGTCCCAAAGAATCAGCC
CENPF-Fw_ex14	GGCTAAAATAAATGGCATGAATATG	*CENPF-Rv_ex14*	CAACAGAAGGTCTTACAGGAACC
CENPF-Fw_ex15	TGAATACATTGCTAAGGATTATAGCAC	*CENPF-Rv_ex15*	ACCTGGGGAAAAGATCGTG
CENPF-Fw_ex16	TTTTAGCTGTGCCTCCTGG	*CENPF-Rv_ex16*	ACCATTAAAATCACTACCACCTG
CENPF-Fw_ex17	AACCACATTGTCAGTTGGGG	*CENPF-Rv_ex17*	TCCCAAGATAAGCACAAAATTC
CENPF-Fw_ex18	TTATCTGCTTCACGATGCCC	*CENPF-Rv_ex18*	GCAGACTATTGCTTATGAACGTC
CENPF-Fw_ex19	TCTTTTCCACTGTAGATAGAATTGG	*CENPF-Rv_ex19*	TCCTCACCCAGGTACTACTGC
CENPF-Fw_ex20-1	TCTCTGTTAGATGGGGCCTG	*CENPF-Rv_ex20-1*	TCAGCTCTGAGAAATTCCCG
CENPF-Fw_ex20-2	CATTGCCATTCCTCTACTGC	*CENPF-Rv_ex20-2*	TTCACGTGTATAAGAGGCTGTTG
MIR205			
MIR205-Fw_ex1	CCTGAAGCTTTGCTGAGAGG	*MIR205-Rv_ex1*	ggtggtgtgcgcttgtagtc
MIR205-Fw_ex2	cccctcaagtagctgggac	*MIR205-Rv_ex2*	GTTTCTTTTGCCCCTAAGGC
MIR205-Fw_ex3	AACTCACCTTTCTGGGGCTC	*MIR205-Rv_ex3*	CCTCCCAACATGTACAGTCC
MIR205-Fw_ex4	TCCTCAGACAATCCATGTGC	*MIR205-Rv_ex4*	CTAAAGGGGAGGTGGCTTAG
CAMK1G			
CAMK1G-Fw_ex1	ACGGTCCTGGGAAAGAAAAC	*CAMK1G-Rv_ex1*	CACGCTTCCCATTAAGCAAG
CAMK1G-Fw_ex2	TGTTTGAGAAAGCAGCTGACC	*CAMK1G-Rv_ex2*	CTAACCTCCCACTGCAGGC
CAMK1G-Fw_ex3	ctggaagtgtaagatgtcTGCC	*CAMK1G-Rv_ex3*	CTTTGGGTGCAGAAATCCAG
CAMK1G-Fw_ex4	TGCATTATTTGTCTGTACAACCTG	*CAMK1G-Rv_ex4*	CAGCGACCAACCTCCTGTAG
CAMK1G-Fw_ex5	GCTTCAAAGCAAACACCTTCTC	*CAMK1G-Rv_ex5*	ATGACCCATGCAGCTTCC
CAMK1G-Fw_ex6	TCAAGGGGAAGGAAAATACTTG	*CAMK1G-Rv_ex6*	GGCATGTGGTTGAATGTAGC

CAMK1G-Fw_ex7	GGCATGTGGTTGAATGTAGC	*CAMK1G-Rv_ex7*	AACGAGAAGGGGTAAGCAGG
CAMK1G-Fw_ex8	CAGAGAGGCTGGCTCAGG	*CAMK1G-Rv_ex8*	AGCTCATTTCTCTTGGCAGC
CAMK1G-Fw_ex9	TTTCCATCAGTCATGAAATCTG	*CAMK1G-Rv_ex9*	GGAAATGCAATATGAATGAGTCAG
CAMK1G-Fw_ex10	CCCTGGATGAGTGTAAATAGCC	*CAMK1G-Rv_ex10*	AGCCCTTATGTCAGCCCC
CAMK1G-Fw_ex11	GTTTAAGCTCCAAGGCCCTC	*CAMK1G-Rv_ex11*	TCCATTTCTTTCAGTCCTGTTG
CAMK1G-Fw_ex12	GAACTACCACCTCTGCCCTG	*CAMK1G-Rv_ex12*	GCATGGGGCTCTCCTCTG
CAMK1G-Fw_ex13-1	TGTGAGTGATGGGGTCTGAG	*CAMK1G-Rv_ex13-1*	AGCTGCCTGTCATTCATGC
CAMK1G-Fw_ex13-2	TACCCAGACTCCCACTCTGC	*CAMK1G-Rv_ex13-2*	TCCAGACATGCTCAGAAGTCC
LAMB3			
LAMB3-Fw_ex1	CTGCCAGATTCCTGAGACCC	*LAMB3-Rv_ex1*	AGATACAGTCCTCCTCCCGC
LAMB3-Fw_ex2	GGGCTAGTGTGCTCTAGGGG	*LAMB3-Rv_ex2*	AAACTCTCCTTCTCCCTGTGG
LAMB3-Fw_ex3	CTGAAAGATACTGGGGTGGC	*LAMB3-Rv_ex3*	acaaatgGCAGCTCACACAC
LAMB3-Fw_ex4	CACAGAGGAGATATGCTGGG	*LAMB3-Rv_ex4*	TAAGAAATGCCTGGGAAACC
LAMB3-Fw_ex5	GCTTCCTTCAGGGCTGG	*LAMB3-Rv_ex5*	GCTCCTCCATGGCTCCAC
LAMB3-Fw_ex6	CTCTGCTTGTGTCTCCACGG	*LAMB3-Rv_ex6*	GTGTGGCTGTGTGAACAGTG
LAMB3-Fw_ex7	AACTCGGGGCACACTATTTG	*LAMB3-Rv_ex7*	ATTTCCATGACCTGGGCTC
LAMB3-Fw_ex8	TCTATGTGCCTCTAACTTCCTTTC	*LAMB3-Rv_ex8*	TCTGTAGTCTGCCCTGGTCC
LAMB3-Fw_ex9	CATCCCTGGTGCCCTTC	*LAMB3-Rv_ex9*	CCAGGTGAGATCATCAAGGC
LAMB3-Fw_ex10	gccagtgagctttgtggag	*LAMB3-Rv_ex10*	GAGGAGATGGGGAGTAACAGAC
LAMB3-Fw_ex11	agcctggctcctggaatc	*LAMB3-Rv_ex11*	AAGCCGTGGGTCTGGTG
LAMB3-Fw_ex12	TGGATGTGTGTCCTCTGGG	*LAMB3-Rv_ex12*	CAGAGAGGCTGGTGCTCAG
LAMB3-Fw_ex13	CAGGGTAAGGCAGAGGCTG	*LAMB3-Rv_ex13*	GCAGACCTACACGCCCC
LAMB3-Fw_ex14	GACAGCTGCTTCTGTCTCCC	*LAMB3-Rv_ex14*	GTACTGGAACCCCTGGAGC
LAMB3-Fw_ex15&16	CTGAGGGGATAGCTGCTTTG	*LAMB3-Rv_ex15&16*	GCAGTTGGTGTGTGACAAGC
LAMB3-Fw_ex17	GGGTGGGATTGGTTATTGAG	*LAMB3-Rv_ex17*	GGGGATCTATCCTGTCTGCC
LAMB3-Fw_ex18	CTGCTGAGAGGTTCCCTTGG	*LAMB3-Rv_ex18*	GCTGGCTGATGCACTGAAC
LAMB3-Fw_ex19	AGAGGCCTCGTATCTCTGGG	*LAMB3-Rv_ex19*	AGGAGCAAAACGCCAGC
LAMB3-Fw_ex20	CGAACCACACATTGATTCAC	*LAMB3-Rv_ex20*	CATTTTCTATGCCTGGTGCC
LAMB3-	GCTGCCTGATGAGTTGGA	*LAMB3-*	AGTGTGCAAAGTCCTCCT

Fw_ex21	G	*Rv_ex21*	CTG
LAMB3-Fw_ex22	TTTGGCCACAGGTCTTTAGG	*LAMB3-Rv_ex22*	TCAGAAGCCTTGGGTTGG
LAMB3-Fw_ex23-1	GTTGGGAGTCTTGGGGAGTC	*LAMB3-Rv_ex23-1*	TGTAACTGTCCCATTGGCTC
LAMB3-Fw_ex23-2	GAATGCTTTCCATCTCCAGG	*LAMB3-Rv_ex23-2*	CCCAAATGGGAGACAGAAAC
G0S2			
G0S2-Fw_ex1	ACACAGTCGGAGCGTGC	*G0S2-Rv_ex1*	AGCAGGAGAGGGGCTTCAG
G0S2-Fw_ex2-1	TGGGAGCCGTTCTTTGG	*G0S2-Rv_ex2-1*	GAGTCAGGCTCCTGTGCG
G0S2-Fw_ex2-2	AGGCAAGCAGCAGGACAC	*G0S2-Rv_ex2-2*	TAAGAACCCATCACCATCGC
HSD11B1			
HSD11B1-Fw_ex1	ATGAAATCCACCACACAGGC	*HSD11B1-Rv_ex1*	AGGAACACTCAAGCACCCC
HSD11B1-Fw_ex2	GAAGGAATTTTGCTGCCAAC	*HSD11B1-Rv_ex2*	TGCGATATGCTAGCTTCTGTG
HSD11B1-Fw_ex3-4	CCGTTACTTCAGAGACTACCCC	*HSD11B1-Rv_ex3-4*	CTGGCATCATCCCTGGC
HSD11B1-Fw_ex5	CCACCCTATGCCTTCGATAC	*HSD11B1-Rv_ex5*	acactgggcctgttggg
HSD11B1-Fw_ex6-1	AGTGGTTGATGTCTCCAGGC	*HSD11B1-Rv_ex6-1*	TTAGAGGAACTCCTTCCATTTG
HSD11B1-Fw_ex6-2	TGTTCTGTCTCATGTTTATCTGAGC	*HSD11B1-Rv_ex6-2*	CACCAAACAAAGATTAGGGTG
TRAF3IP3			
TRAF3IP3-Fw_ex1	GTGCCTACTGATGGGAGGAG	*TRAF3IP3-Rv_ex1*	GAGCATCTTCTGCCCTACCC
TRAF3IP3-Fw_ex2	GGCTCCCTGGAAACTTCTG	*TRAF3IP3-Rv_ex2*	ATTTCAGAGCATCTGGCTCC
TRAF3IP3-Fw_ex3	CCCCTGGTCTAGGGAGATAG	*TRAF3IP3-Rv_ex3*	CTCCTTGTCCAGAGAGGGTC
TRAF3IP3-Fw_ex4&5	AGAGTCTTTCAAGAGGCTCCAG	*TRAF3IP3-Rv_ex4&5*	GGACATGGAAATTCATTGTGTTC
TRAF3IP3-Fw_ex6&7	CACTGCCTCCTGACTTGATTC	*TRAF3IP3-Rv_ex6&7*	GATTAACCTGGAGGATCGGG
TRAF3IP3-Fw_ex8	AGTAAGTGGCATGTGACCCC	*TRAF3IP3-Rv_ex8*	TGAGAAAACCCCTTCTCTTACTG
TRAF3IP3-Fw_ex9	TGTGCTGGGCAGTAGAATAGG	*TRAF3IP3-Rv_ex9*	GGTGGTGTTTCTCGAAAGTAAG
TRAF3IP3-Fw_ex10&11	aggatttggcacacagaagc	*TRAF3IP3-Rv_ex10&11*	TCTTTCATCACCATCCCTCC
TRAF3IP3-Fw_ex12	AACCTCCACCCCAACACTG	*TRAF3IP3-Rv_ex12*	AAAGCCAATCAACCAATCAAG
TRAF3IP3-Fw_ex13	AAGACCATGACAAGACAGCATC	*TRAF3IP3-Rv_ex13*	CCTACTCTGGGTGCCTTTTG
TRAF3IP3-Fw_ex14	ttcaacacacagcaGATGGG	*TRAF3IP3-Rv_ex14*	aggtcatgtagtccatcccg
TRAF3IP3-Fw_ex15	CCTTCCTCAGAGGTGTGGG	*TRAF3IP3-Rv_ex15*	CAGGTTTCTCTGCTATTTCGC
TRAF3IP3-Fw_ex16	CAGTCTCCCCTGCCTGG	*TRAF3IP3-Rv_ex16*	TTGGTTAAAGAAGGAATCGTCC

TRAF3IP3-Fw_ex17	TTTGCCTATATCTTTTCCAATCC	*TRAF3IP3-Rv_ex17*	GACCACCAGTATGCTTAAATATGTG
C1orf74			
C1orf74-Fw_ex1-1	GTGTCTGGCTGAAAGATGCC	*C1orf74-Rv_ex1-1*	GATAAGTCCCTCTGCAGCCC
C1orf74-Fw_ex1-2	AGCGTCACCCTTCTGTCTG	*C1orf74-Rv_ex1-2*	GATGATTATTTGCCATCCCC
IRF6			
IRF6-Fw_ex1	TTCTTGCAACTCGCCAATC	*IRF6-Rv_ex1*	GATGGGTCCTACCCCACC
IRF6-Fw_ex2	CAGACACTGGATTGTTGTATGG	*IRF6-Rv_ex2*	CCACCATGATGAGGGAGAAG
IRF6-Fw_ex3	CCAGTGGCTGGCCTAATTC	*IRF6-Rv_ex3*	TTCTCTCTGTTTCACCAGAGTTTTAG
IRF6-Fw_ex4	TTTGCAGTGGCTCTGGG	*IRF6-Rv_ex4*	CCATCTTAAACTGTGTAAATCAGGC
IRF6-Fw_ex5	GTCAGCTGCAGGGGTGG	*IRF6-Rv_ex5*	TGACAGTCCCAAGGTCAAAAC
IRF6-Fw_ex6	TCTCAGAGCATCAGCAAGTG	*IRF6-Rv_ex6*	TGAAGTTAGAAAGCAGGACAGG
IRF6-Fw_ex7	CTGGTTGAAAGGTGGCTTG	*IRF6-Rv_ex7*	GACAGGGATAGTGGAAGGAATG
IRF6-Fw_ex8	TTGTTGAGTGGATGTGGAATG	*IRF6-Rv_ex8*	GGGGCTTAAGCATTGGC
IRF6-Fw_ex9-1	TGTCCTTCAATCTTGGGTCC	*IRF6-Rv_ex9-1*	aaactcccaggccaaatctc
IRF6-Fw_ex9-2	TTTGTGATTCTCCAAATATGCC	*IRF6-Rv_ex9-2*	TTCTGTGTCAACAGCTTCCC
IRF6-Fw_ex9-3	GGCTGGCTGGTTGCTTAG	*IRF6-Rv_ex9-3*	GATACAGCTGCATTGTTCTATACC

Appendix B

List of Companies Addresses

QIAGEN Inc. 27220 Turnberry Lane, Valencia, CA 91355, USA

Fisher Scientific, 300 Industry Drive, Pittsburgh, PA 15275, USA

Life Technologies, 3175 Staley Road Grand Island, NY 14072, USA

Thermo Scientific 3411 Silverside Rd, Bancroft Building, Wilmington, DE 19810, USA

Zymo Research Corporation, 17062 Murphy Ave., Irvine, CA 92614, U.S.A

Amersham Biosciences Corp. 800 Centennial Ave, P.O. Box 1327 Piscataway, NJ 08855-1327

Corning Life Sciences 31 Mayfield Avenue Edison, New Jersey 08837, United States

Beckman Coulter, Inc. Diagnostics Division Headquarters 250 South Kraemer Boulevard Brea CA 92821-6232

PE Applied Biosystems, 777 Lincoln Centre Dr, Foster City, CA 94404

SoftGenetics, LLC. 100 Oakwood Ave, Suite 350. State College, PA 16803. USA.

Invitrogen (Life Technologies Ltd). 3 Fountain Drive Inchinnan Business Park, Paisley PA4 9RF, UK

Affymetrix 3420 Central Expressway, Santa Clara, CA 95051

Metabion International AG, lena-christ-str. 44/I, 82152 martinsried, Germany

Gene Tools, LLC, 1001 Summerton Way, Philomath, OR 97370 USA

Geneservice Ltd, 2 Cambridge Science Park, Milton Road, Cambridge, CB4 0FE, UK

Mutagenex Inc. P.O. Box 68 Piscataway, NJ 08855, USA

Ambion, 2130 Woodward, Austin, TX 78744-1832, USA

Sigma-Aldrich Company Ltd. The Old Brickyard, New Road, Gillingham, Dorset SP8 4XT, UK

Abcam plc. 330 Cambridge, Science Park, Cambridge, CB4 0FL, UK.

References

2011. SCOT Data: Deceased heart beating donor and organ transplantation in Saudi Arabia. *Saudi J Kidney Dis Transpl,* 22**,** 387-92.

2012. Dialysis in the Kingdom of Saudi Arabia. *Saudi J Kidney Dis Transpl,* 23**,** 881-9.

ABDELHAK, S., KALATZIS, V., HEILIG, R., COMPAIN, S., SAMSON, D., VINCENT, C., LEVI-ACOBAS, F., CRUAUD, C., LE MERRER, M., MATHIEU, M., KONIG, R., VIGNERON, J., WEISSENBACH, J., PETIT, C. & WEIL, D. 1997. Clustering of mutations responsible for branchio-oto-renal (BOR) syndrome in the eyes absent homologous region (eyaHR) of EYA1. *Hum Mol Genet,* 6**,** 2247-55.

ABU-SAFIEH, L., AL-ANAZI, S., AL-ABDI, L., HASHEM, M., ALKURAYA, H., ALAMR, M., SIRELKHATIM, M. O., AL-HASSNAN, Z., ALKURAYA, B., MOHAMED, J. Y., AL-SALEM, A., ALRASHED, M., FAQEIH, E., SOFTAH, A., AL-HASHEM, A., WALI, S., RAHBEENI, Z., ALSAYED, M., KHAN, A. O., AL-GAZALI, L., TASCHNER, P. E., AL-HAZZAA, S. & ALKURAYA, F. S. 2012. In search of triallelism in Bardet-Biedl syndrome. *Eur J Hum Genet,* 20**,** 420-7.

ABU SAFIEH, L., ALDAHMESH, M. A., SHAMSELDIN, H., HASHEM, M., SHAHEEN, R., ALKURAYA, H., AL HAZZAA, S. A., AL-RAJHI, A. & ALKURAYA, F. S. 2010. Clinical and molecular characterisation of Bardet-Biedl syndrome in consanguineous populations: the power of homozygosity mapping. *J Med Genet,* 47**,** 236-41.

AL-ABDULKAREEM, A. A. & BALLAL, S. G. 1998. Consanguineous marriage in an urban area of Saudi Arabia: rates and adverse health effects on the offspring. *J Community Health,* 23**,** 75-83.

AL-GHWERY, S. & AL-ASMARI, A. 2004. Chronic Renal Failure among Children in Riyadh Military Hospital, Riyadh, Saudi Arabia. *Saudi J Kidney Dis Transpl,* 15**,** 75-8.

AL-RASHEED, S. A., AL-MUGEIREN, M. M., AL-SALLOUM, A. A. & AL-SOHAIBANI, M. O. 1996. Childhood renal diseases in Saudi Arabia. A clinicopathological study of 167 cases. *Int Urol Nephrol,* 28**,** 607-13.

AL-SABBAN, E. 1997. Spectrum of glomerular disease among children in saudi arabia. *Saudi J Kidney Dis Transpl,* 8**,** 285-8.

AL HARBI, N. 1997. Chronic renal failure in children in asir region of saudi arabia. *Saudi J Kidney Dis Transpl,* 8**,** 294-7.

AL HUSAIN, M. & AL BUNYAN, M. 1997. Consanguineous marriages in a Saudi population and the effect of inbreeding on prenatal and postnatal mortality. *Ann Trop Paediatr,* 17**,** 155-60.

ALAZAMI, A. M., ALSHAMMARI, M. J., SALIH, M. A., ALZAHRANI, F., HIJAZI, H., SEIDAHMED, M. Z., ABU SAFIEH, L., ALDOSARY, M., KHAN, A. O. & ALKURAYA, F. S. 2012. Molecular characterization of Joubert syndrome in Saudi Arabia. *Hum Mutat,* 33**,** 1423-8.

ALDAHMESH, M. A., HUMEIDAN, A., ALMOJALLI, H. A., KHAN, A. O., RAJAB, M., AA, A. L.-A., MEYER, B. F. & ALKURAYA, F. S. 2009. Characterization of CTNS mutations in Arab patients with cystinosis. *Ophthalmic Genet,* 30**,** 185-9.

ALPER, S. L. 2010. Familial renal tubular acidosis. *J Nephrol,* 23 Suppl 16**,** S57-76.

AMIR, R. E., VAN DEN VEYVER, I. B., WAN, M., TRAN, C. Q., FRANCKE, U. & ZOGHBI, H. Y. 1999. Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2. *Nat Genet,* 23**,** 185-8.

ANSLEY, S. J., BADANO, J. L., BLACQUE, O. E., HILL, J., HOSKINS, B. E., LEITCH, C. C., KIM, J. C., ROSS, A. J., EICHERS, E. R., TESLOVICH, T. M., MAH, A. K., JOHNSEN, R. C., CAVENDER, J. C., LEWIS, R. A., LEROUX, M. R., BEALES, P. L. & KATSANIS, N. 2003. Basal body dysfunction is a likely cause of pleiotropic Bardet-Biedl syndrome. *Nature,* 425**,** 628-33.

ANZENBERGER, U., BIT-AVRAGIM, N., ROHR, S., RUDOLPH, F., DEHMEL, B., WILLNOW, T. E. & ABDELILAH-SEYFRIED, S. 2006. Elucidation of megalin/LRP2-dependent endocytic transport processes in the larval zebrafish pronephros. *J Cell Sci,* 119**,** 2127-37.

BADANO, J. L., ANSLEY, S. J., LEITCH, C. C., LEWIS, R. A., LUPSKI, J. R. & KATSANIS, N. 2003. Identification of a novel Bardet-Biedl syndrome protein, BBS7, that shares structural features with BBS1 and BBS2. *Am J Hum Genet,* 72**,** 650-8.

BADANO, J. L., LEITCH, C. C., ANSLEY, S. J., MAY-SIMERA, H., LAWSON, S., LEWIS, R. A., BEALES, P. L., DIETZ, H. C., FISHER, S. & KATSANIS, N. 2006. Dissection of epistasis in oligogenic Bardet-Biedl syndrome. *Nature,* 439**,** 326-30.

BALOH, R. H., GORODINSKY, A., GOLDEN, J. P., TANSEY, M. G., KECK, C. L., POPESCU, N. C., JOHNSON, E. M., JR. & MILBRANDT, J. 1998. GFRalpha3 is an orphan member of the GDNF/neurturin/persephin receptor family. *Proc Natl Acad Sci U S A,* 95**,** 5801-6.

BANKIER, A., DE CAMPO, M., NEWELL, R., ROGERS, J. G. & DANKS, D. M. 1985. A pedigree study of perinatally lethal renal disease. *J Med Genet,* 22**,** 104-11.

BASHIARDES, S., VEILE, R., HELMS, C., MARDIS, E. R., BOWCOCK, A. M. & LOVETT, M. 2005. Direct genomic selection. *Nat Methods,* 2**,** 63-9.

BATES, C. M. 2000. Kidney development: regulatory molecules crucial to both mice and men. *Mol Genet Metab,* 71**,** 391-6.

BEALES, P. L., ELCIOGLU, N., WOOLF, A. S., PARKER, D. & FLINTER, F. A. 1999. New criteria for improved diagnosis of Bardet-Biedl syndrome: results of a population survey. *J Med Genet,* 36**,** 437-46.

BEALES, P. L., WARNER, A. M., HITMAN, G. A., THAKKER, R. & FLINTER, F. A. 1997. Bardet-Biedl syndrome: a molecular and phenotypic study of 18 families. *J Med Genet,* 34**,** 92-8.

BEDELL, V. M., WESTCOT, S. E. & EKKER, S. C. 2011. Lessons from morpholino-based screening in zebrafish. *Brief Funct Genomics,* 10**,** 181-8.

BELTCHEVA, O., MARTIN, P., LENKKERI, U. & TRYGGVASON, K. 2001. Mutation spectrum in the nephrin gene (NPHS1) in congenital nephrotic syndrome. *Hum Mutat,* 17**,** 368-73.

BENOIT, G., MACHUCA, E. & ANTIGNAC, C. 2010. Hereditary nephrotic syndrome: a systematic approach for genetic testing and a review of associated podocyte gene mutations. *Pediatr Nephrol,* 25**,** 1621-32.

BERDELI, A., MIR, S., YAVASCAN, O., SERDAROGLU, E., BAK, M., AKSU, N., ONER, A., ANARAT, A., DONMEZ, O., YILDIZ, N., SEVER, L., TABEL, Y., DUSUNSEL, R., SONMEZ, F. & CAKAR, N. 2007. NPHS2 (podicin)

mutations in Turkish children with idiopathic nephrotic syndrome. *Pediatr Nephrol,* 22**,** 2031-40.

BETTENCOURT-DIAS, M., HILDEBRANDT, F., PELLMAN, D., WOODS, G. & GODINHO, S. A. 2011. Centrosomes and cilia in human disease. *Trends Genet,* 27**,** 307-15.

BILL, B. R., PETZOLD, A. M., CLARK, K. J., SCHIMMENTI, L. A. & EKKER, S. C. 2009. A primer for morpholino use in zebrafish. *Zebrafish,* 6**,** 69-77.

BINGHAM, C., ELLARD, S., ALLEN, L., BULMAN, M., SHEPHERD, M., FRAYLING, T., BERRY, P. J., CLARK, P. M., LINDNER, T., BELL, G. I., RYFFEL, G. U., NICHOLLS, A. J. & HATTERSLEY, A. T. 2000. Abnormal nephron development associated with a frameshift mutation in the transcription factor hepatocyte nuclear factor-1 beta. *Kidney Int,* 57**,** 898-907.

BIRD, A. P. 1986. CpG-rich islands and the function of DNA methylation. *Nature,* 321**,** 209-13.

BITTLES, A. 2001. Consanguinity and its relevance to clinical genetics. *Clin Genet,* 60**,** 89-98.

BITTLES, A. H. 2003. Consanguineous marriage and childhood health. *Dev Med Child Neurol,* 45**,** 571-6.

BOUTE, N., GRIBOUVAL, O., ROSELLI, S., BENESSY, F., LEE, H., FUCHSHUBER, A., DAHAN, K., GUBLER, M. C., NIAUDET, P. & ANTIGNAC, C. 2000. NPHS2, encoding the glomerular protein podocin, is mutated in autosomal recessive steroid-resistant nephrotic syndrome. *Nat Genet,* 24**,** 349-54.

BOWDEN, D. W. 2003. Genetics of kidney disease. *Kidney Int Suppl***,** S8-12.

BOYER, O., BENOIT, G., GRIBOUVAL, O., NEVO, F., PAWTOWSKI, A., BILGE, I., BIRCAN, Z., DESCHENES, G., GUAY-WOODFORD, L. M., HALL, M., MACHER, M. A., SOULAMI, K., STEFANIDIS, C. J., WEISS, R., LOIRAT, C., GUBLER, M. C. & ANTIGNAC, C. 2010. Mutational analysis of the PLCE1 gene in steroid resistant nephrotic syndrome. *J Med Genet,* 47**,** 445-52.

BRODEHL, J., KROHN, H. P. & EHRICH, J. H. 1982. The treatment of minimal change nephrotic syndrome (lipoid nephrosis): cooperative studies of the Arbeitsgemeinschaft fur Padiatrische Nephrologie (APN). *Klin Padiatr,* 194**,** 162-5.

BROWN, E. J., SCHLONDORFF, J. S., BECKER, D. J., TSUKAGUCHI, H., TONNA, S. J., USCINSKI, A. L., HIGGS, H. N., HENDERSON, J. M. & POLLAK, M. R. 2010. Mutations in the formin gene INF2 cause focal segmental glomerulosclerosis. *Nat Genet,* 42**,** 72-6.

CARIDI, G., BERTELLI, R., DI DUCA, M., DAGNINO, M., EMMA, F., ONETTI MUDA, A., SCOLARI, F., MIGLIETTI, N., MAZZUCCO, G., MURER, L., CARREA, A., MASSELLA, L., RIZZONI, G., PERFUMO, F. & GHIGGERI, G. M. 2003. Broadening the spectrum of diseases related to podocin mutations. *J Am Soc Nephrol,* 14**,** 1278-86.

CARIDI, G., GIGANTE, M., RAVANI, P., TRIVELLI, A., BARBANO, G., SCOLARI, F., DAGNINO, M., MURER, L., MURTAS, C., EDEFONTI, A., ALLEGRI, L., AMORE, A., COPPO, R., EMMA, F., DE PALO, T., PENZA, R., GESUALDO, L. & GHIGGERI, G. M. 2009. Clinical features and long-term outcome of nephrotic syndrome associated with heterozygous NPHS1 and NPHS2 mutations. *Clin J Am Soc Nephrol,* 4**,** 1065-72.

CARROLL, P., AL-MOJALLI, H., AL-ABBAD, A., AL-HASSOUN, I., AL-HAMED, M., AL-AMR, R., BUTT, A. I. & MEYER, B. F. 2006. Novel mutations

underlying nephrogenic diabetes insipidus in Arab families. *Genet Med,* 8**,** 443-7.

CASTRO, D. S., SKOWRONSKA-KRAWCZYK, D., ARMANT, O., DONALDSON, I. J., PARRAS, C., HUNT, C., CRITCHLEY, J. A., NGUYEN, L., GOSSLER, A., GOTTGENS, B., MATTER, J. M. & GUILLEMOT, F. 2006. Proneural bHLH and Brn proteins coregulate a neurogenic program through cooperative binding to a conserved DNA motif. *Dev Cell,* 11**,** 831-44.

CHEN, J., SMAOUI, N., HAMMER, M. B., JIAO, X., RIAZUDDIN, S. A., HARPER, S., KATSANIS, N., RIAZUDDIN, S., CHAABOUNI, H., BERSON, E. L. & HEJTMANCIK, J. F. 2011. Molecular analysis of Bardet-Biedl syndrome families: report of 21 novel mutations in 10 genes. *Invest Ophthalmol Vis Sci,* 52**,** 5317-24.

CHERIAN, M. P. & AL-SANNA'A, N. A. 2009. Clinical spectrum of Bardet-Biedl syndrome among four Saudi Arabian families. *Clin Dysmorphol,* 18**,** 188-94.

CHIANG, A. P., BECK, J. S., YEN, H. J., TAYEH, M. K., SCHEETZ, T. E., SWIDERSKI, R. E., NISHIMURA, D. Y., BRAUN, T. A., KIM, K. Y., HUANG, J., ELBEDOUR, K., CARMI, R., SLUSARSKI, D. C., CASAVANT, T. L., STONE, E. M. & SHEFFIELD, V. C. 2006. Homozygosity mapping with SNP arrays identifies TRIM32, an E3 ubiquitin ligase, as a Bardet-Biedl syndrome gene (BBS11). *Proc Natl Acad Sci U S A,* 103**,** 6287-92.

CHIANG, A. P., NISHIMURA, D., SEARBY, C., ELBEDOUR, K., CARMI, R., FERGUSON, A. L., SECRIST, J., BRAUN, T., CASAVANT, T., STONE, E. M. & SHEFFIELD, V. C. 2004. Comparative genomic analysis identifies an ADP-ribosylation factor-like gene as the cause of Bardet-Biedl syndrome (BBS3). *Am J Hum Genet,* 75**,** 475-84.

CHOI, M., SCHOLL, U. I., JI, W., LIU, T., TIKHONOVA, I. R., ZUMBO, P., NAYIR, A., BAKKALOGLU, A., OZEN, S., SANJAD, S., NELSON-WILLIAMS, C., FARHI, A., MANE, S. & LIFTON, R. P. 2009. Genetic diagnosis by whole exome capture and massively parallel DNA sequencing. *Proc Natl Acad Sci U S A,* 106**,** 19096-101.

CUMMING, W. A. & OHLSSON, A. 1984. Nephrocalcinosis in Bartter's syndrome. Demonstration by ultrasonography. *Pediatr Radiol,* 14**,** 125-6.

DAVIES, J. A. & FISHER, C. E. 2002. Genes and proteins in renal development. *Exp Nephrol,* 10**,** 102-13.

DE BORST, M. H., BENIGNI, A. & REMUZZI, G. 2008. Primer: strategies for identifying genes involved in renal disease. *Nat Clin Pract Nephrol,* 4**,** 265-76.

DETRICH, H. W., WESTERFIELD, M. & ZON, L. I. 2011. *The zebrafish : genetics, genomics and informatics,* Amsterdam Netherlands ; Boston Mass., Elsevier/Academic.

DEVONALD, M. A., SMITH, A. N., POON, J. P., IHRKE, G. & KARET, F. E. 2003. Non-polarized targeting of AE1 causes autosomal dominant distal renal tubular acidosis. *Nat Genet,* 33**,** 125-7.

DRAPER, B. W., MORCOS, P. A. & KIMMEL, C. B. 2001. Inhibition of zebrafish fgf8 pre-mRNA splicing with morpholino oligos: a quantifiable method for gene knockdown. *Genesis,* 30**,** 154-6.

DRUMMOND, I. A. 2005. Kidney development and disease in the zebrafish. *J Am Soc Nephrol,* 16**,** 299-304.

DRUMMOND, I. A. 2008. Some assembly required: renal hypodysplasia and the problem with faulty parts. *J Am Soc Nephrol,* 19**,** 834-6.

DRUMMOND, I. A., MAJUMDAR, A., HENTSCHEL, H., ELGER, M., SOLNICA-KREZEL, L., SCHIER, A. F., NEUHAUSS, S. C., STEMPLE, D. L., ZWARTKRUIS, F., RANGINI, Z., DRIEVER, W. & FISHMAN, M. C. 1998. Early development of the zebrafish pronephros and analysis of mutations affecting pronephric function. *Development,* 125**,** 4655-67.

DRUMMOND, K. N., MICHAEL, A. F., ULSTROM, R. A. & GOOD, R. A. 1964. The Blue Diaper Syndrome: Familial Hypercalcemia with Nephrocalcinosis and Indicanuria; a New Familial Disease, with Definition of the Metabolic Abnormality. *Am J Med,* 37**,** 928-48.

EL-HAZMI, M. A., AL-SWAILEM, A. R., WARSY, A. S., AL-SWAILEM, A. M., SULAIMANI, R. & AL-MESHARI, A. A. 1995. Consanguinity among the Saudi Arabian population. *J Med Genet,* 32**,** 623-6.

FAN, Y., ESMAIL, M. A., ANSLEY, S. J., BLACQUE, O. E., BOROEVICH, K., ROSS, A. J., MOORE, S. J., BADANO, J. L., MAY-SIMERA, H., COMPTON, D. S., GREEN, J. S., LEWIS, R. A., VAN HAELST, M. M., PARFREY, P. S., BAILLIE, D. L., BEALES, P. L., KATSANIS, N., DAVIDSON, W. S. & LEROUX, M. R. 2004. Mutations in a member of the Ras superfamily of small GTP-binding proteins causes Bardet-Biedl syndrome. *Nat Genet,* 36**,** 989-93.

FAQEIH, E., AL-AKASH, S. I., SAKATI, N. & TEEBI, P. A. 2007. Four siblings with distal renal tubular acidosis and nephrocalcinosis, neurobehavioral impairment, short stature, and distinctive facial appearance: a possible new autosomal recessive syndrome. *Am J Med Genet A,* 143A**,** 1951-7.

FARAG, T. I. & TEEBI, A. S. 1989. High incidence of Bardet Biedl syndrome among the Bedouin. *Clin Genet,* 36**,** 463-4.

FORSYTHE, E. & BEALES, P. L. 2012. Bardet-Biedl syndrome. *Eur J Hum Genet*.

FRANCESCHINI, N., NORTH, K. E., KOPP, J. B., MCKENZIE, L. & WINKLER, C. 2006. NPHS2 gene, nephrotic syndrome and focal segmental glomerulosclerosis: a HuGE review. *Genet Med,* 8**,** 63-75.

GERDES, J. M., LIU, Y., ZAGHLOUL, N. A., LEITCH, C. C., LAWSON, S. S., KATO, M., BEACHY, P. A., BEALES, P. L., DEMARTINO, G. N., FISHER, S., BADANO, J. L. & KATSANIS, N. 2007. Disruption of the basal body compromises proteasomal function and perturbs intracellular Wnt response. *Nat Genet,* 39**,** 1350-60.

GIGANTE, M., PONTRELLI, P., MONTEMURNO, E., ROCA, L., AUCELLA, F., PENZA, R., CARIDI, G., RANIERI, E., GHIGGERI, G. M. & GESUALDO, L. 2009. CD2AP mutations are associated with sporadic nephrotic syndrome and focal segmental glomerulosclerosis (FSGS). *Nephrol Dial Transplant,* 24**,** 1858-64.

GODEFROID, N. & DAHAN, K. 2010. Expanding the clinical spectrum of congenital nephrotic syndrome caused by NPHS1 mutations. *Nephrol Dial Transplant,* 25**,** 2837-9.

GONG, K. Q., YALLOWITZ, A. R., SUN, H., DRESSLER, G. R. & WELLIK, D. M. 2007. A Hox-Eya-Pax complex regulates early kidney developmental gene expression. *Mol Cell Biol,* 27**,** 7661-8.

GRAHAM, F. L., SMILEY, J., RUSSELL, W. C. & NAIRN, R. 1977. Characteristics of a human cell line transformed by DNA from human adenovirus type 5. *J Gen Virol,* 36**,** 59-74.

GRESCHIK, H., WURTZ, J. M., SANGLIER, S., BOURGUET, W., VAN DORSSELAER, A., MORAS, D. & RENAUD, J. P. 2002. Structural and

functional evidence for ligand-independent transcriptional activation by the estrogen-related receptor 3. *Mol Cell,* 9**,** 303-13.

GRIBOUVAL, O., GONZALES, M., NEUHAUS, T., AZIZA, J., BIETH, E., LAURENT, N., BOUTON, J. M., FEUILLET, F., MAKNI, S., BEN AMAR, H., LAUBE, G., DELEZOIDE, A. L., BOUVIER, R., DIJOUD, F., OLLAGNON-ROMAN, E., ROUME, J., JOUBERT, M., ANTIGNAC, C. & GUBLER, M. C. 2005. Mutations in genes in the renin-angiotensin system are associated with autosomal recessive renal tubular dysgenesis. *Nat Genet,* 37**,** 964-8.

GUDBJARTSSON, D. F., THORVALDSSON, T., KONG, A., GUNNARSSON, G. & INGOLFSDOTTIR, A. 2005. Allegro version 2. *Nat Genet,* 37**,** 1015-6.

HAREWOOD, L., LIU, M., KEELING, J., HOWATSON, A., WHITEFORD, M., BRANNEY, P., EVANS, M., FANTES, J. & FITZPATRICK, D. R. 2010. Bilateral renal agenesis/hypoplasia/dysplasia (BRAHD): postmortem analysis of 45 cases with breakpoint mapping of two de novo translocations. *PLoS One,* 5**,** e12375.

HEERINGA, S. F., VLANGOS, C. N., CHERNIN, G., HINKES, B., GBADEGESIN, R., LIU, J., HOSKINS, B. E., OZALTIN, F. & HILDEBRANDT, F. 2008. Thirteen novel NPHS1 mutations in a large cohort of children with congenital nephrotic syndrome. *Nephrol Dial Transplant,* 23**,** 3527-33.

HILDEBRANDT, F. 2010. Genetic kidney diseases. *Lancet,* 375**,** 1287-95.

HILDEBRANDT, F., HEERINGA, S. F., RUSCHENDORF, F., ATTANASIO, M., NURNBERG, G., BECKER, C., SEELOW, D., HUEBNER, N., CHERNIN, G., VLANGOS, C. N., ZHOU, W., O'TOOLE, J. F., HOSKINS, B. E., WOLF, M. T., HINKES, B. G., CHAIB, H., ASHRAF, S., SCHOEB, D. S., OVUNC, B., ALLEN, S. J., VEGA-WARNER, V., WISE, E., HARVILLE, H. M., LYONS, R. H., WASHBURN, J., MACDONALD, J., NURNBERG, P. & OTTO, E. A. 2009. A systematic approach to mapping recessive disease genes in individuals from outbred populations. *PLoS Genet,* 5**,** e1000353.

HINKES, B., WIGGINS, R. C., GBADEGESIN, R., VLANGOS, C. N., SEELOW, D., NURNBERG, G., GARG, P., VERMA, R., CHAIB, H., HOSKINS, B. E., ASHRAF, S., BECKER, C., HENNIES, H. C., GOYAL, M., WHARRAM, B. L., SCHACHTER, A. D., MUDUMANA, S., DRUMMOND, I., KERJASCHKI, D., WALDHERR, R., DIETRICH, A., OZALTIN, F., BAKKALOGLU, A., CLEPER, R., BASEL-VANAGAITE, L., POHL, M., GRIEBEL, M., TSYGIN, A. N., SOYLU, A., MULLER, D., SORLI, C. S., BUNNEY, T. D., KATAN, M., LIU, J., ATTANASIO, M., O'TOOLE J, F., HASSELBACHER, K., MUCHA, B., OTTO, E. A., AIRIK, R., KISPERT, A., KELLEY, G. G., SMRCKA, A. V., GUDERMANN, T., HOLZMAN, L. B., NURNBERG, P. & HILDEBRANDT, F. 2006. Positional cloning uncovers mutations in PLCE1 responsible for a nephrotic syndrome variant that may be reversible. *Nat Genet,* 38**,** 1397-405.

HINKES, B. G., MUCHA, B., VLANGOS, C. N., GBADEGESIN, R., LIU, J., HASSELBACHER, K., HANGAN, D., OZALTIN, F., ZENKER, M. & HILDEBRANDT, F. 2007. Nephrotic syndrome in the first year of life: two thirds of cases are caused by mutations in 4 genes (NPHS1, NPHS2, WT1, and LAMB2). *Pediatrics,* 119**,** e907-19.

HJORTSHOJ, T. D., GRONSKOV, K., PHILP, A. R., NISHIMURA, D. Y., ADEYEMO, A., ROTIMI, C. N., SHEFFIELD, V. C., ROSENBERG, T. & BRONDUM-NIELSEN, K. 2008. Novel mutations in BBS5 highlight the

importance of this gene in non-Caucasian Bardet-Biedl syndrome patients. *Am J Med Genet A,* 146A**,** 517-20.

HOEFELE, J., WOLF, M. T., O'TOOLE, J. F., OTTO, E. A., SCHULTHEISS, U., DESCHENES, G., ATTANASIO, M., UTSCH, B., ANTIGNAC, C. & HILDEBRANDT, F. 2007. Evidence of oligogenic inheritance in nephronophthisis. *J Am Soc Nephrol,* 18**,** 2789-95.

HOFFMANN, K. & LINDNER, T. H. 2005. easyLINKAGE-Plus--automated linkage analyses using large-scale SNP data. *Bioinformatics,* 21**,** 3565-7.

HOH, J. & OTT, J. 2000. Scan statistics to scan markers for susceptibility genes. *Proc Natl Acad Sci U S A,* 97**,** 9615-7.

HUANG, L., SZYMANSKA, K., JENSEN, V. L., JANECKE, A. R., INNES, A. M., DAVIS, E. E., FROSK, P., LI, C., WILLER, J. R., CHODIRKER, B. N., GREENBERG, C. R., MCLEOD, D. R., BERNIER, F. P., CHUDLEY, A. E., MULLER, T., SHBOUL, M., LOGAN, C. V., LOUCKS, C. M., BEAULIEU, C. L., BOWIE, R. V., BELL, S. M., ADKINS, J., ZUNIGA, F. I., ROSS, K. D., WANG, J., BAN, M. R., BECKER, C., NURNBERG, P., DOUGLAS, S., CRAFT, C. M., AKIMENKO, M. A., HEGELE, R. A., OBER, C., UTERMANN, G., BOLZ, H. J., BULMAN, D. E., KATSANIS, N., BLACQUE, O. E., DOHERTY, D., PARBOOSINGH, J. S., LEROUX, M. R., JOHNSON, C. A. & BOYCOTT, K. M. 2011. TMEM237 is mutated in individuals with a Joubert syndrome related disorder and expands the role of the TMEM family at the ciliary transition zone. *Am J Hum Genet,* 89**,** 713-30.

HUBER, T. B., SIMONS, M., HARTLEBEN, B., SERNETZ, L., SCHMIDTS, M., GUNDLACH, E., SALEEM, M. A., WALZ, G. & BENZING, T. 2003. Molecular basis of the functional podocin-nephrin complex: mutations in the NPHS2 gene disrupt nephrin targeting to lipid raft microdomains. *Hum Mol Genet,* 12**,** 3397-405.

IBANEZ-TALLON, I., HEINTZ, N. & OMRAN, H. 2003. To beat or not to beat: roles of cilia in development and disease. *Hum Mol Genet,* 12 Spec No 1**,** R27-35.

IGARASHI, T., INATOMI, J., SEKINE, T., CHA, S. H., KANAI, Y., KUNIMI, M., TSUKAMOTO, K., SATOH, H., SHIMADZU, M., TOZAWA, F., MORI, T., SHIOBARA, M., SEKI, G. & ENDOU, H. 1999. Mutations in SLC4A4 cause permanent isolated proximal renal tubular acidosis with ocular abnormalities. *Nat Genet,* 23**,** 264-6.

JALALAH, S. M. & JAMAL, A. A. 2009. Childhood primary glomerular diseases in the western region of Saudi Arabia. *Saudi J Kidney Dis Transpl,* 20**,** 608-12.

JAYE, M., HOWK, R., BURGESS, W., RICCA, G. A., CHIU, I. M., RAVERA, M. W., O'BRIEN, S. J., MODI, W. S., MACIAG, T. & DROHAN, W. N. 1986. Human endothelial cell growth factor: cloning, nucleotide sequence, and chromosome localization. *Science,* 233**,** 541-5.

JEANPIERRE, C., DENAMUR, E., HENRY, I., CABANIS, M. O., LUCE, S., CECILLE, A., ELION, J., PEUCHMAUR, M., LOIRAT, C., NIAUDET, P., GUBLER, M. C. & JUNIEN, C. 1998. Identification of constitutional WT1 mutations, in patients with isolated diffuse mesangial sclerosis, and analysis of genotype/phenotype correlations by use of a computerized mutation database. *Am J Hum Genet,* 62**,** 824-33.

KAPLAN, J. M., KIM, S. H., NORTH, K. N., RENNKE, H., CORREIA, L. A., TONG, H. Q., MATHIS, B. J., RODRIGUEZ-PEREZ, J. C., ALLEN, P. G., BEGGS, A. H. & POLLAK, M. R. 2000. Mutations in ACTN4, encoding alpha-actinin-4, cause familial focal segmental glomerulosclerosis. *Nat Genet,* 24**,** 251-6.

KARET, F. E., FINBERG, K. E., NAYIR, A., BAKKALOGLU, A., OZEN, S., HULTON, S. A., SANJAD, S. A., AL-SABBAN, E. A., MEDINA, J. F. & LIFTON, R. P. 1999a. Localization of a gene for autosomal recessive distal renal tubular acidosis with normal hearing (rdRTA2) to 7q33-34. *Am J Hum Genet,* 65**,** 1656-65.

KARET, F. E., FINBERG, K. E., NELSON, R. D., NAYIR, A., MOCAN, H., SANJAD, S. A., RODRIGUEZ-SORIANO, J., SANTOS, F., CREMERS, C. W., DI PIETRO, A., HOFFBRAND, B. I., WINIARSKI, J., BAKKALOGLU, A., OZEN, S., DUSUNSEL, R., GOODYER, P., HULTON, S. A., WU, D. K., SKVORAK, A. B., MORTON, C. C., CUNNINGHAM, M. J., JHA, V. & LIFTON, R. P. 1999b. Mutations in the gene encoding B1 subunit of H+-ATPase cause renal tubular acidosis with sensorineural deafness. *Nat Genet,* 21**,** 84-90.

KARI, J. A. 2002. Changing trends of histopathology in childhood nephrotic syndrome in western Saudi Arabia. *Saudi Med J,* 23**,** 317-21.

KARI, J. A. 2006. Chronic renal failure in children in the Western area of saudi arabia. *Saudi J Kidney Dis Transpl,* 17**,** 19-24.

KATSANIS, N., ANSLEY, S. J., BADANO, J. L., EICHERS, E. R., LEWIS, R. A., HOSKINS, B. E., SCAMBLER, P. J., DAVIDSON, W. S., BEALES, P. L. & LUPSKI, J. R. 2001. Triallelic inheritance in Bardet-Biedl syndrome, a Mendelian recessive disorder. *Science,* 293**,** 2256-9.

KATZIR, Z., DINOUR, D., REZNIK-WOLF, H., NISSENKORN, A. & HOLTZMAN, E. 2008. Familial pure proximal renal tubular acidosis--a clinical and genetic study. *Nephrol Dial Transplant,* 23**,** 1211-5.

KELALIS, P. P., KING, L. R. & BELMAN, A. B. 1985. *Clinical pediatric urology,* Philadelphia, Saunders.

KERECUK, L., SCHREUDER, M. F. & WOOLF, A. S. 2008. Renal tract malformations: perspectives for nephrologists. *Nat Clin Pract Nephrol,* 4**,** 312-25.

KESTILA, M., LENKKERI, U., MANNIKKO, M., LAMERDIN, J., MCCREADY, P., PUTAALA, H., RUOTSALAINEN, V., MORITA, T., NISSINEN, M., HERVA, R., KASHTAN, C. E., PELTONEN, L., HOLMBERG, C., OLSEN, A. & TRYGGVASON, K. 1998. Positionally cloned gene for a novel glomerular protein--nephrin--is mutated in congenital nephrotic syndrome. *Mol Cell,* 1**,** 575-82.

KIM, J. C., OU, Y. Y., BADANO, J. L., ESMAIL, M. A., LEITCH, C. C., FIEDRICH, E., BEALES, P. L., ARCHIBALD, J. M., KATSANIS, N., RATTNER, J. B. & LEROUX, M. R. 2005. MKKS/BBS6, a divergent chaperonin-like protein linked to the obesity disorder Bardet-Biedl syndrome, is a novel centrosomal component required for cytokinesis. *J Cell Sci,* 118**,** 1007-20.

KIM, J. M., WU, H., GREEN, G., WINKLER, C. A., KOPP, J. B., MINER, J. H., UNANUE, E. R. & SHAW, A. S. 2003. CD2-associated protein haploinsufficiency is linked to glomerular disease susceptibility. *Science,* 300**,** 1298-300.

KIM, S. K., SHINDO, A., PARK, T. J., OH, E. C., GHOSH, S., GRAY, R. S., LEWIS, R. A., JOHNSON, C. A., ATTIE-BITTACH, T., KATSANIS, N. & WALLINGFORD, J. B. 2010. Planar cell polarity acts through septins to control collective cell movement and ciliogenesis. *Science,* 329**,** 1337-40.

KOHLHASE, J., WISCHERMANN, A., REICHENBACH, H., FROSTER, U. & ENGEL, W. 1998. Mutations in the SALL1 putative transcription factor gene cause Townes-Brocks syndrome. *Nat Genet,* 18**,** 81-3.

KONG, A. & COX, N. J. 1997. Allele-sharing models: LOD scores and accurate linkage tests. *Am J Hum Genet,* 61**,** 1179-88.

KONRAD, M., SCHALLER, A., SEELOW, D., PANDEY, A. V., WALDEGGER, S., LESSLAUER, A., VITZTHUM, H., SUZUKI, Y., LUK, J. M., BECKER, C., SCHLINGMANN, K. P., SCHMID, M., RODRIGUEZ-SORIANO, J., ARICETA, G., CANO, F., ENRIQUEZ, R., JUPPNER, H., BAKKALOGLU, S. A., HEDIGER, M. A., GALLATI, S., NEUHAUSS, S. C., NURNBERG, P. & WEBER, S. 2006. Mutations in the tight-junction gene claudin 19 (CLDN19) are associated with renal magnesium wasting, renal failure, and severe ocular involvement. *Am J Hum Genet,* 79**,** 949-57.

KOZIELL, A., GRECH, V., HUSSAIN, S., LEE, G., LENKKERI, U., TRYGGVASON, K. & SCAMBLER, P. 2002. Genotype/phenotype correlations of NPHS1 and NPHS2 mutations in nephrotic syndrome advocate a functional inter-relationship in glomerular filtration. *Hum Mol Genet,* 11**,** 379-88.

KREMER, E. J., PRITCHARD, M., LYNCH, M., YU, S., HOLMAN, K., BAKER, E., WARREN, S. T., SCHLESSINGER, D., SUTHERLAND, G. R. & RICHARDS, R. I. 1991. Mapping of DNA instability at the fragile X to a trinucleotide repeat sequence p(CCG)n. *Science,* 252**,** 1711-4.

KRUGLYAK, L., DALY, M. J., REEVE-DALY, M. P. & LANDER, E. S. 1996. Parametric and nonparametric linkage analysis: a unified multipoint approach. *Am J Hum Genet,* 58**,** 1347-63.

KRUGLYAK, L. & LANDER, E. S. 1998. Faster multipoint linkage analysis using Fourier transforms. *J Comput Biol,* 5**,** 1-7.

LAING, C. M., TOYE, A. M., CAPASSO, G. & UNWIN, R. J. 2005. Renal tubular acidosis: developments in our understanding of the molecular basis. *Int J Biochem Cell Biol,* 37**,** 1151-61.

LANDER, E. S. & BOTSTEIN, D. 1987. Homozygosity mapping: a way to map human recessive traits with the DNA of inbred children. *Science,* 236**,** 1567-70.

LANGE, M., KAYNAK, B., FORSTER, U. B., TONJES, M., FISCHER, J. J., GRIMM, C., SCHLESINGER, J., JUST, S., DUNKEL, I., KRUEGER, T., MEBUS, S., LEHRACH, H., LURZ, R., GOBOM, J., ROTTBAUER, W., ABDELILAH-SEYFRIED, S. & SPERLING, S. 2008. Regulation of muscle development by DPF3, a novel histone acetylation and methylation reader of the BAF chromatin remodeling complex. *Genes Dev,* 22**,** 2370-84.

LECHNER, M. S. & DRESSLER, G. R. 1997. The molecular basis of embryonic kidney development. *Mech Dev,* 62**,** 105-20.

LEITCH, C. C., ZAGHLOUL, N. A., DAVIS, E. E., STOETZEL, C., DIAZ-FONT, A., RIX, S., ALFADHEL, M., LEWIS, R. A., EYAID, W., BANIN, E., DOLLFUS, H., BEALES, P. L., BADANO, J. L. & KATSANIS, N. 2008. Hypomorphic mutations in syndromic encephalocele genes are associated with Bardet-Biedl syndrome. *Nat Genet,* 40**,** 443-8.

LENKKERI, U., MANNIKKO, M., MCCREADY, P., LAMERDIN, J., GRIBOUVAL, O., NIAUDET, P. M., ANTIGNAC, C. K., KASHTAN, C. E., HOMBERG, C., OLSEN, A., KESTILA, M. & TRYGGVASON, K. 1999. Structure of the gene for congenital nephrotic syndrome of the finnish type (NPHS1) and characterization of mutations. *Am J Hum Genet,* 64**,** 51-61.

LI, J. B., GERDES, J. M., HAYCRAFT, C. J., FAN, Y., TESLOVICH, T. M., MAY-SIMERA, H., LI, H., BLACQUE, O. E., LI, L., LEITCH, C. C., LEWIS, R. A., GREEN, J. S., PARFREY, P. S., LEROUX, M. R., DAVIDSON, W. S., BEALES, P. L., GUAY-WOODFORD, L. M., YODER, B. K., STORMO, G. D., KATSANIS, N. & DUTCHER, S. K. 2004. Comparative genomics identifies a flagellar and basal body proteome that includes the BBS5 human disease gene. *Cell,* 117**,** 541-52.

LINDNER, T. H. & HOFFMANN, K. 2005. easyLINKAGE: a PERL script for easy and automated two-/multi-point linkage analyses. *Bioinformatics,* 21**,** 405-7.

LOWIK, M., LEVTCHENKO, E., WESTRA, D., GROENEN, P., STEENBERGEN, E., WEENING, J., LILIEN, M., MONNENS, L. & VAN DEN HEUVEL, L. 2008. Bigenic heterozygosity and the development of steroid-resistant focal segmental glomerulosclerosis. *Nephrol Dial Transplant,* 23**,** 3146-51.

M'HAMDI, O., OUERTANI, I., MAAZOUL, F. & CHAABOUNI-BOUHAMED, H. 2011. Prevalence of Bardet-Biedl syndrome in Tunisia. *J Community Genet,* 2**,** 97-9.

MACHUCA, E., BENOIT, G. & ANTIGNAC, C. 2009. Genetics of nephrotic syndrome: connecting molecular genetics to podocyte physiology. *Hum Mol Genet,* 18**,** R185-94.

MACHUCA, E., BENOIT, G., NEVO, F., TETE, M. J., GRIBOUVAL, O., PAWTOWSKI, A., BRANDSTROM, P., LOIRAT, C., NIAUDET, P., GUBLER, M. C. & ANTIGNAC, C. 2010. Genotype-phenotype correlations in non-Finnish congenital nephrotic syndrome. *J Am Soc Nephrol,* 21**,** 1209-17.

MATSUZAKI, H., LOI, H., DONG, S., TSAI, Y. Y., FANG, J., LAW, J., DI, X., LIU, W. M., YANG, G., LIU, G., HUANG, J., KENNEDY, G. C., RYDER, T. B., MARCUS, G. A., WALSH, P. S., SHRIVER, M. D., PUCK, J. M., JONES, K. W. & MEI, R. 2004. Parallel genotyping of over 10,000 SNPs using a one-primer assay on a high-density oligonucleotide array. *Genome Res,* 14**,** 414-25.

MATTOO, T. K. 1998. Genetically transmitted renal diseases in children: a saudi perspective. *Saudi J Kidney Dis Transpl,* 9**,** 105-9.

MATTOO, T. K., MAHMOOD, M. A. & AL-HARBI, M. S. 1990. Nephrotic syndrome in Saudi children clinicopathological study of 150 cases. *Pediatr Nephrol,* 4**,** 517-9.

MELE, C., IATROPOULOS, P., DONADELLI, R., CALABRIA, A., MARANTA, R., CASSIS, P., BUELLI, S., TOMASONI, S., PIRAS, R., KRENDEL, M., BETTONI, S., MORIGI, M., DELLEDONNE, M., PECORARO, C., ABBATE, I., CAPOBIANCHI, M. R., HILDEBRANDT, F., OTTO, E., SCHAEFER, F., MACCIARDI, F., OZALTIN, F., EMRE, S., IBSIRLIOGLU, T., BENIGNI, A., REMUZZI, G. & NORIS, M. 2011. MYO1E mutations and childhood familial focal segmental glomerulosclerosis. *N Engl J Med,* 365**,** 295-306.

MITWALLI, A. H., AL WAKEEL, J. S., AL MOHAYA, S. S., MALIK, H. G., ABU-AISHA, H., HASSAN, O. S. & AKHTAR, M. 1996. Pattern of glomerular disease in Saudi Arabia. *Am J Kidney Dis,* 27**,** 797-802.

MOORE, S. J., GREEN, J. S., FAN, Y., BHOGAL, A. K., DICKS, E., FERNANDEZ, B. A., STEFANELLI, M., MURPHY, C., CRAMER, B. C., DEAN, J. C., BEALES, P. L., KATSANIS, N., BASSETT, A. S., DAVIDSON, W. S. & PARFREY, P. S. 2005. Clinical and genetic epidemiology of Bardet-Biedl syndrome in Newfoundland: a 22-year prospective, population-based, cohort study. *Am J Med Genet A,* 132**,** 352-60.

MORRIS, M. C., CHAMBERS, T. L., EVANS, P. W., MALLESON, P. N., PINCOTT, J. R. & ROSE, G. A. 1982. Oxalosis in infancy. *Arch Dis Child,* 57**,** 224-8.

MUNDEL, P. & SHANKLAND, S. J. 2002. Podocyte biology and response to injury. *J Am Soc Nephrol,* 13**,** 3005-15.

MYKYTYN, K., NISHIMURA, D. Y., SEARBY, C. C., SHASTRI, M., YEN, H. J., BECK, J. S., BRAUN, T., STREB, L. M., CORNIER, A. S., COX, G. F., FULTON, A. B., CARMI, R., LULECI, G., CHANDRASEKHARAPPA, S. C., COLLINS, F. S., JACOBSON, S. G., HECKENLIVELY, J. R., WELEBER, R. G., STONE, E. M. & SHEFFIELD, V. C. 2002. Identification of the gene (BBS1) most commonly involved in Bardet-Biedl syndrome, a complex human obesity syndrome. *Nat Genet,* 31**,** 435-8.

NACHURY, M. V., LOKTEV, A. V., ZHANG, Q., WESTLAKE, C. J., PERANEN, J., MERDES, A., SLUSARSKI, D. C., SCHELLER, R. H., BAZAN, J. F., SHEFFIELD, V. C. & JACKSON, P. K. 2007. A core complex of BBS proteins cooperates with the GTPase Rab8 to promote ciliary membrane biogenesis. *Cell,* 129**,** 1201-13.

NANNYA, Y., SANADA, M., NAKAZAKI, K., HOSOYA, N., WANG, L., HANGAISHI, A., KUROKAWA, M., CHIBA, S., BAILEY, D. K., KENNEDY, G. C. & OGAWA, S. 2005. A robust algorithm for copy number detection using high-density oligonucleotide single nucleotide polymorphism genotyping arrays. *Cancer Res,* 65**,** 6071-9.

NATIONAL ORGANIZATION FOR RARE DISORDERS. 1992. Physicians' guide to rare diseases. Montvale, N.J.: Dowden Pub. Co.

NISHIMURA, D. Y., SEARBY, C. C., CARMI, R., ELBEDOUR, K., VAN MALDERGEM, L., FULTON, A. B., LAM, B. L., POWELL, B. R., SWIDERSKI, R. E., BUGGE, K. E., HAIDER, N. B., KWITEK-BLACK, A. E., YING, L., DUHL, D. M., GORMAN, S. W., HEON, E., IANNACCONE, A., BONNEAU, D., BIESECKER, L. G., JACOBSON, S. G., STONE, E. M. & SHEFFIELD, V. C. 2001. Positional cloning of a novel gene on chromosome 16q causing Bardet-Biedl syndrome (BBS2). *Hum Mol Genet,* 10**,** 865-74.

NISHIMURA, D. Y., SWIDERSKI, R. E., SEARBY, C. C., BERG, E. M., FERGUSON, A. L., HENNEKAM, R., MERIN, S., WELEBER, R. G., BIESECKER, L. G., STONE, E. M. & SHEFFIELD, V. C. 2005. Comparative genomics and gene expression analysis identifies BBS9, a new Bardet-Biedl syndrome gene. *Am J Hum Genet,* 77**,** 1021-33.

OBEIDOVA, H., MERTA, M., REITEROVA, J., MAIXNEROVA, D., STEKROVA, J., RYSAVA, R. & TESAR, V. 2006. Genetic basis of nephrotic syndrome--review. *Prague Med Rep,* 107**,** 5-16.

OBERLE, I., ROUSSEAU, F., HEITZ, D., KRETZ, C., DEVYS, D., HANAUER, A., BOUE, J., BERTHEAS, M. & MANDEL, J. 1991. Instability of a 550-base pair DNA segment and abnormal methylation in fragile X syndrome. *Science,* 252**,** 1097-102.

OTTO, E. A., HURD, T. W., AIRIK, R., CHAKI, M., ZHOU, W., STOETZEL, C., PATIL, S. B., LEVY, S., GHOSH, A. K., MURGA-ZAMALLOA, C. A., VAN REEUWIJK, J., LETTEBOER, S. J., SANG, L., GILES, R. H., LIU, Q., COENE, K. L., ESTRADA-CUZCANO, A., COLLIN, R. W., MCLAUGHLIN, H. M., HELD, S., KASANUKI, J. M., RAMASWAMI, G., CONTE, J., LOPEZ, I., WASHBURN, J., MACDONALD, J., HU, J., YAMASHITA, Y., MAHER, E. R., GUAY-WOODFORD, L. M., NEUMANN, H. P., OBERMULLER, N., KOENEKOOP, R. K., BERGMANN, C., BEI, X., LEWIS, R. A., KATSANIS,

N., LOPES, V., WILLIAMS, D. S., LYONS, R. H., DANG, C. V., BRITO, D. A., DIAS, M. B., ZHANG, X., CAVALCOLI, J. D., NURNBERG, G., NURNBERG, P., PIERCE, E. A., JACKSON, P. K., ANTIGNAC, C., SAUNIER, S., ROEPMAN, R., DOLLFUS, H., KHANNA, H. & HILDEBRANDT, F. 2010. Candidate exome capture identifies mutation of SDCCAG8 as the cause of a retinal-renal ciliopathy. *Nat Genet,* 42**,** 840-50.

OTTO, E. A., SCHERMER, B., OBARA, T., O'TOOLE, J. F., HILLER, K. S., MUELLER, A. M., RUF, R. G., HOEFELE, J., BEEKMANN, F., LANDAU, D., FOREMAN, J. W., GOODSHIP, J. A., STRACHAN, T., KISPERT, A., WOLF, M. T., GAGNADOUX, M. F., NIVET, H., ANTIGNAC, C., WALZ, G., DRUMMOND, I. A., BENZING, T. & HILDEBRANDT, F. 2003. Mutations in INVS encoding inversin cause nephronophthisis type 2, linking renal cystic disease to the function of primary cilia and left-right axis determination. *Nat Genet,* 34**,** 413-20.

OZALTIN, F., IBSIRLIOGLU, T., TASKIRAN, E. Z., BAYDAR, D. E., KAYMAZ, F., BUYUKCELIK, M., KILIC, B. D., BALAT, A., IATROPOULOS, P., ASAN, E., AKARSU, N. A., SCHAEFER, F., YILMAZ, E. & BAKKALOGLU, A. 2012. Disruption of PTPRO causes childhood-onset nephrotic syndrome. *Am J Hum Genet,* 89**,** 139-47.

PAPIC, L., FISCHER, D., TRAJANOSKI, S., HOFTBERGER, R., FISCHER, C., STROBEL, T., SCHMIDT, W. M., BITTNER, R. E., SCHABHUTTL, M., GRUBER, K., PIEBER, T. R., JANECKE, A. R. & AUER-GRUMBACH, M. 2011. SNP-array based whole genome homozygosity mapping: a quick and powerful tool to achieve an accurate diagnosis in LGMD2 patients. *Eur J Med Genet,* 54**,** 214-9.

PARFITT, D. A., MICHAEL, G. J., VERMEULEN, E. G., PRODROMOU, N. V., WEBB, T. R., GALLO, J. M., CHEETHAM, M. E., NICOLL, W. S., BLATCH, G. L. & CHAPPLE, J. P. 2009. The ataxia protein sacsin is a functional co-chaperone that protects against polyglutamine-expanded ataxin-1. *Hum Mol Genet,* 18**,** 1556-65.

PEREIRO, I., HOSKINS, B. E., MARSHALL, J. D., COLLIN, G. B., NAGGERT, J. K., PINEIRO-GALLEGO, T., OITMAA, E., KATSANIS, N., VALVERDE, D. & BEALES, P. L. 2011. Arrayed primer extension technology simplifies mutation detection in Bardet-Biedl and Alstrom syndrome. *Eur J Hum Genet,* 19**,** 485-8.

POHL, M., BHATNAGAR, V., MENDOZA, S. A. & NIGAM, S. K. 2002. Toward an etiological classification of developmental disorders of the kidney and upper urinary tract. *Kidney Int,* 61**,** 10-9.

PRAGA, M., VARA, J., GONZALEZ-PARRA, E., ANDRES, A., ALAMO, C., ARAQUE, A., ORTIZ, A. & RODICIO, J. L. 1995. Familial hypomagnesemia with hypercalciuria and nephrocalcinosis. *Kidney Int,* 47**,** 1419-25.

RAJAB, A., BAPPAL, B., AL-SHAIKH, H., AL-KHUSAIBI, S. & MOHAMMED, A. J. 2005. Common autosomal recessive diseases in Oman derived from a hospital-based registry. *Community Genet,* 8**,** 27-30.

RAJADHYAKSHA, A. M., ELEMENTO, O., PUFFENBERGER, E. G., SCHIERBERL, K. C., XIANG, J. Z., PUTORTI, M. L., BERCIANO, J., POULIN, C., BRAIS, B., MICHAELIDES, M., WELEBER, R. G. & HIGGINS, J. J. 2010. Mutations in FLVCR1 cause posterior column ataxia and retinitis pigmentosa. *Am J Hum Genet,* 87**,** 643-54.

ROODHOOFT, A. M., BIRNHOLZ, J. C. & HOLMES, L. B. 1984. Familial nature of congenital absence and severe dysgenesis of both kidneys. *N Engl J Med,* 310**,** 1341-5.

ROSS, A. J., MAY-SIMERA, H., EICHERS, E. R., KAI, M., HILL, J., JAGGER, D. J., LEITCH, C. C., CHAPPLE, J. P., MUNRO, P. M., FISHER, S., TAN, P. L., PHILLIPS, H. M., LEROUX, M. R., HENDERSON, D. J., MURDOCH, J. N., COPP, A. J., ELIOT, M. M., LUPSKI, J. R., KEMP, D. T., DOLLFUS, H., TADA, M., KATSANIS, N., FORGE, A. & BEALES, P. L. 2005. Disruption of Bardet-Biedl syndrome ciliary proteins perturbs planar cell polarity in vertebrates. *Nat Genet,* 37**,** 1135-40.

RUF, R. G., LICHTENBERGER, A., KARLE, S. M., HAAS, J. P., ANACLETO, F. E., SCHULTHEISS, M., ZALEWSKI, I., IMM, A., RUF, E. M., MUCHA, B., BAGGA, A., NEUHAUS, T., FUCHSHUBER, A., BAKKALOGLU, A. & HILDEBRANDT, F. 2004a. Patients with mutations in NPHS2 (podocin) do not respond to standard steroid treatment of nephrotic syndrome. *J Am Soc Nephrol,* 15**,** 722-32.

RUF, R. G., XU, P. X., SILVIUS, D., OTTO, E. A., BEEKMANN, F., MUERB, U. T., KUMAR, S., NEUHAUS, T. J., KEMPER, M. J., RAYMOND, R. M., JR., BROPHY, P. D., BERKMAN, J., GATTAS, M., HYLAND, V., RUF, E. M., SCHWARTZ, C., CHANG, E. H., SMITH, R. J., STRATAKIS, C. A., WEIL, D., PETIT, C. & HILDEBRANDT, F. 2004b. SIX1 mutations cause branchio-oto-renal syndrome by disruption of EYA1-SIX1-DNA complexes. *Proc Natl Acad Sci U S A,* 101**,** 8090-5.

SANNA-CHERCHI, S., BURGESS, K. E., NEES, S. N., CARIDI, G., WENG, P. L., DAGNINO, M., BODRIA, M., CARREA, A., ALLEGRETTA, M. A., KIM, H. R., PERRY, B. J., GIGANTE, M., CLARK, L. N., KISSELEV, S., CUSI, D., GESUALDO, L., ALLEGRI, L., SCOLARI, F., D'AGATI, V., SHAPIRO, L. S., PECORARO, C., PALOMERO, T., GHIGGERI, G. M. & GHARAVI, A. G. 2011. Exome sequencing identified MYO1E and NEIL1 as candidate genes for human autosomal recessive steroid-resistant nephrotic syndrome. *Kidney Int,* 80**,** 389-96.

SANNA-CHERCHI, S., CARIDI, G., WENG, P. L., SCOLARI, F., PERFUMO, F., GHARAVI, A. G. & GHIGGERI, G. M. 2007. Genetic approaches to human renal agenesis/hypoplasia and dysplasia. *Pediatr Nephrol,* 22**,** 1675-84.

SANTIN, S., BULLICH, G., TAZON-VEGA, B., GARCIA-MASET, R., GIMENEZ, I., SILVA, I., RUIZ, P., BALLARIN, J., TORRA, R. & ARS, E. 2011. Clinical utility of genetic testing in children and adults with steroid-resistant nephrotic syndrome. *Clin J Am Soc Nephrol,* 6**,** 1139-48.

SANYANUSIN, P., SCHIMMENTI, L. A., MCNOE, L. A., WARD, T. A., PIERPONT, M. E., SULLIVAN, M. J., DOBYNS, W. B. & ECCLES, M. R. 1995. Mutation of the PAX2 gene in a family with optic nerve colobomas, renal anomalies and vesicoureteral reflux. *Nat Genet,* 9**,** 358-64.

SCHLUNEGGER, M. P. & GRUTTER, M. G. 1992. An unusual feature revealed by the crystal structure at 2.2 A resolution of human transforming growth factor-beta 2. *Nature,* 358**,** 430-4.

SCHMITTGEN, T. D. & LIVAK, K. J. 2008. Analyzing real-time PCR data by the comparative C(T) method. *Nat Protoc,* 3**,** 1101-8.

SCHOEB, D. S., CHERNIN, G., HEERINGA, S. F., MATEJAS, V., HELD, S., VEGA-WARNER, V., BOCKENHAUER, D., VLANGOS, C. N., MOORANI, K. N., NEUHAUS, T. J., KARI, J. A., MACDONALD, J., SAISAWAT, P., ASHRAF,

S., OVUNC, B., ZENKER, M. & HILDEBRANDT, F. 2010. Nineteen novel NPHS1 mutations in a worldwide cohort of patients with congenital nephrotic syndrome (CNS). *Nephrol Dial Transplant,* 25**,** 2970-6.

SCHULTHEISS, M., RUF, R. G., MUCHA, B. E., WIGGINS, R., FUCHSHUBER, A., LICHTENBERGER, A. & HILDEBRANDT, F. 2004. No evidence for genotype/phenotype correlation in NPHS1 and NPHS2 mutations. *Pediatr Nephrol,* 19**,** 1340-8.

SELF, M., LAGUTIN, O. V., BOWLING, B., HENDRIX, J., CAI, Y., DRESSLER, G. R. & OLIVER, G. 2006. Six2 is required for suppression of nephrogenesis and progenitor renewal in the developing kidney. *EMBO J,* 25**,** 5214-28.

SHAHEEN, F. A. & AL-KHADER, A. A. 2005. Preventive strategies of renal failure in the Arab world. *Kidney Int Suppl***,** S37-40.

SHAHEEN, R., FAQEIH, E., ALSHAMMARI, M. J., SWAID, A., AL-GAZALI, L., MARDAWI, E., ANSARI, S., SOGATY, S., SEIDAHMED, M. Z., ALMOTAIRI, M. I., FARRA, C., KURDI, W., AL-RASHEED, S. & ALKURAYA, F. S. 2012. Genomic analysis of Meckel-Gruber syndrome in Arabs reveals marked genetic heterogeneity and novel candidate genes. *Eur J Hum Genet*.

SIMMS, R. J., HYNES, A. M., ELEY, L., INGLIS, D., CHAUDHRY, B., DAWE, H. R. & SAYER, J. A. 2012. Modelling a ciliopathy: Ahi1 knockdown in model systems reveals an essential role in brain, retinal, and renal development. *Cell Mol Life Sci,* 69**,** 993-1009.

SIMON, D. B., LU, Y., CHOATE, K. A., VELAZQUEZ, H., AL-SABBAN, E., PRAGA, M., CASARI, G., BETTINELLI, A., COLUSSI, G., RODRIGUEZ-SORIANO, J., MCCREDIE, D., MILFORD, D., SANJAD, S. & LIFTON, R. P. 1999. Paracellin-1, a renal tight junction protein required for paracellular Mg2+ resorption. *Science,* 285**,** 103-6.

SKINNER, M. A., SAFFORD, S. D., REEVES, J. G., JACKSON, M. E. & FREEMERMAN, A. J. 2008. Renal aplasia in humans is associated with RET mutations. *Am J Hum Genet,* 82**,** 344-51.

SLY, W. S., HEWETT-EMMETT, D., WHYTE, M. P., YU, Y. S. & TASHIAN, R. E. 1983. Carbonic anhydrase II deficiency identified as the primary defect in the autosomal recessive syndrome of osteopetrosis with renal tubular acidosis and cerebral calcification. *Proc Natl Acad Sci U S A,* 80**,** 2752-6.

STOETZEL, C., LAURIER, V., DAVIS, E. E., MULLER, J., RIX, S., BADANO, J. L., LEITCH, C. C., SALEM, N., CHOUERY, E., CORBANI, S., JALK, N., VICAIRE, S., SARDA, P., HAMEL, C., LACOMBE, D., HOLDER, M., ODENT, S., HOLDER, S., BROOKS, A. S., ELCIOGLU, N. H., SILVA, E. D., ROSSILLION, B., SIGAUDY, S., DE RAVEL, T. J., LEWIS, R. A., LEHEUP, B., VERLOES, A., AMATI-BONNEAU, P., MEGARBANE, A., POCH, O., BONNEAU, D., BEALES, P. L., MANDEL, J. L., KATSANIS, N. & DOLLFUS, H. 2006. BBS10 encodes a vertebrate-specific chaperonin-like protein and is a major BBS locus. *Nat Genet,* 38**,** 521-4.

STOETZEL, C., MULLER, J., LAURIER, V., DAVIS, E. E., ZAGHLOUL, N. A., VICAIRE, S., JACQUELIN, C., PLEWNIAK, F., LEITCH, C. C., SARDA, P., HAMEL, C., DE RAVEL, T. J., LEWIS, R. A., FRIEDERICH, E., THIBAULT, C., DANSE, J. M., VERLOES, A., BONNEAU, D., KATSANIS, N., POCH, O., MANDEL, J. L. & DOLLFUS, H. 2007. Identification of a novel BBS gene (BBS12) highlights the major role of a vertebrate-specific branch of chaperonin-related proteins in Bardet-Biedl syndrome. *Am J Hum Genet,* 80**,** 1-11.

SUMIYAMA, K., WASHIO-WATANABE, K., SAITOU, N., HAYAKAWA, T. & UEDA, S. 1996. Class III POU genes: generation of homopolymeric amino acid repeats under GC pressure in mammals. *J Mol Evol,* 43**,** 170-8.

SUN, Z., AMSTERDAM, A., PAZOUR, G. J., COLE, D. G., MILLER, M. S. & HOPKINS, N. 2004. A genetic screen in zebrafish identifies cilia genes as a principal cause of cystic kidney. *Development,* 131**,** 4085-93.

SUN, Z. & HOPKINS, N. 2001. vhnf1, the MODY5 and familial GCKD-associated gene, regulates regional specification of the zebrafish gut, pronephros, and hindbrain. *Genes Dev,* 15**,** 3217-29.

SUTHERLAND, G. R. & BAIN, A. D. 1972. Culture of cells from the urine of newborn children. *Nature,* 239**,** 231.

TEEBI, A. S., TEEBI, S. A., PORTER, C. J. & CUTICCHIA, A. J. 2002. Arab genetic disease database (AGDDB): a population-specific clinical and mutation database. *Hum Mutat,* 19**,** 615-21.

WEBB, T. R., PARFITT, D. A., GARDNER, J. C., MARTINEZ, A., BEVILACQUA, D., DAVIDSON, A. E., ZITO, I., THISELTON, D. L., RESSA, J. H., APERGI, M., SCHWARZ, N., KANUGA, N., MICHAELIDES, M., CHEETHAM, M. E., GORIN, M. B. & HARDCASTLE, A. J. 2012. Deep intronic mutation in OFD1, identified by targeted genomic next-generation sequencing, causes a severe form of X-linked retinitis pigmentosa (RP23). *Hum Mol Genet,* 21**,** 3647-54.

WEBER, S., GRIBOUVAL, O., ESQUIVEL, E. L., MORINIERE, V., TETE, M. J., LEGENDRE, C., NIAUDET, P. & ANTIGNAC, C. 2004. NPHS2 mutation analysis shows genetic heterogeneity of steroid-resistant nephrotic syndrome and low post-transplant recurrence. *Kidney Int,* 66**,** 571-9.

WEBER, S., TAYLOR, J. C., WINYARD, P., BAKER, K. F., SULLIVAN-BROWN, J., SCHILD, R., KNUPPEL, T., ZUROWSKA, A. M., CALDAS-ALFONSO, A., LITWIN, M., EMRE, S., GHIGGERI, G. M., BAKKALOGLU, A., MEHLS, O., ANTIGNAC, C., NETWORK, E., SCHAEFER, F. & BURDINE, R. D. 2008. SIX2 and BMP4 mutations associate with anomalous kidney development. *J Am Soc Nephrol,* 19**,** 891-903.

WINN, M. P., CONLON, P. J., LYNN, K. L., FARRINGTON, M. K., CREAZZO, T., HAWKINS, A. F., DASKALAKIS, N., KWAN, S. Y., EBERSVILLER, S., BURCHETTE, J. L., PERICAK-VANCE, M. A., HOWELL, D. N., VANCE, J. M. & ROSENBERG, P. B. 2005. A mutation in the TRPC6 cation channel causes familial focal segmental glomerulosclerosis. *Science,* 308**,** 1801-4.

WU, H., BOACKLE, S. A., HANVIVADHANAKUL, P., ULGIATI, D., GROSSMAN, J. M., LEE, Y., SHEN, N., ABRAHAM, L. J., MERCER, T. R., PARK, E., HEBERT, L. A., ROVIN, B. H., BIRMINGHAM, D. J., CHANG, D. M., CHEN, C. J., MCCURDY, D., BADSHA, H. M., THONG, B. Y., CHNG, H. H., ARNETT, F. C., WALLACE, D. J., YU, C. Y., HAHN, B. H., CANTOR, R. M. & TSAO, B. P. 2007. Association of a common complement receptor 2 haplotype with increased risk of systemic lupus erythematosus. *Proc Natl Acad Sci U S A,* 104**,** 3961-6.

YALAVARTHY, R. & PARIKH, C. R. 2003. Congenital renal agenesis: a review. *Saudi J Kidney Dis Transpl,* 14**,** 336-41.

YEN, H. J., TAYEH, M. K., MULLINS, R. F., STONE, E. M., SHEFFIELD, V. C. & SLUSARSKI, D. C. 2006. Bardet-Biedl syndrome genes are important in retrograde intracellular trafficking and Kupffer's vesicle cilia function. *Hum Mol Genet,* 15**,** 667-77.

ZAGHLOUL, N. A. & KATSANIS, N. 2009. Mechanistic insights into Bardet-Biedl syndrome, a model ciliopathy. *J Clin Invest,* 119**,** 428-37.

ZAGHLOUL, N. A., LIU, Y., GERDES, J. M., GASCUE, C., OH, E. C., LEITCH, C. C., BROMBERG, Y., BINKLEY, J., LEIBEL, R. L., SIDOW, A., BADANO, J. L. & KATSANIS, N. 2010. Functional analyses of variants reveal a significant role for dominant negative and common alleles in oligogenic Bardet-Biedl syndrome. *Proc Natl Acad Sci U S A,* 107**,** 10602-7.

ZAKZOUK, S. 2002. Consanguinity and hearing impairment in developing countries: a custom to be discouraged. *J Laryngol Otol,* 116**,** 811-6.

ZENKER, M., AIGNER, T., WENDLER, O., TRALAU, T., MUNTEFERING, H., FENSKI, R., PITZ, S., SCHUMACHER, V., ROYER-POKORA, B., WUHL, E., COCHAT, P., BOUVIER, R., KRAUS, C., MARK, K., MADLON, H., DOTSCH, J., RASCHER, W., MARUNIAK-CHUDEK, I., LENNERT, T., NEUMANN, L. M. & REIS, A. 2004. Human laminin beta2 deficiency causes congenital nephrosis with mesangial sclerosis and distinct eye abnormalities. *Hum Mol Genet,* 13**,** 2625-32.