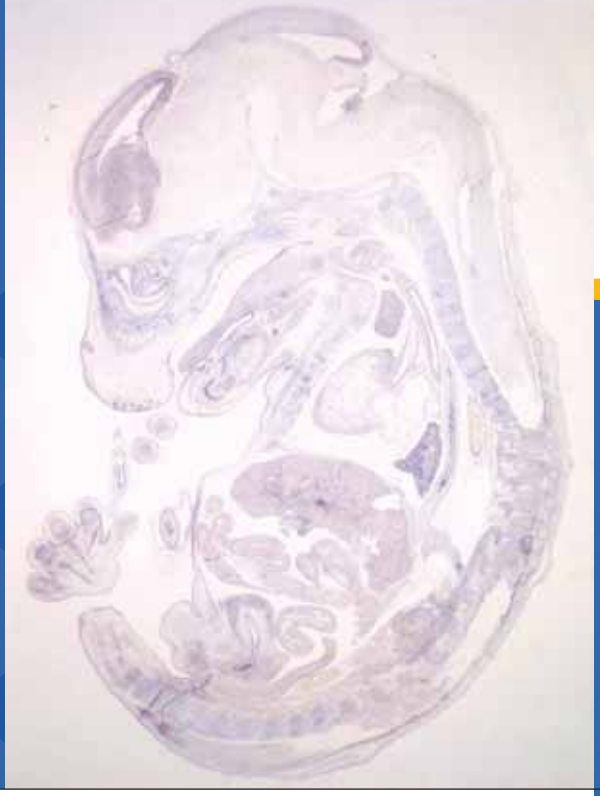

Mira Kyttälä

Identification of the Meckel Syndrome Gene *(MKS1)* Exposes a Novel Ciliopathy

Publications of the National Public Health Institute A 5/2006

Department of Molecular Medicine
National Public Health Institute, Helsinki, Finland
and
Department of Human Genetics
David Geffen School of Medicine at UCLA, Los Angeles, California, USA
and
Department of Medical Genetics
University of Helsinki, Finland

Mira Kyttälä

IDENTIFICATION OF THE MECKEL SYNDROME GENE (*MKS1*) EXPOSES A NOVEL CILIOPATHY.

ACADEMIC DISSERTATION

To be presented with the permission of the Medical Faculty, University of Helsinki, for public examination in the large lecture hall at the Haartman Institute, Haartmaninkatu 3, Helsinki, March 31st, at 12 noon

Department of Molecular Medicine, National Public Health Institute,
Helsinki, Finland
and
Department of Human Genetics, David Geffen School of Medicine at
UCLA, Los Angeles, California, USA.
and
Department of Medical Genetics, University of Helsinki, Finland

Helsinki 2006

Publications of the National Public Health Institute
KTL A5 / 2006

Julkaisija-Utgivare-Publisher

Kansanterveyslaitos (KTL)
Mannerheimintie 166
00300 Helsinki
Puh. vaihde (09) 474 41, telefax (09) 4744 8408

Folkhälsoinstitutet
Mannerheimvägen 166
00300 Helsingfors
Tel. växel (09) 474 41, telefax (09) 4744 8408

National Public Health Institute
Mannerheimintie 166
FIN-00300 Helsinki, Finland
Telephone +358 9 474 41, telefax +358 9 4744 8408

ISBN 951-740-610-X
ISSN 0359-3584
ISBN 951-740-611-8 (pdf)
ISSN 1458-6290 (pdf)

Kannen kuva - cover graphic: The *Mks1* gene expression in mouse embryo at embryonic day 15.5 (RNA *in situ* hybridization).

Painopaikka
Edita Prima Oy
Helsinki 2006

Supervised by
Docent Marjo Kestilä
National Public Health Institute
Department of Molecular Medicine
Helsinki, Finland

Academy professor Leena Peltonen-Palotie
National Public Health Institute
Department of Molecular Medicine and
University of Helsinki
Department of Medical Genetics
Helsinki, Finland

Reviewed by
Docent Kirsi Huoponen
Department of Medical Genetics
University of Turku
Turku, Finland

Docent Elisabeth Widen
The Finnish Genome Center
University of Helsinki
Helsinki, Finland

Opponent
Associate Professor Brendan Lee
Department of Molecular and Human Genetics
Baylor College of Medicine
Houston, TX, USA and
Department of Molecular and Human Genetics
Howard Hughes Medical Institute
Chevy Chase, MD, USA

To my family and friends.

…In the end everything makes sense.

Mira Kyttälä, Identification of the Meckel syndrome gene (*MKS1*) exposes a novel ciliopathy.
Publications of the National Public Health Insitute, A5/2006, 112 Pages
ISBN 951-740-610-X; 951-740-611-8 (pdf-version)
ISSN 0359-3584; 1458-6290 (pdf-version)
http://www.ktl.fi/ portal/4043

ABSTRACT

Meckel syndrome (MKS, MIM 249000) is an autosomal recessive developmental disorder causing death in utero or shortly after birth. The hallmarks of the disease are cystic kidney dysplasia and fibrotic changes of the liver, occipital encephalocele with or without hydrocephalus and polydactyly. Other anomalies frequently seen in the patients are incomplete development of the male genitalia, club feet and cleft lip or palate. The clinical picture has been well characterized in the literature while the molecular pathology underlying the disease has remained unclear until now.

In this study we identified the first MKS gene by utilizing the disease haplotypes in Finnish MKS families linked to the MKS1 locus on chromosome 17q23 (MKS1) locus. Subsequently, the genetic heterogeneity of MKS was established in the Finnish families. Mutations in at least four different genes can cause MKS. These genes have been mapped to the chromosomes 17q23 (MKS1), 11q13 (MKS2), 8q22 (MKS3) and 9q33 (MKS4). Two of these genes have been identified so far: The *MKS1* gene (this work) and the *MKS3* gene (Smith et al., 2006).

The identified *MKS1* gene was initially a novel human gene which is conserved among species. We found three different MKS mutations, one of them being the Finnish founder mutation. The information available from *MKS1* orthologs in other species convinced us that the *MKS1* gene is required for normal ciliogenesis. Defects of the cilial system in other human diseases and model organisms actually cause phenotypic features similar to those seen in MKS patients. The *MKS3 (TMEM67)* gene encodes a transmembrane protein and the gene maps to

the syntenic Wpk locus in the rat, which is a model with polycystic kidney disease, agenesis of the corpus callosum and hydrocephalus. The available information from these two genes suggest that *MKS1* would encode a structural component of the centriole required for normal ciliary functions, and *MKS3* would be a transmembrane component most likely required for normal ciliary sensory signaling.

The MKS4 locus was localized to chromosome 9q32-33 in this study by using an inbred Finnish family with two affected and two healthy children. This fourth locus contains *TRIM32* gene, which is associated to another well characterized human ciliopathy, Bardet Biedl syndrome (BBS). Future studies should identify the *MKS4* gene on chromosome 9q and confirm if there are more than two genes causing MKS Finnish families.

The research on critical signaling pathways in organogenesis have shown that both Wnt and Hedgehog pathways are dependent on functional cilia. The MKS gene products will serve as excellent model molecules for more detailed studies of the functional role of cilia in organogenesis.

Keywords: MKS1, MKS2, MKS3, MKS4, developmental disorder, disease mutation, ciliopathy.

Mira Kyttälä, Identification of the Meckel syndrome gene (*MKS1*) exposes a novel ciliopathy.
Kansanterveyslaitoksen julkaisuja, A5/2006, 112 sivua
ISBN 951-740-610-X; 951-740-611-8 (pdf-versio)
ISSN 0359-3584; 1458-6290 (pdf-versio)
http://www.ktl.fi/ portal/4043

TIIVISTELMÄ

Meckelin oireyhtymä (MKS, MIM 249000) on monioireinen sikiön vakava kehityshäiriö, joka johtaa kuolemaan sikiövaiheessa tai heti syntymän jälkeen. Taudin keskeisimmät oireet ovat munuaisten kystinen dysplasia, maksan fibroottiset muutokset ja keskushermostoputken sulkeutumishäiriö. Näiden oireiden kanssa tavataan aina postaksiaalista polydaktyliaa (pikkusormen/-varpaan puolella ylimääräinen sormi/varvas).

MKS:n on perinteisesti ajateltu kuuluvan nk. suomalaiseen tautiperintöön, joka koostuu 36 lähinnä väistyvästi periytyvästä taudista, jotka esiintyvät runsaimmin Suomessa kuin muualla maailmassa, jos lainkaan. MKS:n esiintyvyys on Suomessa noin yksi 9 000:sta vastasyntyneestä. Epätyypillistä on että MKS:ää on raportoitu esiintyvän jopa korkeammalla esiintyvyydellä muissa geneettisesti suhteellisen homogeenisissä väestöissä. *MKS1* geeni sijaitsee kromosomin 17 pitkässä käsivarressa q23. Useiden alueellisten kandidaattigeenien poissuljennan jälkeen ja uusien perheiden haplotyyppikartoituksen avulla pystyimme lopulta pitävästi rajaamaan kriittisen alueen, jolle sijoittuu vain viisi transkriptiä. Näiden sekvensointi potilaiden DNA:sta osoitti patogeenisen muutoksen ainoastaan tuntemattomassa transkriptissä *FLJ20345* (*MKS1*), Noin 70% suomalaisista MKS perheistä edustaa kromosomin 17q23 (MKS1) valtamutaatiota (FIN_{major}), ja potilaat näissä perheissä ovat homotsygootteja tämän mutaation suhteen. Niistä perheistä (30%), joissa potilaat eivät ole homotsygootteja valtamutaation suhteen ei ole löytynyt mutaatioita potilaiden *MKS1* geeniä koodittavilta alueilta. FIN_{major}-mutaatio löydettiin homotsygoottisena sairailta myös kolmesta muusta perheestä, jotka

edustavat ei-suomalaista valkoihoista eurooppalaista väestöä. Tämä osoittaisi, että suomalainen valtamutaatio on vanha muutos eurooppalaisessa geeniperimässä.

Tämä mutaatio on introninen 29 emäksen deleetio, josta aiheuttuu virheellinen geenin luenta ja tämän seurauksena potilaan transkriptistä puuttuu kokonaan yksi eksoni. Tästä seuraa lukukehyksen siirtymä, joka johtaa virheelliseen proteiinituotteeseen alkaen aminohaposta 470 polypeptidissä. Lisäksi saksalaisesta MKS perheestä löytyi kaksi muuta *MKS1* geenin mutaatiota, jotka ovat viiden emäksen insertio eksonissa yksi ja emästransitio intronin yksi silmukoinnin donorikohdassa (engl. splice donor site).

Geneettinen heterogeenisyys on tyypillistä MKS:lle, ja tautigeenejä onkin kartoitettu jo neljään eri kohtaan genomissa: kromosomiin 17q23 (MKS1), 11q13 (MKS2), 8q24 (MKS3) ja 9q33 (MKS4). Neljäs MKS lokus on kartoitettu toimestamme suomalaisessa perheessä, jossa on kaksi sairasta ja kaksi tervettä lasta. Lähtökohtaisesti genealogian pohjalta tiesimme, että perheen vanhemmilla on yhteisiä esivanhempia arviolta 1700-luvulla. Kun perheessä taudin kytkeytyminen tunnettuihin MKS lokuksiin oli poissuljettu, etenimme koko genomin laajuiseen kartoitukseen, missä käyttämämme keskimääräinen geenimerkkietäisyys oli n. 10 cM. Tutkimme näitä tuloksia tarkemmin genotyyppittämällä 10 K SNP geenisirun, jossa keskimääräinen etäisyys SNP-geenimerkkien välillä on n. 200 kb. Ainoana positiivisena löydöksenä näistä koko genomin laajuisista kartoituksista havaitsimme kromosomissa 9q32-33.1 monipiste LOD arvon 3.6, joka nähtiin yli 5 cM kokoisella alueella. Erityisen mielenkiintoiseksi tämän löydöksen tekee seikka, että samalle kromosomaaliselle alueelle on kartoitettu toinen ihmisen tauti, Bardet Biedl-oireyhtymä (BBS, MIM 209900), joka on tunnettu perinnöllinen tauti. BBS:n tunnetut geenivirheet johtavat värekarvarakenteiden (cilia) toiminnallisiin vikoihin.

MKS1 geenin on osoitettu lajien välisen vertailevan genetiikan avulla kuuluvan ryhmään geenejä, jotka ovat keskeisiä kudosten värekarvarakenteille (engl. cilia), niiden ylläpidolle ja toiminnalle (Flagellar and basal body proteome, FABB proteome). Värekarvat eli ciliat ovat solu-organelleja, jotka liittyvät mm. solun

liikkumiseen, solusykliin ja sensoriseen solusignalointiin, ja niillä on keskeinen rooli solun tasopolarisuuden muodostumisessa. Lisäksi muut taudit sekä eläinmallit, joissa tunnetaan cilian rakenteellisia tai toiminnan kannalta keskeisiä geenivirheitä jakavat paljon yhteisiä piirteitä MKS:n kanssa. Myös alustavat tutkimuksemme MKS1:n solulokalisaatiosta tukevat MKS:n ciliopatia luonnetta. MKS geenituotteet tulevat mitä todennäköisimmin toimimaan tärkeinä mallimolekyyleinä näiden suhteellisen vähän karakterisoitujen soluorganellien toiminnassa.

Avansanat: MKS, lokusheterogeenisyys, suomalainen tautiperintö, sikiön kehityshäiriö, kytkentäepätasapaino, tautimutaatio.

TABLE OF CONTENTS

ABBREVIATIONS

aa	amino acid
BAC	bacterial artificial chromosome
BBS	Bardet Biedl syndrome
bp	basepair
CEPH	Centre d'Etude du Polymorphisme Humain
cM	centiMorgan
cDNA	complementary DNA
chr	chromosome
CNS	central nervous system
DNA	deoxyribonucleic acid
EMSA	electromobility shift assay
EST	expressed sequence tag
FABB	flagellar apparatus basal body
FISH	fluorescence in situ hybridization
FDH	Finnish disease heritage
HGP	Human genome project
kb	kilobase
LD	linkage disequilibrium
LOD	logarithm of odds
Mb	megabase
MIM	mendelian inheritance in man
MKS	Meckel syndrome
mRNA	messenger RNA
ORF	open reading frame
PAC	P1 derived artificial chromosome

PAGE	polyacrylamide gel electrophoresis
PCR	polymerase chain reaction
PFGE	pulsed field gel electrophoresis
PKD	polycystic kidney disorder
RFLP	restriction fragment length polymorphism
RNA	ribonucleic acid
RT	reverse transcriptase
SNP	single nucleotide polymorphism
STR	short tandem repeat
STS	sequence tagged site
Θ	theta, recombination fraction
UTR	untranslated region
YAC	yeast artificial chromosome

In addition, the standard one letter abbreviations of nucleotides and amino acids are used.

LIST OF ORIGINAL PUBLICATIONS

This thesis is based on the following articles, referred to in the text by roman numerals.

I High-resolution physical and genetic mapping of the critical region for Meckel syndrome and Mulibrey Nanism on chromosome 17q22-q23.
Paavola P, Avela K, Horelli-Kuitunen N, Bärlund M, Kallioniemi A, Idänheimo N, Kyttälä M, de la Chapelle A, Palotie A, Lehesjoki AE, Peltonen L.
Genome Res 9:267-76 (1999)

II Comparative physical maps of the human and mouse Meckel syndrome critical regions.
**Hentges KE, *Kyttälä M, Justice MJ, Peltonen L.*
Mamm Genome 15:252-64 (2004)
* Authors equally contributed to the work.

III MKS1, encoding a component of the flagellar apparatus basal body proteome, is mutated in Meckel syndrome
Mira Kyttälä, Jonna Tallila, Riitta Salonen, Outi Kopra, Nicolai Kohlschmidt, Paulina Paavola-Sakki, Leena Peltonen, Marjo Kestilä.
Nat Genet 38(2):155-7 (2006)

IV Genetic analysis of an inbred pedigree exposes the fourth locus for Meckel syndrome, MKS
Mira Kyttälä, Riitta Salonen, Riitta Karikoski, Juha Saharinen, Teppo Varilo, Marjo Kestilä, Leena Peltonen
Submitted.

Publication I has appeared in the thesis by Paulina Paavola (1998) and Kristiina Avela (2002). Some unpublished data are also presented.

INTRODUCTION

Meckel syndrome (MKS) is one of the best characterized human malformation syndromes. The clinical features of MKS were described for the first time in the literature over hundred years ago, but the molecular and cellular pathogenesis of this autosomal recessive developmental defect has remained obscure until now. The aim of this thesis work was to identify the gene underlying the disease by utilizing the founder effect of an isolated population. Our study was greatly aided by the advances of the Human genome project and we successfully identified a novel human gene mutated in MKS. The initial analysis of this gene (*MKS1*) and its gene product has already provided significant new knowledge about the pathogenesis of MKS at the molecular level. In the future the research on the *MKS1* gene will provide more detailed information about the molecular mechanisms critical in normal embryonic development. This should also provide essential information in understanding some more common conditions with related defects.

REVIEW OF THE LITERATURE

1. MECKEL SYNDROME

1.1 History

Meckel syndrome is one of the best characterized malformation syndromes affecting normal embryonic development. The clinical picture of MKS has been described in the literature for the first time already in 1684 by Christopher Krahe describing a monstrous child born in Denmark who possibly suffered from MKS (Kompanje, 2003). In 1822 J. F. Meckel published a detailed pathological report of two siblings (male and female), who died with identical lethal malformations. Meckel´s clinical descriptions still agree with the modern diagnostic criteria. Over hundred years later Gruber (1934) suggested the genetic origin of MKS because the cases were familial. In the 1960´s Opitz and Howe introduced the name Meckel syndrome and delineated the clinical and pathological features based on one observation and presented a thorough review of the literature (Opitz and Howe, 1969). In Finland the careful clinical, epidemiological and pathological studies were performed by Salonen and Norio which expedited the research on the molecular genetics of MKS in Finland (Salonen 1984; Salonen and Norio 1984).

1.2 Meckel syndrome world wide

MKS is reported world wide (Hsia et al., 1971; Mecke and Passarge, 1971; Moerman 1982). There are several reports of the incidences available in the literature, however the true extended reports of the prevelences in distinct populations are rare. The incidence in Northern America and Western Europe varies from the estimates of

1: 3 400- 1: 140 000 newborn (Holmes, 1976; Seller 1978). The lowest incidence is reported from Great Britain and respectively the highest from Belgium. In Finland Salonen and Norio (1984) studied a comprehensive cohort of cases and families and estimated the prevalence in Finland to be 1: 9 000 newborn. This means that there are approximately 5-7 new cases in a year. Several reports are available from

distinct populations throughout the world: Indonesia, Pakistan and Japan (Tan and Thomas, 1970; Crawfurd et al., 1978; Sugiura et al., 1996) as well as in Arabs in Saudi Arabia, Jews and Palestinian Arabs in Israel (Haque and Zaidi 1981; Fried, 1973; Zlotogora, 1997). Surprisingly high incidences have been reported in Gujarati Indians, 1: 1 100; Kuwaitian Bedouins, 1: 3 530 and in Belgians, 1: 3 400 (Young et al., 1985; Teebi et al., 1992; Moerman et al., 1982).

1.3 Clinical picture

1.3.1 Main features

In the most typical cases MKS results in a death in couple hours after birth. In the 67 MKS cases reported by Salonen (1984; 1986) the detected rate of spontaneous abortions (15.2%) is within the normal range. The survival time of 40 liveborn infants was under 2.5 hours. The cesarian section was performed in ten cases because of hydrocephaly, abnormal position or asphyctic signs of the fetus. Oligohydramnious was detected at birth in 30 cases including 15 cases without amniotic fluid. MKS can be diagnosed reliably by the ultrasound examination of the embryo. In most cases the pregnancies will be terminated at 10-14 weeks of gestation if MKS is diagnosed.

The main characteristic features are the CNS (central nervous system) defect that always appears together with cystic dysplasia of the kidneys, fibrotic changes of the liver and polydactyly (Fig. 1).

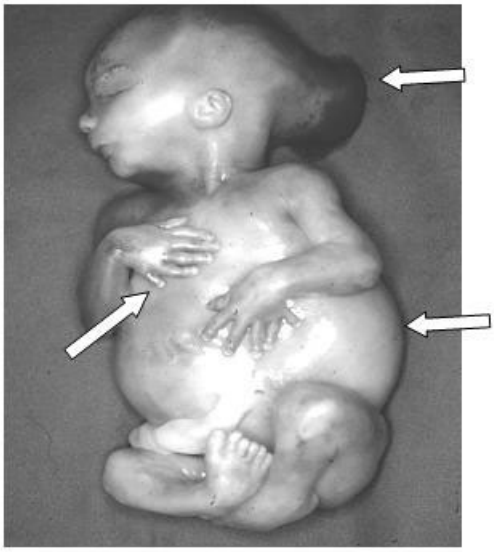

Figure 1. A fetus with MKS having the typical phenotype at 21 weeks of gestation. The arrows are pointing the main hallmarks: Occipital encephalocele, polydactyly and the massive abdomen caused by the distended cystic kidneys. The picture is a courtesy from Dr. Riitta Salonen.

1.3.2 Central nervous system, (CNS)

The occipital meningo encephalocele is seen in 90% of the cases (Aleksic et al., 1984; Paetau et al., 1985; Ahdad-Barmada and Claasen, 1990). In 10-20% of the patients it is accompanied with hydrocephaly. However, the severity of the CNS malformation can vary from a mild partial defect of corpus callosum to a severe total craniorachischisis (open spinal canal) and anencephaly.

The dysgenesis of the forebrain with the absence of the olfactory tracts and bulbs is frequent (Miller and Selden, 1967; Rehder and Labbe, 1981; Anderson, 1982; Paetau et al., 1985; Ahbad-Barmada and Claasen, 1990). A microscopically interesting feature is the reported neuroepithelial rosettes, or as the author himself described the phenomenon: “peculiar ectopic neuroepithelial rosettes” (Paetau et al., 1985). These can be interpreted as ependymal rosettes, which are mainly detected in

the germinal matrix region but they can also be found far from the midline of the brain. The neuroepithelial rosettes could reflect a defect in the neuronal proliferation and migration. Similar findings are also observed in the eye: retinal dysplasia with rosettes of pigmented retinal epithelium (Miller and Selden, 1967; Anderson, 1982).

1.3.3 Kidneys

One of the most consistent malformation in MKS is the enlarged kidneys (with 5-50 fold increase in weight), that cause the bulky abdomen which is typical for an embryo having MKS. In most cases the enlargement involves both kidneys. The macroscopical cysts vary in size from several millimetres to 200 mm (Fig. 2a). Histologically the renal parenchyma is characterized by smaller cysts in the peripheral cortex with a thin zone of normal glomeruli, and thin walled, larger cysts in the medullary part separated by loose connective tissue (Fig. 2b and 2c) (Moerman et al., 1982; Rehder and Labbe, 1981; Anderson, 1982; Rapola and Salonen, 1985; Blankenberg, 1987).

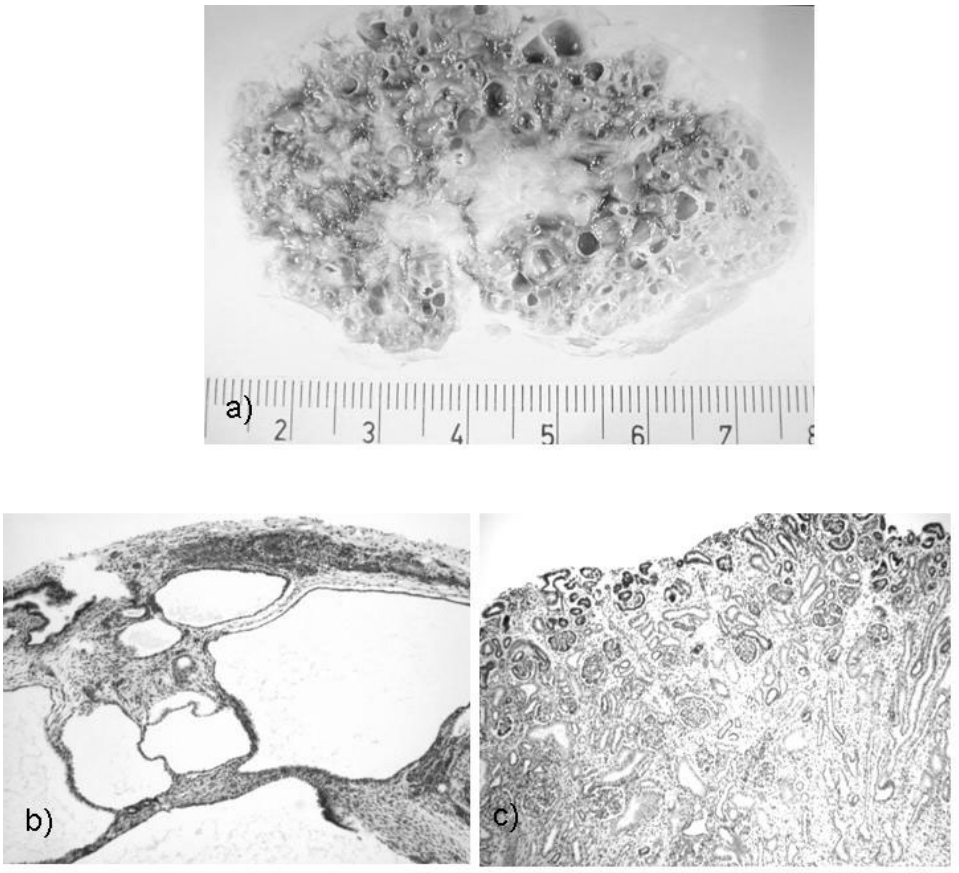

Figure 2 a) Photograph of MKS kidney. The picture is a courtesy from Dr. Riitta Salonen. b) Histology of MKS kidney at 20+4 weeks of gestation. c) Histology of normal control kidney at 13 weeks of gestation.

1.3.4 Liver

Macroscopically the liver often seems normal, but can be somewhat enlarged. Occasionally macroscopic fibrosis and cysts of liver occur. The microscopic findings are detected in a histological examination exposing in all cases liver bile duct anomalies and fibrotic changes ("collagenous tissue", Salonen, 1986),

especially in the portal areas (Fig. 3) (Moerman et al., 1982; Rehder and Labbe, 1981; Anderson, 1982; Rapola and Salonen, 1985; Blankenberg, 1987).

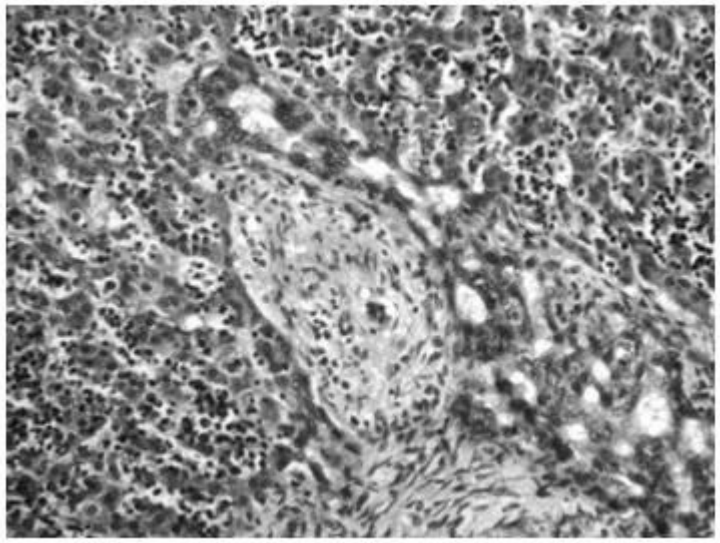

Figure 3. Histology of MKS liver at 20+4 weeks of gestation showing the typical ductal plate malformation.

1.3.5 Situs inversus in MKS

In the initial publication by Salonen (1984) three out of 67 MKS cases are reported with a finding of situs inversus totalis, demonstrating an elevated incidence of this rare condition in the MKS patients. Since MKS is screened efficiently by prenatal ultrasound examination, and the detected MKS pregnancies are normally terminated during the gestational weeks 10-14, it is possible that in the small embryo the abnormal organ situs has not always been detected if it is not an expected finding.

Situs inversus is known to occur with increased frequency in diseases that affect the structure or the function of cilia. Good examples are human diseases like Kartagener syndrome (MIM 244400) and primary ciliary dyskinesia (PCD, MIM 242650). The embryonic nodal cilia have been studied in detailed in mice, and they are involved in the left right situs determination and ciliary dysfunction leads to randomized situs formation (Supp et al., 1997; Okada et al., 1999).

1.3.6 Other frequent symptoms

Epithelial branching morphogenesis is typical for most of the tissues affected in MKS and the pathological findings are considerably similar in different tissue types (Blankenberg et al., 1987). Lungs are frequently reported to be small and hypoplastic in patients. Rapola and Salonen (1985) reported that the weight of the lungs varied 20-80% from the expected in 14 recorded cases and only in one case the weight was within normal limits. Club feet and the ambiguous male genitalia are also typical features of MKS. The other variably associated malformations are: heart defect, cleft lip and/or palate, anomalies and tumours of the tongue, ocular anomalies and polysplenia. (Fried et al., 1971; Mecke and Passarge, 1971; Fraser and Lytwyn, 1981; Rehder and Labbe, 1981; Anderson, 1982; Moerman et al., 1982, Salonen 1984; Ahbad-Barmada and Claasen, 1990).

1.4 Genetics of MKS

The autosomal recessive inheritance and locus heterogeneity have been established in MKS (Hsia et al., 1971; Mecke and Passarge, 1971; Salonen, 1984; Roume et al., 1997; Morgan et al., 2002). Three genetic loci are known to be involved. These loci have been assigned in families representing relatively isolated populations. The three identified MKS loci are located on chromosomes 17q23 (MKS1, Paavola et al., 1995), 11q13 (MKS2, Roume et al., 1998) and 8q22 (MKS3, Morgan et al., 2002). The *MKS3* gene has recently been identified (Smith et al., 2006).

1.4.1 Chromosome 17q, MKS1

The original linkage study assigned the first MKS locus to chromosome 17q21-q24 in a genome scan of five Finnish families with two or more MKS affected siblings (Paavola et al., 1995). This produced a maximum pair wise LOD score of 5.42 for microsatellite marker D17S1606. The initial linkage was observed over a 13 cM genetic region (Paavola et al., 1995).

1.4.2 Chromosome 11q, MKS2

The MKS2 locus was assigned to chromosome 11q13 in Middle Eastern and Northern African MKS families (Roume et al., 1998). Linkage to MKS1 locus had been previously excluded in these families (Roume et al., 1997). A genome wide scan in a total of seven consanguineous families (4 Tunisian, 1 Algerian, 1 Senegalian and 1 Pakistani) resulted in the highest pair wise LOD score of 4.4, which was obtained at two marker loci: D11S906 and D11S911. The affected fetuses of Tunisian ancestry had the same haplotype at marker loci D11S911 and D11S906, suggesting a founder effect in Tunisian MKS families.

1.4.3 Chromosome 8q and MKS3 gene mutated in MKS

The MKS3 locus was mapped in consanguineous English families originating from Pakistan and India. The original genome scan assigned the MKS3 locus to chromosome 8q24 (Morgan et al., 2002). The *MKS3* (*Meckelin*) gene was recently identified and is mutated in MKS3 (Smith et al., 2006). The *MKS3* gene is located on 8q22, which is somewhat more proximal than the region on 8q24, which was originally defined by linkage (Morgan et al., 2002).

The TMEM67/MKS3 gene is mutated in MKS3

The *TMEM67/MKS3* gene encodes a transmembrane receptor protein, which mutations in the rat ortholous gene (*Mks3*) causes polycystic kidney disease, agenesis of the corpus callosum and hydrocephalus. (Smith et al., 2006). The role of the *MKS3* gene in ciliary functions has been hypothesized based on its topological similarity to the *Drosophila melanogaster* FZ (Frizzled) family of receptors (Cadigan and Nusse, 1997; Xu et al., 1998). The Frizzled signaling in fruit fly has an important role in epithelial planar cell polarity (PCP) (Jenny et al., 2005). Disrupted PCP signaling has been implicated in Bardet Biedl syndrome (BBS) (Ross et al., 2005) a disorder associated with primary ciliary and basal body dysfunction (Ansley et al., 2003; Beales 2005). Further evidence of the ciliary expression comes from the *Caenorhabditis elegans* homolog of *MKS3/TMEM67* (*F35D2.4*) which has a previously described X-box motif approximately 75 bp upstream of the probable start codon, a characteristic of proteins found in cilia (Efimenko et al., 2005).

2. THE FINNISH DISEASE HERITAGE

2.1 Concept

MKS belongs to the so called Finnish disease heritage (FDH). The concept of FDH was introduced over 30 years ago in the original publication by Norio and co-workers in 1973: “Hereditary diseases in Finland; A rare flora in a rare soil”. It described 10 genetic disorders more common in Finland than elsewhere in the world. Nowadays total of 36 disorders are included to the FDH (table 1). In 33 of them the disease causing mutation has been identified in a relatively short time period.

The genetic structure of Finnish the population is probably the most important factor that led to the rapid identification of the mutations in the FDH (de la Chapelle and Wright, 1998; Peltonen et al., 1999; Norio, 2003; www.findis.org). In addition, the Finnish healthcare is of high quality and available for all citizens, the common attitude towards research is positive, and families are willing to participate in genetic studies. Furthermore, the church has kept comprehensive records of the population

over centuries, which makes it possible to trace back families in some cases up to the 16th century with the help of taxation records.

The identification of the disease mutations underlying many of the diseases in the FDH has revealed previously unknown biological functions of several human genes. The monogenic human disorders can provide detailed information about novel cellular mechanisms affecting both health and disease. This information can be exploited in medicine in understanding related conditions, some being common in complex traits.

Table 1. The 36 diseases belonging to the Finnish disease heritage (FDH). In 33 of them the disease mutation has been identified. In most of the diseases the patients are homozygous for the same major (founder) mutation. The diseases are presented below in the chronological order of the identification of the gene defect.

Gene defect is identified, disease, MIM number	Defective protein product	Reference for the Finnish mutation
HYPERORNITHINEMIA WITH GYRATE ATROPHY OF CHOROID AND RETINA, (HOGA), 258870	Ornithine ketoacid aminotransferase, (OAT)	Mitchell et al., 1989
MERETOJA TYPE AMYLOIDOSIS, (FAF), 105120, dominant	Gelsolin (GSN)	Levy et al., 1990; Maury et al., 1990
ASPARTYLGLUCOSAMINURIA (AGU), 208400	Aspartyliglucosaminidase, (AGA)	Ikonen et al., 1991
CHOROIDEREMIA, (CHM), 303100, X linked, Recessive	Rab-protein, (REP1)	Sankila et al., 1992
HYPERGLYCINEMIA, NONKETOTIC, (NKH), 605899	Glycine decarboxylase, (GLDC)	Kure et al., 1992
DIASTROPHIC DYSPLASIA, (DTD) 222600	Diastrofic dysplasia sulfate carrier, (DTDST)	Hästbacka et al., 1994
CEROID LIPOFUSCINOSIS, NEURONAL 1, INFANTILE, 256730	palmitoyl-protein thioesterase 1, (PPT)	Vesa et al., 1995
NEURONAL CEROID LIPOFUSCINOSIS, JUVENILE TYPE, (JINCL), 204200	CLN3-protein	The international Batten disease consortium
OVARIAN DYSGENESIS, HYPERGONADOTROPIC, (ODG1), 233300	FSH-receptor, (FSHR)	Aittomäki et al., 1995
CHLORIDE DIARRHEA, FAMILIAL, (CLD), 214700	DRA-protein	Höglund et al., 1996

PROGRESSIVE MYOCLONIC EPILEPSY, (PME1), 254800	Cystatin B, (CSTB)	Pennacchio et al., 1996, Virtaneva et al., 1997
AUTOIMMUNE POLYENDOCRINOPATHY SYNDROME, (APECED), 240300	Autoimmune regulator AIRE-protein	Nagamine et al., 1997; The Finnish German APECED consortium,1997
FINNISH CONGENITAL NEPHROSIS, (CNF), 256300	Nephrin	Kestilä et al., 1998
NEURONAL CEROID LIPOFUSCINOSIS, FINNISH VARIANT, LATE INFANTILE, (vLINCL),256731	CLN5-protein	Savukoski et al., 1998
RETINOSCHISIS, 312700 (X linked recessive)	XLRS1-protein	The retinoschisis consortium, 1998
LYSINURIC PROTEIN INTOLERANCE, (LPI), 222700	SLC7A7-protein	Torrents et al., 1999; Borsani et al., 1999
MEGALOBLASTIC ANEMIA 1, (SMB12), 261100	Cubilin protein (CUBN)	Aminoff et al., 1999
NORTHERN EPILEPSY, (EPMR), 600143	CLN8-protein	Ranta et al., 1999
SIALURIA, FINNISH TYPE, 604369	SLC17A5-protein	Verheijen et al., 1999
CORNEA PLANA 2, (CNA2), 217300	KERA-protein	Pellegata et al., 2000
MULIBREY NANISM, 253250	TRIM37-protein	Avela et al., 2000
NASU-HAKOLA DISEASE, (PLOSL), 221770	TYRO protein tyrosine kinase-binding protein (TYROBP)	Paloneva et al., 2000
CARTILAGE-HAIR HYPOPLASIA, (CHH), 250250	RNA-processing endoribonuclease (RNRP)	Ridanpää et al., 2001
USHER SYNDROME, TYPE III, 276902	USH3A-protein	Joensuu et al., 2001
TIBIAL MUSCULAR DYSTROPHY (TMD), 600334, dominant	Titin (TTN)	Hackman et al., 2002
COHEN SYNDROME, 216550	COH1-protein	Kolehmainen et al., 2003
GRACILE SYNDROME, 603358	BCS1L-protein	Visapää et al., 2003
RAPADILINO SYNDROME, 266280	REQL4-helicase	Siitonen et al., 2003
MUSCLE-EYE-BRAIN DISEASE, (MEB), 253280	POMGnT1	Diesen et al., 2004
HYDROLETHALUS SYNDROME, 236680	HYLS1-protein	Mee et al., 2005
INFANTILE-ONSET SPINOCEREBELLAR ATAXIA; IOSCA, 271245	Twinkle	Nikali et al., 2005
LACTASE DEFICIENCY, CONGENITAL, (CLD), 223000	Lactase	Kuokkanen et al., 2005

MECKEL SYNDROME, (MKS1), 249000	**MKS1 protein**	**Kyttälä et al., 2006 (original publication III)**
The chromosomal locus of the gene has been assigned, disease, MIM number	Chromosomal locus	
LETHAL CONGENITAL CONTRACTURE SYNDROME 1, (LCCS), 253310	9q	Mäkelä-Bengs et al., 1998
The chromosomal location of the defect gene is not known, disease, MIM number		Reference
PEHO SYNDROME, 260565		Salonen et al., 1991
LETHAL ARTHROGRYPOSIS WITH ANTERIOR HORN CELL DISEASE, (LAAHD)		Vuopala et al., 1995

2.2 Inhabitation of Finland and founder effect

One theory of the inhabitation of Finland is called the dual theory and it is supported by studies of the Y chromosomal haplotypes (Kittles et al., 1998; Lahermo et al, 1999). According to this theory there have been two important groups of settlers (Fig. 4).

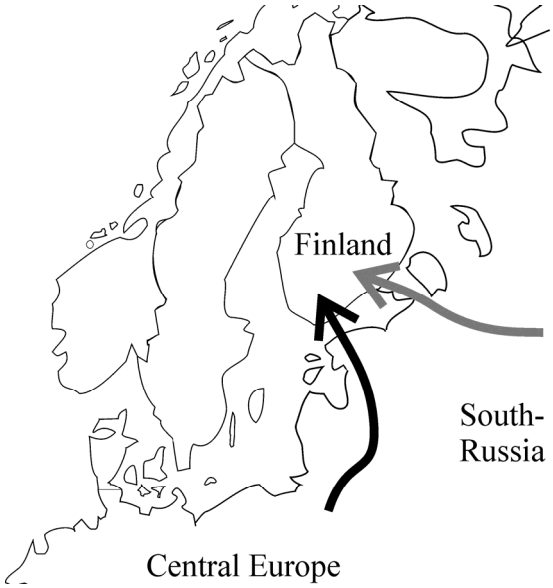

Figure 4. The two important migrations to Finland according to the dual theory: The first settlers came 4000 years ago from South Russia, and the second group 2000 years ago from central Europe (Peltonen et al., 1999).

The assumption is that the first inhabitants notably affecting to the gene pool came to Finland already 4000 years ago from the east (Norio 1981, Eriksson 1973, Peltonen et al. 1999). The second migration some 2000 years ago has still at the present day the most impact on the Finnish gene pool (Nevanlinna 1972, de la Chapelle 1993, Peltonen et al. 1999). The mitochondrial DNA (mDNA) and nuclear DNA variations confirm the European origin of the majority of the gene variations in Finland (Cavalli-Sforza and Piazza 1993, Sajantila and Pääbö 1995, Lahermo et al., 1996, Sajantila et al., 1996, Laan and Pääbö 1997, Torroni et al., 1998). Since there is no explicit evidence of two major immigration waves, it is most probable that the immigration has occurred more gradually in small groups (Peltonen at al., 1999).

The original founder population was small and it has been estimated that the population size was 50 000 people in the 12th century, but it expanded and spread to the country following the coastal lines while the inland remained virtually unpopulated. In the beginning of the 16th century the population size of Finland was approximated to be 250 000 concentrating mainly on the coastal parts of the country. The Swedish king Gustavus of Wasa/Vasa promoted the inhabitation of the central and eastern parts of Finland in 16th century by providing taxation relieves for people moving further to the inland. This movement established the late settlement region in Finland (Fig. 5).

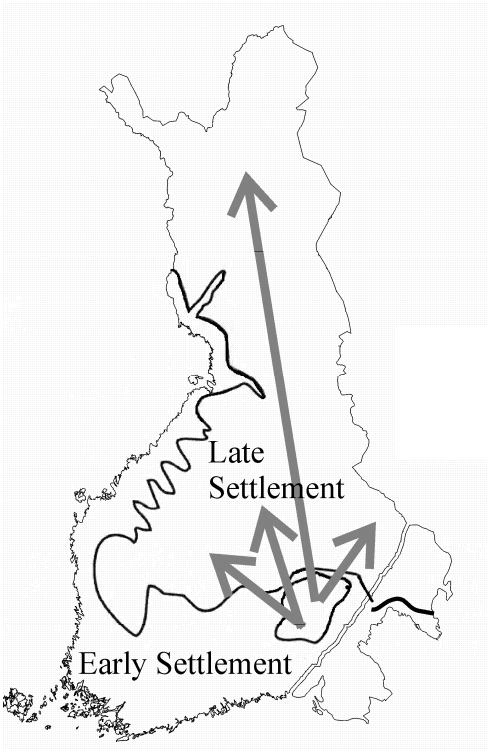

Figure 5.The inhabitation of Finland. The early settlement region follows the coastal line. The late settlement region was formed by the internal migration in the 16th century (Peltonen et al., 1999).

The villages in the late settlement region originating from a small area of Southeastern Finland (South Savo) remained stable for centuries because of the relatively long geographical distances and low population density enabling the formation of genetic subisolates. This as well as wars, epidemics and famines created significant population bottlenecks. The Finnish population has expanded to 5.3 million in just about three centuries from approximately 250 000 inhabitants. The rare disease alleles were enriched in the subisolates and to the whole population. The unique gene pool was moulded by different bottleneck effects in the small founder populations over centuries. Despite the industrialization and the population migration from rural areas to urban regions, the subpopulations particularly in the

late settlement areas have remained unmixed (de la Chapelle 1993; Peltonen et al., 1999; Varilo 1999; Norio 2003).

Many disorders, which are common elsewhere in the world, such as cystic fibrosis (incidence 1: 2 500 in Western Europe) are uncommon in Finland. On the other hand, there are many diseases in FDH which have been found only in Finland, like the variant form of late infantile neuronal ceroid lipofuscinosis (vLINCL) and Northern epilepsy (EPMR). Genetic drift also influences other more common monogenic diseases like familial hypercholesterolemia (FH), in which only two founder mutations are responsible for most of the cases among Finns (Aalto-Setälä et al., 1989; Koivisto et al., 1992). For many of the diseases belonging to FDH the birthplaces of the patients´ grandparents reflect the historical population migration (Varilo 1999; Norio 2000). The distribution of the grandparental birthplaces can be utilized in categorizing the diseases by the estimated age of the founder mutation.

In approximately half of the diseases belonging to the FDH the grandparents´ birthplaces cluster to the regions of the late settlement (Norio, 2000), suggesting that the disease mutations have been spread during the inhabitation of central and eastern parts of the country in the 16th century. These mutations were introduced to the population some 30-40 generations ago (Varilo 1999).

Another group of diseases is composed of those which are the most common in FDH: AGU, INCL, PME, JNCL and CNF. The prevalence of these disorders has been estimated to be 1: 8 000- 1: 19 000 and grandparental birthplaces are distributed evenly through out the country; however some denser regions can be detected on the late settlement regions. This indicates that the mutations have been introduced to the population earlier, approximately 80-120 generations ago (Peltonen 1995, Varilo 1999). vLINCL and EPMR are the youngest diseases in FDH, and the ancestral birthplaces can be seen only in a small, very restricted region on the map (Varilo et al. 1996, Norio 2000). In vLINCL families the grandparental birthplaces are geographically densely located in the Southern Ostrobotnia region on the West coast and the significant LD can be detected across 11 cM genetic region

thus supporting the young age of the mutation in Finland. EPMR can also be geographically positioned to the Nortwestern part of the country to a small region based on the grandparental birthplaces. The length of significant LD in the EPMR disease chromosomes is 10 cM is congruent with the hypothesis of a recent mutation (approximately 20-30 generations) (Varilo, 1999).

3. DISEASE GENE IDENTIFICATION

3.1 Main strategies

There are several strategies to identify human disease genes. Probably the most important of them are the functional cloning, candidate gene analysis and positional cloning. In most cases it is beneficial to use a combination of the strategies.

The functional cloning is the oldest approach to clone a disease gene. It was the only possible method to identify disease genes before there were genetic maps available from the human genome. In this strategy the defective protein product is first isolated and its peptide sequence will indicate the corresponding gene. The identification of Haemophilia A gene is an example wher this strategy was used (Gitschier et al., 1984).

The candidate gene analysis requires knowledge about the molecular level pathogenesis of the disease. This information may be available from functional studies in humans or animal models. Mutations can be screened directly from the candidate gene in the patients. Many times this strategy is not successful alone without positional cloning (positional candidate gene approach).

The human genome maps have enabled *the positional cloning* approach. This strategy depends fully on suitable family material. This approach is most successful when families represent isolated populations with one founder mutation. In this strategy the defective gene is identified by first assigning its chromosomal location with classic linkage mapping. No functional information about the gene is required. This method is sensitive to misdiagnosis, also locus heterogeneity in the family material can substantially hamper the mapping effort. The linked

chromosomal region can be further fine mapped by utilizing linkage disequilibrium (LD) in the disease chromosomes. By utilizing the LD and haplotype mapping the critical chromosomal region can be greatly reduced from that originally detected by linkage. Linkage is typically seen in a vast chromosomal region spanning across several cM region. A combination of positional cloning and candidate gene analysis (positional candidate gene analysis) is at the present time probably the most rational approach for the disease gene identification with the growing knowledge of the function of the human genome. After assignment of the gene locus the regional candidates can be evaluated based on the information of their known functions and their tissue expression profiles.

3.2 Human genome project

The sequence of the 3 billion base pairs human genome has been available since 2003, when the International Human Genome Sequencing Consortium completed this gigantic sequencing project (Collins et al., 2003).

The development of the research on the human genome from the structure of DNA to the finished genomic sequence has proceeded enormously just in fifty years (Watson and Crick, 1953; Collins et al., 2003). At the very early stage leading to the human genome sequencing project a YAC contig map was available which covered approximately 75% of the genome (Chumakov et al., 1992). This map was soon complemented with STS based physical maps (Hudson et al., 1995), which were achieved in part by mapping whole genome radiation panels. The human genome mapping has evolved in only two decades from the genetic maps based on RFLPs (restricted fragment length polymorphism) to the first high resolution maps based on microsatellite markers (Weissenbach et al., 1992), and finally to the finished human sequence.

Since then, the HGP has completed sequencing of a large amount of model organisms (http://www.ncbi.nlm.nih.gov/genomes/static/gpstat.html). Results from the various genome sequencing projects have carried the biomedical research into

the genomic era that utilize the gene conservation among species and develop the ultimate tools for gene mapping in form of high-throughput SNP based techniques.

3.2.1 Comparative genomics

Comparative genomics involves analysis of two or more genomes to identify the extent of similarity of various features, or large scale screening of the genomes to identify sequences present in another genome.

Comparative genomics can be utilized to study evolutionary relationships of different species. This can be further employed to recognize genes behind in specific cellular machineries or functions, which are known to be present/absent in the species that are compared. The comparisons of different complex genomes also enable also identification of gene coding regions as well as potential regulatory regions. An analysis of the of conserved regulatory regions showed that for most of them (81%) there is more than 50% sequence similarity between human and mouse, which is significantly higher than seen in average non-coding intronic sequence (Liu et al., 2004).

3.3 Linkage analysis

Polymorphisms are normal sequence variations in the genome, and can be used as genetic markers in linking a disease to a certain chromosomal region. The closer the marker is to the disease gene the less likely a meiotic recombination has occurred between the two (disease and polymorphism) and hence they are more likely to be inherited together. The distinction between a polymorphism and a rare variant is that the latter one occurs in a frequency less than 1% in the population.

The first generation polymorphic genetic markers were RFLPs. Later the PCR technology made the microsatellite markers as a standard tools for linkage analysis. These are most often di -, tri - or tetra nucleotide repeats. The advantage of the multiallelic microsatellite markers is that they are normally more informative in genetic analyses than the SNP markers. However, the mutation rate is approximately

10^{-3}-10^{-4} per generation for repeat markers (microsatellites), which is much higher than for SNPs (10^{-8} or less) (Drake et al., 1998). The lower mutation rate makes SNPs more optimal for analyses where the cosegregation of a marker allele and disease is monitored for many generations (Kruglyak., 1997). There are nearly 1.4 million SNPs in human genome and some 9 million of them have been identified as a result of the HGP project.

In order to perform linkage analysis the first requirement is to establish genotypes for the individuals in the study sample. The analysis follows the segregation of the condition and marker alleles at consecutive loci. The interpretation of the genotype data requires statistical means and the basis of the statistics is the formula developed by N.E. Morton (1955), which measures the strength of linkage in units of logarithm of the odds (LOD score, Z). This can be written:

(1) $Z(\Theta)=\log_{10} L(\Theta)/L(0.5)$

Θ stands for recombination fraction between two loci where the LOD score peaks. $\Theta \geq 0.5$ indicates that there is no linkage between loci. In logarithmic scale the definition for statistically significant linkage is LOD score ≥ 3.0 for human disorders with classic mendelian inheritance.

In most cases the initial linkage can be detected over a large chromosomal region. This resolution is not sufficient for efficient candidate gene study unless there are some functionally obvious candidates. In most cases the chromosomal region observed with the initial linkage analysis needs to get narrowed down.

3.3.1 Linkage disequilibrium, LD

Linkage disequilibrium is the non random association of alleles at different loci on chromosomes with frequencies higher than expected from the random combinations of their frequencies in the population. Several measures are available for quantifying

this phenomenon. λ, D´ and r^2 are frequently used parameters to measure the amount of LD (Terwilliger, 1995; Zondervan and Cardon, 2004).

The fundamental idea of utilizing linkage disequilibrium and haplotype mapping in gene mapping is that the affected individuals have inherited the gene defect underlying the disease from the same ancestral relative (founder) who has introduced the mutation to the population. As a results the genetic region in the vicinity of the mutation is identical in all the affected individuals carrying the same founder mutation in their chromosomes i.e. the mutation and the marker alleles are in linkage disequilibrium (LD) (Fig. 6). The length of LD reflects the mutation age (number of generations since the mutation was introduced to the population). The mutation age can be mathematically estimated by using the Luria Delbruck based algorithm (Hästbacka et al. 1992).

Isolated populations with relatively small amount of founders have been especially useful in isolating genes in Mendelian disorders due to the reduced genetic heterogeneity (Peltonen, 1999) and often the disease causing mutation is in complete LD with regional genetic marker(s).

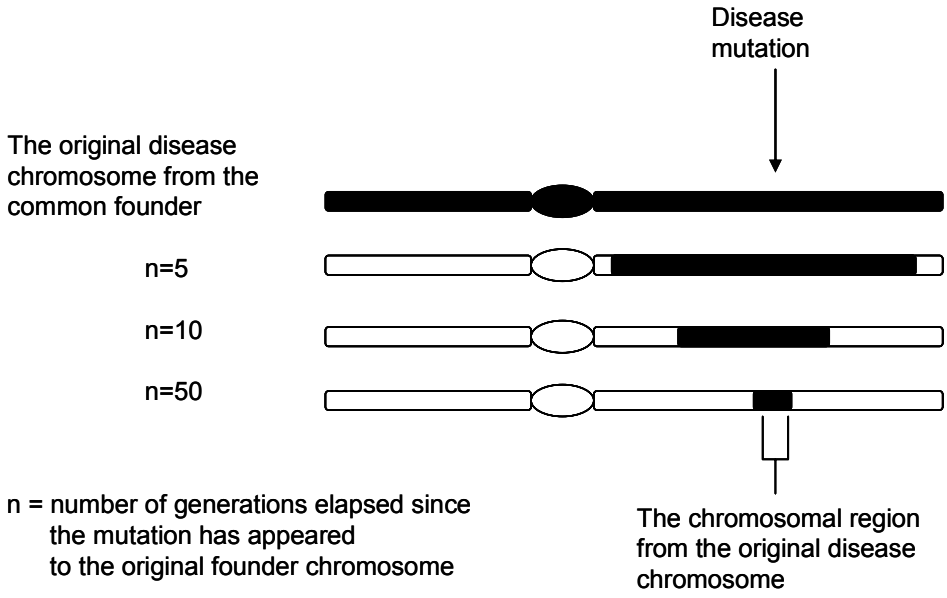

Figure 6. The principle of LD. The black chromosome represents the original ancestral chromosome carrying the disease mutation. After ten generations (n=10) the recombinations have narrowed the chromosomal region around the mutation. After 50 generations (n=50) there is only a relatively short chromosomal region in the closest surrounding area of the mutation originating from the ancestral disease chromosome.

3.3.2 Haplotype mapping

In homogenic and isolated populations patients with a rare genetic disorder are likely to be remotely related. The haplotype conservation of disease chromosomes over a large chromosomal region confirms that they are identical by descent. Haplotype mapping can be utilized after the disease locus has been fine-mapped and where the analysis of the recombination sites in disease chromosomes has enabled significant reduction in the size of the critical chromosomal region. A good example of restriction of the critical DNA region based on the shared haplotype is from EPM1 (Progressive myoclonus epilepsy). The haplotype shared by affected individuals restricted the EPM1 critical chromosomal region to 176 kb (Lehesjoki et al., 1993).

3.3.3 Homozygosity mapping

In rare recessive diseases the affected children of consanguineous parents have inherited two identical defective gene copies from a common ancestor. As a result the affected individuals share the same homozygous alleles near the disease mutation (Smith, 1953). The approach, in which shared homozygous disease haplotypes resulting from consanguineous matings are utilized in disease locus identification, is called homozygosity mapping (Lander and Botstein, 1987).

3.3.4 Identification of a disease mutation

The interpretation of observed sequence variants in the patients´ DNA and their putative role in disease pathogenesis is critical. Mutations causing frameshift alterations in the transcript or premature stop codons (nonsense mutations) significantly alter the gene product and make the evaluation of the disease causing mutations normally less problematic than in cases where the putative disease causing mutation changes only one amino acid to another (missense mutation). In the latter case the pathogenicity of the suspected disease mutation has to be tested. This can be done by utilizing the gene conservation between species since most often the disease causing mutation alter a conserved functional domains in the gene products. The mutation may change the stability of the polypeptide, which can be experimentally studied by comparing the stability of the mutated gene product to a normal control. Another important tool to analyse the potential disease mutation is to establish experimentally its carrier frequency in the population, which should not exceed the estimated carrier frequency. Even if the experimental tests would support the pathogenic role of the missense mutation the final proof of its causative role comes from studies in a model organism, where the missense mutation produces a pathogenic phenotype. However, in some cases species may have evolved differently and complementing pathways may still result in a normal phenotype in the model organism (Engle et al., 1996; Cantor 1998).

4. CILIA

The implications of cilia in the vertebrate organogenesis have only recently been recognized. The establishment of left right asymmetry of the gastrulating embryo is the earliest developmental process associated with cilia. Patients with primary ciliary dyskinesia syndrome (PCD) have abnormalities in the organ situs which are caused by defects in different axonemal dynein proteins affecting the ciliary motility (Afzelius, 1985). Around the time of gastrulation, the embryo forms a node, which carries motile cilia. The cilia beat, producing a leftward flow of fluid, which is detected by immotile sensory cilia on the left side, "telling the left side of the embryo that that's the left side". Paralysis and loss of fluid flow of the embryonic nodal cilia in mouse mutants affecting the cilia functions have indicated the key role of cilia in establishing the organ situs (Brody et al., 2000; Chen et al., 1998; Marszlek et al., 1999, Nonaka et al., 1998). In zebrafish it has also been shown that the loss of the fluid flow in the Kupffer´s vesicle (KV) is associated with laterality defects (Kramer-Zucker et al., 2005).

Known mouse phenotypes associated with abnormal cilial functions are CNS defects (hydrocephalus), bile duct hyperplasia, kidney cysts and skeletal patterning defects (Yoder et al., 1995; Cano et al., 2004). Zebrafish embryos with ciliary dysfunction have similar developmental phenotypes as characterized in mouse models for cilial dysfunction. The cilia driven fluid flow is required for normal organogenesis in zebrafish (Kramer-Zucker et al., 2005). The mechanism by which ciliary dysfunction may lead to the various organ pathologies still remains unclear.

4.1 Virtually ubiquitous organelle in mammalian cells

Cilia and flagella (the terms are interchangeable) are microtubule-based structures nucleated by modified centrioles termed basal bodies. Cilia are predominantly associated with epithelial cells. These small organelles are also present in endothelial cells, neurons, fibroblasts, chondrocytes, and many other cell types, with the few exceptions being avoid from cells of myeloid and lymphoid origin (Wheatley, 1995). The general structure of a cilium consists of a membrane-enclosed tube surrounding a central core of microtubules (axoneme). A cross section of cilia axoneme reveals nine microtubule doublets (9+0) surrounding, in some cases a central pair of microtubules (9+2) (Takahashi, 1984). The motile 9+2 cilia are present on the respiratory tract, the epithelium of the oviduct, the efferent ductules of the testis, and the ependymal lining of the brain. The primary cilia (9+0), found in many cell types, are normally non-motile, however the cilia in the embryonic node with 9+0 structure are motile and are capable in rotational beating (Nonaka, 1998). All eukaryotic cilia and flagella are surrounded by a membrane that is continuos with the plasma membrane, but appears to be a separate domain with a unique complement of membrane proteins (Handel, 1999; Brailov, 2000; Pazour, 2002). Cilium originates from one of the basal bodies, a modified form of centriole. The primary cilia are thought to be sensory organelles involved in chemo-, photo-, and mechanoreception (Rohlich, 1975; Burchell, 1991).

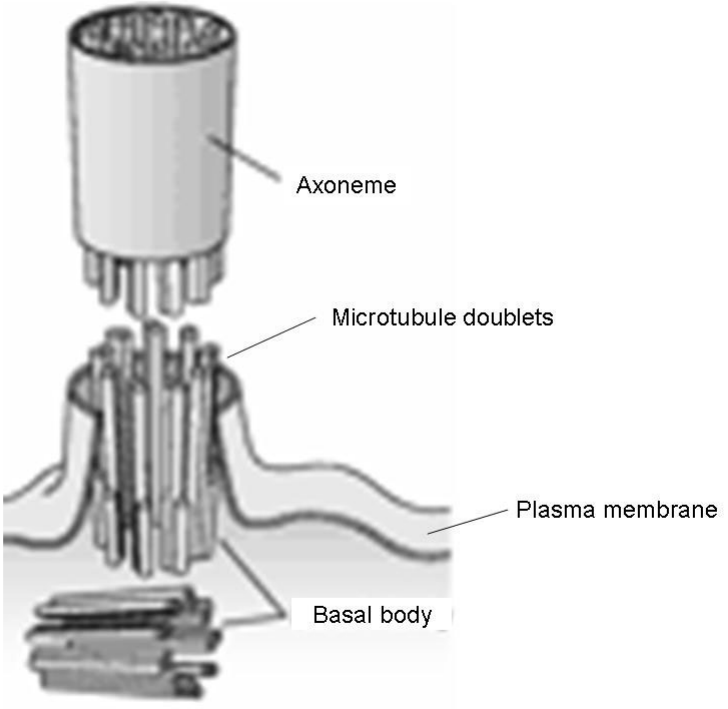

Figure 7. Schematic picture of the structure of a cilium composed of the basal body and cilial axoneme. Modified from Zhang et al., 2004.

4.2 FABB proteome

The assembly and maintenance (ciliogenesis) of 9+2, 9+0 cilia/ flagella are dependent on intraflagellar transport (IFT). IFT is the bidirectional movement of multisubunit protein particles along axonemal microtubules, which is required for assembly and maintenance of eukaryotic flagella and cilia. One posited role of IFT is to transport flagellar precursors to the flagellar tip for assembly. The IFT was first described in *Chlamydomononas* (algae) (Kozminski, 1993) and has subsequently been found to be essential for the assembly of cilia in many organisms. IFT polypeptides are conserved in ciliated organisms from *C. elegans* to humans, but are not found in non-ciliated organisms like *Saccharomyces* and *Arabidopsis* (Cole, 1998; Pazour, 2000). Comparative genomics study in *Chlamydomonas* (having flagella), human (having cilia) and *Arabidopsis* (have no cilia of flagella) has

resulted in the identification of the proposed human ciliary proteome, a group of 688 gene products. These gene products compose so called flagellar (cilial) basal body (FABB) -proteome (Avidor-Reiss et al., 2004; Li et al., 2004).

4.3 *C.elegans´* X box genes

Studies with model organisms such as *C. elegans* and *Chlamydomonas* have provided important information about the well conserved mechanisms in ciliogenesis. *C. elegans* has only 9+0 type of cilia and the regulation of the ciliogenesis in worm has also been characterized on the transcriptional level. In *C. elegans* 60 out of the 302 neurons of the hermaphrodite are ciliated sensory neurons (CSN) (Ward et al., 1975) forming many structurally distinct types of sensory cilia The RFX-type (regulatory factor binding to the X box) transcription factor daf-19 is required for the sensory neuron cilium formation, and loss of daf-19 function results in absence of sensory cilia (Swoboda et al., 2000). The analysis of X box sites in the *C.elegans* genome have enabled to identify genes coding peptides for the worm ciliary proteome that are under the control of cilia specific transcription factor daf-19 (Efimenko et al., 2005). The high conservation of the sensory cilia in worm and mammals makes the nematode a significant model for the cilia functions.

4.4 Cilia in the early development and cell cycle

4.4.1 Conserved nodal cilia

The conserved nodal cilia in different model organisms (mouse, xenopus, zebrafish and chicken) indicate that the activity of nodal cilia probably is a universal mechanism for specifying the left-right axis. The earliest known asymmetric expression patterns are common to all vertebrates although exhibit considerable variability in their time of onset among different vertebrate classes (Capdevila et al., 2000; Wright 2001). The asymmetries are always preceded by the onset left-right dynein heavy chain (*Lrdr*) gene expression, which is a component of the ciliary motor and by the appearance of the cilia, indicating that nodal cilia may be

responsible for initiating the left-right asymmetric gene expression and for establishing the final body plan in all vertebrates (Essner et al., 2002).

In human the mutations in the dynein heavy-chain gene (*DNAH5*) are also associated with immotile cilia syndrome, which is an inherited disorder that has mirror-image reversal of the internal organs in half of the affected individuals (Olbrich et al., 2002).

4.4.2 Cilia in cell cycle

The presence of a cilium is associated with establishment of polarity and differentiation of the cell. In tissue culture cells usually grow primary cilium as they approach confluence, and most ciliated cells are probably in stationary or G_0 phase of the cell cycle (Tucker et al., 1979). In many cells the entry into the cell cycle is preceded by ciliary resorption whereas exit from mitosis is accompanied by ciliary assembly, which may reflect the use of the basal bodies/centrioles as mitotic spindle poles (Rieder et al., 1979; Tucker et al., 1979; Ehler et al., 1995; Wheatley, 1996).

Research on *Chlamydomonas* has provided insights as to how cilia and the cell cycle are co-coordinately regulated. Nek family of cell cycle kinases provides an important general connection between cilia and regulation of the cell cycle progression because of the axonemal location (Quarmby et al., 2005). The many Neks in organisms with cilia have roles in cell cycle regulation. The only connection to the vertebrates comes from studies of PKD models. The causative genes of two mouse PKD models are Nek kinases (*Nek1 and Nek8*). Human Nek1 has been found to bind kinesin-2 and some other proteins that are required for ciliogenesis (Surpili et al., 2003). A kinase domain mutation of the human orthologue of Nek8 affects cell cycle progression, probably at G_2/M checkpoint.

There is a relationship between cilia, and cell cycle progression. Available data indicates that most vertebrate cells are ciliated and that the genesis and disassembly of the cilia is coordinated with progression through the cell cycle.

4.4.3 Cilia and signaling

The kidney morphogenesis and Wnt

The Wnt/Frizzled pathway is a major pathway (Fig. 8) implicated in the specification of cell and tissue polarity and operating in different developmental processes, including heart and neural tissue development, kidney morphogenesis, limb polarity and sex determination. The mouse model with disrupted *inversin* – gene (*Inv/Inv*) has left-right axis defects, hepatobiliary tract anomalies and severe renal cystic disease (Mochizuki et al, 1998; Morgan 1998). Mutations in human the homolog *INVS* result in an infantile form of nephronophthisis (NPHP2; Otto et al., 2003). The inversin peptide contains multiple ankyrin repeats, two D boxes and two calmodulin binding motifs. It associates with nephrocystin-1 (NPHP1) and this protein complex has been localized to the primary cilia-basal body-centrosomal complex (Otto et al., 2003; Morgan et al., 2002; Nurnberg et al., 2002). The kidney phenotype in *inv/inv* mice is similar to mice with dysregulated Wnt-β-catenin signaling proposing that *Inversin* is implicated in the Wnt pathway. Based on the studies in mouse and xenopus it has been shown that inversin inhibits the canonical Wnt signaling upstream of the β-catenin degradation complex and further the study in the zebrafish model showed that inversin functions as a positive regulator of the non-canonical Wnt signaling (Fig. 8). Further, studies on the ciliated MDCK (Madin-Darby canine kidney cells) showed that the passing a flow (acting as a signal that flips the switch from canonical to non canonical Wnt signaling) over the apical surface of these ciliated cells results in increased expression of inversin accompanied by modest reduction in β-catenin levels (Simons et al., 2005). This would suggest a model that any process that impairs this switch would result in unopposed canonical Wnt signaling and renal cystic disease (Simons et al., 2005; Germino 2005). The essential role of unopposed canonical signaling as a central pathogenic failure in cystic renal disease remains still unknown and this model most likely will not explain all forms of polycystic kidney disease. However, the finding

that cilial-basal body-centrosomal complex regulates non-canonical Wnt signaling could explain the various genes implicated in cystic kidney disease (Germino 2005).

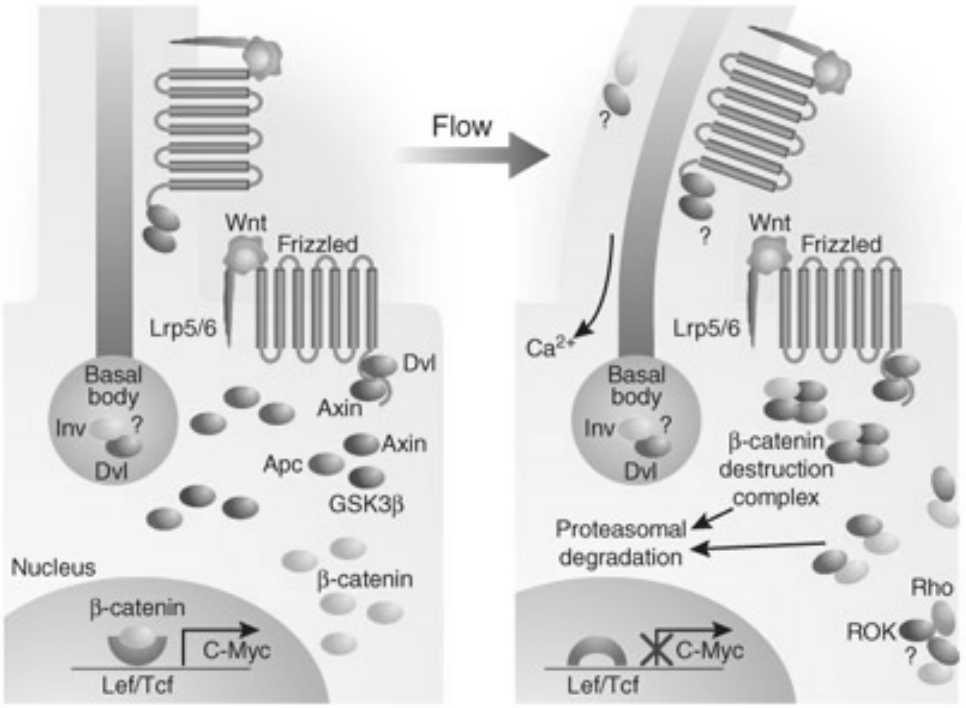

Figure 8. Linking cilia to Wnts. Flow-based model of Wnt regulation in the kidney. In the model proposed by Simons et al., Wnt signaling occurs primarily through -catenin−dependent pathways in the absence of flow. Ligand binding by the Frizzled-LRP complex results in inactivation of the -catenin destruction complex, increased cytoplasmic and nuclear -catenin levels, and upregulation of effector gene expression (left). Stimulation of the primary cilium by flow is postulated to result in increased expression of inversin (Inv), which then reduces levels of cytoplasmic Dvl by increasing its proteasomal degradation. This process is thought to switch off the canonical pathway by allowing reassembly and activation of the -catenin destruction complex. Inversin might also enhance trafficking of Dvl to the plasma membrane, where it could activate noncanonical Wnt pathway components. The question marks indicate uncertainty as to whether the inversin-Dvl or Frizzled-LRP complexes are present in the primary cilia−basal body−centrosomal complex. Modified from Germino, 2005.

Hedgehog: CNS and limb bud morphology

The specification of ventral cell types in mammalian neural tube depends on sonic hedgehog release from the notochord (Chiang et al., 1996; Ingham and McMahon, 2001). Mouse studies have shown that the intraflagellar transport (IFT) proteins Ift172/Wimple and Polaris/Ift88 and the anterograde IFT motor protein Kinesin-2 are required for the ciliogenesis as well as for the activation of targets on the mouse Hedgehog pathway by Gli transcription factors (Huangfu and Anderson, 2005). Gli3 repressor is required for the normal anterior-posterior patterning of the limb, and the reduction of Gli3 can account for the polydactyly phenotype seen in partial loss of *Polaris* mouse mutants (Huangfu et al., 2003; Liu et al., 2005). It has been already noticed that specialized primary cilia can detect the fluid flow (mechanosensory cilia), and also that the somastotatin receptor 3 and serotonin 5-HT_6 receptor localize to the primary cilia on of brain neurons (Händel et al., 1999; Brailov 2000) and thus suggest that the cilia have chemosensory functions. The studies with the IFT proteins in vertebrate model organisms propose a structural role of cilia in cells which is essential for responding to positive and negative Hedgehog signals (Huangfu and Anderson, 2005).

4.5 Human ciliopathies

The main hallmark of the known human ciliopathies is cyst formation in the kidneys, which occurs when the kidneys are not properly formed or the instructions for maintaining the proper kidney diameter are somehow missing. It has been presumed with the present knowledge that in all cystic kidney diseases no matter of the age of the clinical manifestation of the disease, the initial events of cyst formation occur during embryonic development (Simons and Walz, unpublished). Two of the most common and best characterized human disorders that can be considered as ciliopathies are Polycystic kidney disorders and Bardet Biedl syndrome.

Autosomal dominant polycystic kidney disorder (ADPKD) is the most common human monogenic disease affecting 500-1000 in a population. ADPKD leads to end

stage renal disease at the age of 50-60. In ARPKD the patients suffer renal failures and liver fibrosis in the early childhood. Another subgroup of PKD is Nephronophthisis (NPHP); with prevalence of 1: 100 000 being the most frequent genetic cause of chronical renal failure of children. Five different loci have been assigned so far for NPHP that account for approximately 50% of the cases (Otto et al., 2005).

Several mechanisms have been proposed underlying the cyst formation in PKDs, but they still remain poorly understood. The PKD mouse models have shown the significance of the functional cilia in the pathology of the kidney cyst formation.

Bardet Biedl syndrome

Bardet Biedl syndrome (BBS, MIM 209900) is a pleiotropic condition. The phenotype of BBS partially overlaps with MKS. However, the BBS phenotype is notably milder, characterized by postaxial polydactyly, progressive retinal dystrophy, obesity, hypogonadism, renal dysfunction and learning difficulties. Other frequent symptoms are diabetes mellitus, heart disease, hepatic fibrosis, and neurological features. The condition is heterogeneous and already ten disease genes have been identified to be involved in genetics of the BBS (Nishimura et al., 2005; Chiang et al., personal communication and abstract: ASHG meeting 2005). Polydactyly, cystic kidneys, genital-, heart- and liver anomalies can be observed in BBS patients at the birth but not severe CNS defect (occipital encephalocele) as in MKS patients. The ethiology underlying BBS is dysfunction of proteins involved in centriole/cilia assembly and function (Ansley et al., 2003). It has been already suggested in the literature that MKS and BBS might interact genetically, although the authors assumed that it would be unlikely that these disorders could be allelic (Karmous-Beinally et al., 2005).

AIMS OF THE PRESENT STUDY

When this study was initiated the MKS1 locus had been assigned to chromosome 17q and the following aims were set for this study:

1. To produce physical maps for the critical MKS1 chromosomal regions in both human and mouse to enable and facilitate efficient candidate gene study (I, II).

2. To identify the MKS1 disease gene in Finnish families (III).

3. To provide initial information of the molecular pathogenesis in MKS (III).

4. To validate the genetic heterogeneity of MKS in the Finnish families (IV).

SUBJECTS AND METHODS

1. SUBJECTS

1.1 Family material and ethical aspects

Initially we had DNA samples from 24 Finnish MKS families (I), but 16 new families have been included in the study during the years 1999-2006, so at the moment we have samples from 40 Finnish families (III). The DNA samples are available from parents, MKS affected children and most of the healthy siblings. All the Finnish families fulfil the modern diagnostic criteria for MKS (Salonen, 1984). In addition, we have DNA samples from 20 non-Finnish families of European descent.

Fibroblast cell lines have been established for most of the Finnish patients. The DNA samples and patient fibroblast cell lines have been obtained in collaboration with clinicians and genetic counsellors both in Finland and abroad.

This study was carried out with approvals (decision numbers 28/94, 205/E0/04, 57/E0/04) from the ethical committees of the Hospital District of Helsinki and Uusimaa (HUS).

1.1.1 Control chromosomes and normal control tissues

A total of 541 control DNA samples were collected from western, central and eastern Finland. In addition 92 CEPH grandparental DNA samples, two human variation panels (one with 50 Caucasian and one with 100 African American DNAs) from the Coriell Institute were used for carrier screening of the detected sequence alterations.

The normal control tissues samples and fibroblast cell lines were obtained from fetuses aborted for social reasons.

2. METHODS

2.1 Laboratory methods

Table 1: The used laboratory methods and materials.

Materials and methods	Reference	Study
Bacterial artificial chromosomes (BAC)	Shizuya et al., 1992	I
Direct sequencing	Maxam and Gilbert, 1977; Sanger et al., 1977	I, II, III, IV
Nucleic acid extraction	Sambrook et al., 1989	I, II, III, IV
Electromobility shift assay (EMSA)	Samadani and Costa, 1996; Kyttälä masters´ thesis 2003	unpublished
Fluorescence *in situ* hybridization (FISH)	Pinkel et al. 1986; Lichter et al. 1988	I
Haplotype analysis	Ramsay et al., 1993	I, III, IV
Linkage analysis	Botstein et al., 1980; Ott, 1991	IV, unpublished
Linkage disequilibrium (LD) analysis	Hästbacka et al., 1992	I, III, IV
Molecular cloning	Sambrook et al., 1989	unpublished
Northern blot analysis	Sambrook et al., 1989	II
Physical mapping	Green and Olson, 1990	I, II, IV
Polymerase chain reaction (PCR)	Weber and May, 1989	I, II, III, IV
P1 clones (PACs)	Ioannou et al., 1994	I
Reverse transcriptase PCR (RT-PCR)	Ohara et al., 1989	II, III
RNA in Situ hybridization	Breitschopf et al., 1992	III
Western blot analysis	Burnette, 1981	unpublished
Yeast artificial chromosomes (YACs)	Burke et al., 1987	I

2.1.1 Physical mapping

The contiguous genomic contig (I)

Firstly the YAC clones were identified for the MKS-MUL region (I). Clones positive for the previously mapped STSs and ESTs (Hudson et al., 1995; Dib et al., 1996) between markers D17S957 and D17S1604 were provided by CEPH. Individual clones were cultured and total yeast DNA was extracted and stored in agarose beads or extracted by the alkaline lysis method. The chimerism of YAC clones was analyzed by the metaphase FISH technique (Bray-Ward et al., 1996), and clones mapping only to 17q22-q23 were used for further mapping. The presence of STSs and ESTs (Hudson et al., 1995; Schuler et al., 1996) in the YACs was tested

by PCR. The STSs and ESTs present in the YACs were used for screening of PAC and BAC libraries by PCR. The PAC library consisting of 321 multiwell microtiter plates was kindly provided by Professor Pieter J. De Jong (Roswell Park Institute, Buffalo, NY). The commercially available P1 library (Genome Systems, Inc., St. Louis, MO) and BAC library DNA pools (release III, Research Genetics, Inc., Huntsville, AL) were used for the P1 and BAC screenings. PCR-positive clones were picked and cultured in Terrific Broth media of Luria-Bertani media, supplemented with 25 mg/ml of kanamycin (PACs) (Ioannou et al., 1994) or 12.5 mg/ml of chloramphenicol (BACs) (Shizuya et al., 1992). DNA was extracted using a Plasmid Maxiprep (tip 500) kit (Qiagen).

FISH, Fluorescence in situ hybridization (I)

The metaphase chromosome targets were achieved by culturing human peripheral blood lymphocytes according to standard protocols (Lemieux et al., 1992; Tenhunen et al. 1995). Agarose-embedded lymphocytes were used as a source of free DNA fibers according to the method described earlier (Heiskanen et al., 1994, 1996). All clones used for the FISH experiments (PACs and BACs) were labeled with biotin 11-dUTP (Sigma) and digoxigenin 11-dUTP (Boehringer Mannheim) by nick translation according to standard protocols. The FISH procedure was performed as described earlier (Pinkel et al. 1986; Lichter et al. 1988; Aaltonen et al. 1997). Briefly, the hybridization mixture contained 50% formamide and 10% dextran sulfate in 2× SSC. Repetitive sequences were suppressed with 10-fold excess of Cot-1 DNA (BRL, Gaithersburg, MD). After overnight incubation at 37°C the slides were washed at 44°C three times in 50% formamide and 2× SSC, twice in 2× SSC, and once in 0.5× SSC. The slides were counterstained with DAPI (Sigma).

Digital Imaging A multicolor image analysis was used for acquisition, display and quantitative analysis of fiber FISH as described elsewhere (Heiskanen et al. 1996; Aaltonen et al. 1997). In fiber FISH, the clone sizes determined by pulse field gel electrophoresis (PFGE) were used as a standard for calibration in each individual

fiber FISH image. The distance measurements were done by applying the distance measurement option of the IPLab software (Signal Analytics Corp., Vienna, VA). In each fiber FISH image, at least one clone was used as a calibration standard to measure the sizes of other clones, gaps separating different clones, and overlaps between these clones. The known clone sizes enabled compatible and therefore reliable measurements to be made. If two known clone sizes were in the same hybridization, the results of the distance measurements were verified with each other.

Isolation of Novel CA-Repeats (I, III)

Before the genomic sequence was available novel CA-repeats were isolated from the PAC and BAC clones (I, Klockars et al., 1996; Ranta et al., 1997). Shortly, PAC DNA was digested with *Sau*3A1 and ligated to pGEM7, transformed into DH5α cells and plated onto selective agar plates. BAC DNA was digested with *Rsa*I and ligated to *Sma*I-cut pUC18 vector (Pharmacia, Biotech), transformed into Epicurian Coli SURE competent cells (Stratagene), and plated onto selective agar plates. The colonies were transferred onto nylon filters (Hybond-N, Amersham or Magna Chart, MSI) and hybridized with a radioactively labeled $(GT)_{16}$ oligonucleotide at 65°C overnight. After autoradiography, positive colonies were selected, and DNA was isolated using a Plasmid Miniprep (tip 20) kit (Qiagen) or Wizard Plus Miniprep kit (Promega). The inserts were sequenced with SP6 and T7 vector primers. Oligonucleotide primers flanking the CA-repeats were selected with Primer3 software (Whitehead Institute). Nowadays the repeats are identified directly from the genomic sequence and the putative polymorphism is tested in chromosomes representing the population (III).

2.1.2 Candidate gene studies in the patients

The positional candidate genes for MKS1 were selected by monitoring the shortest region of shared homozygosity in the Finnish MKS chromosomes (I, III). We chose three Finnish patients with the common Finnish disease haplotype in the marker interval

(17q: D17S1290-132-CA (I), and later D17S2263-D17S1290 (III) for the candidate gene sequencing. In addition, we extracted RNA for northern blot and RT-PCR analyses from fibroblast cell lines and from frozen tissue biopsies (in liquid Nitrogen) from patients with the common Finnish MKS haplotype and normal controls.

Candidate gene sequencing

The coding regions and intron-exon boundaries of the candidate genes were sequenced from genomic DNA and RT-PCR products. PCR primers were designed with the Primer3 program (Whitehead Institute).

PCR products were enzymatically purified (schrimp alkaline phosphatase and EXOI, Amersham Pharmasia) and bidirectionally sequenced with BigDye Terminator chemistry on 377/3730 DNA Analyzer (Applied Biosystems). We analyzed the sequences with Sequencher software (Gene Codes; v. 4.0.5).

Electromobility shift assay, EMSA (Kyttälä, 2003)

The nuclear extract was extracted from the human embryonic kidney cell line HEK293T by following the protocol by Andrews and Faller (1991). The EMSA assay conditions reported by Samadani and Costa (1996) were slightly modified. The DNA-protein hybridization contained 2ng labeled DNA probe (PCR product/ 61-mer oligo), 20 mM HEPES (pH 7.9), 4% Ficoll, 2mM $MgCl_2$, 40 mM KCl, 0.1 mM EGTA, 0.5 mM DDT (dithiotreiol), 4 ng sonicated salmon sperm and 40 ng of the nuclear extract from the HEK293T cells.

Two different types of DNA probes were utilized: ds 61-mer oligos and PCR products. Both type of DNA probes were tested for the wild type (normal control) sequence and for the patient sequence (homozygous for the *SEPT4* gene G-allele; mRNA GI_17986248, bp 594).

The used ds 61-mer oligos for patient and wild type alleles were (only the forward strand sequence shown): G GAG TGT GTG GTG GTG GTG GTG GTG GTG GTG GTG GTG GTG TGT **T/G** TG TGT GTG TGT GTA TCT. The underlined **T/G**

bases show the site of observed sequence alteration in the patients (T→G). The results, which were obtained with the ds oligos, were verified with PCR amplified DNA probes (146 bp in size) from patient and control DNA templates. The PAGE and blotting were performed in standard non-denaturating conditions. The ds 61-mer oligos and PCR amplified DNA probes were both detected in ECL (biotinylated DNA probe) (Amersham Biosciences) and in autoradiography with ^{32}P labeling (Redivue γ-^{32}P-ATP, Amersham Biosciences).

Northern blot (II, unpublished)

The transcripts were analyzed on blots generated from 3 μg of mRNA extracted from fetal tissue (MKS and normal control) by using the Direct mRNA midi kit (Qiagen) following the manufacturer's protocol. Probes: 400 to 500 bp DNA fragments were generated by RT-PCR amplification, and ^{32}P labeled with a random priming kit (Invitrogen). Blots were hybridized with the Express-Hyb solution (Clontech), as recommended by the manufacturer.

2.1.3 Genotyping

Polymorphic markers were amplified by PCR and the fragments were analyzed on 377 (I) or 3730 (III, IV) DNA Analyzers (Applied Biosystems). The allele sizes were determined using the size standard provided by the manufacturer with the standard software package (GeneMapper v.3.7).

10K SNP array (IV)

The genome wide SNP analysis was undertaken by using the GeneChip® Human Mapping 10K Array Xba 131 and assay kit (Affymetrix, Santa Clara, CA). The 10K array contains nearly 11 500 SNPs and the median intermarker distance is approximately 200 kb. An average heterozygosity of 0.37 for these SNPs is reported by the manufacturer. The assay was performed according to the Affymetrix GeneChip® Mapping Assay Manual. SNP-genotypes were obtained by following

the protocol for the GeneChip® mapping 10K Xba Array. All analyzed DNA-samples achieved over 95% call rate.

Genome-wide scan with microsatellite-markers (IV)

The genome-wide scan with 380 multiallelic markers (Applied Biosystems) was performed by The Finnish Genome Center (FGC) (http://www.genome.helsinki.fi/).

2.2 Biocomputing programs and statistical analyses

Table 2: Biocomputing programs and databases utilized in computational analysis.

Program/ Database	Brief description	URL/ Reference	Study
Blast	Sequence alignments and comparisons.	http://www.ncbi.nlm.nih.gov/BLAST/	I, II, III
Boxshade	Organizes the results from multiple sequence alignments to a visually	http://www.ch.embnet.org/software/BOX_form.html	III
ClustalW	Multiple sequence alignment tool.	http://searchlauncher.bcm.tmc.edu/multi-align/multi-align.html	III
GENEHUNTER	Multpoint calculations.	Kruglyak et al., 1996	IV, unpublished
Genotator	Gene prediction: an annotation workbench that runs various sequence analysis programs	Harris, 1997	II
Genscan	Gene prediction: Predicts exon-intron structures of genes in genomic sequences	Burge and Karlin, 1997	II, III
Linkage package	Basic linkage calculations.	Lathrop et al., 1985	IV, unpublished
MAPMAKER/ HOMOZ	Rapid multipoint analysis in nuclear pedigrees.	Kruglyak et al., 1995	IV
MatInspector	Transcription factor binding site prediction program.	http://www.genomatix.de/online_help/help_matinspector/matinspector_help.html	unpublished
NCBI main page	Genome data from different species and tools to analyze them. Contains links to a variety of other genome and related resourses.	http://www.ncbi.nlm.nih.gov/	I, II, III, IV

Pfam	Predicts domains based on the peptide sequence.	http://www.sanger.ac.uk/Software/Pfam/	III
Primer3	Primer design.	http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi; Rozen and Skaletsky, 2000	I,II, III, IV
Prosite	Proteomic tools.	http://au.expasy.org/prosite/; Hofmann et al., 1999	III
PSORT	Predicts cellular localization for peptide sequence.	http://www.psort.org/	unpublished
UCSC genome bioinformatics	Genome data from different species and tools to analyze them. Contains links to a variety of other genome and related resourses.	http://hgdownload.cse.ucsc.edu/downloads.html	II, III, IV

2.2.1 Statistical analysis

Linkage

The two point LOD-scores were calculated using the linkage package program MLINK (Lathrop et al., 1985). Autosomal recessive inheritance with full penetrance and estimated disease allele frequency 0.001 were used in the calculations.

LD analysis

The significance of LD between the disease and marker loci was analyzed using the DISLAMB program (Terwilliger, 1995). This program applies a likelihood-based test for linkage disequilibrium and has only one degree of freedom, irrespective of the number of alleles at any given marker. The calculations are based on the parameter λ, which expresses the proportion of the increase of a certain allele in the disease chromosomes, relative to its population frequency. Confidence intervals for λ were calculated for 1-lod units. The most likely haplotypes were constructed manually, assuming a minimum number of recombinations in each family.

Multipoint analysis of the 10K SNP array with MAPMAKER/HOMOZ program (IV)

We analyzed the genotypes produced with the 10K SNP (Affymetrix) array with the MAPMAKER/HOMOZ program (Kruglyak et al., 1995) for finding the regions of the shared homozygosity by the affected siblings. This program has been developed

for fast multipoint linkage analysis in small nuclear families. The fundamental assumption was that the parents have common ancestors (IV). The SNP markers on the 10K array were localized on the genetic map by using the data provided by Affymetrix. The missing genetic map information was estimated for some of the markers by utilizing their known physical position, and using linear interpolation with the adjacent SNP markers for which both physical and genetic map information was known. The markers were arranged into sets of 70 adjacent markers with 20 markers overlap between marker sets on a given chromosome. The genetic map positions of SNP-markers were converted to the Haldane units. The allelic frequencies were used as provided by Affymetrix in the Caucasian population, derived from the larger sample size. Five out of 11 229 SNP markers were void of allelic frequency information (SNP_A_1511539, on Chr5; SNP_A_1513409, on Chr11; SNP_A_1512847, on Chr12; SNP_A_1508640, on Chr15; and SNP_A_1511759, on Chr16) and were left out of the analysis. The SNP multipoint linkage analysis was performed by utilizing MAPMAKER/HOMOZ (v. 0.9) program with increments of 1 cM and 20 cM beyond the ends of the chromosomes.

2.3 RNA *in situ* hybridization of *Mks1 (I)*

The RNA *In situ* hybridization was performed to sections of mouse embryos (C57BL/National Public Health Institute, Helsinki, Finland) at embryonic day 15.5 (E15.5). The embryos were fixed by overnight immersion in 4% PFA in 0.1 M PB, pH 7.4. Samples were embedded in paraffin and cut into 5 µm thick sagittal sections. Mounting was performed onto SuperFrost/Plus microscope slides (Merck, Germany). The study is approved by the animal care and use committee of the National Public Health Institute, Helsinki, Finland. The work has been carried out following the good practice in handling of laboratory animals. The probe was produced by amplifying cDNA derived from a mouse fibroblast cell line with the *Mks1* specific primers: Insitu-F: AAGGGTTCAGCCAGCAGAGT; Insitu-R: TGGTTGCCAAACTCCCTTT. The 500 bp probe fragment was cloned into the

pGEM-T Easy Vector (Promega) and the anti-sense and sense probes with the digoxigenin label were created with a DIG RNA labeling kit (Roche), according to the manufacturer's protocol. The probes were diluted in 1:400 in a hybridization buffer containing 2x SSC, 10% dextran sulfate, 0.01% sheared salmon sperm DNA, 0.02% SDS and 50% formamide. A method described by Breitschopf and coworkers (1992) was used for the hybridization with slight modifications. Photographing and analysis of the samples were performed using a Leica microscope equipped with DC300F camera and IM1000 software and a Zeiss Axioplan 2 imaging microscope (Zeiss). The final figures were prepared by using CorelDraw v.12 and Paint Shop Pro v.7.04 softwares.

2.4 Transient expression of MKS1 (unpublished)

Plasmid construct

The protein coding region of the *MKS1* cDNA was amplified by RT-PCR with primers coding-F: ATGGCGGAGACCGTCTGGA and coding-R: CTAGGAGACCAGGGTTCCA. The RT-PCR product was cloned into pCR-BluntII-TOPO vector by using Zero blunt TOPO-PCR Cloning kit (Invitrogen) following manufacturers´ guidelines. The MKS1 insert was restricted from pCR-BluntII-TOPO vector with *Nsi*I (NEB) and ligated with T4-ligase (NEB) to pEGFp-C1 vector restricted with *Pst*I (NEB). The N-terminally GFP-tagged MKS1 construct was verified by sequencing and in western blot with GFP-antibody, which produced a band agreeing with the size of GFP-MKS1 fusion protein. The cell lysate for the western blot was made of HEK293 cells transiently transfected with Fugene 6 (Roche).

Cell culture

HEK293 and fibroblast cell lines were maintained in DMEM supplemented with 10% fetal bovine serum (FBS), 2 mM glutamine, 100 IU/ml penicillin and 100 µg/ml streptomycin at 37ºC degrees in 5% CO_2.

Immunofluorescence

Cells were grown on cover slips washed in phosphate buffered saline (PBS) and fixed with 4% paraformaldehyde (PFA). The cover slips were rinsed with PBS and permeabilized with 0.1% Triton x-100/0.5% bovine serum albumin (BSA) in PBS for 15 min in RT. Washed twice with PBS and blocked with 0.5% BSA in PBS for 30 min in RT. The cover slips were incubated for 1 h in RT with monoclonal anti-polyglutamylated tubulin mouse immunoglobulin (T9822, Sigma-Aldrich), which was diluted 1:16 000 into the blocking solution. Same procedure was repeated with the TRITC conjugated secondary antibody. Cover slips were finally washed twice with PBS and with dH_2O prior mounting.

RESULTS AND DISCUSSION

1. PHYSICAL MAP OF THE MKS1 LOCUS ON 17q23

The original linkage study by Paavola and colleagues (1995) assigned the first MKS locus (MKS1) to a 13 cM region on chromosome 17q in Finnish families. At that time, the public genetic maps for the region were inadequate for LD based mapping and direct candidate gene studies. Soon after, the MKS1 region was restricted to 5 cM between markers D17S1606 and D17S807 (Paavola et al., unpublished), but the exact physical size and gene content of the chromosomal region was missing. When the gene defect causing Mulibrey Nanism (MUL, MIM 253250) was assigned to the same chromosomal region on 17q (Avela et al., 1997), a close collaboration was initiated for physical mapping of the overlapping MKS-MUL loci together with the MUL research group.

In the beginning of physical mapping project a contig of genomic YAC clones was available for the human genome covering approximately 75% of the genome (Chumakov et al., 1995). This provided the basis of the MKS-MUL physical map. The YAC clones were utilized for STS-content mapping to confirm the order of the clones and the available STSs. We screened PAC and BAC libraries (Shizuya et al., 1992; Ioannou et al., 1994) with the available STSs and expressed sequence tags (ESTs) to construct a multiple coverage of genomic PAC and BAC clones and to generate additional microsatellite markers, which enabled the LD and haplotype mapping in the critical disease gene regions.

1.1 Initially restricted critical chromosomal region

Six novel dinucleotide repeat markers were isolated from three PAC (95i19, 52i20, and 58p18) and two BAC (272g3 and 132i10) clones on the 17q23 region. These were the markers: 95-CA, 272a-CA, 272b-CA, 132-CA, 58-CA and 52-CA) (Fig. 1).

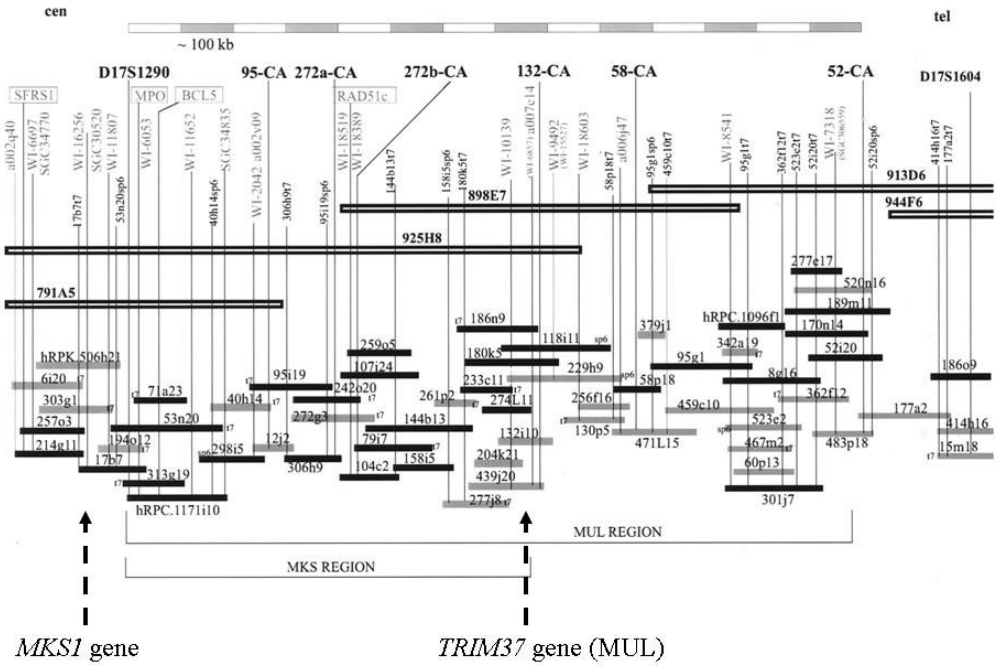

Figure 1. The physical map of the 17q MKS-MUL critical region. The horizontal bars stand for the multiple coverage YAC (with black line on top), PAC (black) and BAC (grey) clones. At the time only two sequenced genomic clones were available in the Genbank (NCBI) for the region: hRPK.506h21 and hRPC.1171i10. The dashed arrow lines indicate the genes which were later found to be defective in MUL (Avela et al. 2000) and MKS (III).

The original assumption was that all the 24 Finnish families available for us at the time would have mutations in the same gene on chromosome 17q. Due to this all the disease chromosomes were included in the LD and haplotype analyses. This defined a minimum 1 cM genetic region for the MKS locus between the markers D17S1290 and 132-CA (Fig. 2, on page 64). Five patients were thought to be compound heterozygous for the common MKS haplotype and another haplotype, and in two patients, both chromosomes differed from the common MKS haplotype. In seven of

the patients, four different putatively rare MKS haplotypes were observed and those were interpreted to represent probable "minor" MKS mutations.

The insert sizes of the PAC and BAC clones constituting the minimum tiling path of the contig were determined by using restriction digestion followed by pulsed-field gel electrophoresis. Fiber FISH analyses were used to verify and confirm these size estimations. In fiber FISH, one or two bacterial clones, whose sizes were determined by PFGE, were used as a standard for calibration in each hybridization image to estimate the sizes of other clones. According to the PFGE data and fiber FISH, the size of the contig from D17S1290 to 52-CA, that is also the critical region for MUL, was estimated to be ~1.4 Mb. Furthermore, the total physical length of the critical region for MKS was determined to be ~800 kb between the markers D17S1290 and 132-CA (Fig. 1).

The detailed physical map enabled us to precisely localize a total of 20 ESTs of both known and unknown homology, which were previously only roughly assigned either by radiation hybrid mapping or by STS-content mapping to the YAC clones in the region (Fig. 1) These ESTs were derived from the gene map of the human genome (Schuler et al., 1996).

Markers	MKS-associated haplotpyes from unrelated patients																				
D17S1606	3	3	×	3	×	3	3	5	4	5	8	7	10	7	N.D.	3	N.D.	5	13	7	2
D17S1290	5	6[a]	5	5	5	5	5	5	5[b]	5	2	3	8	8	6	2	1	6	6	6	6
9S-CA	3	3	3	3	3	3	3	3	3	3	3	3	2	2	2	4	4	3[b]	3	3[b]	6
272a-CA	3	3	3	3	3	3	3	3	3	2[a]	3	3	5	5	5	4	4	2	2	2	4
272b-CA	5	5	5	5	5	5	5	5	5	5	5	5	4	4	4	2	2	1	7	6	3
132-CA	5	5	5	5	5	5	5	5	5	3	4	4	4	5	5	4	6	4	5	5[b]	3
58-CA	7	7	7	7	7	7	6	7	5	6	3	N.D.	2	6	6	N.D.	5	6	5	6	5
52-CA	3	3	3	3	3	6	5	6	6	3[b]	2	3	5	6	N.D.	2	N.D.	6	2	6	2
D17S1604	6	6	6	×	×	5	9	9	5	8	6[b]	4	9	14	N.D.	2	N.D.	4	1	6	2
	Number of chromosomes detected:																				
	13	1	3	4	11	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	Total number of unrelated MKS chromosomes 48																				

1.2 Discussion of the impacts of the physical map

Before the genomic sequence and transcript maps were available physical mapping was a necessity prior to positional candidate gene analysis. It also enabled the LD and haplotype analysis with novel dense markers in disease chromosomes. The physical map made it possible to exclude functionally significant candidate genes based on their physical position

The assigned gene loci for MKS and MUL overlapped but it was not clear whether these two disorders were allelic or not. The *MUL* gene was found in the distal part of the MKS region by Avela and colleagues in 2000 (Fig. 1). The gene mutated in MUL did not contain any pathogenic sequence alterations in direct sequencing of the genomic- and RT-PCR products in the MKS patients. The northern analysis of the *MUL* gene in patients´ tissue samples did not detect any changes in the expression levels nor in the transcript sizes when compared to a normal controls (Kyttälä et al., unpublished).

The size of the MKS1 critical chromosomal region (D171290-132-CA) was estimated to be 800 kb by pulsed field gel electrophoresis and fiber FISH analysis. This agrees very well with the exact size of 782 254 bp (USCS, genome browser, May 2004).

2. HUMAN AND MOUSE COMPARATIVE MAPS OF MKS REGIONS

Comparative genomics in human and mouse was utilized to facilitate the positional cloning of the MKS gene on the human chromosome 17q. The critical MKS region was restricted in the LD and haplotype analyses between the marker interval D171290-132-CA, which is 800 kb in size. We identified the mouse syntenic regions for the three known human MKS loci and the gene content and order in them. Initially the available public genomic sequences and transcript maps in both species were incomplete alone. The comparative genomics provided an alternative tool to extract more information from the regional candidate genes. Reciprocal comparisons of the known transcripts, ESTs and genomic sequences in human and

mouse enabled us to predict the correct gene structures and the regional gene order in both species. This was done by utilizing the blast tools (UCSC, NCBI), but for the human and mouse MKS1 loci the results were also verified experimentally by PCR and RT-PCR.

Ideally we could have identified a novel mouse gene in one of the three syntenic MKS loci, which function and gene expression pattern would have agreed with the MKS phenotype. This type of mouse ortholog would have represented an excellent functional and positional candidate gene for MKS. The identification of a mouse ortholog for the human MKS gene at the MKS2 or MKS3 loci could have provided essential functional information for the identification of the *MKS1* gene on human chromosome 17q. In the most optimal circumstances we could have identified a mouse phenotype mimicking the patient phenotype caused by a mutation in a gene at the mouse syntenic MKS1 locus. In addition, the identification of the conserved non coding sequences could have indicated important regulatory elements in the human MKS1 region. The identification of the conserved regulatory regions was relevant, since no coding region mutations had not been detected in the regional candidates in the patients by the time Original publication II was published (2004). At that time, the finished genomic sequence was already available and the transcript map for the chromosome 17q23 locus was complete.

2.1 Syntenic mouse MKS1 locus

The comparative mapping project was conducted in a collaboration with a research group screening recessive mouse phenotypes which were produced with the point mutagen ENU (N-ethyl-N-nitrosourea) and a balancer chromosome, inversion on mouse chromosome 11, which covered the mouse syntenic MKS1 region (Kile et al. 2003). The mouse chromosome 11C is syntenic to human chromosome 17q23 (Fig. 3).

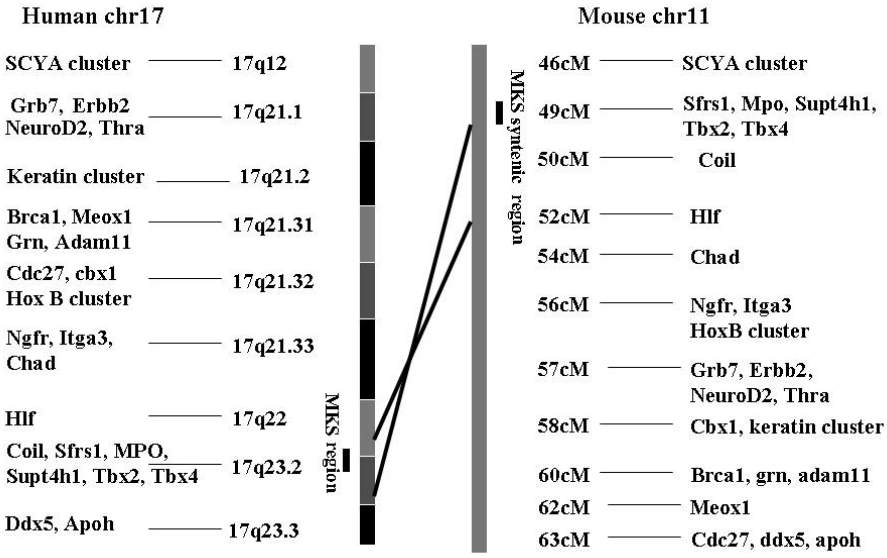

Figure 3: The human and mouse MKS1 critical regions. On the left is the original MKS1 linkage region on human Chr17, spanning 17q12 – 17q23.3. On the right is the corresponding mouse genomic region, from 46 –63 cM on mouse Chr11. Note, the MKS1 critical region is shown as a box next to the human and mouse chromosomes. Original publication II.

We observed only small differences in human and mouse MKS loci on the gene coding regions: The order of *sept4* (P*NUTL*)-gene (NM_004574) and *flj12338* (AK022400) transcript differed between species and the homolog for the putative human transcript *FLJ20264* (AK000271) was missing completely from the mouse MKS syntenic region on chromosome 11C (Fig. 4). The mouse syntenic MKS1 locus was slightly smaller in size (500 kb) compared to the 800 kb MKS1 locus on human chromosome 17q23.

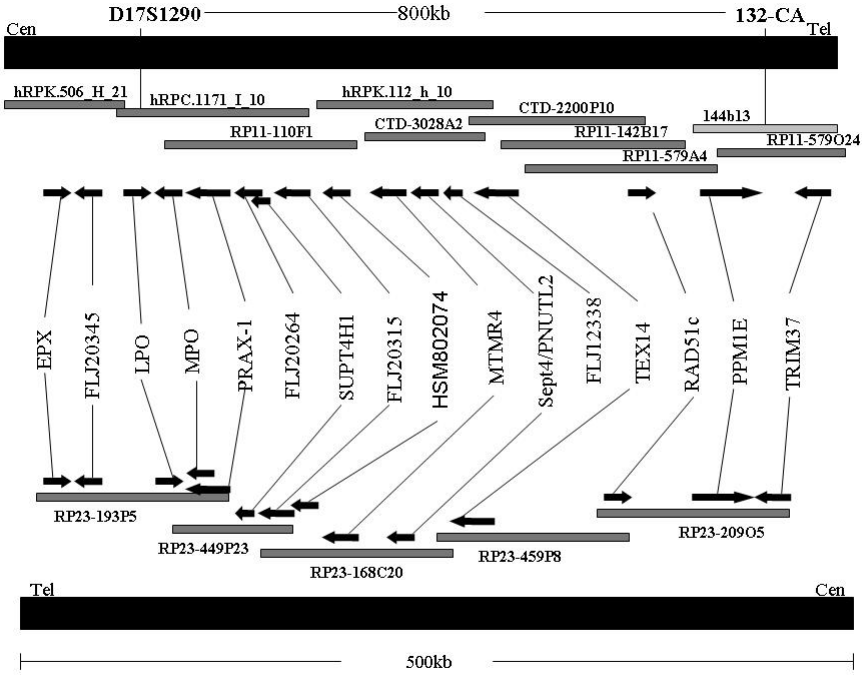

Figure 4. The human MKS1 critical region extends from D17S1290 – 132CA. BAC clones are shown as grey boxes. Genes in the MKS1 interval are shown in the middle. The transcript direction is shown as black arrows. The corresponding mouse genomic region is shown below the human. Mouse BAC clones are shown as gray boxes, and mouse transcript direction is shown as black arrows. The identified *MKS1* gene is the second transcript on the left (*FLJ20345*). Original publication II.

2.2 Discussion of the comparative mapping

The comparative mapping demonstrated high conservation of the syntenic MKS1 regions in human and mouse, both in gene order and content. This supports the idea of mouse being a good animal model for human MKS as a mammal and allowing targeted mutagenesis in its genome.

Unfortunately the ENU-mutagenesis project conducted by our collaborators did not produce phenotypes similar to human MKS, which could have potentially represented a mutation in the mouse *Mks1* gene. This could have provided an

alternative tool for the MKS gene identification in human. The comparative genomics of the two other MKS loci (MKS2 and MKS3) in human (11q and 8q) and mouse (chr7 and chr15) did not either provide any resolving information of the MKS gene on human chromosome 17q.

The MKS3 locus was initially mapped too far to the distal part of chromosome 8q24 (Morgan et al., 2002). This was recently recognized when the *MKS3* gene was identified to be mutated in MKS on 8q22 (Smith et al., 2006). The future studies will show if the MKS2 locus have been assigned correctly to chromosome 11q13 (Roume et al., 1998). The original linkage is often pointing to a vast chromosomal region and a careful analysis of the disease chromosomes with densely distributed markers is required in most cases to establish the minimum chromosomal region for the disease gene. In the case of MKS the locus heterogeneity can cause major disturbances in this type of studies.

3. EXCLUSION OF THE CANDIDATE GENES

3.1 Analysis of a rare sequence variant in the disease chromosomes

The scrutiny of the 800 kb critical MKS region, which was restricted between the markers D17S1290-132-CA on chromosome 17q23.2, demonstrated that it contains total of 13 transcripts (Fig. 4; on page 68: MPO-TRIM37). The direct sequencing of these transcripts in the Finnish MKS patients revealed only non pathogenic polymorphisms and one rare variant in the mRNA variant 4 (NM_080417) of the *SEPT4* (PNUTL)-gene. The variant was bp 594 T→G (NM_080417) on non coding region of the 5′ UTR. This allele was observed to be in complete LD in the 23 unrelated patients (p= 0.00) with the common MKS haplotype on chromosome 17q23. We screened the carrier frequency of this rare variant in over 1000 Finnish control chromosomes and it correlated well with the estimated carrier frequency of Finnish founder mutation (common disease associated haplotype) in the families. We were not able to detect any alterations in the patients' transcript of *SEPT4* gene in the RT-PCR and northern analysis when compared to normal controls.

MatInspector program predicted for the genomic sequence on this particular region a transcription factor binding site for HFH-transcription factor, which was missing in the patients´ sequence with the T→G transition. However, we were not able to confirm this in electro mobility shift assay (EMSA). In the EMSA analysis we utilized nuclear extract from human embryonic kidney cell line (HEK 293), but no specific shift was observed in this analysis. This indicates that the prediction by the MatInspector program might be invalid (Kyttälä, 2003, Masters´ thesis).

Finally, all the 13 transcripts on chromosome 17q23.2: bp 53,686,396-54,468,630 (UCSC May 2004 assembly) were excluded in the candidate gene analysis based on direct sequencing and RT-PCR sequencing and/or northern analysis in patient and control samples (Kyttälä et al., unpublished; Table 1.).

Table 1. The excluded regional transcripts on the marker interval D17S1290- 132-CA (Kyttälä et al., unpublished). The dark colour indicates the methods which have been used for the exclusion of the regional candidate genes. The Transcript *FLJ20264* (Fig. 4) is not included to the table since it exists no longer in Genbank.

Transcript Name; Genbank acc#	Sequencing of the patient DNA		Northern analysis
	Genomic	cDNA	
LPO, NM_006151			
MPO, NM_000250			
PRAX-1, NM_004758			
SUPT4H1, NM_003168			
FJL20315, NM_017763			
HSM802074, BC033020			
MTMR4, NM_004687			
SEPT4, NM_004574			
FJL12338, BC022189			
TEX14, NM_198393			
RAD51c, NM_002876			
PPM1E, NM_014906			
TRIM37, NM_001005207			

3.2 Discussion and the exclusion of the entire initial candidate region (D17S1290-132-CA)

The analysis of the regional candidate genes did not detect any mutations in the Finnish MKS patients. The rare variant in the *SEPT4* gene was excluded as the disease causing mutation. This was based on the observation of normal transcripts in the patients and the functional analysis of the variant. The exclusion was confirmed after additional MKS disease chromosomes with more informative meiotic recombinations were available for us, since the *SEPT4* gene is located outside of the refined MKS locus (see 4.1 Final restriction of MKS1 locus).

However, the rare variant (T/G) of the *SEPT4* gene provided a very specific marker for MKS chromosomes since the allele G was originally detected in complete LD in 23 Finnish families with the common MKS haplotype. The genetic heterogeneity had been established in MKS and we began to use this rare variant as the criterion for Finnish families to have linkage to chromosome 17q23 region. By

analysing only the disease chromosomes of the patients who were either homozygous or heterozygous for the G-allele, we observed that there was only one disease chromosome, which restricted the critical region proximally to the marker D17S1290 (Fig. 5a: The haplotype on the second row from top). The genotypes with the novel markers detected significant LD on proximal region of the marker D17S1290 (Fig. 5a). The analyses of the disease haplotypes in the 26 Finnish MKS families excluded eventually completely the region that was initially restricted between the markers D17S1290-132-CA (Fig. 5).

4. IDENTIFICATION OF THE *MKS1* GENE

4.1 Final restriction of MKS1 locus

The new families which had participated in the study after construction of the original physical map (Fig. 1) demonstrated the deficient restriction of the critical chromosomal region on the proximal site of marker D17S1290. There were no genotype data available from the Finnish disease chromosomes on this region except the original linkage marker D17S1606 (Paavola et al., 1995), which locates proximally 1 cM apart from the D17S1290. Unluckily the allelic mutation in the disease chromosome with D17S1290 (Fig. 5a: the haplotype on the 2nd row) had an impact on the initial disease locus restriction. The locus heterogeneity in Finnish MKS families interfered the haplotype mapping of the MKS1 locus.

Since there were no markers available in the database (UCSC, human genome browser, May 2004) on the near proximal region to the marker D17S1290, we constructed four new polymorphic markers (D17S2261-D17S2264) in this particular region in the close vicinity of D17S1290 by utilizing polymorphic short tandem repeats in the genomic sequence. The genotypes with the novel markers in Finnish disease chromosomes greatly facilitated the final restriction of the chromosome 17q MKS locus to 99 kb (Fig. 5b). This newly assigned MKS critical chromosomal region was restricted between the markers D17S2262-D17S1290 and it contains only five transcripts (UCSC May 2004 assembly).

a)

Marker name	D17S2261	D17S2262	D17S2263	MKS1Δ	D17S2264	D17S1290	D17S2265	272a-CA	272b-CA	132-CA	# of chrs
	244	**245**	**222**	**Δ**	**144**	**187**	**249**	**209**	**92**	**117**	40
	244	**245**	**222**	**Δ**	*146*	*191*	**249**	**209**	**92**	**117**	1
	244	**245**	**222**	**Δ**	*148*	**187**	**249**	**209**	**92**	**117**	1
	238	**245**	**222**	**Δ**	**144**	**187**	**249**	**209**	**92**	**117**	1
	242	**245**	**222**	**Δ**	**144**	**187**	**249**	**209**	**92**	**117**	1
	238	**245**	228	**Δ**	**144**	**187**	**249**	**209**	**92**	**117**	1
	244	**245**	**222**	**Δ**	**144**	**187**	**249**	**209**	**92**	113	1
	246	247	**222**	**Δ**	**144**	191	252	207	84	115	1
	244	**245**	**222**	**Δ**	**144**	191	252	207	94	**117**	1
	244	**245**	**222**	**Δ**	**144**	**187**	252	**209**	94	**117**	1
	244	**245**	**222**	**Δ**	**144**	**187**	252	207	94	115	2
	244	**245**	**222**	**Δ**	**144**	**187**	252	211	94	**117**	1
											tot. 52

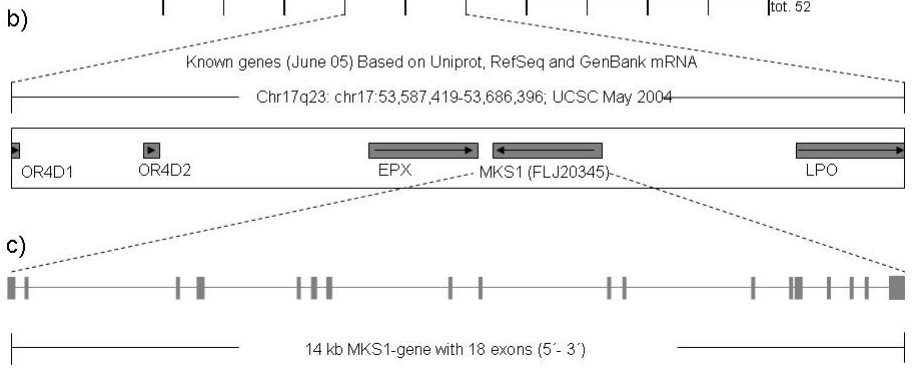

Figure 5.
a) The markers and marker alleles in the disease chromosomes of 26 Finnish unrelated MKS families with confirmed linkage to chr17q23. MKS1Δ marks the 29 bp intronic deletion (Δ) of MKS1 gene. This allele is in complete LD in the disease chromosomes (p= 0.00). The alleles marked with bold font are the same size as alleles which are seen in the full length disease haplotype (the 1st row). The alleles

marked with Italics represent alleles with probable allele mutation in the ancient disease chromosome (haplotypes on the 2nd and 3rd row). Grey color shows recombination haplotypes.
b) The transcript map of the restricted 99 kb MKS1 region on chr17q23. We sequenced the coding regions and exon-intron boundaries of the transcripts in the patients. The only pathogenic sequence alteration was found in the *MKS1* gene.
c) The genomic structure of the *MKS1* gene. The longest open reading frame is coded by the exons 1-18, which is the region between bp: 76-1755, of the 2267 bp long cDNA (DQ185029). The exon 16 is missing from the MKS1 transcript in the Finnish patients homozygous for the 29 bp intronic deletion. This is the founder mutation in Finland (FIN_{major} mutation). The black arrows on the transcripts are pointing the 5´-3´ orientation. Original publication III.

4.2 *MKS1* mutations in patients

The sequencing of the four of the regional (D17S2262-D17S1290) genes revealed only non pathogenic polymorphism in direct sequencing of three Finnish patients and one German patient with strong evidence of linkage to MKS1 locus. The sequencing of the novel gene *MKS1* (*FLJ20345)* detected a 29 bp deletion in intron 15 only four bp apart from the splice acceptor site (Fig. 5c). We sequenced the *MKS1* gene in 20 non-Finnish MKS families available to us and identified disease causing mutations in four of them. In three of the non-Finnish families the affected individuals were homozygous for the Finnish founder mutation. One MKS case of German ancestry was found to be a compound heterozygote for two different mutations in *MKS1*. Both mutations are located at the very beginning of the transcript: a five bp insertion in exon 1 (50insCCGGG) causing a frame shift and a T→C substitution at the splice donor site in intron 1 (IVS1+2T→C) (Table 2).

Table 2. The identified patient mutations in *MKS1* gene. Grey highlighting marks the FIN_{major} mutation. All the patients in the Finnish families were homozygous for this mutation. The FIN_{major} mutation was found in homozygous form also in four non-Finnish families of European descent.

Mutation	Consequence in the gene product	exon	intron
IVS15-7_35del	P470fsX562	-	15
50insCCGGG	P17fsX163	1	-
IVS1+2T→C	splicing	-	1

All the patients in the 26 Finnish families were homozygous for the FIN_{major} mutation and we could not identify other mutations in the sequence analysis of the 14 kb *MKS1* gene in the remaining 14 (30%) Finnish MKS families. However, in these families the common disease haplotype is not observed and the definitive evidence for the linkage to 17q23 is missing in these families with one affected child. This would suggest that mutations in the other two known or in some still uncharacterized MKS loci would explain the molecular background in these families.

The consequence of the FIN_{major} mutation in the patient transcript

To investigate the effect of the intronic deletion on the mRNA level, we analyzed the RT-PCR products from MKS (n= 3) and control fibroblasts (n= 6). A size difference was observed in agarose gel electrophoresis and sequence analysis of the RT-PCR products demonstrated that the patients´ transcript misses 83 bp encompassing exon 16 (Figs 6a and b). To confirm the presence of this exon in various control tissues, we performed PCR analysis of the cDNAs from the fetal multiple tissue panel (BD Biosciences). All the tested tissue samples produced one major product containing exon 16, which was missing from the patients´ transcript. The FIN_{major} mutation is a 29 bp intronic deletion which locates only four bp apart from the splice acceptor site, and most likely grounds to a lacking splice branching site in intron 15, which leads to incorrect spliced patient transcript missing completely exon 16.

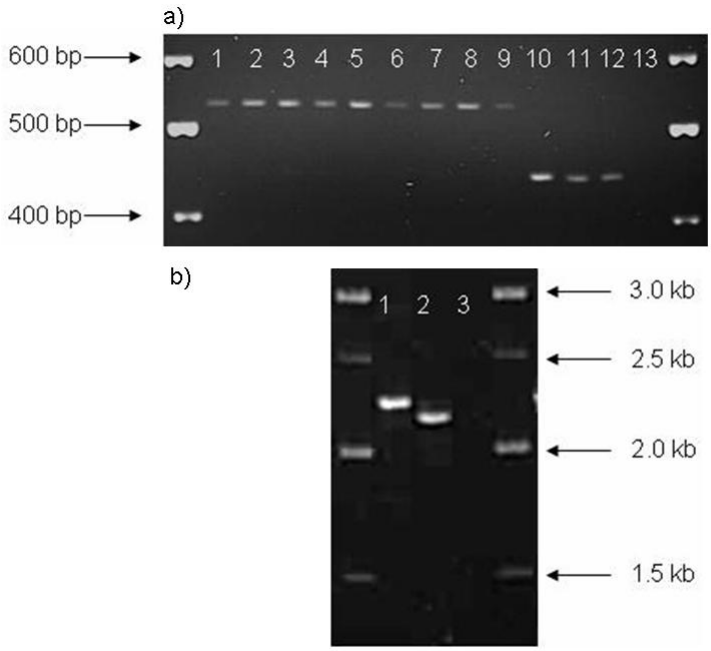

Figure 6.

a) RT-PCR products which are made with primers flanking exon 16 of the *MKS1* gene. On lanes 1-8 are the cDNAs from the embryonic MTC panel (Clontech): 1. brain, 2. lung, 3. liver, 4. kidney, 5. heart, 6. spleen, 7. thymus and 8. skeletal muscle. Lane 9: control embryonic skin fibroblast cDNA. Lanes 10-12 skin fibroblast cDNAs of patients homozygous for the FIN_{major} mutation. The size difference of healthy control transcript and the FIN_{major} homozygous patient is 83 bp, which is the size of exon 16. Lane 13. water.
b) RT-PCR from skin fibroblast RNA across the 18 exons of *MKS1* including the first methionine and STOP codons. Lanes: 1. Control (WT/WT), 2. FIN_{major} homozygous patient 3. Water.

4.3 FIN_{major} -mutation: Age and carrier frequency in Finland

The carrier frequency of the identified 29 bp intronic deletion was found to be 1% among 1082 Finnish control chromosomes, the prevalence well correlating with the estimated 1.5% carrier frequency of the Finnish MKS founder mutation (FIN_{major}). Based on the available MKS samples and Hardy Weinberg (HW) calculations, we would estimate that approximately 70% of the Finnish cases are homozygous for the founder mutation. MKS is enriched in the Finnish population, the prevalence being one in 9000 newborn (Salonen and Norio, 1984), but it does not show geographically regional clustering, unlike most of the diseases belonging to FDH. Not even when dividing the families into two separate groups. Group1: Families with FIN_{major} mutation (70%) and group2: Families with mutation(s) in other gene(s) (30%). The grandparental birthplaces in these both groups reflect geographically the population density rather than showing regional clustering. Also the interval of significant linkage disequilibrium (LD) in Finnish MKS alleles is exceptionally short spanning only some 1 cM region on 17q, the average LD interval in 16 other diseases belonging to the FDH being 6.3 cM (Peltonen et al., 1999). These findings imply that the *MKS1*-FIN_{major} mutation is older than the most FIN_{major} mutations of FDH. Since our mutation detection in *MKS1* gene identified non-Finnish patients of European descent who were homozygous for the FIN_{major} mutation, it is very likely that the *MKS1*-FIN_{major} mutation has been introduced early to the European gene pool and the mutation has entered to the area of Finland with the first inhabitants immigrating from south.

4.4 Discussion of the mutations in the *MKS1* gene

The old age of the FIN_{major}-mutation is supported by the small chromosomal region of significant LD the disease chromosomes and the geographical distribution of the disease throughout the country, opposite to most of the diseases in FDH. Also the identification of this mutation in homozygous form in patients in non-Finnish families of European descent suggests that the mutation has been introduced early to

the European gene pool. The homozygous FIN_{major} mutations result in the same outcome than the compound heterozygous mutations at the very beginning of the gene. A frame shift in exon one and a splice donor site mutation in intron one would support the idea that a FIN_{major} homozygous patient most likely represents a null mutation of *MKS1* lacking completely the gene product. In theory a patient homozygous for FIN_{major} –mutation would have a peptide with frame shift error at position 470 aa (P470fsX562), the size of the wild type peptide being 559 aa. However, we could not detect expression level alterations in the *MKS1* transcript in expression array analysis (Affymetrix Genechip® HG U133 Plus 2.0 Array) in the total RNA samples from patients´ skin fibroblast cell lines when compared to normal controls (Kyttälä et al., unpublished). However, it requires further investigitaion to confirm that this applies to all the tissue types.

5. INITIAL CHRACTERIZATION OF THE NOVEL HUMAN GENE, *MKS1*

The *MKS1* represented a novel human gene with unidentified biological function. It contains an open reading frame (DQ185029: bp 76-1755, full length cDNA 2.3 kb) for a 559 aa polypeptide containing a conserved B9 domain (pfam07162.1). There are two other human proteins containing the conserved B9 domain. One is encoded by the *EPBB9* (AB030506) gene on 17p and one by the *LOC80776* (BC004157) sequence on 19q. However, the biological and molecular properties of the corresponding polypeptides are so far poorly understood. All the three human segments that are coding peptides with B9 domains can be found from the list of the human orthologs for genes coding peptides in the flagellar (cilia in human) basal body proteome (Li et al., 2004). This would propose a role of the less characterized B9 protein domain in the human ciliary basal body proteome.

5.1 *Mks1* gene expression in mouse embryo

The mouse homologue of *MKS1*, the *Mks1* gene locates in mouse syntenic MKS1 locus on chromosome 11qC. This particular DNA region is well conserved between man and mouse both in gene content and order. We confirmed the structure of the previously unidentified mouse *Mks1* gene (DQ177342) by sequencing the cDNA derived from mouse fibroblast cell line. Similarly to the human gene, the mouse gene was found to be composed of 18 exons. Mouse and human coding regions are 86- 88% similar at the nucleotide level and 89% at the amino acid level.

In situ hybridization detected ubiquitous *Mks1* expression in mouse embryonic tissues at embryonic day 15.5. The prominent expression was detected in all the tissues which malformations are characteristic to MKS (Fig. 7). Intriguingly the strongest staining can be seen in the esophageal tract and in bronchioles of the lungs.

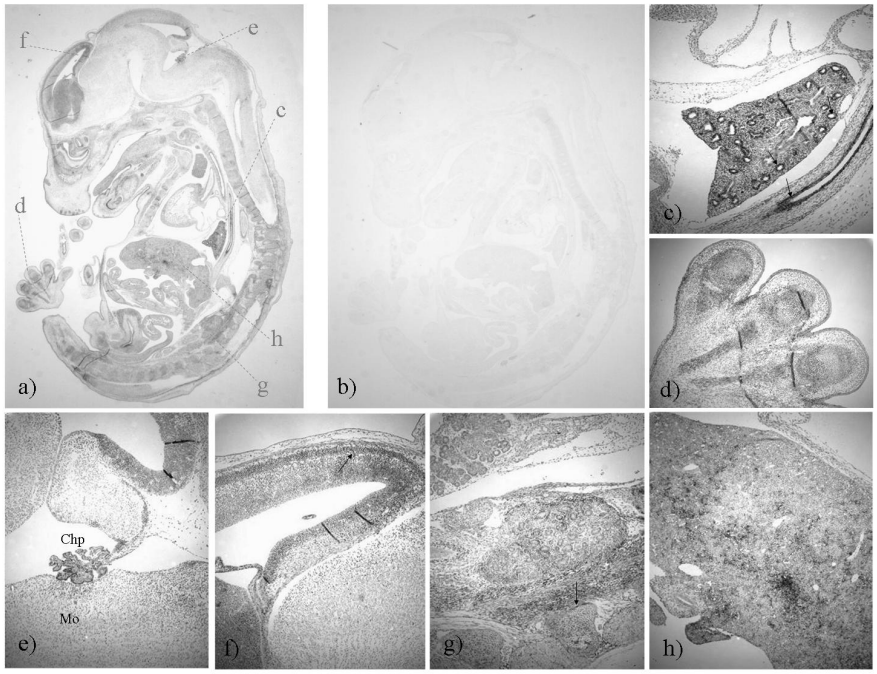

Figure 7. *In situ* hybridization expression profile of *Mks1* in mouse embryo. Sagittal sections of the whole embryo at embryonic day (E) 15.5 with antisense (a) and sense

(b) RNA probes. The expression of *Mks1* is ubiquitous. The expression is prominent in all the tissues which malformations characterize the disease phenotype: brain, kidney, liver and cartilage tissue of developing bone. A strong signal is detected also in the lungs. Higher magnification figures show expression in epithelial cells of the bronchiolus (arrows) and the staining is also prominent in the esophageal epithelium (arrow) **(c);** Front digits. Strong expression is seen in the cartilage tissue of the developing digits **(d);** Selected brain areas. Chp, choroid plexus. Mo, medulla oblongata. **(e)** Prominent signal is also seen from the neopallial cortex (arrow) **(f);** Cortical region of the kidney. Arrow is pointing the position of vertebra **(g)**; Liver **(h).** Original magnifications: (a)-(b) 1x, (c)-(h) 20x.

5.2 MKS1 is conserved among species

Comparison of the MKS1 sequence across the species reveals high conservation (Fig. 8). The MKS1 polypeptide identifies in the reciprocal blastp search only one result in *C. elegans* proteome identifying the xbx-7 gene (R148.1) that contains the X-box promoter element. This element is regulated by the transcription factor daf-19, which is a member of the RFX protein family and known to be required for cilia formation in *C. elegans* (Swoboda et al., 2000; Efimenko et al., 2005). Cilia and flagella are essentially identical organelles and highly conserved among species. A recent study used comparative genomics of human (having cilia), *Chlamydomonas* (having flagella) and *Arabidopsis* (lacking both) to identify genes specific for ciliary proteins and proposed that the *MKS1* sequence would encode a polypeptide that is conserved in the flagellar (cilia in human) and basal body proteome (flagellar apparatus basal body proteome, FABB proteome) (Li et al., 2004; Keller et al., 2005). More specifically, the *Chlamydomonas* ortholog of the *MKS1* gene has been suggested to encode a core structural component of the centriole (Keller et al., 2005). Two other *Chlamydomonas* orthologs in human are suggested to encode centriole components (POC, proteome of centriole). These genes are mutated in OFD1 (oral-facial-digital syndrome, MIM 258860) and NPHP-4 (nephronophthisis-4, MIM_606966). The authors of the proteomic analysis of *Chlamydomonas* centrioles suggest that OFD1 and NPH-4 are basal body or centriole defects which lead to ciliary dysfunction (Keller et al., 2005). Both of these diseases share similar phenotypic features with MKS.

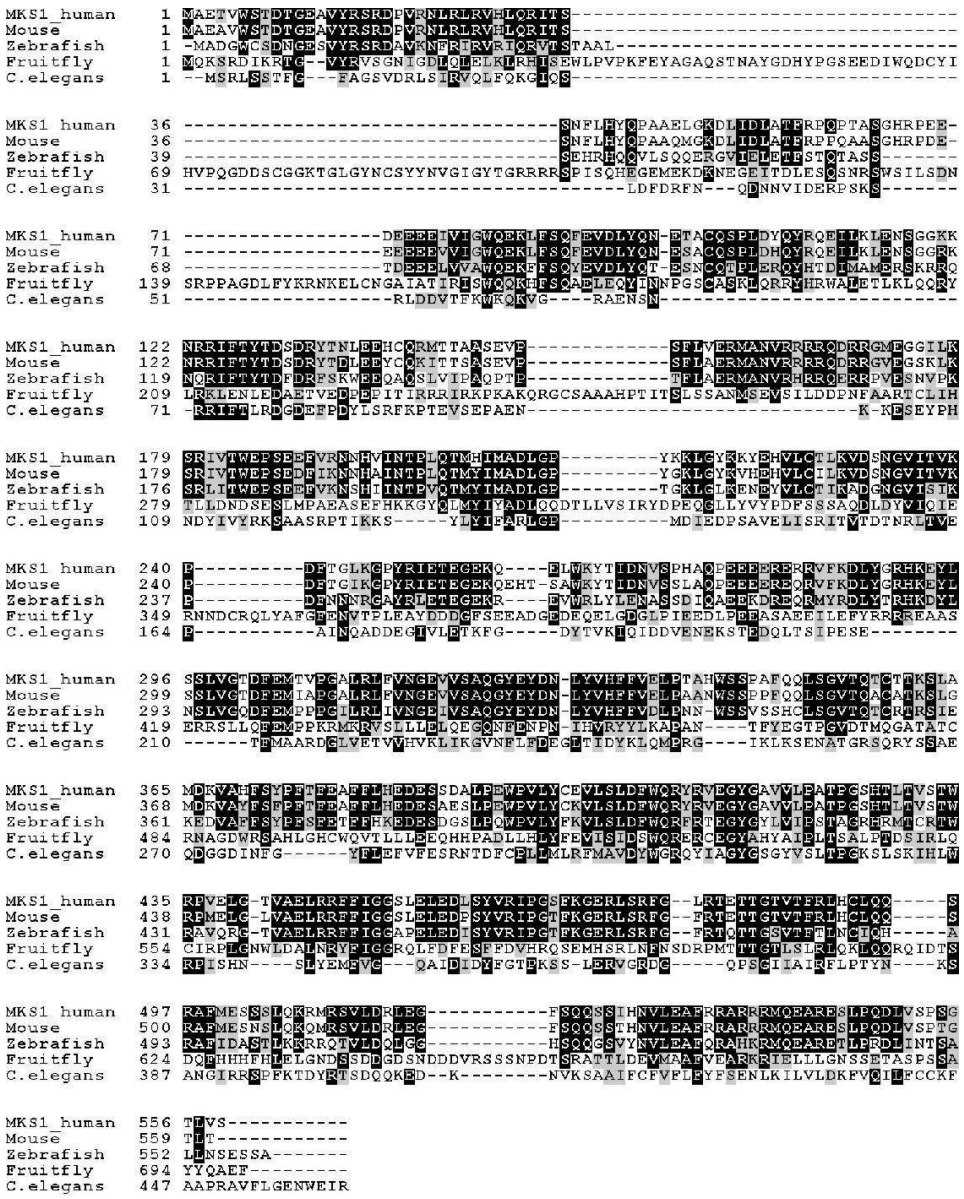

Figure 8. Comparison of human (DQ185029), mouse (CAI25723), zebrafish

(XP_707300), *Drosophila* (NP_572804) and *C. elegans* (NP_497669) MKS1 orthologs using ClustalW1.8 (BCM) followed by Boxshade. The orthologs were identified in reciprocal blastp searches. Original publication III.

Intriguingly, all three human genes (*MKS1*, *LOC80776* and *EPPB9*) encoding polypeptides with B9 domains are listed in the FABB proteome genes and all worm orthologs (R148.1, Y38F2AL.2 and K03E6.4) for the human B9 proteins have X-box promoter elements and are also both proposed to belong to the ciliary proteome based on comparative genomics and their exclusive expression in ciliated cells (Blacque et al., 2005).

5.3 Putative cellular localization of MKS1

Bioinformatics-based analysis (PSORTII) predicts cytoplasmic localization for the deduced MKS1 polypeptide. Our immunofluorescence analysis of transiently transfected HEK293 cells with the GFP-MKS1 agrees with the predicted cytoplasmic location of the peptide. To address the potential cilial localization of MKS1 polypeptide, we compared the in vitro expression pattern of the MKS1 with the immunostaining of polyglutamylated tubulin, a known molecular component of cilia (Bobinnec et al., 1998). Although there are problems associated with transient expression systems for intracellular localization studies, we could observe co-localization of polyglutamylated tubulin and fluorescent MKS1 protein providing further support for the ciliopathy-character of MKS (data not shown).

5.4 Discussion of the initial characterization of MKS1

The identification of the *MKS1* gene exposed a previously unknown human gene, which function is essential for the normal development of many tissue types and organs. MKS1 is conserved among species which has greatly facilitated the initial characterization of its cellular role in ciliary and basal body proteome (FABB proteome) (Fig. 9).

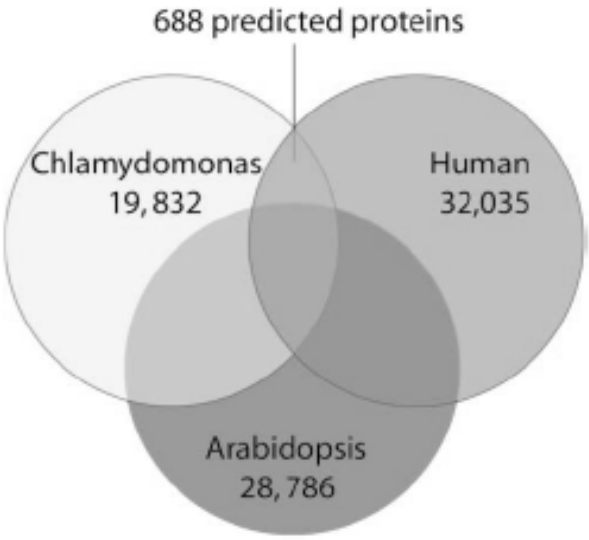

Figure 9. The concept of the FABB proteome: 688 proteins in the flagellar apparatus basal body proteome. Venn diagram of the comparative approach to enrich for flagellar and basal body proteins. The proteome of *Chlamydomonas* (upper left) was compared to the proteome of human (upper right) by WU-BLASTP to find all the matches with a cutoff E value 10^{-10}. These 4 348 matches were then compared to the proteome of Arabidopsis (below) by WU-BLASTP to remove all matches a cutoff E value of 10^{-18}. The 688 proteins remaining are designated the FABB proteome (The middle overlapping section). Li et al., 2004. The human *MKS1 (FLJ20345)* belongs to the FABB proteome.

The 559 aa MKS1 peptide contains a conserved B9 domain, which cellular function is poorly understood. In the human proteome there are two other gene products with B9 domain, and moreover both of them belong to the suggested flagellar (ciliary) basal body proteome. In addition, the *C.elegans´* orthologs for the human genes encoding B9 protein are known to have a binding site for the cilia specific transcription factor daf-19.

We identified and characterized also the genomic structure of the well conserved mouse *Mks1* gene, which is highly similar to human gene, and provides the possibility in future studies for targeted mutagenesis in this good model organism for mammalian organogenesis. RNA *in situ* hybridization of *Mks1* on the

mouse embryonic tissues (embryonic day 15.5) established the expression profile for the gene, which is congruent with the patients´ tissue phenotype. Although the strong staining in the esophageal tract and lungs was highly unexpected. It is known that the epithels of the bronchioles in the lungs and esophageal tract are rich in cilia. Hypoplastic lungs are frequently reported in the patients, but this has generally been thought to be caused by the mechanical pressure of the vast kidneys and lack of amniotic fluid. Another important clinical observation is the increased risk of MKS patients for situs inversus totalis. It is known to be caused by dysfunction of the nodal cilia and thus supports the putative role of MKS1 in ciliogenesis. The cilia are known to be critical in signaling pathways which are involved in the early embryonic development. The genes encoding molecular components for the WNT and Hedgehog pathways were originally suggested to be excellent candidate genes for MKS. The mouse phenotypes with defects in Wnt signaling leads to hepatobiliary tract anomalies with cystic kidneys and respectively the defects in the Hh pathway causes CNS malformation seen together with dysmorphic limb buds (polydactyly). The current data demonstrate a functional structural role of the cilia in both signaling pathways critical for normal embryonic development, which agrees completely with the initial identification of the MKS1 peptide being a putative structural component of the cilia.

6. LOCUS HETEROGENEITY IN FINNISH FAMILIES, MKS4

We found the primary evidence of the further locus heterogeneity by analyzing a Finnish family with two MKS affected children and two healthy siblings. The affected siblings fulfilled the diagnostic criteria and no phenotypic differences were observed in the careful clinical examination to MKS patients linked to other MKS loci. The genealogical study revealed a consanguinity between the maternal grandmothers of the parents: six and seven generations back in the middle of the 18th century their ancestors lived in the same rural municipality and shared a rare family name, which is known to originate from the same household in 90 % of cases.

Originally in this family linkage was excluded to the three known MKS loci with distinct lack of cosegregation of the diseased haplotypes on MKS1, MKS2 and MKS3 loci.

The genomescan with average 10 cM marker distance in our family showing exclusion to known MKS loci produced the maximum two point LOD scores (0.85) in six positions in the genome, which were all novel for MKS: D9S1776 (9q33), D14S280 (14q32), D16S516 (16q23), D17S799 (17p12), D17S798 (17q11), D22S280 (22q12) and D22S283 (22q12). (Fig. 9a). Further the 10K SNP array (Affymetrix) revealed a multipoint LOD score of 3.6 (MAPMAKER/HOMOZ) on 9q33-34 agreeing with the maximum two point LOD score with STR marker D9S1776 (Fig. 9b). No other significant LOD scores were observed elsewhere in the genome in the analysis of the SNP genotypes. The 5 cM homozygous haplotype shared by the affected siblings on 9q was confirmed by genotyping a dense STR marker set covering this region with intermarker distance ≤ 500 kb.

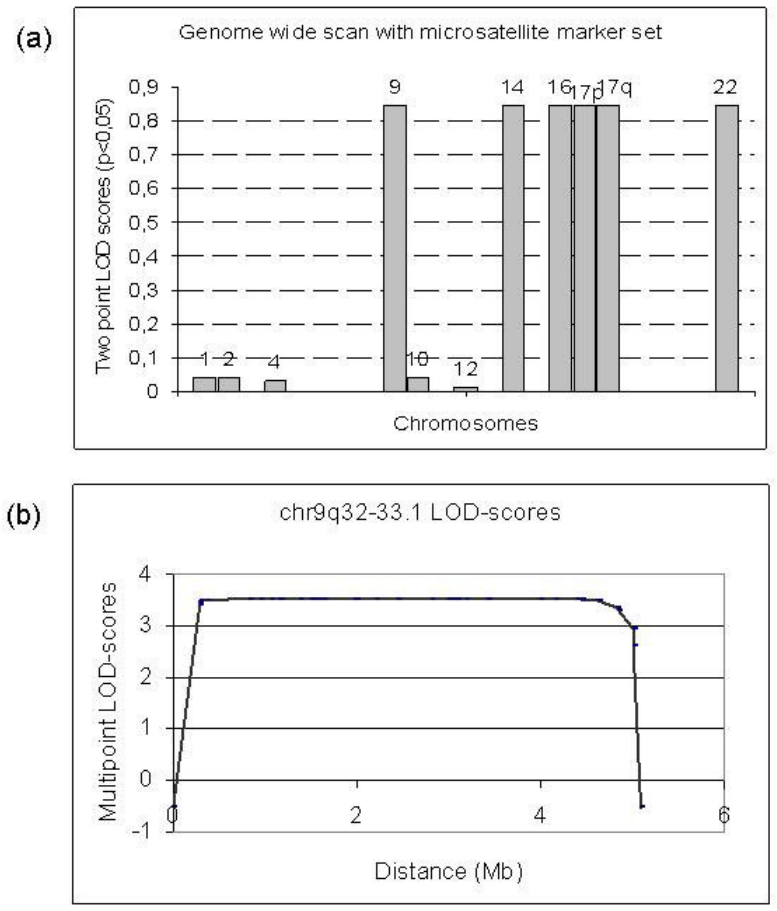

Figure 9. a) Two point LOD scores (MLINK) in the family. The mean intermarker distance is 10 cM in the scan with multiallelic markers. The maximum LOD-score of 0.85 ($p<0.02$) was obtained for six marker loci. None of them overlap any of the three previously established MKS loci. The chromosome 17q maximum LOD-score was obtained with marker D17S798 (17q11) that is distinct to MKS1 locus on 17q23. b) The multipoint LOD scores on chromosome 9q32-33.1 (HOMOZ/MAPMAKER). The microsatellite marker D9S1776 gets assigned on this region. The region of shared homozygosity in the MKS affected fetuses is flanked by the SNP markers: SNP_A-1508357-SNP_A-1510335 (positions on chromosome 9: 113527724-118614930; USCS, May 2004). We confirmed the shared homozygosity on this

chromosomal region by genotyping eight additional multiallelic markers (STR) with intermarker distance of 500 kb. Original publication IV.

6.1 Novel locus on 9q

The observed putative shared homozygosity by the affected fetuses with the 9q32-33.1 SNP genotypes (MAPMAKER/HOMOZ) was confirmed by genotyping eight regional multiallelic markers. The longest region of shared homozygosity can be restricted between markers D9S1824 and SNP_A-1510335, which demarcate a 4.6 Mb region on the basepair positions 113 971 790-118 614 930 (UCSC, May 2004) (Fig. 10).

Interestingly the identified MKS4 locus overlaps a novel Bardet Biedl syndrome locus (BBS10), which has been mapped in a consanguineous Bedouin family. The BBS in this Bedouin family is associated to a missense mutation in the *TRIM32* gene. This finding is supported with additional functional data in a zebrafish model (Chiang et al., personal communication and abstract: ASHG meeting 2005). This gene is also associated to the autosomal recessive limb girdle muscular dystrophy type 2H (LGMD2H) (MIM 254110).

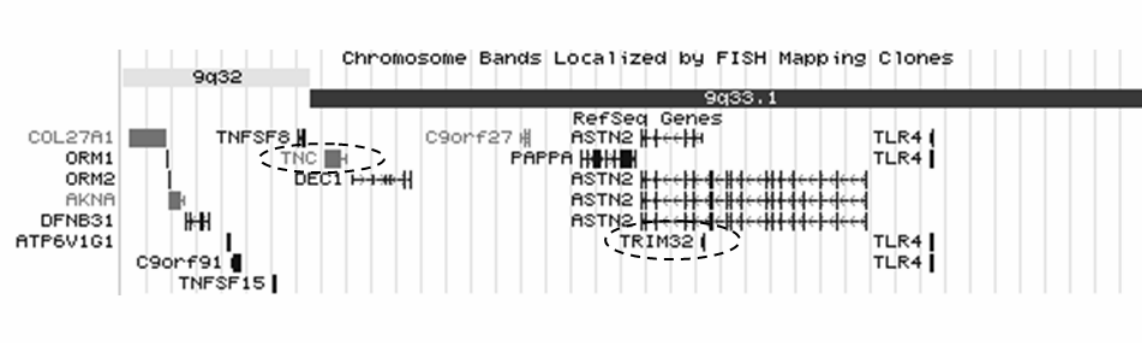

Figure 10. The physical map of the critical chromosomal region on MKS4 locus on 9q32-33.1 (UCSC May 2004; Refseq genes). The sequencing of the circled genes TNC and TRIM32 in the patients did not detect any potential pathogenic sequence alteration in the coding regions.

6.1.1 Candidate gene analysis on MKS4 locus

There are 17 regional candidates on 9q32-33.1. The most prominent candidates are *TNC* and *TRIM32* genes based on the functional information available. *TNC* gene is also a human ortholog for the FABB proteome similarly to the *MKS1* gene. *TRIM32* is associated to Bardet Biedl syndrome (Chiang et al., personal communication, ASHG 2005).

The direct sequencing of *TNC* and *TRIM32* coding regions in the two MKS affected siblings did not detect any potentially pathogenic sequence alterations. *TRIM32* is a gene coded by an intron of *ASTN2* gene (Fig. 10), which is positionally an excellent candidate for MKS4 as well. We are currently analyzing this transcript in patients.

6.2 Discussion of the observed locus heterogeneity and the MKS4 locus

MKS shares characteristic clinical features with other disorders which are known to be caused by ciliary defects. Typical for these disorders are locus heterogeneity and the autosomally recessive inheritance mode. An excellent example being Bardet Biedl syndrome (BBS), which can be caused by mutations in a minimum of ten genes.

We identified the fourth MKS gene locus to chromosome 9q32-33.1 in an inbred Finnish family. It is not known yet if all the 14 Finnish families in which we have not been able to detect coding region mutations in *MKS1* gene, would be linked to chromosome 9q. Unfortunately most of these families are small in size and have only one affected child. Our initial haplotype analyses with multiallelic marker set with intermarkers distance 0.5 kb in these families do not support the idea of a common founder mutation on chromosome 9q (Kyttälä et al., unpublished data). However, if the mutation is very old it is probable in theory that the marker density is not sufficient. The final disease gene identification will confirm this.

The MKS4 locus might provide a genetic link to MKS and BBS since one of the BBS genes (BBS10, Chiang et al., 2005 ASHG personal communication) has been mapped to this particular region, which is at the moment under investigation in the MKS patients.

CONCLUDING REMARKS AND FUTURE PROSPECTS

MKS1 is a novel human gene which mutations cause a severe multisymptomic embryonic developmental disorder, Meckel syndrome (MKS). The current data available from the conserved *MKS1* orthologs as well as from other human diseases and animal models with ciliary defects all imply that MKS represents a novel ciliopathy. There are at least four genes which defects can cause MKS. The characterized human ciliopathies demonstrate locus heterogeneity suggesting that it is a typical feature for the disorders of ciliary origin. The same phenotype can be caused by defects in several molecular components implicated in the ciliogenesis and ciliary functions. To date there are four genetic loci identified in MKS and the future studies should demonstrate the true extent of the genetic heterogeneity. The identification of the disease genes and the extent of the locus heterogeneity will provide more accurate prenatal and preimplantation diagnosis for MKS families. MKS can not be cured, which makes the precise diagnosis important for the families with history of MKS.

The existence of cilia had been almost ignored for decades, but during the past years the research on these cell organelles has shown their implications in many disease pathologies. MKS should serve as an excellent model to understand and study the role of ciliogenesis in the early embryonic development and provide information of its developmental role in very distinct type of tissues. The recent research data from Wnt and Hh signaling pathways have shown that they are both dependent on functional cilia, which would propose that the secondary effect in the cellular pathogenesis of MKS most likely leads to defects in these pathways.

The loss of normal ciliary function in mammals is responsible for cystic and noncystic pathology in the kidney, liver, brain, and pancreas, as well as in severe developmental patterning abnormalities. However, the physiological role of cilia in most tissues remains obscure. The gene defects underlying in Bardet Biedl syndrome have proposed that ciliary functions may have roles also in more common human disorders such as hypertension, obesity and diabetes, thus making the

research on basal body and cilia highly meaningful also in this aspect. The *MKS1* gene will most likely provide general information of the ciliary functions that can be also essential in understanding these common human conditions currently under growing interest of pharmaceutical research to provide more specific and efficient therapeutic medication for these epidemics in the modern world. Therefore the main challenges in the future functional studies will be to provide the information about the specific functional role of MKS1 and the other MKS gene products in basal body and ciliary proteome. They are critical for the normal function of this cell organelle. The identification of the *MKS1* gene proposes that the research on MKS genes should offer novel molecular level information for better understanding the physiological role of cilia in many tissue types in which it has not yet been recognized.

ACKNOWLEDGEMENTS

This study was carried out in the department of Molecular Medicine, National Public Health Institute (August 1998-April 1999; September 2002-March 2006) and department of Human Genetics, University of California Los Angeles (UCLA), David Geffen School of medicine (May 1999-August 2002). I wish to express my gratitude to the former and the present Heads of the Institute: Jussi Huttunen and Pekka Puska as well as to Anu Jalanko the head of the Molecular Medicine at NPHI for providing the excellent research facilities.

I thank my supervisor Leena Peltonen Palotie for providing the opportunity to work with an interesting and challenging project. I am especially grateful for all the time she has always found from her extremely busy schedule for advising even with the smallest details. In the fall 2002 I returned back to Helsinki from UCLA and my second supervisor Marjo Kestilä joined the project. I wish to thank Marjo for all the advices, help and her admirable patience. This type of study requires a clinical specialist as well. I wish to thank Riitta Salonen for her clinical expertise and being the third unofficial mentor to me.

I wish to warmly thank Brendan Lee for accepting the role as the opponent in my thesis defense. The reviewers of my thesis Elisabeth Widen and Kirsi Huoponen are thanked for reviewing this thesis in a short time and for their valuable comments.

Paulina Paavola-Sakki is thanked for guiding me into the secrets of the molecular genetics of MKS and also for answering to all the questions during the years after she has been already working elsewhere. Kristiina Avela is thanked for collaboration in the physical map project and also for all the advices and conversations. Kathryn Hentges is thanked for the collaboration in the comparative mapping project. Jouni Vesa is thanked for advising in the EMSA method. Hilde M.F.R. Stensland is thanked for her help, advises and commenting the thesis. Juha

Saharinen is thanked for his expertise in bioinformatics and help with the multipoint analysis of the SNP haplotypes. Massimiliano Gentile is thanked for all his advices in the expression array analysis and for analyzing the data. Teppo Varilo is thanked for his expertise in the genealogical studies. Outi Kopra is thanked for her advices with the *in situ* hybridization.

Aija Kyttälä is thanked for introducing the lab to me for the first time in the spring 1996 when I did my graduation work to the laboratory school. Aija, Nina Aula and Laura Ahtiainen are also thanked for their kind help with the immunofluorescence microscope. The sequencing core is thanked for the excellent service: Pekka, Heli, Susanna and Tiina H. Jari and Juri are thanked for their help with computer related problems. Ritva, Paula and Katta are thanked for their excellent technical assistance. Sari K., Sari M. and Tuija are thanked for their kind help with several practical matters. Jonna, I am glad that you are going to continue the project – I wish you best of luck for your research on this fascinating project.

The years at UCLA were fantastic. Kathrin, Mark, Lisa, Ali, Jenny Lee, Atila, Tina, Kathy N., Chris R., Tutsa, Tuula, Lennu, Jenny E., Heidi L., Kisse, Niklas, Leila, Juha, Petra, Outi, Aune, Jouni, Päivi, Jaana, Mikko, Tero H., Pauliina, Anne J, Anne , Ed, Denis, Tony, Brian, Mervi, Maikki, Greg, Daniel, Janna, Elina S. and Joni are thanked for all their help and support in the lab as well as for their company on the leisure time (there was some!) in the sunny LA. Special thanks to "LA tennis gang": Joyce, Linda, Jia, Hilde and Manny. The tennis girls are also thanked for all the nice trips, hikes and dinners together. I miss doubles games on Sunday mornings in Sunset Canyon.

In the NPHI Helsinki there are many great people who I wish to thank for many fun moments and also for all the help and support during the years: Emma P., Will, Marika, Heidi M., Liisa, Juho, Sisko, Minna, Olli, Minttu, Anu L., Iiris, Kaisa, Kaisu L., Kaisu K., Anna K., Mikko K., Heli R., Nabil, Karola, Tero Y., Heli H., Annika, Tintti, Annina, Elli, Tiina P., Iski, Annina, Laura, Kristiina, Kaija, Anne N., Anne V., Arja T., Jussi, Johannes, Tuuli, Elina P., Sampo, Heidi N., Auli, Outi T., Minttu, Arja T., Suvi, Tuuli, Emma N., Mervi, Kati and Kirsi. I have probably forgotten some important names –my apologies for this.

.

I thank all the MKS families who have participated in the study. You have made this study possible and I wish that these results will benefit at least some of you.

I thank my family for their support –it is way too seldom when I express my appreciation for all the things you have done for me. I wish to thank my friends Mikaela, Tuuli, Satu and Jokke, as well as my friends from the University Chemistry department Henriikka, Noora and other chemists and former chemists.

This thesis has been financially supported by Finnish cultural foundation, Research Funds of the University of Helsinki and Biomedicum Helsinki foundation.

Helsinki 6th of March 2006.

Mira Kyttälä

REFERENCES

Aaltonen J, Horelli-Kuitunen N, Fan JB, Bjorses P, Perheentupa J, Myers R, Palotie A, Peltonen L. High-resolution physical and transcriptional mapping of the autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy locus on chromosome 21q22.3 by FISH. Genome Res. 1997;7(8):820-9.

Aalto-Setala K, Helve E, Kovanen PT, Kontula K. Finnish type of low density lipoprotein receptor gene mutation (FH-Helsinki) deletes exons encoding the carboxy-terminal part of the receptor and creates an internalization-defective phenotype. J Clin Invest. 1989;84(2):499-505.

Afzelius BA. The immotile-cilia syndrome: a microtubule-associated defect. CRC Crit Rev Biochem. 1985;19(1):63-87.

Ahdab-Barmada M, Claassen D. A distinctive triad of malformations of the central nervous system in the Meckel-Gruber syndrome.J Neuropathol Exp Neurol. 1990;49(6):610-20.

Aittomaki K, Lucena JL, Pakarinen P, Sistonen P, Tapanainen J, Gromoll J, Kaskikari R, Sankila EM, Lehvaslaiho H, Engel AR, Nieschlag E, Huhtaniemi I, de la Chapelle A. Mutation in the follicle-stimulating hormone receptor gene causes hereditary hypergonadotropic ovarian failure. Cell. 1995;82(6):959-68.

Aleksic S, Budzilovich G, Greco MA, Reuben R, Feigin I, Pearson J, Epstein F. Cerebellocele and associated central nervous system anomalies in the Meckel syndrome.Childs Brain. 1984;11(2):99-111.

Aminoff M, Carter JE, Chadwick RB, Johnson C, Grasbeck R, Abdelaal MA, Broch H, Jenner LB, Verroust PJ, Moestrup SK, de la Chapelle A, Krahe R. Mutations in CUBN, encoding the intrinsic factor-vitamin B12 receptor, cubilin, cause hereditary megaloblastic anaemia 1. Nat Genet. 1999;21(3):309-13.

Anderson VM. Meckel syndrome: morphologic considerations. Birth Defects Orig Artic Ser. 1982;18(3B):145-60.

Andrews NC, Faller DV. A rapid micropreparation technique for extraction of DNA-binding proteins from limiting numbers of mammalian cells. Nucleic Acids Res. 1991;19(9):2499.

Ansley SJ, Badano JL, Blacque OE, Hill J, Hoskins BE, Leitch CC, Kim JC, Ross AJ, Eichers ER, Teslovich TM, Mah AK, Johnsen RC, Cavender JC, Lewis RA, Leroux MR, Beales PL, Katsanis N. Basal body dysfunction is a likely cause of pleiotropic Bardet-Biedl syndrome. Nature. 2003;425(6958):628-33.

Aula P, Autio S, Raivio KO, Rapola J, Thoden CJ, Koskela SL, Yamashina I. "Salla disease": a new lysosomal storage disorder. Arch Neurol. 1979;36(2):88-94.

Avela K, Lipsanen-Nyman M, Perheentupa J, Wallgren-Pettersson C, Marchand S, Faure S, Sistonen P, de la Chapelle A, Lehesjoki AE. Assignment of the mulibrey nanism gene to 17q by linkage and linkage-disequilibrium analysis. Am J Hum Genet. 1997;60(4):896-902.

Avela K, Lipsanen-Nyman M, Idanheimo N, Seemanova E, Rosengren S, Makela TP, Perheentupa J, Chapelle AD, Lehesjoki AE. Gene encoding a new RING-B-box-Coiled-coil protein is mutated in mulibrey nanism. Nat Genet. 2000;25(3):298-301.

Avidor-Reiss T, Maer AM, Koundakjian E, Polyanovsky A, Keil T, Subramaniam S, Zuker CS. Decoding cilia function: defining specialized genes required for compartmentalized cilia biogenesis. Cell. 2004;117(4):527-39.

Beales PL. Lifting the lid on Pandora's box: the Bardet-Biedl syndrome. Curr Opin Genet Dev. 2005;15(3):315-23.

Blacque OE, Perens EA, Boroevich KA, Inglis PN, Li C, Warner A, Khattra J, Holt RA, Ou G, Mah AK, McKay SJ, Huang P, Swoboda P, Jones SJ, Marra MA, Baillie DL, Moerman DG, Shaham S, Leroux MR. Functional genomics of the cilium, a sensory organelle. Curr Biol. 2005:15(10):935-41.

Blankenberg TA, Ruebner BH, Ellis WG, Bernstein J, Dimmick JE. Pathology of renal and hepatic anomalies in Meckel syndrome. Am J Med Genet Suppl. 1987;3:395-410.

Borsani G, Bassi MT, Sperandeo MP, De Grandi A, Buoninconti A, Riboni M, Manzoni M, Incerti B, Pepe A, Andria G, Ballabio A, Sebastio G. SLC7A7, encoding a putative permease-related protein, is mutated in patients with lysinuric protein intolerance. Nat Genet. 1999;21(3):297-301.

Botstein D, White RL, Skolnick M, Davis RW. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. Am J Hum Genet. 1980;32(3):314-31.

Brailov I, Bancila M, Brisorgueil MJ, Miquel MC, Hamon M, Verge D. Localization of 5-HT(6) receptors at the plasma membrane of neuronal cilia in the rat brain. Brain Res. 2000 28;872(1-2):271-5.

Bray-Ward P, Menninger J, Lieman J, Desai T, Mokady N, Banks A, Ward DC. Integration of the cytogenetic, genetic, and physical maps of the human genome by FISH mapping of CEPH YAC clones.Genomics. 1996;32(1):1-14.

Breitschopf H, Suchanek G, Gould RM, Colman DR, Lassmann H. In situ hybridization with digoxigenin-labeled probes: sensitive and reliable detection method applied to myelinating rat brain. Acta Neuropathol (Berl). 1992;84(6):581-7.

Brody SL, Yan XH, Wuerffel MK, Song SK, Shapiro SD. Ciliogenesis and left-right axis defects in forkhead factor HFH-4-null mice. Am J Respir Cell Mol Biol. 2000;23(1):45-51.

Burchell B. Turning on and turning off the sense of smell. Nature. 1991;350(6313):16-7.

Burge C, Karlin S. Prediction of complete gene structures in human genomic DNA.
J Mol Biol. 1997;268(1):78-94.

Burke DT, Carle GF, Olson MV. Cloning of large segments of exogenous DNA into yeast by means of artificial chromosome vectors. Science. 1987;236(4803):806-12.

Burnette WN. "Western blotting": electrophoretic transfer of proteins from sodium dodecyl sulfate--polyacrylamide gels to unmodified nitrocellulose and radiographic detection with antibody and radioiodinated protein A. Anal Biochem. 1981;112(2):195-203.

Bobinnec Y, Moudjou M, Fouquet JP, Desbruyeres E, Edde B, Bornens M. Glutamylation of centriole and cytoplasmic tubulin in proliferating non-neuronal cells. Cell Motil Cytoskeleton. 1998;39(3):223-32.

Cadigan KM, Nusse R. Wnt signaling: a common theme in animal development. Genes Dev. 1997;11(24):3286-305.

Cano DA, Murcia NS, Pazour GJ, Hebrok M. Orpk mouse model of polycystic kidney disease reveals essential role of primary cilia in pancreatic tissue organization. Development. 2004;131(14):3457-67.

Cantor CR. How will the Human Genome Project improve our quality of life? Nat Biotechnol. 1998;16(3):212-3.

Capdevila J, Johnson RL. Hedgehog signaling in vertebrate and invertebrate limb patterning.Cell Mol Life Sci. 2000 ;57(12):1682-94. Review.

Cavalli-Sforza LL, Piazza A. Human genomic diversity in Europe: a summary of recent research and prospects for the future. Eur J Hum Genet. 1993;1(1):3-18.

Chelly J, Concordet JP, Kaplan JC, Kahn A. Illegitimate transcription: transcription of any gene in any cell type. Proc Natl Acad Sci U S A. 1989;86(8):2617-21.

Chen J, Knowles HJ, Hebert JL, Hackett BP. Mutation of the mouse hepatocyte nuclear factor/forkhead homologue 4 gene results in an absence of cilia and random left-right asymmetry. J Clin Invest. 1998;102(6):1077-82.

Chiang C, Litingtung Y, Lee E, Young KE, Corden JL, Westphal H, Beachy PA. Cyclopia and defective axial patterning in mice lacking Sonic hedgehog gene function.Nature. 1996;383(6599):407-13.

Chumakov IM, Rigault P, Le Gall I, Bellanne-Chantelot C, Billault A, Guillou S, Soularue P, Guasconi G, Poullier E, Gros I, et al. A YAC contig map of the human genome. Nature. 1995;377(6547 Suppl):175-297.

Cohen D, Chumakov I, Weissenbach J. A first-generation physical map of the human genome. Nature. 1993;366(6456):698-701.

Cole DG, Diener DR, Himelblau AL, Beech PL, Fuster JC, Rosenbaum JL. Chlamydomonas kinesin-II-dependent intraflagellar transport (IFT): IFT particles contain proteins required for ciliary assembly in Caenorhabditis elegans sensory neurons. J Cell Biol. 1998;141(4):993-1008.

Collins FS, Morgan M, Patrinos A. The Human Genome Project: lessons from large-scale biology. Science. 2003;300(5617):286-90.

Cormand B, Avela K, Pihko H, Santavuori P, Talim B, Topaloglu H, de la Chapelle A, Lehesjoki AE. Assignment of the muscle-eye-brain disease gene to 1p32-p34 by linkage analysis and homozygosity mapping. Am J Hum Genet. 1999;64(1):126-35.

Crawfurd MdA, Jackson P, Kohler HG. Meckel's syndrome (dysencephalia splanchno-cystica) in two Pakistani sibs. J Med Genet. 1978;15(3):242-5.

de la Chapelle A. Disease gene mapping in isolated human populations: the example of Finland. J Med Genet. 1993;30(10):857-65.

de la Chapelle A, Wright FA. Linkage disequilibrium mapping in isolated populations: the example of Finland revisited. Proc Natl Acad Sci U S A. 1998;95(21):12416-23.

Dib C, Faure S, Fizames C, Samson D, Drouot N, Vignal A, Millasseau P, Marc S, Hazan J, Seboun E, Lathrop M, Gyapay G, Morissette J, Weissenbach J. A comprehensive genetic map of the human genome based on 5,264 microsatellites. Nature. 1996;380(6570):152-4.

Diesen C, Saarinen A, Pihko H, Rosenlew C, Cormand B, Dobyns WB, Dieguez J, Valanne L, Joensuu T, Lehesjoki AE. POMGnT1 mutation and phenotypic spectrum in muscle-eye-brain disease. J Med Genet. 2004;41(10):e115

Drake JW, Charlesworth B, Charlesworth D, Crow JF. Rates of spontaneous mutation. Genetics. 1998;148(4):1667-86.

Efimenko E, Bubb K, Mak HY, Holzman T, Leroux MR, Ruvkun G, Thomas JH, Swoboda P. Analysis of xbx genes in C. elegans. Development. 2005;132(8):1923-34.

Ehler LL, Holmes JA, Dutcher SK. Loss of spatial control of the mitotic spindle apparatus in a Chlamydomonas reinhardtii mutant strain lacking basal bodies. Genetics. 1995;141(3):945-60.

Engle SJ, Womer DE, Davies PM, Boivin G, Sahota A, Simmonds HA, Stambrook PJ, Tischfield JA. HPRT-APRT-deficient mice are not a model for lesch-nyhan syndrome. Hum Mol Genet. 1996;5(10):1607-10.

Essner JJ, Vogan KJ, Wagner MK, Tabin CJ, Yost HJ, Brueckner M. Conserved function for embryonic nodal cilia. Nature. 2002;418(6893):37-8.

Fraser FC, Lytwyn A. Spectrum of anomalies in the Meckel syndrome, or: "Maybe there is a malformation syndrome with at least one constant anomaly". Am J Med Genet. 1981;9(1):67-73.

Fried K. Relatively high prevalence of the Meckel syndrome among Jews.Isr J Med Sci. 1973;9(9):1399.

Germino GG. Linking cilia to Wnts. Nat Genet. 2005 ;37(5):455-7

Gitschier J, Wood WI, Goralka TM, Wion KL, Chen EY, Eaton DH, Vehar GA, Capon DJ, Lawn RM. Characterization of the human factor VIII gene. Nature. 1984;312(5992):326-30.

Green ED, Olson MV. Systematic screening of yeast artificial-chromosome libraries by use of the polymerase chain reaction. Proc Natl Acad Sci U S A. 1990;87(3):1213-7.

Gruber BG. Beiträge zur frage “gekoppelter” missbildungen. (Acrocephalo-Syndactylie und Dysencephalia splanchocystica). Beitrage zur Pathologishchen Anatomie. 1936;93:459-476.

Hackman P, Vihola A, Haravuori H, Marchand S, Sarparanta J, De Seze J, Labeit S, Witt C, Peltonen L, Richard I, Udd B. Tibial muscular dystrophy is a titinopathy caused by mutations in TTN, the gene encoding the giant skeletal-muscle protein titin. Am J Hum Genet. 2002;71(3):492-500.

Handel M, Schulz S, Stanarius A, Schreff M, Erdtmann-Vourliotis M, Schmidt H, Wolf G, Hollt V. Selective targeting of somatostatin receptor 3 to neuronal cilia. Neuroscience. 1999;89(3):909-26.

Haque KN, Zaidi M.Dysecephalia splanchnocystica (Meckel´s syndrome). Report of two cases and review of the literature. Saudi Medical journal. 1981;2:61-64.

Harris NL. Genotator: a workbench for sequence annotation.Genome Res. 1997;7(7):754-62.

Heiskanen, M., R. Karhu, E. Hellsten, L. Peltonen, O. Kallioniemi, and A. Palotie. High resolution mapping using fluorescence in situ hybridization (FISH) to extended DNA fibers prepared from agarose-embedded cells. BioTechniques. 1994;17: 928-934

Heiskanen, M., O.-P. Kallioniemi, and A. Palotie. Fiber-FISH: Experiences and a refined protocol. 1996;Genet. Anal. Biomol. Eng. 12: 179-184.

Hoglund P, Haila S, Socha J, Tomaszewski L, Saarialho-Kere U, Karjalainen-Lindsberg ML, Airola K, Holmberg C, de la Chapelle A, Kere J. Mutations of the Down-regulated in adenoma (DRA) gene cause congenital chloride diarrhoea. Nat Genet. 1996;14(3):316-9.

Holmes LB, Driscoll SG, Atkins L. Etiologic heterogeneity of neural-tube defects. N Engl J Med. 1976;294(7):365-9.

Hsia YE, Bratu M, Herbordt A. Genetics of the Meckel syndrome (dysencephalia splanchnocystica). Pediatrics. 1971;48(2):237-47.

Huangfu D, Liu A, Rakeman AS, Murcia NS, Niswander L, Anderson KV. Hedgehog signalling in the mouse requires intraflagellar transport proteins. Nature. 2003;426(6962):83-7.

Huangfu D, Anderson KV. Cilia and Hedgehog responsiveness in the mouse. Proc Natl Acad Sci U S A. 2005;102(32):11325-30.

Hudson, TJ, LD Stein, SS Gerety, J Ma, AB Castle, J Silva, DK Slonim, R Baptista, L Kruglyak, SH Xu An STS-based map of the human genomeScience 1995;270(5244):1945-54.

Hästbacka J, de la Chapelle A, Kaitila I, Sistonen P, Weaver A, Lander E. Linkage disequilibrium mapping in isolated founder populations: diastrophic dysplasia in Finland.Nat Genet. 1992;2(3):204-11. Erratum in: Nat Genet 1992;2(4):343.

Hästbacka J, de la Chapelle A, Mahtani MM, Clines G, Reeve-Daly MP, Daly M, Hamilton BA, Kusumi K, Trivedi B, Weaver A, et al. The diastrophic dysplasia gene encodes a novel sulfate transporter: positional cloning by fine-structure linkage disequilibrium mapping.Cell. 1994 23;78(6):1073-87.

Ingham PW, McMahon AP. Hedgehog signaling in animal development: paradigms and principles. Genes Dev. 2001;15(23):3059-87.

Ioannou PA, Amemiya CT, Garnes J, Kroisel PM, Shizuya H, Chen C, Batzer MA, de Jong PJ. A new bacteriophage P1-derived vector for the propagation of large human DNA fragments. Nat Genet. 1994;6(1):84-9.

Ikonen E, Baumann M, Gron K, Syvanen AC, Enomaa N, Halila R, Aula P, Peltonen L. Aspartylglucosaminuria: cDNA encoding human aspartylglucosaminidase and the missense mutation causing the disease. EMBO J. 1991;10(1):51-8.

Jarvela I, Sabri Enattah N, Kokkonen J, Varilo T, Savilahti E, Peltonen L. Assignment of the locus for congenital lactase deficiency to 2q21, in the vicinity of but separate from the lactase-phlorizin hydrolase gene. Am J Hum Genet. 1998;63(4):1078-85.

Jenny A, Reynolds-Kenneally J, Das G, Burnett M, Mlodzik M. Diego and Prickle regulate Frizzled planar cell polarity signalling by competing for Dishevelled binding. Nat Cell Biol. 2005;7(7):691-7.

Joensuu T, Hamalainen R, Yuan B, Johnson C, Tegelberg S, Gasparini P, Zelante L, Pirvola U, Pakarinen L, Lehesjoki AE, de la Chapelle A, Sankila EM. Mutations in a novel gene with transmembrane domains underlie Usher syndrome type 3. Am J Hum Genet. 2001;69(4):673-84. Erratum in: Am J Hum Genet 2001;69(5):1160.

Karmous-Benailly H, Martinovic J, Gubler MC, Sirot Y, Clech L, Ozilou C, Auge J, Brahimi N, Etchevers H, Detrait E, Esculpavit C, Audollent S, Goudefroye G, Gonzales M, Tantau J, Loget P, Joubert M, Gaillard D, Jeanne-Pasquier C, Delezoide AL, Peter MO, Plessis G, Simon-Bouy B, Dollfus H, Le Merrer M, Munnich A, Encha-Razavi F, Vekemans M, Attie-Bitach T. Antenatal presentation of Bardet-Biedl syndrome may mimic Meckel syndrome. Am J Hum Genet. 2005;76(3):493-504.

Keller LC, Romijn EP, Zamora I, Yates JR 3rd, Marshall WF. Proteomic analysis of isolated chlamydomonas centrioles reveals orthologs of ciliary-disease genes. Curr Biol. 2005;15(12):1090-8.

Kestila M, Lenkkeri U, Mannikko M, Lamerdin J, McCready P, Putaala H, Ruotsalainen V, Morita T, Nissinen M, Herva R, Kashtan CE, Peltonen L, Holmberg C, Olsen A, Tryggvason K. Positionally cloned gene for a novel glomerular protein--nephrin--is mutated in congenital nephrotic syndrome. Mol Cell. 1998;1(4):575-82.

Kittles RA, Perola M, Peltonen L, Bergen AW, Aragon RA, Virkkunen M, Linnoila M, Goldman D, Long JC. Dual origins of Finns revealed by Y chromosome haplotype variation. Am J Hum Genet. 1998;62(5):1171-9.

Klockars T, Savukoski M, Isosomppi J, Laan M, Jarvela I, Petrukhin K, Palotie A, Peltonen L. Efficient construction of a physical map by fiber-FISH of the CLN5 region: refined assignment and long-range contig covering the critical region on 13q22.Genomics. 1996;35(1):71-8.

Koivisto UM, Turtola H, Aalto-Setala K, Top B, Frants RR, Kovanen PT, Syvanen AC, Kontula K. The familial hypercholesterolemia (FH)-North Karelia mutation of the low density lipoprotein receptor gene deletes seven nucleotides of exon 6 and is a common cause of FH in Finland. J Clin Invest. 1992;90(1):219-28.

Kolehmainen J, Black GC, Saarinen A, Chandler K, Clayton-Smith J, Traskelin AL, Perveen R, Kivitie-Kallio S, Norio R, Warburg M, Fryns JP, de la Chapelle A, Lehesjoki AE. Cohen syndrome is caused by mutations in a novel gene, COH1, encoding a transmembrane protein with a presumed role in vesicle-mediated sorting and intracellular protein transport. Am J Hum Genet. 2003;72(6):1359-69.

Kompanje EJ. Features described and illustrated in 1684 suggesting Meckel-Gruber syndrome. Pediatr Dev Pathol. 2003;6(6):595-8.

Kozminski KG, Johnson KA, Forscher P, Rosenbaum JL. A motility in the eukaryotic flagellum unrelated to flagellar beating. Proc Natl Acad Sci U S A. 1993;90(12):5519-23.

Kramer-Zucker AG, Olale F, Haycraft CJ, Yoder BK, Schier AF, Drummond IA. Cilia-driven fluid flow in the zebrafish pronephros, brain and Kupffer's vesicle is required for normal organogenesis. Development. 2005;132(8):1907-21.

Kruglyak L, Daly MJ, Lander ES. Rapid multipoint linkage analysis of recessive traits in nuclear families, including homozygosity mapping. Am J Hum Genet. 1995;56(2):519-27.

Kruglyak L, Daly MJ, Reeve-Daly MP, Lander ES. Parametric and nonparametric linkage analysis: a unified multipoint approach. Am J Hum Genet. 1996;58(6):1347-63.

Kruglyak L. The use of a genetic map of biallelic markers in linkage studies. Nat Genet. 1997;17(1):21-4

Kure S, Takayanagi M, Narisawa K, Tada K, Leisti J. Identification of a common mutation in Finnish patients with nonketotic hyperglycinemia. J Clin Invest. 1992 ;90(1):160-4.

Kyttälä M. PNUTL2-geenin mRNA variantti 4:n (GI:17986248) potilasmuutoksen poissuljenta Meckelin oireyhtymän tautimutaationa (MKS1) ja tautilokuksen rajauksen tarkennus. Masters´ theses (pro gradu work). Dept. of biochemistry. Faculty of Biosciencies, University of Helsinki. 2003. (In Finnish)

Laan M, Paabo S. Demographic history and linkage disequilibrium in human populations. Nat Genet. 1997;17(4):435-8.

Lahermo P, Sajantila A, Sistonen P, Lukka M, Aula P, Peltonen L, Savontaus ML. The genetic relationship between the Finns and the Finnish Saami (Lapps): analysis of nuclear DNA and mtDNA. Am J Hum Genet. 1996;58(6):1309-22.

Lahermo P, Savontaus ML, Sistonen P, Beres J, de Knijff P, Aula P, Sajantila A. Y chromosomal polymorphisms reveal founding lineages in the Finns and the Saami. Eur J Hum Genet. 1999;7(4):447-58.

Lander ES, Botstein D. Homozygosity mapping: a way to map human recessive traits with the DNA of inbred children. Science. 1987 Jun 19;236(4808):1567-70.
Lathrop GM, Lalouel JM, Julier C, Ott J. Multilocus linkage analysis in humans: detection of linkage and estimation of recombination. Am J Hum Genet. 1985;37(3):482-98.

Lehesjoki AE, Koskiniemi M, Norio R, Tirrito S, Sistonen P, Lander E, de la Chapelle A. Localization of the EPM1 gene for progressive myoclonus epilepsy on chromosome 21: linkage disequilibrium allows high resolution mapping. Hum Mol Genet. 1993;2(8):1229-34.

Lemieux N, Dutrillaux B, Viegas-Pequignot E. A simple method for simultaneous R- or G-banding and fluorescence in situ hybridization of small single-copy genes. Cytogenet Cell Genet. 1992;59(4):311-2.

Levy E, Haltia M, Fernandez-Madrid I, Koivunen O, Ghiso J, Prelli F, Frangione B. Mutation in gelsolin gene in Finnish hereditary amyloidosis. J Exp Med. 1990;172(6):1865-7.

Li JB, Gerdes JM, Haycraft CJ, Fan Y, Teslovich TM, May-Simera H, Li H, Blacque OE, Li L, Leitch CC, Lewis RA, Green JS, Parfrey PS, Leroux MR, Davidson WS, Beales PL, Guay-Woodford LM, Yoder BK, Stormo GD, Katsanis N, Dutcher SK. Comparative genomics identifies a flagellar and basal body proteome that includes the BBS5 human disease gene. Cell. 2004;117(4):541-52.

Liu A, Niswander LA. Signalling in development: Bone morphogenetic protein signalling and vertebrate nervous system development. Nat Rev Neurosci. 2005;6(12):945-954.

Liu Y, Liu XS, Wei L, Altman RB, Batzoglou S. Eukaryotic regulatory element conservation analysis and identification using comparative genomics. Genome Res. 2004;14(3):451-8.

Lichter P, Cremer T, Borden J, Manuelidis L, Ward DC. Delineation of individual human chromosomes in metaphase and interphase cells by in situ suppression hybridization using recombinant DNA libraries. Hum Genet. 1988;80(3):224-34.

Makela-Bengs P, Jarvinen N, Vuopala K, Suomalainen A, Ignatius J, Sipila M, Herva R, Palotie A, Peltonen L. Assignment of the disease locus for lethal congenital contracture syndrome to a restricted region of chromosome 9q34, by genome scan using five affected individuals. Am J Hum Genet. 1998;63(2):506-16.

Marszalek JR, Ruiz-Lozano P, Roberts E, Chien KR, Goldstein LS. Situs inversus and embryonic ciliary morphogenesis defects in mouse mutants lacking the KIF3A subunit of kinesin-II. Proc Natl Acad Sci U S A. 1999;96(9):5043-8.

Maury CP, Kere J, Tolvanen R, de la Chapelle A. Finnish hereditary amyloidosis is caused by a single nucleotide substitution in the gelsolin gene. FEBS Lett. 1990;276(1-2):75-7.

Maxam AM, Gilbert W. A new method for sequencing DNA. Proc Natl Acad Sci U S A. 1977;74(2):560-4.

Mecke S, Passarge E. Encephalocele, polycystic kidneys, and polydactyly as an autosomal recessive trait simulating certain other disorders: the Meckel syndrome. 1971;14(2):97-103.

Mee L, Honkala H, Kopra O, Vesa J, Finnila S, Visapaa I, Sang TK, Jackson GR, Salonen R, Kestila M, Peltonen L. Hydrolethalus syndrome is caused by a missense mutation in a novel gene HYLS1. Hum Mol Genet. 2005;14(11):1475-88.

Miller JQ, Selden RF. Arhinencephaly, encephalocele, and 13-15 trisomy syndrome with normal chromosomes.Neurology. 1967;17(11):1087-91.

Mitchell GA, Brody LC, Sipila I, Looney JE, Wong C, Engelhardt JF, Patel AS, Steel G, Obie C, Kaiser-Kupfer M, et al. At least two mutant alleles of ornithine delta-aminotransferase cause gyrate atrophy of the choroid and retina in Finns. Proc Natl Acad Sci U S A. 1989;86(1):197-201.

Mochizuki T, Saijoh Y, Tsuchiya K, Shirayoshi Y, Takai S, Taya C, Yonekawa H, Yamada K, Nihei H, Nakatsuji N, Overbeek PA, Hamada H, Yokoyama T. Cloning of inv, a gene that controls left/right asymmetry and kidney development. Nature. 1998;395(6698):177-81.

Moerman P, Verbeken E, Fryns JP, Goddeeris P, Lauweryns JM. The Meckel Syndrome. Pathological and cytogenetic observations in eight cases. Hum Genet. 1982;62(3):240-5.

Morgan D, Turnpenny L, Goodship J, Dai W, Majumder K, Matthews L, Gardner A, Schuster G, Vien L, Harrison W, Elder FF, Penman-Splitt M, Overbeek P, Strachan T. Inversin, a novel gene in the vertebrate left-right axis pathway, is partially deleted in the inv mouse.Nat Genet. 1998 Oct;20(2):149-56. Erratum in: Nat Genet 1998;20(3):312.

Morgan NV, Gissen P, Sharif SM, Baumber L, Sutherland J, Kelly DA, Aminu K, Bennett CP, Woods CG, Mueller RF, Trembath RC, Maher ER, Johnson CA. A novel locus for Meckel-Gruber syndrome, MKS3, maps to chromosome 8q24. Hum Genet. 2002;111(4-5):456-61.

Morton NE. Sequential tests for the detection of linkage. Am J Hum Genet. 1955;7(3):277-318.

Murcia NS, Richards WG, Yoder BK, Mucenski ML, Dunlap JR, Woychik RP. The Oak Ridge Polycystic Kidney (orpk) disease gene is required for left-right axis determination. Development. 2000;127(11):2347-55.

Nagamine K, Peterson P, Scott HS, Kudoh J, Minoshima S, Heino M, Krohn KJ, Lalioti MD, Mullis PE, Antonarakis SE, Kawasaki K, Asakawa S, Ito F, Shimizu N. Positional cloning of the APECED gene. Nat Genet. 1997;17(4):393-8.

Nauli SM, Alenghat FJ, Luo Y, Williams E, Vassilev P, Li X, Elia AE, Lu W, Brown EM, Quinn SJ, Ingber DE, Zhou J. Polycystins 1 and 2 mediate mechanosensation in the primary cilium of kidney cells.Nat Genet. 2003;33(2):129-37.

Nikali K, Suomalainen A, Saharinen J, Kuokkanen M, Spelbrink JN, Lonnqvist T, Peltonen L. Infantile onset spinocerebellar ataxia is caused by recessive mutations in mitochondrial proteins Twinkle and Twinky. Hum Mol Genet. 2005 15;14(20):2981-90.

Nishimura DY, Swiderski RE, Searby CC, Berg EM, Ferguson AL, Hennekam R, Merin S, Weleber RG, Biesecker LG, Stone EM, Sheffield VC. Comparative genomics and gene expression analysis identifies BBS9, a new Bardet-Biedl syndrome gene. Am J Hum Genet. 2005;77(6):1021-33.

Nonaka S, Tanaka Y, Okada Y, Takeda S, Harada A, Kanai Y, Kido M, Hirokawa N. Randomization of left-right asymmetry due to loss of nodal cilia generating leftward flow of extraembryonic fluid in mice lacking KIF3B motor protein. Cell. 1998;95(6):829-37. Erratum in: Cell 1999;99(1):117.

Norio R, Nevanlinna HR, Perheentupa J. Hereditary diseases in Finland; rare flora in rare soul. Ann Clin Res. 1973;5(3):109-41.

Norio R. Suomi-neidon geenit –tautiperinnön takana juurillemme johtamassa. Otava. Keuruu. Finland. 2000. (The book available only in Finnish)

Norio R. Finnish Disease Heritage I: characteristics, causes, background. Hum Genet. 2003;112(5-6):441-56.

Nurnberger J, Bacallao RL, Phillips CL. Inversin forms a complex with catenins and N-cadherin in polarized epithelial cells. Mol Biol Cell. 2002;13(9):3096-106.

Ohara O, Dorit RL, Gilbert W. One-sided polymerase chain reaction: the amplification of cDNA. Proc Natl Acad Sci U S A. 1989;86(15):5673-7.

Okada Y, Nonaka S, Tanaka Y, Saijoh Y, Hamada H, Hirokawa N. Abnormal nodal flow precedes situs inversus in iv and inv mice. Mol Cell. 1999;4(4):459-68.

Olbrich H, Haffner K, Kispert A, Volkel A, Volz A, Sasmaz G, Reinhardt R, Hennig S, Lehrach H, Konietzko N, Zariwala M, Noone PG, Knowles M, Mitchison HM, Meeks M, Chung EM, Hildebrandt F, Sudbrak R, Omran H. Mutations in DNAH5 cause primary ciliary dyskinesia and randomization of left-right asymmetry. Nat Genet. 2002;30(2):143-4.

Opitz JM, Howe JJ. The Meckel syndrome (Dysencephalia splanchnocystica, The Gruber syndrome). In malformation syndromes (ed. D. Bergsma). 1969, pp. 167-79. The national foundation Williams & Wilkins, New York.

Ott J. Analysis of human genetic linkage. Johns Hopkins University Press. Baltimore. 1991.

Otto EA, Schermer B, Obara T, O'Toole JF, Hiller KS, Mueller AM, Ruf RG, Hoefele J, Beekmann F, Landau D, Foreman JW, Goodship JA, Strachan T, Kispert A, Wolf MT, Gagnadoux MF, Nivet H, Antignac C, Walz G, Drummond IA, Benzing T, Hildebrandt F. 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. 2003;34(4):413-20.

Otto EA, Loeys B, Khanna H, Hellemans J, Sudbrak R, Fan S, Muerb U, O'Toole JF, Helou J, Attanasio M, Utsch B, Sayer JA, Lillo C, Jimeno D, Coucke P, De Paepe A, Reinhardt R, Klages S, Tsuda M, Kawakami I, Kusakabe T, Omran H, Imm A, Tippens M, Raymond PA, Hill J, Beales P, He S, Kispert A, Margolis B, Williams DS, Swaroop A, Hildebrandt F. Nephrocystin-5, a ciliary IQ domain protein, is mutated in Senior-Loken syndrome and interacts with RPGR and calmodulin. Nat Genet. 2005;37(3):282-8.

Paavola P, Salonen R, Weissenbach J, Peltonen L. The locus for Meckel syndrome with multiple congenital anomalies maps to chromosome 17q21-q24. Nat Genet. 1995;11(2):213-5.
Palo J. Prevalence of phenylketonuria and some other metabolic disorders among mentally retarded patients in Finland. Acta Neurol Scand. 1967;43(5):573-9.

Paetau A, Salonen R, Haltia M. Brain pathology in the Meckel syndrome: a study of 59 cases.Clin Neuropathol. 1985;4(2):56-62.

Paloneva J, Kestila M, Wu J, Salminen A, Bohling T, Ruotsalainen V, Hakola P, Bakker AB, Phillips JH, Pekkarinen P, Lanier LL, Timonen T, Peltonen L. Loss-of-function mutations in TYROBP (DAP12) result in a presenile dementia with bone cysts. Nat Genet. 2000;25(3):357-61.

Pazour GJ, Dickert BL, Vucica Y, Seeley ES, Rosenbaum JL, Witman GB, Cole DG. Chlamydomonas IFT88 and its mouse homologue, polycystic kidney disease gene tg737, are required for assembly of cilia and flagella. J Cell Biol. 2000;151(3):709-18.

Pazour GJ, Rosenbaum JL. Intraflagellar transport and cilia-dependent diseases. Trends Cell Biol. 2002;12(12):551-5.

Pellegata NS, Dieguez-Lucena JL, Joensuu T, Lau S, Montgomery KT, Krahe R, Kivela T, Kucherlapati R, Forsius H, de la Chapelle A. Mutations in KERA, encoding keratocan, cause cornea plana. Nat Genet. 2000;25(1):91-5.

Peltonen L, Pekkarinen P, Aaltonen J. Messages from an isolate: lessons from the Finnish gene pool. Biol Chem Hoppe Seyler. 1995 Dec;376(12):697-704.

Peltonen L, Jalanko A, Varilo T. Molecular genetics of the Finnish disease heritage.
Hum Mol Genet. 1999;8(10):1913-23.

Pennacchio LA, Lehesjoki AE, Stone NE, Willour VL, Virtaneva K, Miao J, D'Amato E, Ramirez L, Faham M, Koskiniemi M, Warrington JA, Norio R, de la Chapelle A, Cox DR, Myers RM. Mutations in the gene encoding cystatin B in progressive myoclonus epilepsy (EPM1). Science. 1996;271(5256):1731-4.

Pinkel D, Straume T, Gray JW. Cytogenetic analysis using quantitative, high-sensitivity, fluorescence hybridization. Proc Natl Acad Sci U S A. 1986;83(9):2934-8.

Quarmby LM, Mahjoub MR. Caught Nek-ing: cilia and centrioles. J Cell Sci. 2005;118(Pt 22):5161-9.

Ramsay M, Williamson R, Estivill X, Wainwright BJ, Ho MF, Halford S, Kere J, Savilahti E, de la Chapelle A, Schwartz M, et al. Haplotype analysis to determine the position of a mutation among closely linked DNA markers. Hum Mol Genet. 1993;2(7):1007-14. Erratum in: Hum Mol Genet 1993;2(9):1523.

Ranta, S, A-E Lehesjoki, M de Fatima Bonaldo, JA Knowles, A Hirvasniemi, B Ross, PJ de Jong, M Bento Soares, A de la Chapelle, and TC Gilliam High-resolution mapping and transcript identification at the progressive epilepsy with mental retardation locus on chromosome 8p 1997; Genome Res 7: 887-896.

Ranta S, Zhang Y, Ross B, Lonka L, Takkunen E, Messer A, Sharp J, Wheeler R, Kusumi K, Mole S, Liu W, Soares MB, Bonaldo MF, Hirvasniemi A, de la Chapelle A, Gilliam TC, Lehesjoki AE.The neuronal ceroid lipofuscinoses in human EPMR and mnd mutant mice are associated with mutations in CLN8. Nat Genet. 1999;23(2):233-6.

Rapola J, Salonen R. Visceral anomalies in the Meckel syndrome. Teratology. 1985;31(2):193-201.

Rehder H, Labbe F. Prenatal morphology in Meckel's syndrome. Prenat Diagn. 1981;1(3):161-71.

Ridanpaa M, van Eenennaam H, Pelin K, Chadwick R, Johnson C, Yuan B, vanVenrooij W, Pruijn G, Salmela R, Rockas S, Makitie O, Kaitila I, de la Chapelle A. Mutations in the RNA component of RNase MRP cause a pleiotropic human disease, cartilage-hair hypoplasia. Cell. 2001 26;104(2):195-203.

Rieder CL, Jensen CG, Jensen LC. The resorption of primary cilia during mitosis in a vertebrate (PtK1) cell line. J Ultrastruct Res. 1979;68(2):173-85.

Rohlich P. The sensory cilium of retinal rods is analogous to the transitional zone of motile cilia. Cell Tissue Res. 1975;161(3):421-30.

Ross AJ, May-Simera H, Eichers ER, Kai M, Hill J, Jagger DJ, Leitch CC, Chapple JP, Munro PM, Fisher S, Tan PL, Phillips HM, Leroux MR, Henderson DJ, Murdoch JN, Copp AJ, Eliot MM, Lupski JR, Kemp DT, Dollfus H, Tada M, Katsanis N, Forge A, Beales PL. Disruption of Bardet-Biedl syndrome ciliary proteins perturbs planar cell polarity in vertebrates.Nat Genet. 2005;37(10):1135-40. Erratum in: Nat Genet. 2005;37(12):1381

Roume J, Genin E, Cormier-Daire V, Ma HW, Mehaye B, Attie T, Razavi-Encha F, Fallet-Bianco C, Buenerd A, Clerget-Darpoux F, Munnich A, Le Merrer M. A gene for Meckel syndrome maps to chromosome 11q13. Am J Hum Genet. 1998;63(4):1095-101.

Steve Rozen and Helen J. Skaletsky. Primer3 on the WWW for general users and for biologist programmers. In: Krawetz S, Misener S (eds) *Bioinformatics Methods and Protocols: Methods in Molecular Biology.* Humana Press, Totowa, NJ. 2000; pp 365-386

Sajantila A, Paabo S. Language replacement in Scandinavia. Nat Genet. 1995;11(4):359-60.

Sajantila A, Salem AH, Savolainen P, Bauer K, Gierig C, Paabo S. Paternal and maternal DNA lineages reveal a bottleneck in the founding of the Finnish population. Proc Natl Acad Sci U S A. 1996;93(21):12035-9.

Salonen R. The Meckel syndrome: clinicopathological findings in 67 patients.
Am J Med Genet. 1984;18(4):671-89.

Salonen R, Norio R. The Meckel syndrome in Finland: epidemiologic and genetic aspects. Am J Med Genet. 1984;18(4):691-8.

Salonen R. The Meckel and Hydrolethalus syndromes diagnostics, genetics, and epidemiological studies. Ph.D. Theses. The department of Medical genetics, Väestöliitto; Department of Medical genetics, University of Helsinki. 1986.

Salonen R, Somer M, Haltia M, Lorentz M, Norio R. Progressive encephalopathy with edema, hypsarrhythmia, and optic atrophy (PEHO syndrome). Clin Genet. 1991;39(4):287-93.
Sambrook, J., E.F. Fritsch, and T. Maniatis. 1989. Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press.

Samadani U, Costa RH. The transcriptional activator hepatocyte nuclear factor 6 regulates liver gene expression. Mol Cell Biol. 1996;16(11):6273-84.

Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain-terminating inhibitors.
Proc Natl Acad Sci U S A. 1977;74(12):5463-7.

Sankila EM, Tolvanen R, van den Hurk JA, Cremers FP, de la Chapelle A. Aberrant splicing of the CHM gene is a significant cause of choroideremia. Nat Genet. 1992;1(2):109-13.

Savukoski M, Klockars T, Holmberg V, Santavuori P, Lander ES, Peltonen L. CLN5, a novel gene encoding a putative transmembrane protein mutated in Finnish variant late infantile neuronal ceroid lipofuscinosis. Nat Genet. 1998;19(3):286-8.

Schuler GD, Boguski MS, Stewart EA, Stein LD, Gyapay G, Rice K, White RE, Rodriguez-Tome P, Aggarwal A, Bajorek E, Bentolila S, Birren BB, Butler A, Castle AB, Chiannilkulchai N, Chu A, Clee C, Cowles S, Day PJ, Dibling T, Drouot N, Dunham I, Duprat S, East C, Edwards C, Fan JB, Fang N, Fizames C, Garrett C, Green L, Hadley D, Harris M, Harrison P, Brady S, Hicks A, Holloway E, Hui L, Hussain S, Louis-Dit-Sully C, Ma J, MacGilvery A, Mader C, Maratukulam A, Matise TC, McKusick KB, Morissette J, Mungall A, Muselet D, Nusbaum HC, Page DC, Peck A, Perkins S, Piercy M, Qin F, Quackenbush J, Ranby S, Reif T, Rozen S, Sanders C, She X, Silva J, Slonim DK, Soderlund C, Sun WL, Tabar P, Thangarajah T, Vega-Czarny N, Vollrath D, Voyticky S, Wilmer T, Wu X, Adams MD, Auffray C, Walter NA, Brandon R, Dehejia A, Goodfellow PN, Houlgatte R, Hudson JR Jr, Ide SE, Iorio KR, Lee WY, Seki N, Nagase T, Ishikawa K, Nomura N, Phillips C, Polymeropoulos MH, Sandusky M, Schmitt K, Berry R, Swanson K, Torres R, Venter JC, Sikela JM, Beckmann JS, Weissenbach J, Myers RM, Cox DR, James MR, Bentley D, Deloukas P, Lander ES, Hudson TJ. A gene map of the human genome. Science. 1996 25;274(5287):540-6.

Seller MJ. Meckel syndrome and the prenatal diagnosis of neural tube defects. J Med Genet. 1978;15(6):462-5.

Shizuya H, Birren B, Kim UJ, Mancino V, Slepak T, Tachiiri Y, Simon M. Cloning and stable maintenance of 300-kilobase-pair fragments of human DNA in Escherichia coli using an F-factor-based vector. Proc Natl Acad Sci U S A. 1992;89(18):8794-7.

Siitonen HA, Kopra O, Kaariainen H, Haravuori H, Winter RM, Saamanen AM, Peltonen L, Kestila M. Molecular defect of RAPADILINO syndrome expands the phenotype spectrum of RECQL diseases. Hum Mol Genet. 2003;12(21):2837-44.

Simons M, Gloy J, Ganner A, Bullerkotte A, Bashkurov M, Kronig C, Schermer B, Benzing T, Cabello OA, Jenny A, Mlodzik M, Polok B, Driever W, Obara T, Walz G. Inversin, the gene product mutated in nephronophthisis type II, functions as a molecular switch between Wnt signaling pathways. Nat Genet. 2005;37(5):537-43.

Simons M, Waltz G. Planar cell polarity and the cilia-centrosomal connection. Unpublished.

Smith CAB. The detection of linkage in human genetics. J Royal Stat Soc B. 1953;15(2):153-184.
Smith UM, Consugar M, Tee LJ, McKee BM, Maina EN, Whelan S, Morgan NV, Goranson E, Gissen P, Lilliquist S, Aligianis IA, Ward CJ, Pasha S, Punyashthiti R, Malik Sharif S, Batman PA, Bennett CP, Woods CG, McKeown C, Bucourt M, Miller CA, Cox P, Algazali L, Trembath RC, Torres VE, Attie-Bitach T, Kelly DA, Maher ER, Gattone VH 2nd, Harris PC, Johnson CA. The transmembrane protein meckelin (MKS3) is mutated in Meckel-Gruber syndrome and the wpk rat.Nat Genet. 2006;38(2):191-6.

Strachan T & Read AP. Human molecular genetics. Bioscientific Publisher Ltd. Oxford. UK. 1999

Sugiura Y, Suzuki Y, Kobayashi M. The Meckel syndrome: report of two Japanese sibs and a review of literature. Am J Med Genet. 1996;67(3):312-4.

Supp DM, Witte DP, Potter SS, Brueckner M. Mutation of an axonemal dynein affects left-right asymmetry in inversus viscerum mice. Nature. 1997. 30;389(6654):963-6.

Surpili MJ, Delben TM, Kobarg J. Identification of proteins that interact with the central coiled-coil region of the human protein kinase NEK1. Biochemistry. 2003;42(51):15369-76.

Swoboda P, Adler HT, Thomas JH. The RFX-type transcription factor DAF-19 regulates sensory neuron cilium formation in C. elegans. Mol Cell. 2000;5(3):411-21.

Takahashi K. Cilia and flagella. Cell Struct Funct. 1984;9 Suppl:s87-90.

Tan KL, Thomas MA. Polycystic kidneys and liver in two siblings with other severe congenital abnormalities. Med J Malaya. 1970;25(1):46-9.

Taulman PD, Haycraft CJ, Balkovetz DF, Yoder BK. Polaris, a protein involved in left-right axis patterning, localizes to basal bodies and cilia. Mol Biol Cell. 2001 Mar;12(3):589-99.

Teebi AS, al Saleh QA, Odeh H. Meckel syndrome and neural tube defects in Kuwait.
J Med Genet. 1992;29(2):140.

Tenhunen K, Laan M, Manninen T, Palotie A, Peltonen L, Jalanko A. Molecular cloning, chromosomal assignment, and expression of the mouse aspartylglucosaminidase gene.Genomics. 1995;30(2):244-50.

Terwilliger JD. A powerful likelihood method for the analysis of linkage disequilibrium between trait loci and one or more polymorphic marker loci.Am J Hum Genet. 1995;56(3):777-87.

The Finnish-German APECED Consortium. An autoimmune disease, APECED, caused by mutations in a novel gene featuring two PHD-type zinc-finger domains. Autoimmune Polyendocrinopathy-Candidiasis-Ectodermal Dystrophy. Nat Genet. 1997;17(4):399-403.

The International Batten Disease Consortium. Isolation of a novel gene underlying Batten disease, CLN3. Cell. 1995;82(6):949-57.

The Retinoschisis Consortium. Functional implications of the spectrum of mutations found in 234 cases with X-linked juvenile retinoschisis. Hum Mol Genet. 1998;7(7):1185-92.

Torrents D, Mykkanen J, Pineda M, Feliubadalo L, Estevez R, de Cid R, Sanjurjo P, Zorzano A, Nunes V, Huoponen K, Reinikainen A, Simell O, Savontaus ML, Aula P, Palacin M. Identification of SLC7A7, encoding y+LAT-1, as the lysinuric protein intolerance gene. Nat Genet. 1999;21(3):293-6.

Torroni A, Bandelt HJ, D'Urbano L, Lahermo P, Moral P, Sellitto D, Rengo C, Forster P, Savontaus ML, Bonne-Tamir B, Scozzari R. mtDNA analysis reveals a major late Paleolithic population expansion from southwestern to northeastern Europe. Am J Hum Genet. 1998;62(5):1137-52.

Tucker RW, Pardee AB, Fujiwara K. Centriole ciliation is related to quiescence and DNA synthesis in 3T3 cells. Cell. 1979;17(3):527-35.

Varilo T. The age of the mutations in the Finnish disease heritage; a genealogical and linkage disequilibrium study. PhD thesis, National Public Health Institute and University of Helsinki, Helsinki; (Available at URL: http://ethesis.helsinki.fi/english.html). 1999.

Verheijen FW, Verbeek E, Aula N, Beerens CE, Havelaar AC, Joosse M, Peltonen L, Aula P, Galjaard H, van der Spek PJ, Mancini GM. A new gene, encoding an anion transporter, is mutated in sialic acid storage diseases. Nat Genet. 1999;23(4):462-5.

Vesa J, Hellsten E, Verkruyse LA, Camp LA, Rapola J, Santavuori P, Hofmann SL, Peltonen L. Mutations in the palmitoyl protein thioesterase gene causing infantile neuronal ceroid lipofuscinosis. Nature. 1995;376(6541):584-7.

Virtaneva K, D'Amato E, Miao J, Koskiniemi M, Norio R, Avanzini G, Franceschetti S, Michelucci R, Tassinari CA, Omer S, Pennacchio LA, Myers RM, Dieguez-Lucena JL, Krahe R, de la Chapelle A, Lehesjoki AE. Unstable minisatellite expansion causing recessively inherited myoclonus epilepsy, EPM1.Nat Genet. 1997;15(4):393-6.

Visapaa I, Fellman V, Vesa J, Dasvarma A, Hutton JL, Kumar V, Payne GS, Makarow M, Van Coster R, Taylor RW, Turnbull DM, Suomalainen A, Peltonen L. GRACILE syndrome, a lethal metabolic disorder with iron overload, is caused by a point mutation in BCS1L. Am J Hum Genet. 2002;71(4):863-76.

Vuopala K, Ignatius J, Herva R. Lethal arthrogryposis with anterior horn cell disease.
Hum Pathol. 1995;26(1):12-9.

Ward S, Thomson N, White JG, Brenner S. Electron microscopical reconstruction of the anterior sensory anatomy of the nematode Caenorhabditis elegans.?2UU. J Comp Neurol. 1975;160(3):313-37.

WATSON JD, CRICK FH. Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. Nature. 1953 Apr 25;171(4356):737-8.

Weber JL, May PE. Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. Am J Hum Genet. 1989;44(3):388-96.

Weissenbach J, Gyapay G, Dib C, Vignal A, Morissette J, Millasseau P, Vaysseix G, Lathrop M. A second-generation linkage map of the human genome. Nature. 1992;359(6398):794-801.

Wheatley DN. Primary cilia in normal and pathological tissues. Pathobiology. 1995;63(4):222-38.

Wheatley DN, Wang AM, Strugnell GE. Expression of primary cilia in mammalian cells. Cell Biol Int. 1996 ;20(1):73-81. Review.

Wright CV. Mechanisms of left-right asymmetry: what's right and what's left? Dev Cell. 2001 Aug;1(2):179-86

Xu YK, Nusse R. The Frizzled CRD domain is conserved in diverse proteins including several receptor tyrosine kinases. Curr Biol. 1998;8(12):R405-6.

Yoder BK, Richards WG, Sweeney WE, Wilkinson JE, Avener ED, Woychik RP. Insertional mutagenesis and molecular analysis of a new gene associated with polycystic kidney disease. Proc Assoc Am Physicians. 1995;107(3):314-23.

Young ID, Rickett AB, Clarke M. High incidence of Meckel's syndrome in Gujarati Indians. J Med Genet. 1985;22(4):301-4.

Zhang Q, Taulman PD, Yoder BK. Cystic kidney diseases: all roads lead to the cilium. Physiology (Bethesda). 2004;19:225-30.

Zlotogora J. Genetic disorders among Palestinian Arabs. 2. Hydrocephalus and neural tube defects. Am J Med Genet. 1997 11;71(1):33-5.

Zondervan KT, Cardon LR. The complex interplay among factors that influence allelic association. Nat Rev Genet. 2004;5(2):89-100.

ORIGINAL PUBLICATIONS