

## ONLINE MUTATION REPORT

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

*J Med Genet* 2005;42:e63 (<http://www.jmedgenet.com/cgi/content/full/42/10/e63>). doi: 10.1136/jmg.2005.032318

**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.

Autosomal 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.<sup>1–3</sup> 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.<sup>2</sup> 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.<sup>4–5</sup>

The recently characterised *PKHD1* gene (MIM 606702) on chromosome 6p12 is an exceptionally large gene.<sup>6–7</sup> 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.<sup>7–8</sup> 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.<sup>9–14</sup>

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.<sup>15</sup>

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>).<sup>5–9–10–16–21</sup> (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.<sup>5–20–21</sup> 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.<sup>5–20</sup>

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 by quantitative real time PCR (QRT-PCR) 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

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 found<sup>5 16 20</sup> (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.<sup>5</sup> 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>).<sup>20</sup> 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, 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](http://www.jmedgenet.com/supplemental)).

### Quantitative real time PCR

For detection of fragment copy number deviations in *PKHD1*, we developed QRT-PCR assays using SYBR<sup>®</sup> 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 in parallel. As reference locus, we used exon 3 of the factor VIII gene.<sup>22</sup> Primers were designed using the Primer Express<sup>™</sup> 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/supplemental](http://www.jmedgenet.com/supplemental)). QRT-PCR was carried out using the ABI PRISM<sup>®</sup> 7000 sequence detection system, 96-well optical reaction plates, and optical adhesive covers (Applied Biosystems). Reactions were carried out in a volume of 25  $\mu$ l using the qPCR<sup>™</sup> core kit for SYBR<sup>®</sup> Green I according to the manufacturer's protocol (Eurogentec, Seraing, Belgium). All samples were analysed in triplicate, using 15 ng of genomic DNA in each reaction. Reaction conditions corresponded to those described recently.<sup>23</sup> Data evaluation was carried out using the ABI PRISM<sup>®</sup> 7000 sequence detection software according to Wilke *et al.*<sup>22</sup> The DNA concentrations of target and reference loci were quantified using standard curves created for each fragment. We calculated the ratios of test and reference fragments to determine the copy numbers of the test fragments.

### Confinement and verification of *PKHD1* deletions

After detection of deletions of a genomic region spanning in minimum the fragments included in the initial QRT-PCR screen, we aimed at identifying 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 two copies, 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<sup>®</sup> 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.

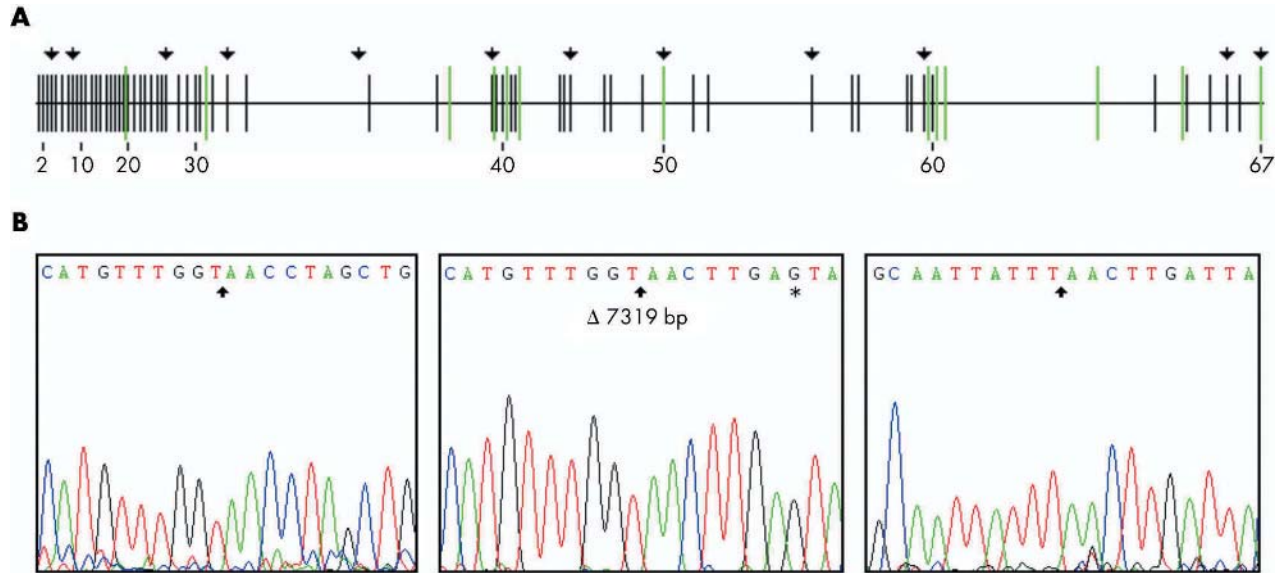
### RESULTS

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).

### 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 at least three independent runs for confirmation. The copy number of all control samples matched in the range for two copies. Thus the observed genomic rearrangements could be excluded as being mere polymorphisms. Among the 58 patients screened, three repeatedly showed decreased copy numbers in favour of a deletion. After screening adjacent exons for copy number deviations, we used a PCR based strategy with direct sequencing of the deletion breakpoints. Finally, all patients of the current study without any copy number deviation were tested by PCR for the three characterised deletions. However, none of those 55 patients nor any of 100 controls showed a deletion, thus excluding a mutation "hotspot" or a polymorphism.

In a previous study, moderately affected patient 309 was found to carry the *PKHD1* missense mutation c.3089C→T (p.Ala1030Val) in exon 27 on his maternal allele.<sup>5</sup> 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).



**Figure 1** (A) Simplified structure of full length *PKHD1* and its alternative exons in approximate scale. In all, 66 exons (ATG initiation codon located in exon 2) encode the 4074 amino acid polyductin protein. Predicted alternatively spliced exons are indicated by green bars (in part overlapping owing to size restrictions) (table S1). Arrows denote the 12 fragments chosen for detection of copy number deviations in *PKHD1* by quantitative real time polymerase chain reaction (QRT-PCR). (B) Genomic sequence encompassing the deletion breakpoints in introns 56 and 58 observed in patient 309. Intron 56 ended at g.332887 (nucleotide position corresponds to that in GenBank AY129465); the sequence of intron 58 broke at position g.340206. Thus in all the deletion was shown to comprise 7319 base pairs (bp) of genomic DNA. Arrows indicate the site of the breakpoints. Left: WT sequence of intron 56; middle: sequence of patient 309 flanking the deletion breakpoints; right: WT sequence of intron 58. The asterisk in the patient's electropherogram marks the polymorphism T→G at position g.340214.

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→T (p.Ser1156Leu) in exon 30 of the *PKHD1* gene on his paternal allele.<sup>5</sup> 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 with primers in exons 26 and 28 was carried out, yielding a product of about 8.5 kb of genomic WT DNA. However, even under optimised electrophoresis conditions we could not

detect any difference between controls' and patient's DNA. Thus the PCR products were tested by restriction enzyme analysis with *EcoRI* (New England Biolabs, Beverly, Massachusetts, USA). For WT DNA this was expected to result in 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/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→T (p. Thr36Met) on their paternal allele.<sup>5 16</sup> 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 proximally 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,

**Table 1** Changes in alternate *PKHD1* exons not included in the longest open reading frame

Exon	Change	Frequency among:	
		Patients	Controls
38/38a (38.1)	g.157552 C→T	1/58	1/100
39b (39.1)	g.175388 A→C	2/58	1/100
39b (39.1)	g.175493 C→G	3/58	2/100
62.1	g.344048 A→T	1/58	1/150
62.1	g.344279 T→C	19/58	31/100
62.2	g.344411 C→G	17/58	21/100
62.3	g.344640 G→C	7/58	33/100
62.3	g.344703 C→T	9/58	13/100
64.1	g.405493 A→T	3/58	6/100
64.4	g.406199 G→A	1/58	2/100
64.4	g.406265 C→T	2/58	1/200
64.4	g.406288 C→G	16/58	21/100
64.5	g.406651 G→C	4/58	5/100
64.6	g.406651 G→C	4/58	5/100

GenBank NM\_138694.2, AY129465.

suggesting 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 (TTGATACTAAG) 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 DHLPC; 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 sites<sup>24-26</sup>; 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.<sup>27-30</sup> 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.<sup>31</sup> 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://au.expasy.org/>) both deletions disrupt the C-terminal PbH1 (parallel  $\beta$ -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*.<sup>7</sup> 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.<sup>8</sup> In cases where various mRNAs are translated, the *PKHD1* gene may encode numerous distinct polypeptides differing in size and amino acid sequence. This hypothesis is further supported by the detection of multiple bands in western blot analysis.<sup>9-13</sup> 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 606702; 263200 (ARPKD); GDB: 433910; GenBank: NM\_138694.2, AF480064, AY074797, AY129465;

<http://www.humgen.rwth-aachen.de> (PKHD1 Mutation Database)

## 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 Golz-Staggemeyer, Edith Bünger, 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 ([www.http:jmedgenet.com/supplemental](http://www.http:jmedgenet.com/supplemental)).

## Authors' affiliations

**C Bergmann, F Küpper, J Senderek, K Zerres**, Department of Human Genetics, Aachen University, Germany

**C P Schmitt**, Division of Paediatric Nephrology, University Children's Hospital Heidelberg, Germany

**U Vester**, Division of Paediatric Nephrology, University Children's Hospital Essen, Germany

**T J Neuhaus**, Division of Paediatric Nephrology, University Children's Hospital, Zurich, Switzerland

Competing interests: none declared

Correspondence to: Dr Carsten Bergmann, Department of Human Genetics, Aachen University, Pauwelsstrasse 30, D-52074 Aachen, Germany; [cbergmann@ukaachen.de](mailto:cbergmann@ukaachen.de)

Received 22 February 2005

Revised version received 28 March 2005

Accepted for publication 31 March 2005

## REFERENCES

- Guay-Woodford LM.** Autosomal recessive polycystic kidney disease: clinical and genetic profiles. In: Watson ML, Torres VE, eds. *Polycystic kidney disease*. Oxford: Oxford University Press, 1996:237–66.
- Zerres K, Rudnik-Schoneborn S, Steinkamm C, Becker J, Mucher G.** Autosomal recessive polycystic kidney disease. *J Mol Med* 1998;**76**:303–9.
- Zerres K, Rudnik-Schoneborn S, Senderek J, Eggermann T, Bergmann C.** Autosomal recessive polycystic kidney disease (ARPKD). *J Nephrol* 2003;**16**:453–8.
- Guay-Woodford LM, Desmond RA.** Autosomal recessive polycystic kidney disease: the clinical experience in North America. *Pediatrics* 2003;**111**:1072–80.
- Bergmann C, Senderek J, Windelen E, Küpper F, Middeldorf I, Schneider F, Dornia C, Rudnik-Schoneborn S, Konrad M, Schmitt CP, Seeman T, Neuhaus TJ, Vester U, Kirfel J, Buttner R, Zerres K, the Arbeitsgemeinschaft für Pädiatrische Nephrologie (APN).** Clinical consequences of *PKHD1* mutations in 164 patients with autosomal recessive polycystic kidney disease (ARPKD). *Kidney Int* 2005;**67**:829–48.
- Ward CJ, Hogan MC, Rossetti S, Walker D, Sneddon T, Wang X, Kubly V, Cunningham JM, Bacallao R, Ishibashi M, Milliner DS, Torres VE, Harris PC.** The gene mutated in autosomal recessive polycystic kidney disease encodes a large, receptor-like protein. *Nat Genet* 2002;**30**:259–69.
- Onuchic LF, Furu L, Nagasawa Y, Hou X, Eggermann T, Ren Z, Bergmann C, Senderek J, Esquivel E, Zeltner R, Rudnik-Schoneborn S, Mrug M, Sweeney W, Avner ED, Zerres K, Guay-Woodford LM, Somlo S, Germino GG.** *PKHD1*, the polycystic kidney and hepatic disease 1 gene, encodes a novel large protein containing multiple immunoglobulin-like plexin-transcription-factor domains and parallel beta-helix 1 repeats. *Am J Hum Genet* 2002;**70**:1305–17.
- Nagasawa Y, Matthiesen S, Onuchic L, Hou X, Bergmann C, Esquivel E, Senderek J, Ren Z, Zeltner R, Furu L, Avner E, Moser M, Somlo S, Guay-Woodford L, Buttner R, Zerres K, Germino GG.** Identification and characterization of *Pkhd1*, the mouse orthologue of the human ARPKD gene. *J Am Soc Nephrol* 2002;**13**:2246–58.
- Masyuk TV, Huang BQ, Ward CJ, Masyuk AI, Yuan D, Splinter PL, Punyashthiti R, Ritman EL, Torres VE, Harris PC, LaRusso NF.** Defects in cholangiocyte fibrocystin expression and ciliary structure in the PCK rat. *Gastroenterology* 2003;**125**:1303–10.
- Ward CJ, Yuan D, Masyuk TV, Wang X, Punyashthiti R, Whelan S, Bacallao R, Torra R, LaRusso NF, Torres VE, Harris PC.** Cellular and subcellular localization of the ARPKD protein; fibrocystin is expressed on primary cilia. *Hum Mol Genet* 2003;**12**:2703–10.
- Zhang MZ, Mai W, Li C, Cho SY, Hao C, Moeckel G, Zhao R, Kim I, Wang J, Xiong H, Wang H, Sato Y, Wu Y, Nakanuma Y, Lilova M, Pei Y, Harris RC, Li S, Coffey RJ, Sun L, Wu D, Chen XZ, Breyer MD, Zhao ZJ, McKanna JA, Wu G.** *PKHD1* protein encoded by the gene for autosomal recessive polycystic kidney disease associates with basal bodies and primary cilia in renal epithelial cells. *Proc Natl Acad Sci USA* 2004;**101**:2311–16.
- Wang S, Luo Y, Wilson PD, Witman GB, Zhou J.** The autosomal recessive polycystic kidney disease protein is localized to primary cilia, with concentration in the basal body area. *J Am Soc Nephrol* 2004;**15**:592–602.
- Menezes LF, Cai Y, Nagasawa Y, Silva AM, Watkins ML, Da Silva AM, Somlo S, Guay-Woodford LM, Germino GG, Onuchic LF.** Polyductin, the *PKHD1* gene product, comprises isoforms expressed in plasma membrane, primary cilium, and cytoplasm. *Kidney Int* 2004;**66**:1345–55.
- Watnick T, Germino G.** From cilia to cyst. *Nat Genet* 2003;**34**:355–56.
- Bergmann C, Kupper F, Dornia C, Schneider F, Senderek J, Zerres K.** Algorithm for efficient *PKHD1* mutation screening in autosomal recessive polycystic kidney disease (ARPKD). *Hum Mutat* 2005;**25**:225–31.
- Bergmann C, Senderek J, Sedlacek B, Pegiazoglou I, Puglia P, Eggermann T, Rudnik-Schoneborn S, Furu L, Onuchic LF, De Baca M, Germino GG, Guay-Woodford L, Somlo S, Moser M, Buttner R, Zerres K.** Spectrum of mutations in the gene for autosomal recessive polycystic kidney disease (ARPKD/*PKHD1*). *J Am Soc Nephrol* 2003;**13**:76–89.
- Furu L, Onuchic LF, Gharavi A, Hou X, Esquivel EL, Nagasawa Y, Bergmann C, Senderek J, Avner E, Zerres K, Germino GG, Guay-Woodford LM, Somlo S.** Milder presentation of recessive polycystic kidney disease requires presence of amino acid substitution mutations. *J Am Soc Nephrol* 2003;**14**:2004–14.
- Rossetti S, Torra R, Coto E, Consugar M, Kubly V, Malaga S, Navarro M, El-Youssef M, Torres VE, Harris PC.** A complete mutation screen of *PKHD1* in autosomal-recessive polycystic kidney disease (ARPKD) pedigrees. *Kidney Int* 2003;**64**:391–403.
- Bergmann C, Senderek J, Kupper F, Schneider F, Dornia C, Windelen E, Eggermann T, Rudnik-Schoneborn S, Kirfel J, Furu L, Onuchic LF, Rossetti S, Harris PC, Somlo S, Guay-Woodford L, Germino GG, Moser M, Buttner R, Zerres K.** *PKHD1* mutations in autosomal recessive polycystic kidney disease (ARPKD). *Hum Mutat* 2004;**23**:453–63.
- Bergmann C, Senderek J, Schneider F, Dornia C, Kupper F, Eggermann T, Rudnik-Schoneborn S, Kirfel J, Moser M, Buttner R, Zerres K.** *PKHD1* mutations in families requesting prenatal diagnosis for autosomal recessive polycystic kidney disease (ARPKD). *Hum Mutat* 2004;**23**:487–95.
- Sharp AM, Messiaen LM, Page G, Antignac C, Gubler MC, Onuchic LF, Somlo S, Germino GG, Guay-Woodford LM.** Comprehensive genomic analysis for *PKHD1* mutations in ARPKD cohorts. *J Med Genet* 2005;**42**:336–49.
- Wilke K, Duman B, Horst J.** Diagnosis of haploidy and triploidy based on measurement of gene copy number by real-time PCR. *Hum Mutat* 2000;**16**:431–6.
- Anhuf D, Eggermann T, Rudnik-Schoneborn S, Zerres K.** Determination of *SMN1* and *SMN2* copy number using TaqMan technology. *Hum Mutat* 2003;**22**:74–8.
- King K, Flinter FA, Nihalani V, Green PM.** Unusual deep intronic mutations in the *COL4A5* gene cause X linked Alport syndrome. *Hum Genet* 2002;**111**:548–54.
- Modrek B, Lee C.** A genomic view of alternative splicing. *Nat Genet* 2002;**30**:13–19.
- Modrek B, Lee CJ.** Alternative splicing in the human, mouse and rat genomes is associated with an increased frequency of exon creation and/or loss. *Nat Genet* 2003;**34**:177–80.
- Stoppa-Lyonnet D, Carter PE, Meo T, Tosi M.** Clusters of intragenic Alu repeats predispose the human *C1* inhibitor locus to deleterious rearrangements. *Proc Natl Acad Sci USA* 1990;**87**:1551–5.
- Morgan NV, Tipping AJ, Joenje H, Mathew CG.** High frequency of large intragenic deletions in the Fanconi anemia group A gene. *Am J Hum Genet* 1999;**65**:1330–41.
- Kutsche K, Ressler B, Katzera HG, Orth U, Gillissen-Kaesbach G, Morlot S, Schwinger E, Gal A.** Characterization of breakpoint sequences of five rearrangements in *L1CAM* and *ABCD1* (ALD) genes. *Hum Mutat* 2002;**19**:526–35.
- McNeil N.** AluElements: repetitive DNA as facilitators of chromosomal rearrangement. *J Assoc Genet Technol* 2004;**30**:41–7.
- Black DL.** Mechanisms of alternative pre-messenger RNA splicing. *Annu Rev Biochem* 2003;**72**:291–336.

**Table S2. Primers Quantitative Real-Time PCR**

Fragment	Size (bp)	Primer sequences		Genomic DNA nucleotide	
		Forward (5'-3')	Reward (5'-3')	From	To
4	118	TTACCCCAACAATGGCTCTCA	CACCACAGGCAAATCCAAGAA	5049	5166
8	155	TCAGCCTTGGTCTGGGATAC	CATGGCTCATGCTTTCACAC	15044	15198
27	102	CCATCAGTGCCACTGGAGAA	AGACCCTCCCAGATTACCAA	44630	44731
34	143	TGTTTCCTGCAGTGTCTTG	GAGGTAAGGCCACGATTCAA	69931	70073
35	121	TTGTCAGCCAGGAGACTG	ATCAGCAAGTTGGAACACCC	123260	123380
38	111	AACTGCCCATGCCCTAGACA	ATCGGTTTGGCACCTTTAACA	175074	175184
46	155	AAGCTACAGGAAGTCCGCAA	TTGCATCACTTTCCAAGACG	204319	204473
50	94	GTGGGAGTCGAGTGGGTCTGT	AGTTGCCTCAGCTGGCTATTG	239734	239827
53	86	ATGTTCTGAGTGTGGCATGCA	CTCAACATGCTCGCAATCCTT	296327	296412
58	110	ATCAACCTCCATGGCAACGT	GAATGCGCCACATTGTCAGA	339262	339371
65	123	ATCCAAGCCATTTGCTGTCT	TTAAGCCAACAGCACACCAG	454909	455031
67	134	GGCAAAGTCGGCTGAGTAAG	CCCTGATCACAGTTGCTCCT	468405	468538
Factor VIII (exon 3)	87	CTACCATCCAGGCTGAGGTTTATG	CACCAACAGCATGAAGACTGACA		

GenBank NM\_138694.2, AY129465

**Table S1. Primers and DHPLC conditions for alternate *PKHD1* exons**

Exon	Size (bp)	Primer sequences		Genomic DNA nucleotide		DHPLC conditions	
		Forward (5'-3')	Reward (5'-3')	From	To	Temp.	Initial % buffer B
20a (20)	261	GGATGTGGACTCCTCACTACG	AATCCTCCCAGCTGACTGAA	33325	33585	58.6, 62	56, 54
32a (32A)	353	CCACAGGTGCTATGAATTC	GTGACTACTGGTGTGCTG	61321	61673	61, 63.7	57, 56
32a (32B)	365	AAGGCTTCACTTTTCATGTATGAA	CCTCACAGTAAGTATGGTCC	61628	61992	59, 60.6	61, 59
32b (32D)	367	CTTCCTTGCAATGTAACTTTT	GATACCAAAGTCCATCTACCTC	62332	62698	57.9, 58.9	59, 59
32b (32E)	441	CTCATCCGGTGCATTGTTC	CCAGAAGTGAAAGGAGCTACCA	62626	63066	58, 62	60, 55
32c (32E)	441	CTCATCCGGTGCATTGTTC	CCAGAAGTGAAAGGAGCTACCA	62626	63066	58, 62	60, 55
<b>38/38a (38.1)</b>	327	TGGGGTTTGGGTAGCTGTAG	CCAAGGCCATGAGACTGTTT	157391	157717	57.6, 59.6	58, 55
<b>38/38a (38.2)</b>	348	GCAGGATGGTGGAGCTTTTA	GCAGACACTTGGCATTTTCA	157589	157936	56.3, 57	60, 59
39a/39b (38)	332	AAATTATCTGGACAACTTTTCTC	AGACCCCAATACAAATGTCCA	175001	175332	59, 60.5	58, 56
<b>39b (39.1)</b>	445	GGTACTGAGTGTGACGTTGGC	CAGCCACCCAATTCTCTGTAA	175247	175691	54.9, 55.9	60, 59
39b (39)	278	TTTGAGTGATGTCCCTCAGTTCT	GCAATGCCATCTATCATCAGAC	175612	175889	57, 59.5	57, 55
<b>39b (39.2)</b>	328	CTCACTAATGAGAGGGAGAAGC	TTCTTTAATAGAAATATTTTCCCCAGA	175757	176084	57.3	57
41a (40)	286	CATGCTTTAGTTCTCTGGACTTT	TGCCTTAAACATGGGAGAAA	178101	178386	58.5, 60.5, 61.5	57, 56, 54
<b>41a (41.1)</b>	400	GAGCTATGAGAGGTAAGGGGGT	CACAAAGCTAGTAAGTGGTAGAGAGACA	178319	178718	56.4	60
<b>41a (41.2)</b>	381	GATTCTGGAACAGGCAGTCC	TTCTGTTTTTCTGATTCTGATCTTG	178598	178978	57.8, 58.8	59, 58
44a (43)	258	TCAAGGGTGTGAGTTGAGTACA	TTGAGAAAGAACTTTATGCCCTC	183819	184076	56.9, 58.9	56, 54
51a/51b (50)	300	TGATGGGGTTCCTTACTAAATGA	TGGTGATTCTGACTATGTGCTC	239591	239890	58.3, 59.3, 60.3	57, 56, 55
60a (59)	285	GTGGCTGGTGGTTTATATG	GGGTTTGAAGAATTGCCAAG	340670	340954	54.1, 57.4, 58.7	57, 55, 52
<b>62.1</b>	378	TGTCCTGAGACCTTTTTCTTC	TAAAGGGAATGGTGAAGGGTG	343981	344358	55.7, 57.2	60, 58
<b>62.2</b>	377	CCTGGGAGTCATCTTTTATTTTC	CAAACGCCCTGATGTAGTGA	344268	344644	55.5, 58.1	60, 57
<b>62.3</b>	440	CGCCTTTTACTTACTTATCCTGG	CAGCAAACCTTCTGGCTGG	344570	345009	54.7, 57.5	60, 57
<b>63</b>	328	CCACAGTCTGTTGGATGAAAAA	TCTGGGTGAGTAGTTTTGTACG	346942	347269	54.2, 56.5	58, 56
<b>64.1</b>	341	AAACTTGTGTTTAGAGAGAAAATTGTG	TCTCTTGAGACAGTGACCAGAA	405320	405660	56.7, 57.2	58, 57
<b>64.2</b>	323	TTCTTGTCATCTTTGGTTGG	GAGGTGGTGAACATAAGCCA	405578	405900	57.5, 58	58.57
<b>64.3</b>	349	TGCGTATAGCTGTGTTCCAA	TGAGTTAAAGAGAACACTGAAGGG	405815	406163	55.8, 56.7, 58.5	58, 57, 55
<b>64.4</b>	339	CATGCTTGATTAAGGCTACTGG	TTCAGTCTGTGGAATGTCA	406098	406436	56, 57.2, 58.8	59, 58, 56
<b>64.5</b>	326	AAGCTGATCTTGGGAAAGGA	GAACCAGTAGGGGACAGAA	406365	406690	56.1, 57.4	58, 57
<b>64.6</b>	346	AACCCCAATTCACAACCAAC	TGGACAGTAATCAACATGAACAA	406630	406975	53.5, 55.5	58, 56
66a (62)	273	AATGGATTGTGGAAAATTGCT	GGCTGAATGCTACATGCTACTT	438315	438587	55.3, 57.3	56, 55
71a (67B)	394	TGCAAGAGACTGGGAAGTGG	CTGGGAACATTCTGCCTTTC	468285	468678	61.1, 62.4	57, 56

GenBank NM\_138694.2, AY129465

Primers of the longest *PKHD1* ORF amplifying (parts of) alternate exons are given in parentheses.

Predicted exons not included in the longest ORF are shown in bold type.