European Journal of Human Genetics (2008) 16, 688–695

www.nature.com/ejhg

ARTICLE

Sequence analysis of 21 genes located in the Kartagener syndrome linkage region on chromosome 15q

Maciej Geremek[1,2], Frederieke Schoenmaker[1], Ewa Zietkiewicz[2], Andrzej Pogorzelski[3], Scott Diehl[4], Cisca Wijmenga*[,1,5,7] and Michal Witt[2,6,7]

[1]*Complex Genetics Group, Department of Biomedical Genetics, University Medical Center Utrecht, Utrecht, the Netherlands;* [2]*Department of Molecular and Clinical Genetics, Institute of Human Genetics, Poznan, Poland;* [3]*Institute of Tuberculosis and Lung Diseases, Rabka, Poland;* [4]*New Jersey Dental School, University of Medicine and Dentistry New Jersey, Newark, NJ, USA;* [5]*Department of Genetics, University Medical Center Groningen, University of Groningen, Groningen, the Netherlands;* [6]*International Institute of Molecular and Cell Biology, Warsaw, Poland*

Primary ciliary dyskinesia (PCD) is a rare genetic disorder, which shows extensive genetic heterogeneity and is mostly inherited in an autosomal recessive fashion. There are four genes with a proven pathogenetic role in PCD. DNAH5 and DNAI1 are involved in 28 and 10% of PCD cases, respectively, while two other genes, DNAH11 and TXNDC3, have been identified as causal in one PCD family each. We have previously identified a 3.5 cM (2.82 Mb) region on chromosome 15q linked to Kartagener syndrome (KS), a subtype of PCD characterized by the randomization of body organ positioning. We have now refined the KS candidate region to a 1.8 Mb segment containing 18 known genes. The coding regions of these genes and three neighboring genes were subjected to sequence analysis in seven KS probands, and we were able to identify 60 single nucleotide sequence variants, 35 of which resided in mRNA coding sequences. However, none of the variations alone could explain the occurrence of the disease in these patients.
European Journal of Human Genetics (2008) **16,** 688–695; doi:10.1038/ejhg.2008.5; published online 13 February 2008

Keywords: Kartagener syndrome; PCD; gene mapping; sequencing; mutation screening

Introduction

Primary ciliary dyskinesia (PCD, MIM no. 242650) is a rare genetic disorder caused – as far as is known – by mutations in genes encoding proteins important for ciliary beating.[1] The major clinical consequences fall into three categories: (1) in the respiratory tract, immotility or dyskinetic beating of cilia in epithelial cells impairs the mucociliary clearance, leading to recurrent infections, sinusitis and bronchiectasis; (2) in the urogenital tract, infertility occurs in some patients because of the dysmotility of sperm tails; (3) mirror reversal of body organ positioning (situs inversus) is seen in approximately half of PCD patients. It is believed to occur at random and to be caused by the immotility of primary cilia of cells on the ventral surface of the embryonic node. The presence of situs inversus in an affected family defines a PCD subtype, known as Kartagener syndrome (KS, MIM no. 244400).

For diagnostic purposes, cilia in bronchial scrapings can be visualized under a light microscope, whereas the clinical diagnosis of PCD is routinely verified by electron microscopy analysis of respiratory cilia ultrastructure. About 20 different ultrastructural defects have been described

*Correspondence: Professor C Wijmenga, Department of Genetics, University Medical Center Groningen, University of Groningen, PO Box 30001, RB Groningen 9700, the Netherlands.
Tel: +31 50 3617 100; Fax: +31 50 3617 230;
E-mail: c.wijmenga@umcutrecht.nl
[7]These authors have contributed equally to this work.
Received 8 June 2007; revised 11 December 2007; accepted 6 January 2008; published online 13 February 2008

in PCD patients, with lesions of outer and/or inner dynein arms being the most frequent defects of the internal anatomy of cilia.[2,3] Recently, it has been shown that immunofluorescence staining of ciliated epithelium with antibodies targeting DNAH5 can detect outer dynein arms defects and therefore aid diagnosis of PCD.[4]

Inheritance of PCD is autosomal recessive in most cases, although pedigrees with an X-linked mode of inheritance have also been described.[5,6] To date, mutations in four genes have been found as causative for PCD, exclusively in patients with outer dynein arm defects. *DNAI1*, coding for intermediate dynein chain 1, was selected for mutation screening based on homology with the *Chlamydomonas reinhardtii* gene, which causes a slow-swimming phenotype when mutated,[7,8] whereas *DNAH5*, coding for heavy dynein chain 5, was identified by homozygosity mapping in a large family and subsequent sequence analysis.[9] *DNAH5* and *DNAI1* are responsible for 28 and 10% of PCD cases, respectively,[10,11] clearly indicating that other genes are involved in PCD/KS etiology. *DNAH11*,[12] coding for dynein heavy chain, and *TXNDC3*,[13] coding for a thioredoxin family member, were also found to be responsible for PCD, but so far each has been found in only one family.

Analysis of the ciliary proteome in unicellular alga *C. reinhardtii* led to the identification of a number of proteins potentially involved in cilia formation and maintenance: 360 proteins with high and 293 with moderate confidence.[14] Besides the known components of ciliary ultrastructure, there were also 90 signal transduction proteins, and many membrane proteins and metabolic enzymes. Given that our knowledge of ciliary genes comes mostly from studies of simple organisms and that the functional spectrum of the possible candidate genes is very wide, it is difficult to predict which of these should be considered as candidate genes for PCD.

On the basis of genetic linkage studies, extensive genetic heterogeneity of PCD is postulated. A total-genome scan performed in 31 multiplex families did not reveal any predominant locus, rather it showed several peaks with suggestive and indicative LOD scores.[15] Another genome scan, performed in five families of Arabic origin and with reported consanguinity, revealed linkage with chromosome 19q13.3 in only three of the families, thus confirming locus heterogeneity.[16] Two additional loci, on 16p12 and 15q, were indicated by studies on genetically isolated, but heterogeneous, populations from the Faroe Islands and the Israeli Druze.[17]

We earlier performed genome-wide linkage analysis in 52 families with KS and reported a region on chromosome 15q24–25, between D15S973 and D15S1037, linked to KS.[18] After obtaining a significant LOD score of 4.34 with D15S154 (using a strictly defined disease model), we defined a region of 3.5 cM (2.82 Mb) as containing the candidate gene.[18] In the present study, we have performed further fine-mapping of the region of interest and screened all 18 genes in the linked region and three neighboring genes for mutations in seven KS patients from the families that contributed most to the linkage results.[18]

Materials and methods

Biological material

DNA samples from the existing PCD collection[18] came from 31 Caucasian families (25 of Polish and 6 of Slovak origin) classified as KS families; none was from a genetically isolated population. Each family had at least one member diagnosed with PCD and exhibiting situs inversus, but with no other major anomalies or dysmorphologies present. The group studied consisted of 38 affected individuals and 99 unaffected family members. The primary complaints in these patients were derived from the respiratory tract and the clinical picture included symptoms of sinusitis, nasal polyps, bronchiectasis, recurrent infections of the upper respiratory tract. All cases were confirmed by a low concentration of nitric oxide (NO) measured in exhaled air.[19] NO was measured from the nasal cavity by a chemiluminescence analyzer with a threshold value of 200 ppb for diagnosing PCD. In 60% of the families, the clinical diagnosis was confirmed by transmission electron microscopy analysis of bronchial cilia ultrastructure; the analysis revealed defects of outer or outer and inner dynein arms; a representative EM picture is shown in Supplementary Figure 1. Direct microscopy of bronchial scrapings was used to confirm the clinical diagnosis of PCD in the patients that did not have electron microscopy imaging of cilia. KS inheritance in these pedigrees was in agreement with general assumptions for an autosomal recessive disease model. KS in these families has previously been shown to be linked with the region between D15S973 and D15S1037, assuming full penetrance, 0.0001 phenocopies and the disease allele frequency of 0.005. The genome-wide significant LOD score of 4.34 was obtained using a cohort of 52 families.[18] Pair-wise LOD score analyses were performed using the FASTLINK program, and multipoint LOD scores for markers on chromosome 15 were calculated using GENEHUNTER. Multipoint LOD scores allowing for locus heterogeneity were calculated using Simwalk2. The 95% confidence interval localized the KS locus to a 3.5 cM region between D15S973 and D15S1037.

Recombination mapping using microsatellite markers

To refine the candidate gene region between D15S973 and D15S1037, seven markers (D15S1027–380 kb–D15S524–33 kb–AFMA085AWB9–580 kb–D15S989–219 kb–D15S1005–58 kb–D15S969–695 kb–D15S551), were genotyped in all 38 affected individuals and 99 unaffected family members. Amplification products, obtained using 5′ end fluorescent-dye-labeled primers, were pooled and

analyzed on an automated ABI Prism 3730 DNA sequencer (Applied Biosystems, Foster City, CA, USA). Genotypes were scored using GENMAPPER (v3.5NT) software. Inheritance was checked using PEDCHECK[20] and haplotypes in the 31 pedigrees were constructed from microsatellite data using SIMWALK2.[21]

Analysis of candidate genes

Sequence analysis was performed in seven affected individuals from the families with the highest LOD score for marker D15S1005.[18] Sequences of 135 exons and surrounding intron sequences of the 20 genes located in the region were obtained using Ensembl genome browser version 42 (Figure 1; Table 1). AL109678, not present in Ensembl, was also sequenced as it is an mRNA coding sequence annotated in the UCSC genome browser. At least 200 bp upstream of the transcription start site were also analyzed and the 3′UTR was sequenced in all genes except for *TMC3*, *KIAA1199*, *ARNT2* and *TMED3*. The Primer3 program was used for designing PCR primers yielding products of 400–500 bp;[22] primers were in the intronic sequences flanking the exons to be analyzed. PCR reactions were performed using standard protocols (data available upon request). Amplification products were purified, quantified on a 2% agarose gel, and diluted for direct sequencing on an automated ABI Prism 3730 DNA sequencer using BigDye Terminator Sequencing Standards (Applied Biosystems). Sequences were assembled using ContigExpress (Vector NTI suite v8.0, Informax Inc.).

Sequencing reactions were performed first with the forward primer, but if the quality of the sequencing electropherogram was unsatisfactory, or if single-nucleotide polymorphisms (SNPs) were found, they were repeated with the reverse primer. Each of the newly identified variants was validated through an independent PCR and sequencing reaction. The sequenced fragments were only included in the results if data were obtained for at least six of the seven patients. Seven reactions could not be amplified despite at least three attempts with two different sets of primers (nucleotides in mRNA sequence: BCL2A1, 602–710; FAM108C1, 1–343; C15ORF37, 421–640; TMED3, 260–299; ENSG00000180725, 457–618; FAH, 683–783, 991–1037). Reactions failed to use control as well as patients' DNA, indicating PCR problems related to the local sequence composition of those genomic fragments. All variations found in the protein-coding sequence and not previously described in public databases were sequenced in a control population group consisting of 48 unrelated, healthy individuals of Polish origin.

Results

Seven microsatellite markers located in the 3.5 cM region reported to be linked to KS were genotyped in 31 families that had positive LOD scores for chromosome 15q in our genome linkage scan[18] (Figure 1). Haplotype analysis was concordant with linkage to chromosome 15q24–25 (data not shown). A recombination was mapped in family numbers 114 and 126 (Figure 2), 1 Mb more telomerically from the previous location defined by marker D15S937, thus narrowing the minimal gene-containing region down to 1.8 Mb, between markers AFMA085AWB9 and D15S1037.

The analysis of haplotype sharing between nonrelated KS patients was inconclusive. The most promising block, encompassing three markers (D15S524, AFMA085WB9 and D15S989) and spanning 0.6 Mb, was present on five nonrelated, affected chromosomes (families 103, 117, 136, 142 and 147) and on one untransmitted parental chromosome (data not shown).

Seven families with the highest positive LOD scores were selected for sequence analysis of all the genes located in

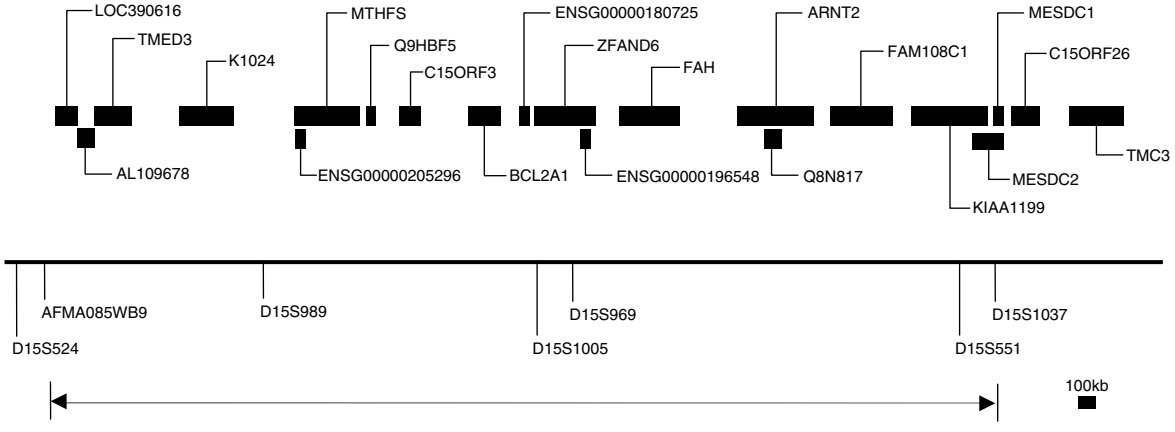

Figure 1 The location of 21 genes in the linkage region on chromosome 15q24-25. The minimal KS gene-containing region is indicated by an arrow.

Table 1 Genes included in the Kartagener syndrome sequencing study

Gene symbol	*Gene name*	*GO description*	*No. of exons*	*Transcript length (bp)*
LOC390616	Unknown	Unknown	1	1608
AL109678	Unknown	Unknown	1	2540
TMED3	Transmembrane emp24 domain-containing protein 3 precursor (membrane protein p24B)	Protein transport; membrane; membrane; protein carrier activity	3	1420
K1024	UPF0258 protein KIAA1024	Membrane; integral to membrane	4	6732
MTHFS	5-formyltetrahydrofolate cyclo-ligase (EC 6.3.3.2) (5,10-methenyl- tetrahydrofolate synthetase) (methenyl-THF synthetase) (MTHFS)	Magnesium ion binding; folic acid binding; ligase activity; metabolism	3	857
ENS-G00000205296	Unknown	Unknown	2	153
Q9HBF5	Cervical cancer suppressor-1	Unknown	2	240
C15ORF37	C15orf37 protein	Unknown	1	2082
BCL2A1	Bcl-2-related protein A1 (protein BFL-1) (hemopoietic-specific early response protein) (protein GRS)	Regulation of apoptosis	3	624
ENS-G00000180725	Unknown	Unknown	2	698
ZFAND6	Zinc-finger A20 domain-containing protein 3 (associated with PRK1 protein)	DNA binding; zinc ion binding; signal transduction; protein binding	7	1680
ENS-G00000196548	Unknown	Unknown	3	123
FAH	Fumarylacetoacetase (EC 3.7.1.2) (fumarylacetoacetate hydrolase) (beta-diketonase) (FAA)	Fumarylacetoacetase activity; magnesium ion binding; regulation of transcription, DNA-dependent; metabolism	15	1471
ARNT2	Aryl hydrocarbon receptor nuclear translocator 2 (ARNT protein 2)	Transcription factor activity; signal transducer activity; aryl hydrocarbon receptor nuclear translocator activity; response to hypoxia; embryonic development	19	6550
Q8N817	CDNA FLJ40133 fis, clone TESTI2012231	Unknown	4	675
FAM108C1	Unknown	Unknown	3	2326
KIAA1199	KIAA1199	Sensory perception of sound	29	6602
MESDC2	Mesoderm development candidate 2 (NY-REN-61 antigen)	Mesoderm development	5	2938
MESDC1	Mesoderm development candidate 1	Unknown	1	2373
C15ORF26	Uncharacterized protein C15orf26	Unknown	7	1574
TMC3	Transmembrane channel-like protein 3	Integral to membrane	19	2443

the 1.8 Mb region (Figure 2). These families contributed 70% of the LOD score for marker D15S1005 on chromosome 15q,[18] and it was interesting that the *DNAI1* mutation analysis was negative in all seven families. The *DNAH5* contribution was excluded in four families (114, 117, 136 and 147; Figure 2), based on the analysis of SNPs located within the gene (discordant genotypes in the affected sibs), while in the three remaining families, the analyzed SNPs in *DNAH5* were uninformative (unpublished data). As the microsatellite data confirmed linkage of the disease to 15q24–25 in all seven families, only one affected individual, representative of each pedigree, was subjected to DNA sequencing. Supplementary Table 2 summarizes the clinical findings of the seven families included for sequencing.

We analyzed the 18 coding sequences located between markers AFMA085WB9 and D15S1037 (Table 1). We also sequenced three genes located in the vicinity of the region, which could be functionally related to cilia: *MESDC2*, which is important in embryonic development; *C15ORF26*, which is expressed in lung and testis and displays homology to a flagellar protein identified in *Chlamydomonas*;[14] and *TMC3*, a membrane protein with partial homology to nephrocystine, as indicated by Blast search of the NCBI nr database (Figure 1). Except for *C15ORF26*, none of the genes have been reported to be a part of cilium.

In total, we analyzed 83 kb of genomic sequence in seven individuals (35 kb of coding sequences and 48 kb of intronic/flanking DNA). We found 60 SNPs, 45 of which were already present in public SNP databases. Twelve SNPs, including the four previously unknown ones, were located in the protein coding regions. None of the identified SNPs in the coding regions resulted in a stop codon change or

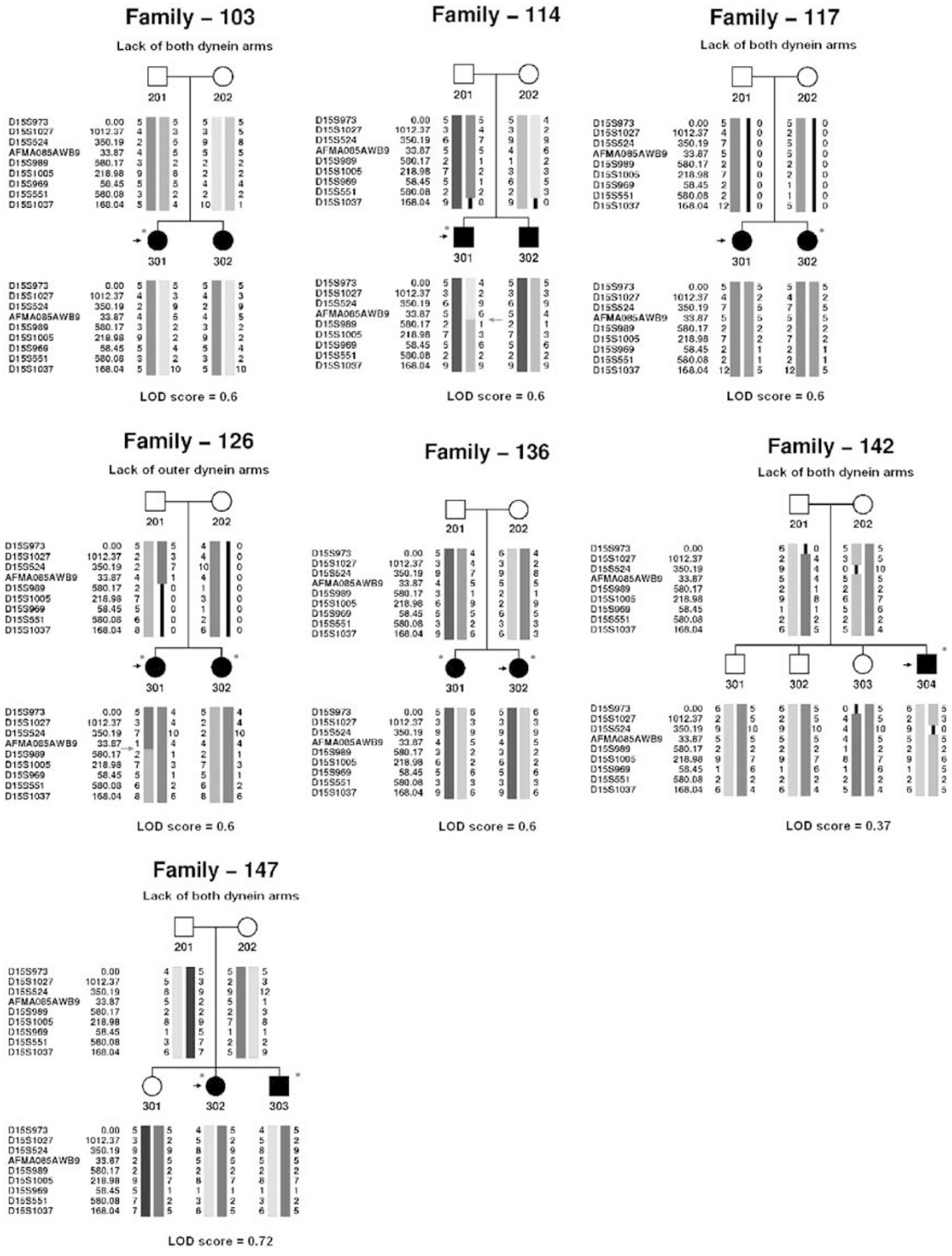

Figure 2 Pedigrees of the KS families included in the sequencing. The results of the pair-wise LOD score for marker D15S1005 are indicated below each family. The recombination (marked by a red arrow) defining the minimal KS gene-containing region can be observed in pedigrees 114 and 126. Individuals included in the sequencing are indicated by black arrows. Red dots indicate individuals exhibiting situs inversus. The numbers beside marker names indicate the intermarker distances in kilobase pairs. Missing data are denoted by zeros. The genotypes of the parents in pedigrees 117 and 126 were deduced from the genotypes of the siblings.

Table 2 List of all sequence variations found in 7 Kartagener syndrome patients

Gene	*Variation*	*Amino acid*	*Exon\intron*	*ID in SNP database*	*Minor allele frequency in NCBI genome database 35*	*Minor allele frequency in 48 control individuals of Polish origin*	*Number of patient chromosomes with the minor allele*
LOC39016	c.735C>T	P245P	Exon 1	rs8038778	T=0.367		7/14
TMED	c.659G>A	A172A	Exon 3	rs906439		A=0.0	0/14
KIAA1024	c.847A>C	N258H[a]	Exon 2	—		C=0.06	2/14
	c.918C>A	A281A	Exon 2	—			1/14
	c.2569A>G	I832V[b]	Exon 3	rs2297773	G=0.208		9/14
	c*1799A>C	—	3′UTR	rs17266017	C=0.375		2/14
	c.*3771_3772insATG	—	3′UTR	—		Ins=0.06	2/14
BCL2A1	c.238G>A	C19Y[c]	Exon 1	rs1138357	A=0.146		3/14
	c.299T>G	N39K[d]	Exon 1	rs1138358	G=0.146		3/14
MTHFS	g.411A>G	—	Upstream	rs2865825	—		1/14
	g.421C>G	—	Upstream	rs2903105	—		3/14
	g.483delG	—	Upstream	rs35401897	—		2/14
C15ORF37	c.903C>A	P166Q[f]	Exon 1	rs3803540	A=0.217		1/14
	c.*46G>T	—	3′UTR	rs17214656	T=0.167		2/14
	c.*124A>C	—	3′UTR	rs36070199	—		2/14
	c.*206A>G	—	3′UTR	rs12442408	G=0.250		1/14
	c.*413A>T	—	3′UTR	rs2733101	T=0.283		2/14
AWP1	c.295−90T>C	—	Intron 2	rs1916048	C=0.217		5/12
	C574+11C>G	—	Intron 5	rs1522636	G=0.261		9/14
	c.466−60A>T	—	Intron 4	rs12915043	—		1/14
	c.506A>G	Q65Q	Exon 4	—		G=0.0	1/14
FAH	c.915−173delC	—	Intron 10	rs3835063	—		6/14
	c.1133C>T	S352S	Exon 13	rs1801374	T=0.058		2/14
	c.1038−35A>C	—	Intron 12	rs2043691	C=0.08		1/14
	c.1038−223C>T	—	Intron 12	rs2043692	T=0.15		1/14
	c.*37A>C	—	3′UTR	rs1049181	—		2/14
	c.*38C>T	—	3′UTR	—	—		2/14
	c.*39C>T	—	3′UTR	rs3210172	—		2/14
	c.*40C>T	—	3′UTR	—	—		2/14
	c.*41C>G	—	3′UTR	—	—		2/14
	c.*50delC	—	3′UTR	—	—		2/14
	c. *93T>C	—	3′UTR	rs1049194	—		2/14
ARNT2	c.312+15C>T	—	Intron 3	rs2278708	T=0.058		1/14
	c.361−28insC	—	Intron 3	rs5814026	—		3/14
	C574+36C>G	—	Intron 4	rs8033507	G=0.017		1/14
	c.892−74C>T	—	Intron 7	rs11072922	T=0.175		2/14
	c.1255+21C>T	—	Intron 11	rs3924894	T=0.405		8/14
	c.1255+22T>C	—	Intron 11	rs3924893	C=0.4		6/14
	c.1279T>A	W399R[e]	Exon 12	-		A=0.0	1/14
FAM108C1	c.*692T>C	—	3′UTR	rs1046417	C=0.130		2/14
	c.*1115C>A	—	3′UTR	rs11072940	A=0.104		2/14
KIAA1199	c.−255T>A	—	5′UTR	rs2273886	A=0.295		4/14
	c.1029+26C>G	—	Intron 7	—	—		7/14
	c.1346+240C>T	—	Intron 9	—	—		1/14
	c.1479+83T>A	—	Intron 10	—	—		1/14
	c.2057+57A>G	—	Intron 13	rs2271164	G=0.417		5/12
	c.2142T>C	C28C	Exon 14	rs2271160	C=0.261		3/14
	c.2462A>C	—	Intron 16	rs2271162	C=0.167		3/14
MESDC2	c.*3G>T	—	3′UTR	rs8039607	T=0.121		1/14
	c.*53G>T	—	3′UTR	rs8039384	T=0.142		1/14
	c.1369−20_1369−18del-CTT	—	Intron 4	—		Del=0.08	1/14
	c.*1416C>T	—	3′UTR	—		T=0.069	1/14
	c.*2153T>C	—	3′UTR	rs10519305	C=0.017		1/14
C15ORF26	G496C>G	—	Upstream	rs2683255	—		1/14
	c.505−105insA	—	Intron 3	—	—		3/14
	c.505−87C>G		Intron 3	—	—		3/14
	c.664−50insA		Intron 5	rs11395672	—		7/14
TMC3	c.−13T>C	—	5′UTR	—		C=0.08	1/14
	c.580+55T>C		Intron 6	rs12903128	C=0.345		1/14
	c.823−66A>G		Intron 7	rs7167034	A=0.153		7/14

Nucleotide variations are presented according to the HUGO nomenclature. Translation start site was used for each gene as the +1 nucleotide. Putative effect on protein:
[a]Small, polar>polar, charged, positive, aromatic;
[b]Aliphatic>aliphatic, small;
[c]Tiny>aromatic;
[d]Small>positive;
[e]Aromatic, polar>polar, positive;
[f]Small>polar.

a frame shift; six were amino-acid substitutions and seven were synonymous codon changes. Twenty-one SNPs, including seven novel ones, were located in the untranslated regions. Four SNPs were located in regions upstream of the transcription start site. Twenty-seven variations were located in noncoding regions, six of which were previously unknown. None of the SNPs found were closer than 10 bp to a splicing site. Variations in the protein-coding sequence, previously not described in SNP databases, were sequenced in 48 individuals of Polish origin to estimate population frequencies. Except for two SNPs (W399R in *ARNT2* and A506G in *AWP1*), all the variations were found in the healthy controls. A complete list of the SNPs found is given in Table 2 and the distribution of all SNPs in the patients can be found in Supplementary Table 1.

Discussion

We have performed a follow-up study to our previous finding[18] of linkage to chromosome 15q24–25 in families with KS. Fine-mapping of the 3.5 cM (2.82 Mb) linked region in 31 KS families allowed us to narrow down the minimal candidate gene region to 1.8 Mb. Analysis of haplotypes composed of microsatellite markers revealed that there was no extensive haplotype sharing among the affected individuals, arguing against the presence of only one or a few frequent mutations occurring on a common haplotype background. On the other hand, the absence of haplotype sharing does not exclude the possibility of the presence of a causative gene or genes in the linked region. The linkage analysis is independent of a founder effect. There are examples in literature of genes involved in rare recessive diseases, but not showing evidence of haplotype sharing. This happens because most of the patients are compound heterozygous for unique mutations in the given gene.[23]

In searching for the KS gene in the 15q24-25 region, we found 37 SNPs in the mRNA-coding sequences and promoters of 14 genes, including 6 nonsynonymous codon changes. No evidently pathogenic mutations, such as the introduction of stop codons or frame-shift mutations, were detected.

Given the rare occurrence of the disease (approximately 1:40 000), the expected cumulative frequency of all the disease variant(s) in the general population should be below 1%. We also sequenced fragments harboring variants in mRNA-coding sequence in 96 control chromosomes (Table 2). One nonsynonymous amino-acid change (W399R in *ARNT2*), and one silent codon change (A506G in *AWP1*) were not seen in the controls, suggesting that their frequencies might be as low as 1%. However, each of these two variants (in two different genes) was found in heterozygous form in only a single patient, with no other possibly pathogenic complementary allele; these are therefore unlikely to represent the disease-causing mutations. None of the remaining SNPs were particularly rare and their occurrence in patients did not significantly deviate from population frequencies (Table 2).

In conclusion, we did not find any variants that directly fulfilled the criteria of a pathogenic mutation for KS in any of the 21 genes located in the minimal gene-containing region, as indicated by linkage analysis (while assuming the rare occurrence of the disease and the recessive mode of inheritance). The analyzed region may, however, contain pathogenic variants that we failed to find. We did not sequence introns that could contain regulatory elements or harbor gain-of-splicing site mutations. It is also possible that some of the genes have alternative exons not annotated in the current databases and thus not analyzed in our study, and we cannot exclude the possibility that some of the genes harbor heterozygous deletion(s) of one or more exons. We did not detect any extended homozygous regions, in neither the microsatellite data nor in the SNP data, but the relatively low density of heterozygous SNPs could have left some regions noninformative with respect to their purported hemizygosity (see Supplementary Table 1).

As the pedigree structures in the cohort we studied were simple, we cannot exclude the possibility that some of the families are false positives, although none of the families was crucial for the significant LOD score. Hence, despite the statistical significance of the LOD score value, it cannot entirely be excluded that the 15q locus is a false positive locus. The current location of the linkage interval is based on two families (114 and 126; Figure 2) that contribute to the left recombination and on one family[18] that contributes to the right recombination. This last family contains only one affected individual and, therefore, contributed very little to the overall LOD score. Under less stringent disease models, the linkage peak would still be significant with a 95% confidence interval matching the minimal gene-containing region. However, the linkage curve would be positive over a larger region with more than 200 genes. Hence, we cannot exclude that the KS gene is located closer to the 15q telomere. Given that this larger region does not contain obvious candidate genes, more families will be needed to establish the exact KS region.

Acknowledgements

We thank all the Polish and Slovak PCD families who participated in this study for their invaluable cooperation, and J Pawlik (Rabka, Poland) and A Kapellerova (Bratislava, Slovakia) for providing study material and clinical data. We gratefully acknowledge the help of Ewa Rutkiewicz in collecting biological material, of Jackie Senior in editing the manuscript and of Behrooz Alizadeh in interpreting the statistical data. This research was partly financed by Grant PBZ-KBN-122/P05-1 from the Polish Ministry of Science and Higher Education. MG is supported by the International PhD Program of the University of Utrecht.

References

1 Geremek M, Witt M: Primary ciliary dyskinesia: genes, candidate genes and chromosomal regions. *J Appl Genet* 2004; **3**: 347–361.
2 Zariwala MA, Knowles MR, Omran H: Genetic defects in ciliary structure and function. *Annu Rev Physiol* 2007; **69**: 423–450.
3 Van's Gravesande KS, Omran H: Primary ciliary dyskinesia: clinical presentation, diagnosis and genetics. *Ann Med* 2005; **37**: 439–449.
4 Fliegauf M, Olbrich H, Horvath J *et al*: Mislocalization of DNAH5 and DNAH9 in respiratory cells from patients with primary ciliary dyskinesia. *Am J Respir Crit Care Med* 2005; **171**: 1343–1349.
5 Krawczynski MR, Witt M: PCD and RP: X-linked inheritance of both disorders? *Pediatr Pulmonol* 2004; **38**: 88–89.
6 Moore A, Escudier E, Roger G *et al*: RPGR is mutated in patients with a complex X linked phenotype combining primary ciliary dyskinesia and retinitis pigmentosa. *J Med Genet* 2006; **43**: 326–333.
7 Guichard C, Harricane M-C, Lafitte J-J *et al*: Axonemal dynein intermediate-chain gene (DNAI1) mutations result in situs inversus and primary ciliary dyskinesia (Kartagener syndrome). *Am J Hum Genet* 2001; **68**: 1030–1035.
8 Pennarun G, Escudier E, Chapelin C *et al*: Loss-of-function mutations in a human gene related to *Chlamydomonas reinhardtii* dynein IC78 result in primary ciliary dyskinesia. *Am J Hum Genet* 1999; **65**: 1508–1519.
9 Olbrich H, Häffner K, Kispert A *et al*: Mutations in DNAH5 cause primary ciliary dyskinesia and randomization of left-right asymmetry. *Nat Genet* 2002; **30**: 143–144.
10 Hornef N, Olbrich H, Horvath J: DNAH5 mutations are a common cause of primary ciliary dyskinesia with outer dynein arm defects. *Am J Respir Crit Care Med* 2006; **174**: 120–126.
11 Zariwala MA, Leigh MW, Ceppa F: Mutations of DNAI1 in primary ciliary dyskinesia: evidence of founder effect in a common mutation. *Am J Respir Crit Care Med* 2006; **174**: 858–866.
12 Bartoloni L, Blouin JL, Pan Y *et al*: Mutations in the DNAH11 (axonemal heavy chain dynein type 11) gene cause one form of situs inversus totalis and most likely primary ciliary dyskinesia. *Proc Natl Acad Sci* 2002; **99**: 10282–10286.
13 Duriez B, Duquesnoy P, Escudier E *et al*: A common variant in combination with a nonsense mutation in a member of the thioredoxin family causes primary ciliary dyskinesia. *Proc Natl Acad Sci* 2007; **104**: 3336–3341.
14 Pazour GJ, Agrin N, Leszyk J, Witman GB: Proteomic analysis of a eukaryotic cilium. *J Cell Biol* 2005; **170**: 103–113.
15 Blouin JL, Meeks M, Radhakrishna U *et al*: Primary ciliary dyskinesia: a genome-wide linkage analysis reveals extensive locus heterogeneity. *Eur J Hum Genet* 2000; **8**: 109–118.
16 Meeks M, Walne A, Spiden S *et al*: A locus for primary ciliary dyskinesia maps to chromosome 19q. *J Med Genet* 2000; **37**: 241–244.
17 Jeganathan D, Chodhari R, Meeks M *et al*: Loci for primary ciliary dyskinesia map to chromosome 16p12.1–12.2 and 15q13.1–15.1 in Faroe Islands and Israeli Druze genetic isolates. *J Med Genet* 2004; **41**: 233–240.
18 Geremek M, Zietkiewicz E, Diehl SR, Alizadeh BZ, Wijmenga C, Witt M: Linkage analysis localises a Kartagener syndrome gene to a 3.5 cM region on chromosome 15q24–25. *J Med Genet* 2006; **43**: e1.
19 Karadag B, James AJ, Gultekin E *et al*: Nasal and lower airway level of nitric oxide in children with primary ciliary dyskinesia. *Eur Respir J* 1999; **13**: 1402–1405.
20 O'Connell JR, Weeks DE: PedCheck: a program for identifying genotype incompatibilities in linkage analysis. *Am J Hum Genet* 1998; **63**: 259–266.
21 Sobel E, Lange K: Descent graphs in pedigree analysis: applications to haplotyping, location scores, and marker sharing statistics. *Am J Hum Genet* 1996; **58**: 1323–1337.
22 Rozen S, Skaletsky HJ: 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*. Totowa, NJ: Humana Press, 2000, pp 365–386.
23 Wijmenga C, Hansen RS, Gimelli G *et al*: Genetic variation in ICF syndrome: evidence for genetic heterogeneity. *Hum Mutat* 2000; **16**: 509–517.

Supplementary Information accompanies the paper on European Journal of Human Genetics website (http://www.nature.com/ejhg)