Multi-exon deletions of the PKHD1 gene cause autosomal recessive polycystic kidney disease (ARPKD)

C Bergmann, F Küpper, C P Schmitt, U Vester, T J Neuhaus, J Senderek, K Zerres

Background: Autosomal recessive polycystic kidney disease (ARPKD) is caused by mutations in the PKHD1 (polycystic kidney and hepatic disease 1) gene on chromosome 6p12, a large gene spanning 470 kb of genomic DNA. So far, only micromutations in the 66 exons encoding the longest open reading frame (ORF) have been described, and account for about 80% of mutations.

Objective: To test the hypothesis that gross genomic rearrangements and mutations in alternatively spliced exons contribute to a subset of the remaining disease alleles.

Methods: Using DHPLC for alternatively spliced exons and quantitative real time polymerase chain reaction to detect genomic imbalances, 58 ARPKD patients were screened, of whom 55 were known to harbour one PKHD1 point mutation in the longest ORF.

Results: Three different heterozygous PKHD1 deletions and several single nucleotide changes in alternatively spliced exons were identified. The detected partial gene deletions are most likely pathogenic, while a potential biological function of the alterations identified in alternatively spliced exons must await the definition of transcripts containing alternative exons and their predicted reading frames.

Conclusions: Gross PKHD1 deletions account for a detectable proportion of ARPKD cases. Screening for major genomic PKHD1 rearrangements will further improve mutation analysis in ARPKD.

A utosomal recessive polycystic kidney disease (ARPKD (MIM 263200)) is an important cause of renal and liver related morbidity and mortality in neonates and infants, occurring 1 in 20 000–40 000 live births.1–3 Principal histological manifestations involve the fusiform dilatation of renal collecting ducts and hepatobiliary ductal plate malformation. Severely affected neonates display massively enlarged echogenic kidneys and a “Potter” oligohydramnios sequence. About 30–50% of affected neonates die shortly after birth from respiratory insufficiency.2 Those who survive the neonatal period or present later in life express widely variable disease phenotypes with systemic hypertension, end stage renal disease (ESRD), and sequelae of portal hypertension.4–5

The recently characterised PKHD1 gene (MIM 606702) on chromosome 6p12 is an exceptionally large gene.6–7 The longest open reading frame (ORF) comprises 66 exons encoding a protein of 4074 amino acids. Expression analysis suggests that PKHD1 and its murine orthologue undergo a complex pattern of alternative splicing.7–8 The predicted full length protein (polyductin/fibrocystin) represents a novel putative integral membrane protein with so far unknown function. Protein expression in kidney and liver has been shown to be consistent with the sites affected by the disorder.

Recent studies have demonstrated that polyductin, similar to other cystoproteins, is localised to primary cilia with concentration in the basal body area.6–14

The large size of PKHD1, its supposed complicated pattern of splicing, and lack of knowledge of the encoded protein’s function pose significant challenges to DNA based diagnostic testing. Moreover, the wide variety of different PKHD1 mutations, with the majority of changes unique to single families in “non-isolate” populations, sets further requirements for investigation and makes PKHD1 molecular testing a time consuming and labour intensive process. However, we recently set up an algorithm for PKHD1 screening that allows for detection of most mutations by analysis of only a subset of fragments and facilitates robust PKHD1 mutation analysis in a routine diagnostic setting.15

Currently, 276 different PKHD1 micromutations (point mutations and small deletions/duplications/insertions) on 670 mutated alleles are listed in the locus specific database (as of 26 March 2005) (http://www.humgen.rwth-aachen.de).16–22 (Michel-Calemard/database submission and own unpublished data). The mutation detection rate obtained in the separate studies varied substantially, from 42% to 85%, most probably depending primarily on the cohort of patients screened within the respective studies. However, most recent surveys have shown detection rates of about 80% for the entire clinical spectrum of ARPKD patients.16–22 The power of PKHD1 mutation analysis is further strengthened by the observation that in more than 95% of families screened at least one mutation could be identified.5 20

However, the molecular defect remains to be determined in a considerable proportion of chromosomes. While genetic heterogeneity is unlikely for the vast majority, the major cause of missing mutations may be limited sensitivity of the screening methods applied. Moreover, genomic rearrangements would have been missed by polymerase chain reaction (PCR) based strategies. We thus analysed 55 ARPKD patients with one point mutation in the longest ORF and three mutation negative ARPKD cases with necropsy proven diagnosis to elucidate the contribution of large genomic alterations in PKHD1. Furthermore, we screened the same cohort of patients for changes in alternatively spliced PKHD1 exons by denaturing high performance liquid chromatography (DHPLC).

METHODS

Study population

A cohort of 58 unrelated families (17 multiplex pedigrees) of different ethnic background (mainly European) was included in the present study. Fifty five of these were previously

Abbreviations: ARPKD, autosomal recessive polycystic kidney disease; DHPLC, denaturing high performance liquid chromatography; ESRD, end stage renal disease; ORF, open reading frame; WT, wild-type

www.jmedgenet.com
shown to carry one mutation in the 66 exons encoding the longest PKHD1 ORF (exons 2–67, fig 1A); however, the second change on the other parental allele had not been found11–20 (and our own unpublished data). In addition, three individuals without identifiable mutation within the longest ORF were enclosed. In the latter cases necropsy confirmed the diagnosis of ARPKD. The majority of index patients were moderately affected and survived the neonatal period (n = 37), whereas 19 individuals died either perinatally or neonatally. Two families were characterised by gross intrafamilial phenotypic variability with peri-/neonatal demise in one and survival into childhood in another affected sibling. DNA or blood samples were obtained with informed consent.

Haplotype analysis
As previously described, at least two informative markers flanking PKHD1 were typed in each family with more than one child.3 All families were compatible with linkage to the PKHD1 locus on chromosome 6p12.

Mutation analysis
All patients included in the present study were initially screened for mutations in the 66 exons encoding the 4074 amino acid polyductin protein (GenBank: NM_138694, AY129465) by DHPLC on a wave fragment analysis system (Transgenomic, Crewe, UK) (http://www.humgen.rwth-aachen.de).21 Many of the primer sets of the longest PKHD1 ORF entirely or partly amplified the predicted alternative exons 20a, 32a, 32b, 32c, 39a, 39b, 41a, 41a, 44a, 51a, 51b, 60a, 66a, and 71a. As regards exons 38, 38a, 39b, 41a, 62, 63, and 64 not or not entirely included in the longest ORF, primers were designed using the Primer3 program (http://www-genome.wi.mit.edu/cgi-bin/primer/primer3.cgi). To limit the size of the amplicons exons 38/38a, 39b, 41a, 62, and 64 were split and amplified in two (38/38a, 39b, and 41a), three (62), and six (64) overlapping fragments, respectively (see supplementary table S1). This can be viewed on the journal website: www.jmedgenet.com/supplemental.

Quantitative real time PCR
For detection of fragment copy number deviations in PKHD1, we developed QRT-PCR assays using SYBR® Green. We chose 12 fragments (exons 4, 8, 27, 34, intron 35, 38, 46, 50, 53, 58, 65, and 67) equally distributed over the PKHD1 gene (fig 1A). To determine the copy number of the screened fragments and to exclude the influence of imprecise amounts and qualities of genomic DNA, we quantified an external reference fragment yielding two products in patient’s DNA: one fragment in parallel. As reference locus, we used exon 3 of the factor VIII gene.22 Primers were designed using the Primer Express™ software, version 2.0.0 (Applied Biosystems, Darmstadt, Germany). They were purchased from MWG (Ebersberg, Germany); sequences are listed in supplementary table S2 (www.jmedgenet.com/supple-
micromutations in the longest ORF. Moreover, we observed several single nucleotide changes involving predicted alternatively spliced exons (table 1).

Validation of quantitative real time PCR assays and characterisation of deletions
Based on the QRT PCR results, the ratios of target and reference gene were calculated. Patients and controls with ratios ranging from 0.75 to 1.3 were considered to carry two fragment copies (wild-type, WT). Ratios below 0.75 and above 1.3 were considered suggestive of a deletion or duplication, respectively. For every DNA samples revealing deviations from the WT copy number, we undertook to identify the extent of the deletions and the genomic sequences in which the rearrangements occurred. We constructed primers flanking the fragment found to be deleted and carried out additional QRT-PCR analyses corresponding to conditions described above. In case these fragments were also shown to be deleted, the next flanking exon was chosen for QRT-PCR. As soon as an exonic fragment yielded results in favour of the underlying deletion was further confined by QRT-PCR analyses and primers located within the intron believed to harbour the genomic rearrangement. Finally, a PCR based strategy was chosen to search for the putative deletion. Primers from the fragments found to be situated just outside the deletion were used to amplify the breakpoint region. Subsequently, the junction fragments were directly sequenced with these PCR primers and the exact positions of the breakpoints were identified. The PCR products were sequenced using the ABI PRISM® BigDye TM terminator cycle sequencing ready reaction kit, version 2.0 (Applied Biosystems) according to standard protocols, and samples were run and analysed on an ABI PRISM 377 Fluorescent DNA sequencer (Applied Biosystems). All primers used either for QRT-PCR or subsequent PCR and sequencing are available on request. Segregation analysis was carried out for any mutation identified.

RESULTS
Partial deletions of the PKHD1 gene were identified in three of 58 ARPKD patients known to harbour one or no PKHD1 micromutations in the longest ORF. Moreover, we observed several single nucleotide changes involving predicted alternatively spliced exons (table 1).

In a previous study, moderately affected patient 309 was found to carry the PKHD1 missense mutation c.309C→T (p.Ala103Val) in exon 27 on his maternal allele.13 We have now identified a deletion encompassing exon 58 on the patient’s paternal allele. QRT PCR assays revealed values around 1.0 for exons 56 and 59, while exon 57 also seemed to be deleted. To identify the presumed deletion breakpoints in introns 56 and 58, we carried out a PCR with primers aligning exons 56 and 59. PCR yielded two products in patient’s DNA: an 8.3 kb band corresponding to the WT allele and a product including the deletion with a length of 1030 base pairs (bp).
As expected, in controls only the WT allele was visible (data not shown). By direct sequencing of the PCR product, we were able to identify the deletion breakpoints in introns 56 and 58 (fig 1B). Intron 56 ended at g.332887 (nucleotide position corresponding to that in GenBank AY129465), and the sequence of intron 58 broke at position g.340206. Thus in all the deletion was shown to comprise 7319 bp of genomic DNA.

Patient 696, also with a moderate phenotype, was previously found to harbour the missense mutation c.3467C\(\rightarrow\)T (p.Ser1156Leu) in exon 30 of the \(PKHD1\) gene on his paternal allele.\(^5\) In the present study, this patient showed decreased ratios for exon 27 in QRT-PCR, whereas the adjacent exons had normal values. Thus a long range PCR of about 8.5 kb of genomic WT DNA. However, even with primers in exons 26 and 28 was carried out, yielding a result of fragments of 3700, 2050, 1338, and 893 bp, respectively. The patient’s DNA additionally revealed a band about 100 bp lower than the smallest band seen in WT DNA. These findings pointed to a deletion of about this size with putative breakpoints close to exon 27. PCR with newly designed primers allowed us to define the exact deletion breakpoints within exon 27 and intron 27 encompassing 103 bp of genomic DNA in total (del c.3007_3097+12\(\rightarrow\)p.Gly1003fs). Segregation analysis revealed the deletion on the patient’s maternal allele.

Both affected sisters of family 27 were known to harbour the most common \(PKHD1\) mutation c.107C\(\rightarrow\)T (p. Thr36Met) on their paternal allele.\(^6\) The younger sister died at the age of three years from ESRD with sepsis, while the elder one died at 14 from ESRD and peritonitis. In the current study, QRT PCR showed decreased values for exon 58. While the proximal flanking exon 57 yielded normal results, the distally flanking exons 59 and 60 revealed copy numbers around 0.5; only exon 61 was shown to be wild-type. Thus putative deletion breakpoints were assumed to be located in introns 57 and 60. The large size of intron 60 of more than 84 kb further complicated the characterisation of the breakpoint’s exact position. We narrowed down the precise extent of the deletion by creating primers for QRT amplifying short sequences spreading over this genomic region. A fragment placed about 7 kb downstream of exon 60 was found to lie outside the deletion, whereas a fragment 3 kb downstream of exon 60 was shown to be deleted. Taqman primers from the identified flanking segments were combined to amplify the putative breakpoint region. These oligonucleotides bracketed a segment of almost 17 kb on the reference allele. As expected, no product was obtained when using DNA of controls under standard PCR conditions. However, amplification of patient DNA yielded a fragment of almost 4 kb,
suggested a deletion of about 13 kb of genomic DNA. Sequencing of the junction fragment allowed the precise identification of the deletion breakpoints. This deletion spans from g.337078 to g.350281 and was found to be maternally inherited. Next to the intron 60 breakpoint, three nucleotides (ACA) were found to be inserted while the following sequence (TTGATACGAAG) was duplicated.

DISCUSSION

While almost 300 different PKHD1 micromutations are encountered in the locus specific database (as of 26 March 2005) (http://www.humgen.rwth-aachen.de), no major genomic PKHD1 rearrangement has been identified so far. Thus this is the first study describing gross deletions in the PKHD1 gene.

Although the detection rates yielded by most recent studies were comparatively high (about 80%), many mutations remain undetermined. The major cause of missing mutations will probably be the limited sensitivity of DHPLC; however, this is most unlikely to account for the entire remaining percentage. There may be alternative explanations: missing mutations may reside in regulatory elements or introns distant from the splice donor and acceptor sites24–26; some of the changes currently categorised as non-pathogenic may ultimately turn out to be disease associated; and major genomic rearrangements may occur in the PKHD1 gene. Thus we analysed by QRT-PCR and DHPLC 58 ARPKD patients previously found to carry one or no PKHD1 micromutation for gross deletions/duplications and for mutations in alternative exons.

Three of our patients were shown to bear a large PKHD1 deletion. While the extent of the deletions differs and seems to be unique for each patient, it may be hypothesised that the underlying pathological mechanism is similar (at least in patients 27 and 309). Owing to repetitive sequences including Alu and other interspersed repetitive elements, as well as simple repeats known to be hotspots for rearrangements of genomic DNA, the affected regions might be prone to recombination, as reported for many other genes and diseases.25–26 Analysis of the sequences spanning the deletion breakpoints carried out by the RepeatMasker documentation program (http://repeatmasker.genome.washington.edu) suggested that the deletion mechanism in patient 309 is MIR (mammalian interspersed repeat) mediated. The 3′ breakpoint lies within a MIRb sequence next to a FLAM_C Alu repeat, while a MIRm SINE and a simple TG repeat are located next to the 5′ transition site. In the immediate vicinity of the breakpoints identified in patients of family 27, AluSx, MER21B, and TG repeats are present. This may contribute to the complex rearrangement of a deletion of more than 13 kb with insertion of three nucleotides at the 3′ breakpoint and duplication of the following 11 base pairs. We also used the repeat masker program to identify whether the genomic region encompassing the deletion in patient 696 had any homology to known human repetitive elements. Homology with several motifs detected was only weak, and cleavage did not occur there. There was no sequence homology between the 5′ and the 3′ junction fragments suggesting a non-homologous recombination event.

From what is known from compiling human mutation data, it is most likely that the exons encompassed in the deletion events in patients 27 and 309 are simply lacking, and flanking exons are fused in mature mRNAs.27 In both cases the corresponding transcripts (missing exons 57 and 58 in patient 309, and exons 58, 59, and 60 in patient 27) would restore the reading frame. It is theoretically possible that an in-frame deletion may just represent a rare apathogenic genetic variant. However, from current knowledge, deletion of exons encoding non-repetitive parts of a protein are not expected to be without clinical consequences (Krawczak M, personal communication). In contrast, it may rather be hypothesised that those missing exons in individuals 27 and 309 are crucial for polyductin function. Given its predicted protein structure (http://expasy.org/) both deletions disrupt the C-terminal PbH1 (parallel β-helix 1) repeats. Although the exact function of these PbH1 motifs remains elusive the consequences of their deletion may indicate their functional importance. The deletion c.3007_3097+12 in patient 696 disrupts the fourth immunoglobulin-like, plexin, transcription factor (IPT) domain and is predicted to cause a frameshift as it affects the first nucleotide of codon 1003 (and exon 27 does not end with a nucleotide in a third codon position). This mechanism was much less likely in patients 27 and 309. Nevertheless, it cannot be excluded that both deletions severely affect and disrupt the splicing machinery and cause considerable changes at a transcript and protein level, finally leading to the premature stop of translation. Substantial clarification would result from investigations on transcript level; however, these could not be carried out as tissues with sufficient PKHD1 expression were not available.

In addition to QRT-PCR analyses, the cohort of 58 ARPKD patients was screened for changes in the predicted alternatively spliced PKHD1 exons by DHPLC. A complex and extensive array of splice variants was previously shown to be encoded by PKHD1.28 However, it is presently unknown how many of these transcripts are actually translated into protein and do have biological functions. In particular, the complicated pattern of splicing was shown to be conserved in the mouse orthologue, pointing to its potential impact.29 In cases were various mRNAs are translated, the PKHD1 gene may encode numerous distinct polyductin like differing in size and amino acid sequence. This hypothesis is further supported by the detection of multiple bands in western blot analysis.30,31 It will be of the utmost importance to establish which isoforms are essential for renal and hepatobiliary integrity to gain a better understanding of the role of polyductin in the aetiology of ARPKD. It may be proposed that mutations disrupt a critical functional stoichiometric or temporal balance between the different protein products that is usually maintained by elaborate, tightly regulated splicing patterns. The identification of PKHD1 mutations is therefore one means of deciphering those exons whose presence in a transcript is essential for the function of polyductin. In the current study, we were able to identify several single nucleotide changes in exons predicted to be alternatively spliced (table 1). While most of these alterations were found to have comparable allele distributions among patients and controls, a few changes (g.157552C→T, g.175388A→C, g.344048A→T, g.406265C→T) were much more prevalent among patients. Thus it may be warranted to hypothesise a potential biological function (at least a modifying effect) for these changes. However, to refer to changes in alternative exons as pathogenic may not be justified. We propose that a putative pathogenic impact of each of those changes must be assessed carefully, given the so far undefined pattern of translated PKHD1 transcripts and their predicted reading frames.

Conclusion

In summary, we identified three large PKHD1 deletions and several single nucleotide changes in alternatively spliced exons. This study shows that gross PKHD1 deletions account for a detectable proportion of ARPKD cases and that screening for major genomic PKHD1 rearrangements will further improve mutation analysis in ARPKD.

DATA ACCESS

PKHD1 – OMIM 600672; 263200 (ARPKD); GDB: 433910; GenBank: NM_138694.2, AF480064, AY074797, AY129465;
ACKNOWLEDGEMENTS

We thank the patients and families involved in these studies for their cooperation. We are grateful to Professor Michael Krawczak for sharing information on human mutation data. The technical assistance of Elvira Goltz-Staggsegmeyer, Edith Burger, Edith von Heel, and Claudia Canales S de Bergmann is gratefully acknowledged. This work was supported by the Deutsche Forschungsgemeinschaft (DFG) (KZ), the German-Israeli Foundation (GIF) (CB), and the START program of the medical faculty of Aachen University (CB). CB is a recipient of a scholarship of the German Kidney Foundation (Deutsche Nierenstiftung).

The supplementary tables can be viewed on the journal website (http://www.humgen.rwth-aachen.de) (PKHD1 Mutation Database).

REFERENCES


Multi-exon deletions of the PKHD1 gene cause autosomal recessive polycystic kidney disease (ARPKD)

C Bergmann, F Küpper, C P Schmitt, U Vester, T J Neuhaus, J Senderek and K Zerres

*J Med Genet* 2005 42: e63
doi: 10.1136/jmg.2005.032318

Updated information and services can be found at:
http://jmg.bmj.com/content/42/10/e63

These include:

**Supplementary Material**

Supplementary material can be found at:
http://jmg.bmj.com/content/suppl/2005/09/29/42.10.e63.DC1

**References**

This article cites 29 articles, 8 of which you can access for free at:
http://jmg.bmj.com/content/42/10/e63#BIBL

**Email alerting service**

Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.

**Topic Collections**

Articles on similar topics can be found in the following collections

- Epidemiology (630)
- Genetic screening / counselling (886)
- Immunology (including allergy) (604)
- Liver disease (51)
- Molecular genetics (1254)
- JMG Online mutation reports (168)

**Notes**

To request permissions go to:
http://group.bmj.com/group/rights-licensing/permissions

To order reprints go to:
http://journals.bmj.com/cgi/reprintform

To subscribe to BMJ go to:
http://group.bmj.com/subscribe/