Comprehensive Molecular Diagnostics in Autosomal Dominant Polycystic Kidney Disease

Sandro Rossetti,* Mark B. Consugar,* Arlene B. Chapman,† Vicente E. Torres,* Lisa M. Guay-Woodford,‡ Jared J. Grantham,§ William M. Bennett,‖ Catherine M. Meyers,¶ Denise L. Walker,* Kyongtae Bae,** Qin (Jean) Zhang,†† Paul A. Thompson,†† J. Philip Miller,†† Peter C. Harris;* and the CRISP Consortium

*Division of Nephrology and Hypertension, Mayo Clinic College of Medicine, Rochester, Minnesota; †Division of Nephrology, Emory University, Atlanta, Georgia; ‡Division of Genetics and Translational Medicine, University of Alabama at Birmingham, Birmingham, Alabama; §Kidney Institute and the Department of Internal Medicine, Kansas University Medical Center, Kansas City, Kansas; ‖Northwest Renal Clinic, Portland, Oregon; ¶National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland; **Department of Radiology, University of Pittsburgh Medical Center, Pittsburgh, Pennsylvania; and ††Department of Medicine and Division of Biostatistics, Washington University School of Medicine, St. Louis, Missouri

ABSTRACT

Mutation-based molecular diagnostics of autosomal dominant polycystic kidney disease (ADPKD) is complicated by genetic and allelic heterogeneity, large multi-exon genes, duplication of *PKD1*, and a high level of unclassified variants (UCV). Present mutation detection levels are 60 to 70%, and *PKD1* and *PKD2* UCV have not been systematically classified. This study analyzed the uniquely characterized Consortium for Radiologic Imaging Study of PKD (CRISP) ADPKD population by molecular analysis. A cohort of 202 probands was screened by denaturing HPLC, followed by direct sequencing using a clinical test of 121 with no definite mutation (plus controls). A subset was also screened for larger deletions, and reverse transcription–PCR was used to test abnormal splicing. Definite mutations were identified in 127 (62.9%) probands, and all UCV were assessed for their potential pathogenicity. The Grantham Matrix Score was used to score the significance of the substitution and the conservation of the residue in orthologs and defined domains. The likelihood for aberrant splicing and contextual information about the UCV within the patient (including segregation analysis) was used in combination to define a variant score. From this analysis, 44 missense plus two atypical splicing and seven small in-frame changes were defined as probably pathogenic and assigned to a mutation group. Mutations were thus defined in 180 (89.1%) probands: 153 (85.0%) *PKD1* and 27 (15.0%) *PKD2*. The majority were unique to a single family, but recurrent mutations accounted for 30.0% of the total. A total of 190 polymorphic variants were identified in *PKD1* (average of 10.1 per patient) and eight in *PKD2*. Although nondefinite mutation data must be treated with care in the clinical setting, this study shows the potential for molecular diagnostics in ADPKD that is likely to become increasingly important as therapies become available.

J Am Soc Nephrol 18: 2143–2160, 2007. doi: 10.1681/ASN.2006121387

Autosomal dominant polycystic kidney disease (ADPKD) is the most common inherited kidney disease, with an incidence of 1 in 400 to 1000 (accounting for approximately 5% of ESRD), and is characterized by the development and progressive enlargement of cysts in the kidney. ADPKD is genetically heterogeneous, with two genes identified: *PKD1* (16p13.3) and *PKD2* (4q21).[1–4] In linkage-characterized populations, *PKD1* accounts for approximately 85% of cases and *PKD2* accounts for

Received December 20, 2006. Accepted April 1, 2007.

Published online ahead of print. Publication date available at www.jasn.org.

Correspondence: Dr. Peter C. Harris, Division of Nephrology, Mayo Clinic, 200 First Street SW, Rochester, MN 55905. Phone: 507-266-0541; Fax: 507-266-9315; E-mail: harris.peter@mayo.edu

most of the remainder,[5,6] but further heterogeneity is possible.[7] *PKD1* has an average age at ESRD of 54.3 yr, compared with 74.0 yr for *PKD2*.[8] *PKD1* and *PKD2* encode polycystin-1 (PC1) and polycystin-2 (PC2), respectively. PC2 is a TRP channel that may be involved in regulating intracellular Ca^{2+}.[9,10] PC1 and PC2 interact and, similar to other cystogenic proteins, have been localized to primary cilia.[11,12] This complex may act as a flow-dependent mechanosensor that regulates the differentiated state of tubular epithelial cells.[13]

The diagnosis of ADPKD is typically determined by renal imaging with age-related cyst number criteria established for a diagnosis by ultrasound.[14] Computed tomography and magnetic resonance imaging can also quantify renal cystic disease, and a recent trial, the Consortium for Radiologic Imaging Study of PKD (CRISP), showed that magnetic resonance imaging is a reliable means to monitor disease progression through renal enlargement.[15,16] Image-based diagnostics is highly reliable in older individuals (>30 yr) but less certain in young adults and can be equivocal in the case of the young, living-related kidney donor. Identification of the ADPKD genes has allowed molecular diagnostics. Linkage studies are of limited utility because of the genetic heterogeneity and common cases with a negative family history. Diagnostics by mutation analysis has been challenging because of the large size and multi-exon structures of the genes, genomic duplication of *PKD1*, marked allelic heterogeneity, and common missense variants.[17]

PKD1 has 46 exons and a coding region of approximately 13 kb; the area that contains exons 1 to 33 is duplicated six times more proximally on chromosome 16 (the *HG* loci).[1,2] Because of the high level of homology between *PKD1* and the *HG*, protocols for locus-specific amplification of *PKD1* are required for analysis.[17] *PKD2* has 15 exons and a coding region of approximately 3 kb. Because of these complications, few complete screens of both genes have been described, and detection levels are at 60 to 70%.[17] A total of 270 different *PKD1* and 73 *PKD2* mutations are described in the Human Gene Mutation Database (HGMD), illustrating the high level of allelic heterogeneity; most mutations are unique to a single family.[17–19] The majority of changes are predicted to truncate the proteins, although a significant number of missense changes have been described.

Whereas the pathogenicity of frame-shifting, nonsense, typical splicing, or large re-arrangements is usually clear, missense variants, atypical splicing changes, or small in-frame deletions (unclassified variants [UCV]), need further evaluation. The Grantham Matrix Score (GMS)[20] has been used to score the significance of amino acid substitutions (the Grantham Distance [GD]) and the conservation of the residue in a multisequence alignment (MSA) of orthologous proteins (Grantham Variation [GV]).[21–24] This evaluation, integrated with family, population (contextual), and histopathologic data, has been used to score UCV at *BRCA1* and *BRCA2* for diagnostics.[25] Similar approaches have been applied to the autosomal recessive PKD gene.[26]

Here, we describe a comprehensive mutation screen and evaluation of UCV in the CRISP ADPKD population. These data are essential to analyze the results of the study,[15,27] and similar approaches will be required to interpret future ADPKD therapeutic trials. Although the demand for molecular diagnostics in ADPKD is limited, the development of therapies is likely to transform the situation as a firm diagnosis in young adults before significant renal changes have occurred will be required.[10]

RESULTS

Mutation Screening

DNA for mutation screening was available from 202 families who had at least one individual enrolled in CRISP. Thirty-two families were multiplexed within CRISP, giving a total of 239 patients. The initial *PKD1* and *PKD2* screening of the 202 probands by denaturing HPLC (DHPLC) identified 81 definite mutations (nonsense, frame-shifting deletions or insertions, typical splicing, or in-frame changes of five or more amino acids [aa]); defined as mutation group (MG) = A in Table IA). A second round of screening by direct sequencing (DS) of both genes in all patients in whom no definite mutation was detected, plus 29 control subjects (a total of 150), revealed an additional 37 definite mutations. In addition, analysis of a subset by field inversion gel electrophoresis (FIGE) and visualization of *PKD1* long-range PCR (LR-PCR) products revealed four large deletion mutations, giving 122 probands with MG = A (Table IA).

In four atypical splicing cases, RNA was available to test the splicing predictions, and in a fifth, IVS43+14del20, skipping of exon 43 had been previously demonstrated.[28] In two patients, substitutions close to the end of IVS15 generated novel AG dinucleotides predicted to form novel splice acceptors and cause frame-shifting mutations. Reverse transcription–PCR (RT-PCR) and sequencing confirmed this abnormal splicing (Figure 1A). A similar mutation, IVS37-10C→A, was predicted to abolish the normal acceptor site by weakening the poly-pyrimidine tract (Figure 1B) and RT-PCR showed skipping of exon 38, plus a product that included the final 180 bp of IVS37. In the final case, *PKD2* IVS4+3delAAGT, the splice donor site was predicted to be destroyed (Figure 1C). No abnormality was seen with RT-PCR across the region, but amplification with a primer in IVS4 showed that all or part of IVS4 is incorporated into the sequence. These mutations were therefore classed as MG = A, giving a total of 127 families with definite mutations.

Scoring of UCV

All UCV that were not previously described as polymorphisms were analyzed and evaluated as possible disease-associated mutations. Those variants that ultimately were scored as probably pathogenic are summarized in Table II. The detailed scoring of UCV is shown in Table II (probably pathogenic), Table III (indeterminate), and Table IV (likely neutral polymorphisms). The algorithm that was used to evaluate the UCV is described in detail in the Materials and Methods section. To illustrate how this algorithm works, two examples are provided here. The *PKD1* change

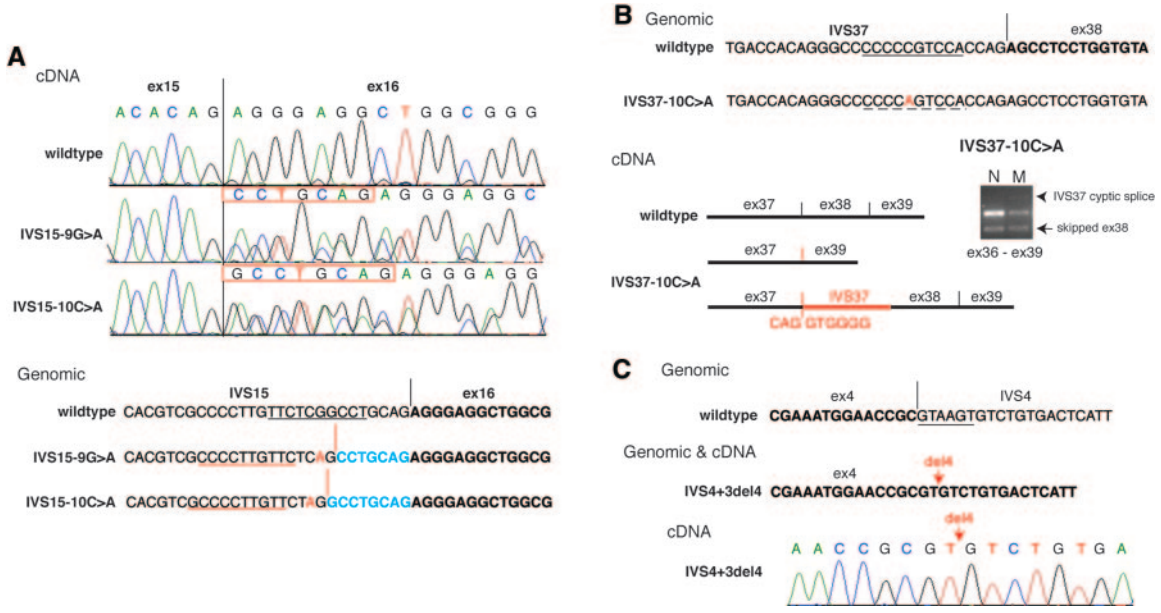

Figure 1. Examples of atypical splicing events in *PKD1* (A and B) and *PKD2* (C). (A, top) Diagram of wild-type and mutant cDNA sequence illustrating the effects of two IVS15 substitutions. Novel sequence included in the transcript is boxed in red. (Bottom) Genomic sequence from these cases. Exonic sequence is in boldface, intron/exon boundaries are indicated with vertical lines, and polypyrimidine tracts are underlined. Substituted nucleotides are in red, and novel acceptor sites and polypyrimidine tracts are shown with red vertical and horizontal lines, respectively. Novel exonic sequence is in blue boldface. (B) Wild-type and mutated sequence of IVS37-10C→A. This substitution weakens the normal polypyrimidine tract (7/10; dashed black line) and eliminates the normal acceptor site. Reverse transcription–PCR (bottom right: M, mutant; N, normal) shows that exon 38 is skipped (the skipped product is also found at a low level normally) and a product including part of IVS37 is also seen. A diagram shows the wild-type and mutant transcripts (bottom left); the genomic sequence of the novel donor site in IVS37 is in red. (C) Diagram of the wild-type and mutant sequence of the *PKD2* mutant IVS4+3del4 (top) and mutant cDNA sequence (bottom). This deletion destroys the normal donor site (underlined), and cDNA sequence (bottom) shows that a product including part of IVS4 is present in the mutant.

G381S is moderately conservative but at a highly conserved site, giving a GD/GV score of +5. The change is not in a conserved domain, and no abnormal splicing was predicted (+0). This change was found in three different families, two of which had no other likely mutation and one had the indeterminate variant R3063C (+5). Segregation was demonstrated in one family with nine affected individuals (+4); and another mutation, G381C, has been described at this residue (+1) giving a variant score (VS) = +15. This is a high score (MG = B) reflecting the recurrent nature of the change, demonstrated segregation, and highly conserved residue that is mutated. In contrast, the same substitution at another site, G3651S, is less clearly pathogenic. The GD/GV score is the same (+5), but it is not in a conserved domain, it is not predicted to change splicing, and segregation has not been demonstrated (+0). No other variant was found in this individual (+2), giving a VS = +7 (MG = C). This lower level of certainty reflects that the substitution is relatively conservative and that the mutation is novel.

As a result of this analysis, 44 missense, two atypical splicing, and seven small in-frame deletions were defined as probably pathogenic, giving a total 180 (89.1%) of 202 with a probable mutation characterized. In 37 probands, a highly likely mutation (MG = B) was defined, with a likely mutation (MG = C) in another 16 (Table II). The pathologic significance of 22 additional variants was not clear, and they were defined as indeterminate (I; Table III). The remainder of novel variants were defined as likely neutral polymorphisms and are shown in Table IV.

The most important factor in determining whether a missense change was likely pathogenic was the degree to which it was conserved in orthologs and in other proteins with the same domain, with the GD almost always substantially greater than the GV (Table II). This contrasted with the 48 newly described neutral variants (Table IV), where the residue was rarely conserved and the GV was often higher than the GD. Recurrence of a variant in two or more patients with no other clear mutation also strongly supported a pathogenic likelihood. A factor that was important in excluding UCV as likely pathogenic was the finding of a more likely mutation in the same patient (Table IV). Four *PKD1* UCV that previously were described as mutations were defined as indeterminate, R1340W and R4276W, or neutral polymorphisms, R324L and R2200C (Tables III and IV). The most interesting polymorphism was a 3-aa deletion (*PKD1*: 2894delANS) that did not segregate with the disease in a family with a segregating *PKD2* missense change, R325Q. Analysis of orthologs in this area (Figure 2A) shows

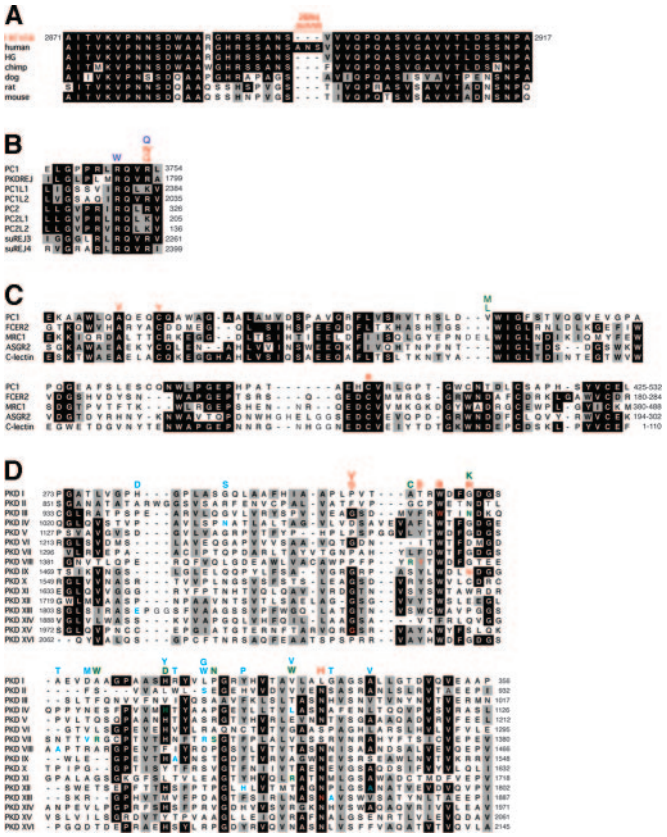

Figure 2. Multisequence alignment illustrating how comparative analysis can inform about the likely pathogenicity of variants. (A) Polycystin (PC1) sequence from CRISP case 187456 (pedigree 120010), mammalian orthologs, as indicated, and an example of the HG sequence, showing the position of the 2894delANS variant. This 3–amino acid region is present just in the human sequence. (B) Alignment of human polycystin-1 PC1 and PC2 homologs and sea urchin REJ sequences, as indicated, in the region of the polycystin-A motif (PC-A).[29] Four missense mutations, two PC1 (red; R3753W and R3753Q) and two PC2 (blue; R322W and R325Q), affect highly conserved basic residues in this motif. (C) C-type lectin domains of human PC1, low-affinity Ig epsilon Fc receptor (FCER2), macrophage mannose receptor 1 (MRC1), asialoglycoprotein receptor 2 (ASGR2), and C-lectin consensus (Pfam). Mutations are shown in red (A432V, C436Y, and C508R), and indeterminate changes are shown in green (V466L and V466M). (D) Alignment of the PKD repeats of human PC1. The position of pathogenic changes are shown in red (W967R, Y1412D, G1503R, N1870H, G1999S, and G1999V), indeterminate changes are shown in green (N970K, H1093D, R1340W, S1352N, R1411C, and R1698W), and neutral changes are shown in blue (S903G, N10345, H1093Y, L1106V, V1339M, R1351W, A1422T, A1516T, H1777P, A1790V, E1811D, and A1871T). The position of the residue changed is shown colored.

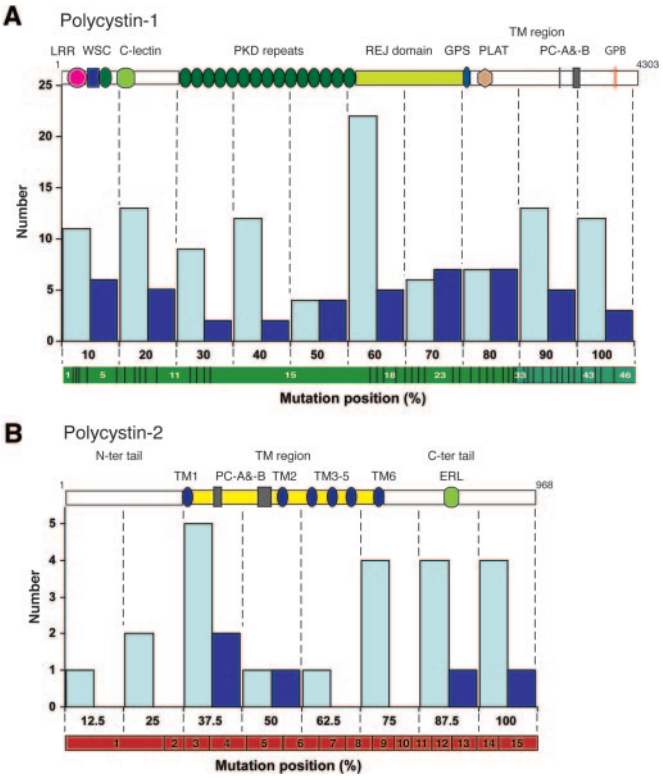

Figure 3. Positions of mutations in *PKD1* (A) and *PKD2* (B). Graph of number of mutations, truncating (light blue) or in-frame (dark blue), in intervals of the genes (*PKD1*, 10 divisions; *PKD2*, eight divisions). Corresponding exonic regions of the transcript are shown below each graph (*PKD1*, green; *PKD2*, red) with exon number indicated. Protein structures with domain regions indicated are shown above the graphs for PC1 (A) and PC2 (B). Only truncating mutations in *PKD1* were distributed significantly differently from uniform (expected per interval 10.7 mutations).

that this region is present only in the human sequence, not in the *HG* pseudogenes or other mammals. The presence of these three residues is, therefore, the result of a recent duplication in the human sequence. Segregation analysis also helped to identify likely mutations; for instance, in case 380166, two possible changes were found (F3168L and R2477C), but only the former segregated with the disease.

Characteristics of Mutations

Mutation analysis of the CRISP population defined 85.0% as PKD1 and 15.0% as PKD2. These values correspond closely to previous studies,[5,6] although these are the first data that are based on mutation detection. The breakdown of mutation type in the two genes is summarized in Table 1. Overall, >70% of mutations were predicted to truncate the protein, >80% in PKD2. Although the majority of mutations were novel, 27 had previously been described and seven others were seen more than once in the study; 30% of families had a recurrent mutation (Table 1). The most common mutation in the study was 5014delAG (five [2.5%] families), followed by Q2556X (four [2.0%] families); two truncating and two missense changes were found three times, and six mutations were found twice (see Table I for details).

Table 1. Summary of mutations in the CRISP families

Mutation Type	*PKD1*	*PKD2*	Total
Definite (MG = A)	106 (69.3%)	21 (77.8%)	127 (70.6%) (62.9%)[c]
Highly likely (MG = B)	34 (22.2%)	3 (11.1%)	37 (20.6%) (18.3%)[c]
Likely (MG = C)	13 (8.5%)	3 (11.1%)	16 (8.9%) (7.9%)[c]
FS deletion/insertion	49 (32.0%)	7 (25.9%)	56 (31.1%)
Nonsense	38 (24.8%)	9 (33.3%)	47 (26.1%)
Splicing	16 (10.5%)	6 (22.2%)	22 (12.2%)
IF deletion/insertion	9 (5.9%)	2 (7.4%)	11 (6.1%)
Missense	41 (26.8%)	3 (11.1%)	44 (24.4%)
Truncating[a]	107 (69.9%)	22 (81.5%)	129 (71.7%)
Small IF	46 (30.1%)	5 (18.5%)	51 (28.3%)
Recurrent within study[b]	7 (16)	0	7 (16)
Recurrent[b]	19 (27)	8 (11)	27 (38)
Total recurrent[b]	26 (43; 28.1%)	8 (11; 40.7%)	34 (54; 30.0%)
Novel	110 (71.9%)	16 (59.3%)	126 (70.0%)
Total different mutations	136	24	160
Total mutations	153 (85.0%)	27 (15.0%)	180 (89.1%)[c]
No mutation defined			22 (10.9%)[c]

[a]Including IF changes of five amino acids or greater and atypical splicing.
[b]Different mutations: Total families in parenthesis.
[c]Percentage of total families in study.

Important Functional Residues in PC1 and PC2

The large number of missense changes identified several residues or regions that were disrupted more than once, suggesting particular functional significances. In the PC-A domain (in the first extracellular loop of PC2 and third extracellular loop of PC1[29]), two basic residues were substituted a total of four times (two PC1 and two PC2; Figure 2B) and four additional times in the literature (see Table I). Several mutations were also seen in the C-type lectin with three highly conserved residues substituted with a nonconservative or, in one case, a relatively conservative change (Figure 2C). Two other interesting substitutions, L4137P and L4139P, are predicted to disrupt the helical structure of the G-protein binding region.[30] A clear indication of how conservation within a domain helped determine whether a substitution is likely pathogenic is shown in the PKD repeats (Figure 2D). Likely pathogenic changes were usually nonconservative in nature and at well-conserved residues, whereas the likely neutral or indeterminate changes were most often found at nonconserved residues and were often conservative changes (in some cases matching the residue found in other PKD repeats).

Mutation Location

Mutations were spread throughout both the *PKD1* and *PKD2* genes with at least three mutations found in each 5% interval across *PKD1* (except one with a single event) and at least one in each 12.5% interval across *PKD2* (Figure 3). Of the 46 *PKD1* exons (including immediate flanking regions), 40 harbored mutations, as did nine of the 15 in *PKD2*. Analysis of the distributions show that they were NS differently from uniform in *PKD1* ($P > 0.05$) or *PKD2* ($P > 0.10$). The median mutation position in *PKD1* was 6873nt (median gene position 6454.5nt) with 55.6% of mutation in the 3′ half of the gene. Missense changes were relatively evenly distributed in *PKD1* ($P > 0.10$), but truncating changes were nonuniform ($P < 0.01$; Figure 3A). In particular, there was a cluster of truncations at the end of exon 15 to exon 19, corresponding to the junction of the PKD repeats and the REJ domain. This clustering was not due to recurrent mutations (Table I). IVS21 has previously been suggested as a possible hotspot for mutations as a result of a long polypyrimidine tract,[31] and although this could be a factor, the peak did not correspond precisely to that region. Too few missense changes were seen in *PKD2* to determine the significance of the distribution (Figure 3B).

Mutation-Negative Cases

Twenty-two families did not have a probable mutation defined (Table V). In 10 of these, no variants other than known polymorphisms were detected, whereas in five others, only a single unique synonymous change was detected, which was not predicted to change splicing significantly. Two cases had a substitution of the same residue, V466, suggesting significance; however, both changes (to leucine or methionine) were highly conservative, and this residue is not well conserved (leucine in frog) and variable in the C-type lectin (Figure 2C). The unique change S1352N was similarly not scored as a mutation, because the substitution matched that found in fish, whereas another, L2696P, was rejected because of the highly nonconservative nature of this site in orthologs (see Table IIIA). Two cases, 103227 (V609G and N970K) and 118641 (V690G and R1340W), had two indeterminate changes, and a third had a single such change (476972; H1093D). V690G is at a residue previously described as the site of a pathogenic change (V690D), whereas H1093D is a nonconservative change at the site of a likely polymorphism (H1093Y). Although these both

are suggestive of pathogenic changes, they did not meet the threshold that we set in this article (Table IIIA). Some cases may have larger deletions because not all were screened by FIGE, and, indeed, suggestive changes (dosage or possible aberrant fragments) or regions of apparent homozygosity (hemizygosity?) were seen in four cases, but the DNA changes have not been characterized (Table V).

Polymorphic Variants

PKD1 is a highly polymorphic gene with 194 different neutral or indeterminate variants detected within the coding and immediate IVS flanking regions. The corresponding figure for *PKD2* was 8 (see Table IV for novel nonsynonymous changes). The average number of variant alleles per patient in *PKD1* was 10.1, with a range from 0 to 55, whereas the corresponding figures for *PKD2* were 0.8 (range 0 to 3). Only one polymorphic variant in *PKD2* had a minor allele frequency >5% (R28P, 30.9% of alleles). In *PKD1*, the majority of variants were rare; 94 (51.1%) were found just once within the study, with 24 (12.6%) having a minor allele frequency >5%. The greatest variability in *PKD1* was in the 26 black patients, in whom an average of 20.6 variant alleles was found. The high level of polymorphisms at *PKD1* compared with *PKD2* is likely explained by the larger coding region (four to five times greater) and GC richness of the region, resulting in a much higher level of hypermutable CpG dinucleotides.[18,32] Additional special mechanisms are probably not required to explain the higher polymorphism rate in *PKD1*.[31,33]

DISCUSSION

We describe here a scheme for comprehensive mutation analysis in ADPKD. A definite mutation was detected in 62.9% of cases; in an additional 26.2%, a UCV was found to be probably mutagenic, giving a total detection level of 89.1%. These values compare favorably with previous mutation studies using DHPLC, with a mutation detection level of approximately 67%,[17] or other screenings that found *PKD1* truncating mutations in <50% of cases.[18] The higher detection level was achieved because of DS; DHPLC alone detected mutations in 63.9% (similar to previous studies), but DS identified an additional 51 mutations, 28.3% of the total cases. These variants did not generate abnormal DHPLC profiles under the conditions used or mimicked one of the many polymorphisms in *PKD1* and so were not characterized.[34] Although the majority of missed mutations were to *PKD1*, seven were to *PKD2*. It is worth noting, however, that two control mutations were not found by DS, and one other required specific re-analysis. In two additional cases, changes were initially seen as homozygous but later shown as heterozygous with different primers. These findings illustrate the complexity of mutation analysis in ADPKD, especially in *PKD1* with locus-specific LR-PCR, and suggest that allele dropout can be a problem, even with carefully designed primers.

Screening for larger deletions revealed four mutations (2.0% of the total), consistent with the level (1 to 3%) detected in other studies,[1,35] with suggestions that some missed mutations may also be larger rearrangements (Table V). However, FIGE is time-consuming, and a rapid screening method, such as multiplex ligation probe-dependent amplification,[36] is required to detect these mutations, although the duplicated area of *PKD1* will prove a challenge to establishing multiplex ligation probe-dependent amplification at this locus.

A major factor in achieving this high detection level was establishing a framework for scoring UCV. Building on experience gained in analysis of the *BRCA1* and *2* genes,[22,25] we used a combination of scoring the potential detrimental effect of the change with contextual information about the variant. Although it would be ideal to test the functional significance of UCV, this is impractical for large, complex proteins such as PC1 and PC2 in a routine diagnostic setting (and at least for PC1, no such functional test is yet available). However, data obtained in the research setting—for instance, channel activity of a PC2 mutant,[37] or effect on GPS cleavage of a PC1 mutant,[38]—may be useful to include in calculations of likely pathogenicity.

The value of an ADPKD molecular test depends on the reliability to which the pathogenicity of UCV can be determined. Factors that were found critical were to include distant related vertebrates (chicken, frog, and fish) in the comparative analysis and to analyze conservation in domain structures. There is scope to improve these predictions as more three-dimensional structures become available.[24] Contextual information was gained by complete analysis of both genes; when a UCV was the only variant, it was a strong candidate to be pathogenic. No patient was found in this study with two definitely pathogenic mutations. One previous example has been described of a PKD1 patient with two nonsense mutations in cis[17] (either occurring simultaneously or a second mutation on an already mutant allele), and rare similar examples have been described in other diseases. However, because these cases are rare, the finding of a UCV in a patient with a definite mutation is highly suggestive that the change is not pathogenic. Data regarding the contextual finding of variants are enhanced by accumulated knowledge on disease-associated and neutral variants being collected in ADPKD databases (HGMD and ADPKD Mutation Database).

The high level of novel variants limits the number of cases in which previous data are useful, but additional information can be obtained by segregation analysis (lack of segregation proves that it is not pathogenic). In this study population, segregation analysis was limited by the family samples that were available, although segregation of 20 UCV was demonstrated (Table II). Furthermore, segregation was crucial for showing that some *PKD1* changes (R2477C and 2894delANS) were not pathogenic (Figure 2A and Table IV). Some variants were classed indeterminate, with a possibility that they may be pathogenic; segregation analysis would be helpful to

define their status, especially in families with more than one such variant. When possible, the family should be the unit of analysis for molecular diagnostics in ADPKD.

From this study, approximately 11% of cases had no probable mutation defined. Further analysis is needed to screen for larger DNA rearrangements (see Discussion). Studies of other large, multi-exon genes have illustrated how exonic and intronic variants can have unpredicted consequences on splicing by influencing the normal splice junction or branch point or by creating a cryptic site[39]; our RT-PCR data have illustrated some examples in ADPKD. A category of changes that have not been analyzed in this study are those that alter splice enhancer and silencing sites.[40] These are difficult to predict because the consensus sequences are short and not highly conserved[41]; RT-PCR analysis in patients in whom no mutation has been defined is likely to be productive and may also identify mutations that are embedded deep within introns that can influence splicing.[42] Other possible sites of missed mutations are the promoter regions.[43] Once this next level of mutation analysis has been completed, it will be possible to assess the representation of unlinked "PKD3" families in this cohort. The level of such families is likely to be low (certainly $<10\%$), but they cannot be excluded altogether at this stage.

Recurrent mutations account for 30% of the families and have diagnostic implications because prescreening regions that contain these changes may simplify molecular diagnostics; ultimately, a focused approach that uses a DNA array may be productive. The most common mutation, 5014delAG (also the most commonly described in the literature; Table I), is not flanked by sequences that are likely to promote recurrent deletion and raises the possibility that it is an ancestral change. Haplotype analysis would be helpful to determine whether a common origin is likely. Of the other five mutations found at least three times in this study, three have been described elsewhere (see Table I for details) and two represent C→T (or A→G) changes at CpG dinucleotides (*PKD1*:E2771K and *PKD2*:R872X), suggesting that they may be recurrent.[44] Others may be population specific; all three 3814delG cases were identified at the same center and may reflect unknown relatedness. The finding of some missense changes multiple times, notably E2771K (seven times in total), giving it a high VS, and to a lesser extent G381S (three times), suggests that a group of missense changes with almost certain pathogenicity will be defined over time.

Our studies have important implications for molecular diagnostics of ADPKD. Sequencing is the most reliable means to identify base-pair variants, but screens for larger changes have revealed mutations in a few additional cases. Definite truncating mutations can be determined in 60 to 65% of cases, and careful evaluation of UCV in the remainder can push the level of detection of probable mutations to approximately 90%. It is particularly important that the pathogenic potential of all UCV be assessed and explained in a mutation report so that the likely significance of the described variants can be judged. Although using the mutation information on the majority of patients with definite mutations in the clinical arena should be straightforward, caution is still required when using information from MG = B or C patients in the setting of transplantation decisions or starting therapy. ADPKD mutation databases will be a guide for those previously described UCV, but many novel changes can be expected so that analysis of a family, especially if large enough to determine the gene responsible, will allow the use of more nondefinite mutations in the clinic. It is likely that improved mutation screening will become an increasingly important part of obtaining a firm diagnosis in the patient with ADPKD.

CONCISE METHODS

Details of the CRISP Study Cohort

The CRISP study cohort consists of 239 patients who have ADPKD, are aged 15 to 46 yr, and have a GFR of >70 ml/mm at enrollment.[15,16] For each proband, a blood sample was obtained for DNA extraction, and Epstein Barr virus (EBV)-transformed lymphoblast cell lines were established. Samples are stored at the National Institute of Diabetes and Digestive and Kidney Disease Center for Genetic Studies.

Mutation Analysis of *PKD1* and *PKD2* by DHPLC

DHPLC was performed as described previously.[17] The duplicated region of *PKD1* was amplified as five *PKD1*-specific fragments by LR-PCR, and exons were amplified from these fragments after dilution 1:1000 to avoid genomic DNA carryover. LR-PCR products were checked on 0.8% agarose gels before the nested PCR to detect large deletions/insertions. DHPLC was performed using a Wave System HT (Transgenomic, Omaha, NE), as described previously.[17] Each amplicon was run at two temperatures, typically the melting temperature and +2°C, as determined by Wavemaker version 4.0.32 (Transgenomic). Chromatograms were subjectively grouped, depending on the differences in the profile from normals and known polymorphic variants.[45] Novel profiles were sequenced, plus a representative profile that resembled a known variant; the remainder were identified by signature-based genotyping.[34]

DS of *PKD1* and *PKD2* and Confirmation of Mutations

DS was performed on a fee-for-service basis using a commercial diagnostic test (Athena Diagnostics, Worcester, MA), and a subset were analyzed at Emory for *PKD2*. This sequence data are now stored at the CRISP web site.

All likely pathogenic changes were double-checked in a different aliquot of the same sample, and, when samples were available, the segregation of variants was tested by DS. A subset of samples were also analyzed for large genomic rearrangements using FIGE, as described previously.[1] The precise mutation in these cases was determined by amplifying across the deletion and sequencing the aberrant product. The designation of all variants in both genes was numbered from their translational start. The *PKD1* sequence also included the alternatively spliced codon at the start of exon 32 so that the total length of the protein was 4303 aa.

RT-PCR to Analyze Atypical Splicing

RNA was isolated from lymphoblast cell lines using TRizol (Invitrogen, Carlsbad, CA), and cDNA was generated with the Superscript III cDNA synthesis kit (Invitrogen). For the IVS15 variants, LR-RT-PCR was performed with the SPEC-4 primers,[18] followed by nested PCR with primers in exons 15 and 18. The IVS37 mutation was analyzed with primers in exons 36 and 39, and the *PKD2* IVS4 mutation was analyzed with primers in exons 3, 7, and IVS4 (primer sequences available on request). All PCR products were sequenced by standard methods.

Classification of UCV

The GMS[20] was used to determine the GD of each substitution. An MSA was generated using the software ClustalW (MacVector Inc, Cary, NC) of PC1 orthologs from human, rat, mouse, chicken, frog (*Xenopus tropicalis*), and fish (*Fugu rubripes*). Other fish species, *Tetraodon nigroviridis* and *Danio rerio*, were also used in some cases. For PC2, human, dog, mouse, chicken, *Danio*, and *Fugu* were used, plus in some cases *Drosophila melanogaster*, sea urchin, and *Caenorhabditis elegans*. Sequences were obtained from NCBI or Ensembl. Predicted sequences for chicken, frog, and fish species were edited as necessary. The GV was determined as the largest GMS between vertebrate orthologs. Conservation within defined motifs used published MSA: WSC[46]; PKD repeat[2,3,47]; REJ domain[48]; GPS[38,46]; PLAT domain[49]; and polycystin motifs, PC-A and B[29]; plus MSA of PC1 and PC2 homologs (PC1L1, PC1L2, PC1L3, PKDREJ, PC2L1, and PC2L2) and sea urchin REJ1 to 4 homologs. Changes within transmembrane regions and the G-protein binding region were also assessed for disruption of secondary structure. Predicted splicing changes were scored using the Splicing Predictor by Neural Network (http://www.fruitfly.org/seq_tools/splice.html).[50] Details of previous descriptions of the UCV were obtained from the HGMD (http://www.hgmd-cf.ac.uk) and/or the ADPKD Mutation Database (http:/pkdb.mayo.edu).

For all of the UVC, a VS was calculated. A matrix comparing the GD and GV was used with the GMS divided into six groups (0, 1 to 35, 36 to 80, 81 to 130, 131 to 179, and >180) and scores from +8 (*e.g.*, GD >131, GV <35) to −4 (*e.g.*, GD <35, GV >81) given. An additional −4 was allotted when the substitution matched the residue in an ortholog. Conservation of the reside in a domain scored +2 to +4 (highly conserved to invariant) with −2 when substitution matched the residue in an ortholog. Splicing changes were scored from +20 when an abnormal splice product was demonstrated by RT-PCR to −5 when an intronic change was not predicted to alter splicing. Additional points were given when a variant segregated in siblings or a family (+2 to +4); failure to segregate = −20. Other descriptions of the change as a mutation added +2 per time described. If another likely mutation (groups A through C) or an indeterminate (I) change was found with the variant, −5 (A) to −1 (I) was subtracted; otherwise +2 was added. The sum of these scores gave the VS. VS ≥11 were classed very likely mutations (MG = B); 5 to 10 as likely mutations (MG = C); −4 to 4 as variants of indeterminate (I) pathogenicity; and ≤5 as neutral polymorphisms (P). The χ^2 test was used to examine whether the mutation location in each gene was uniform.

ACKNOWLEDGMENTS

The study was supported by National Institute of Diabetes and Digestive and Kidney Diseases cooperative agreements (DK56934, DK56956, DK56957, and DK56961), with additional support from an ancillary study for genetic analysis (DK56957-S1) and grant DK58816.

The study coordinators who collected samples are thanked: Jody Mahan, Beth Stafford, Lorna Stevens, Kristin Cornwell, Vickie Kubly, Diane Watkins, Sharon Langley, and Pam Trull. We also thank Mary Virginia Gaines for management support. Jim Calvet is thanked for discussions of specific mutations. John McAuliffe, William Seltzer, Lynne Leclair, and Mark Smith at Athena Diagnostics are thanked for assistance in the fee-for-service DS mutation analysis.

DISCLOSURES

None.

APPENDIX

Table IA. Details of probable mutations in the CRISP cohort: Definite, mutation group A[a]

CRISP ID[b]	N[c]	Exon/ Intron	Codon	Mutation Designation	cDNA Change	Protein Change	Variant Effect	Other[d] Indeterminate Variants	Previous Description
PKD1									
201800	1	1	10	29del18	29_46del	A10-L15del	Deletion		Novel
159106	1	1	33	99delC	99delC	C33fsX72	Truncation		Novel
325290	1	1	55	165del7	165_171del	G55fsX70	Truncation		Novel
298515	1	4	128	383del4	383_386del	E128fsX273	Truncation		Novel
459051	1	4	138	R138X	412C→T	R138X	Truncation		Novel
271684	1	IVS4	177	IVS4+1G→A	IVS4+1G→A	N121fs[e]	Truncation[e]		Novel
223534	1	5	211	632dup4	628_631dup	211fsX659	Truncation		2X
465485	1	5	211	632dup4	628_631dup	211fsX659	Truncation		2X
232174	1	5	227	Q227X	679C→T	Q227X	Truncation		1X[f]
124952	1	5	351	Q351X	1051C→T	Q351X	Truncation		Novel
271460	1	5	369	1105delAG	1105_1106del	S369fsX369	Truncation		Novel
126133	1	6	450	1350delC	1350delC	P450fsX464	Truncation		Novel
293598	1	7	467	W467X	1401G→A	W467X	Truncation		Novel
400002	2	7	504	1510dup19	1492_1510dup	T504fsX524	Truncation		Novel
111722	1	7	522	1564delT	1564delT	C522fsX558	Truncation		Novel
283935	1	8	555	1664delGA	1664_1665del	G555fsX586	Truncation		Novel
101585	1	8	562	Q562X	1684C→T	Q562X	Truncation		2X
327055	1	8	562	Q562X	1684C→T	Q562X	Truncation		2X
424816	1	8	586	1755insT	1755_1756insT	E586fsX586	Truncation		Novel
303868	1	10	693	2079delG	2079delG	G693fsX784	Truncation		Novel
119833	1	10	696	2085insC	2085_2086insC	A696fsX713	Truncation		2X[f]
200002	2	10 + IVS10	699	2096del25	2096_2097del: IVS10+1del23	S699fsX776[e]	Truncation[e]		Novel
120395	1	11 to 15	734	2200del5604, insAC	2200_5946del, insAC	G734fsX867	Truncation		Novel
244831	1	11	827	2477dup5	2473_2477dup5	I827fsX899	Truncation		Novel
392316	1	11	874	C874X	2622C→A	C874X	Truncation		Novel
234053	1	11	896	2687insT	2687_2688insT	V896fs704	Truncation		Novel
320957	1	IVS13	1054	IVS13–2A→G	IVS13-2A→G	L1054fs[e]	Truncation[e]		Novel
321611	1	14	1081	S1081X	3242C→A	S1081X	Truncation		Novel
190066	1	14	1095	Y1095X	3285C→G	Y1095X	Truncation		Novel
400003	2	15	1272	3814delG	3814delG	V1272fsX1272	Truncation		3X
400538	1	15	1272	3814delG	3814delG	V1272fsX1272	Truncation		3X
483553	1	15	1272	3814delG	3814delG	V1272fsX1272	Truncation		3X
100001	3	15	1272	3815del1076	3815_4890del	V1272fsX1657	Truncation		Novel
200922	1	15	1294	3880delC	3880delC	L1294fsX1345	Truncation		Novel
120801	1	15	1332	3994dup38	3957_3994dup38	D1332fsX1358	Truncation		Novel
168106	1	15	1357	4070delT	4070delT	L1357fsX1365	Truncation		Novel
139486	1	15	1370	4109del12insG	4109_4120del,insG	Y1370fsx1426	Truncation		Novel
100005	2	15	1582	W1582X	4746G→A	W1582X	Truncation		Novel
100008	2	15	1672	5014delAG	5014_5015del	R1672fsX1769	Truncation		5X + 9X[f]
201877	1	15	1672	5014delAG	5014_5015del	R1672fsX1769	Truncation		5X + 9X[f]
300911	1	15	1672	5014delAG	5014_5015del	R1672fsX1769	Truncation		5X + 9X[f]
300003	2	15	1672	5014delAG	5014_5015del	R1672fsX1769	Truncation		5X + 9X[f]
400004	2	15	1672	5014delAG	5014_5015del	R1672fsX1769	Truncation	V2267M	5X + 9X[f]
141411	1	15	1696	Q1696X	5086C→T	Q1696X	Truncation		Novel
281977	1	15	1707	W1707X	5120G→A	W1707X	Truncation		Novel
432376	1	15	1956	5866insG	5866_5867insG	V1956fsX1989	Truncation		Novel
226640	1	15	2004	6010delT	6010delT	Y2004fsX2115	Truncation		Novel
178750	1	15	2056	E2056X	6166G→T	E2056X	Truncation		Novel
337315	1	15	2067	Q2067X	6199C→T	Q2067X	Truncation		Novel
423411	1	15	2153	6457insG	6457_6458insG	V2153fsX2175	Truncation		Novel
208280	1	15	2184	6549insT	6549_6550insT	E2184fsX2184	Truncation		Novel
268455	1	15	2217	6650del15	6650_6664del	V2217_L2221del	Deletion		Novel

(Table IA continued on next page)

Table IA. *(continued)*

CRISP ID[b]	N[c]	Exon/ Intron	Codon	Mutation Designation	cDNA Change	Protein Change	Variant Effect	Other[d] Indeterminate Variants	Previous Description
354494	1	15	2219	6656dup15	6642_6656dup	P2219insRLVLP	Insertion		Novel
234650	1	15	2264	R2264X	6791C→G	R2264X	Truncation		Novel
331318	1	15	2269	S2269X	6806C→G	S2269X	Truncation		1X[f]
333524	1	15	2291	Q2291X	6871C→T	Q2291X	Truncation		Novel
166508	1	IVS15	2306	IVS15+1G→T	IVS15+1G→T	R2306fs[e]	Truncation[e]		Novel
100007	2	IVS15	2306	IVS15–10C→A	IVS15-10C→A	R2306fsX2421	Truncation		Novel
154595	1	IVS15	2306	IVS15–9G→A	IVS15-9G→A	R2306fsX2316	Truncation		Novel
139126	1	16	2307	E2307X	6919G→T	E2307X	Truncation		Novel
200005	2	16	2318	6952delC	6952delC	R2318fsX2340	Truncation		Novel
376252	1	16	2343	W2343X	7029G→A	W2343X	Truncation		Novel
323837	1	18	2430	R2430X	7288C→T	R2430X	Truncation		3X[f]
220068	1	18	2435	7303del15	7303_7317del	R2435_R2439del	Deletion		Novel
239960	1	19	2523	E2523X	7567G→T	E2523X	Truncation		Novel
100505	1	19	2556	Q2556X	7666C→T	Q2556X	Truncation		4X, + 1X[f]
271043	1	19	2556	Q2556X	7666C→T	Q2556X	Truncation		4X, + 1X[f]
300001	3	19	2556	Q2556X	7666C→T	Q2556X	Truncation		4X, + 1X[f]
430543	1	19	2556	Q2556X	7666C→T	Q2556X	Truncation		4X, + 1X[f]
247880	1	20	2568	7704delG	7704delG	R2568fsX2619	Truncation		Novel
214376	1	21	2639	R2639X	7915C→T	R2639X	Truncation		1X[f]
387658	1	IVS21	2673	IVS21–2delAG	IVS21-2delAG	G2673fs[e]	Truncation[e]		Novel
397661	1	IVS21	2673	IVS21–2A→G	IVS21-2A→G	G2673fs[e]	Truncation[e]		Novel
128596	1	23	2869	8606delAG	8606_8607del	E2869fsX2935	Truncation		Novel
336167	1	IVS23	2931	IVS23–2A→G	IVS23-2A→G	G2931fs[e]	Truncation[e]		Novel
393936	1	IVS24 to IVS30	2983	IVS24–28del3065	8949_10050del	S2983fs	Truncation		Novel
137708	1	26	3109	9324dup5	9320_9324dup	I3109fsX3320	Truncation		Novel
301157	1	IVS26 to IVS38	3133	IVS26–243del9214	9398_11156del	G3133fs	Truncation		Novel
100003	2	28	3218	W3218X	9654G→A	W3218X	Truncation		Novel
168286	1	30	3326	9978del8ins3	9978_9985del, insACA	G3326fsX3387	Truncation		Novel
295106	1	33	3420	W3420X	10260G→A	W3420X	Truncation		Novel
237192	1	33	3468	S3468X	10403C→G	S3468X	Truncation		Novel
143806	1	IVS35	3540	IVS35–2A→G	IVS35-2A→G	G3540fs[e]	Truncation[e]		Novel
209281	1	36	3582	10745delC	10745delC	P3582fsX3584	Truncation		Novel
300005	2	36	3594	10780insG	10780_10781insG	A3594fsX3626	Truncation		Novel
223343	1	IVS36	3608	IVS36+2G→T	IVS36+2G→T	L3541fs[e]	Truncation[e]		Novel
124300	1	IVS37	3673	IVS37–1G→T	IVS37-1G→T	S3673fs[e]	Truncation[e]		Novel
170367	1	IVS37	3673	IVS37–10C→A	IVS37-10C→A	S3673fsX3674	Truncation	R4276W	1X[f]
304860	1	39	3729	11185delC	11185delC	H3729fsX3825	Truncation		Novel
258950	1	40	3767	11299del23	11299_11321del	V3767fsX3808	Truncation		Novel
281000	1	40	3781	Y3781X	11343C→G	Y3781X	Truncation		1X[f]
188920	1	IVS40	3804	IVS40–1G→A	IVS40-1G→A	G3804fs[e]	Deletion[e]		1X[f]
300004	2	41	3842	W3842X	11525G→A	W3842X	Truncation		2X
376004	1	41	3842	W3842X	11525G→A	W3842X	Truncation		2X
300008	2	42	3874	11621delC	11621delC	P3874fsX3944	Truncation		Novel
358230	1	42	3892	11673dup5	11669_11673dup	R3892fsX3946	Truncation		Novel
439240	1	43	3922	W3922X	11766G→A	W3922X	Truncation		Novel
342131	1	43	3933	11799delC	11799delC	L3933fsX3944	Truncation		Novel
114207	1	43	3955	Q3955X	11863C→T	Q3955X	Truncation	R1411C	Novel
394588	1	43	3982	Q3982X	11944C→T	Q3982X	Truncation		Novel
397931	1	IVS43	4002	IVS43+2T→G	IVS43+2T→G	V3905_K4001[e]	Deletion[e]		Novel
175453	1	IVS43	4002	IVS43+14del20	IVS43+14del20	V3905_K4001	Deletion		2X[f]
100004	2	44	4021	R4021X	12061C→T	R4021X	Truncation		2X[f]
386488	1	46	4161	12483del11	12483_12493del	L4161fsX4205	Truncation		Novel
385151	1	46	4220	E4220X	12658G→T	E4220X	Truncation		Novel
264348	1	46	4228	R4228X	12682C→T	R4228X	Truncation		3X[f]

(Table IA continued on next page)

Table IA. *(continued)*

CRISP ID[b]	N[c]	Exon/ Intron	Codon	Mutation Designation	cDNA Change	Protein Change	Variant Effect	Other[d] Indeterminate Variants	Previous Description
PKD2									
121543	1	1	85	Q85X	253C→T	Q85X	Truncation		Novel
200006	2	1	143	428delG	428delG	G143fsX232	Truncation		Novel
192261	1	2	223	667delG	667delG	E223fsX232	Truncation		Novel
277490	1	4	306	R306X	916C→T	R306X	Truncation		7X[f]
132635	1	4	316	948delA	948delA	L316fsX337	Truncation		Novel
419620	1	4	320	R320X	958C→T	R320X	Truncation		2X + 1X[f]
438283	1	4	320	R320X	958C→T	R320X	Truncation		2X + 1X[f]
140005	2	4	325	R325X	973C→T	R325X	Truncation	I2463T, S4144T	Novel
301372	1	IVS4	365	IVS4+3del4	IVS4+3delAAGT	A365	Truncation		1X[f]
200008	2	6	495	1483delA	1483delA	I495fsX513	Truncation		Novel
300002	2	IVS8	633	IVS8+1G→A	IVS8+1G→A	L573fs[e]	Truncation[e]		Novel
203328	1	11	720	2159delA	2159delA	N720fsX736	Truncation		3X[f]
224502	1	11	720	2159dupA	2159_2160insA	N720fsX724	Truncation		3X[f]
300006	2	IVS11	747	IVS11–1G→T	IVS11-1G→T	K748fs[e]	Truncation[e]		Novel
208324	1	IVS12	787	IVS12+1G→A	IVS12+1G→A	K748fs[e]	Truncation[e]		Novel
312317	1	13	807	R807X	2419C→T	R807X	Truncation		1X[f]
294511	1	13	833	2498delG	2498delG	G833fsX843	Truncation		Novel
140353	1	14	872	R872X	2614C→T	R872X	Truncation		3X + 6X[f]
151030	1	14	872	R872X	2614C→T	R872X	Truncation		3X + 6X[f]
161547	1	14	872	R872X	2614C→T	R872X	Truncation		3X + 6X[f]
300641	1	14	891	IVS14–2A→G	IVS14-2A→G	D891fs[e]	Truncation[e]		Novel

Table IB. Details of probable mutations in the CRISP cohort: Highly likely, mutation group B

CRISP ID[b]	n[c]	Exon/ Intron	Codon	Mutation Designation	cDNA Change	Protein Change	Variant Effect	Other[d] Indeterminate Variants	Previous Description
PKD1									
200001	2	5	230	C230S	688T→A	C230S	Substitution	IVS7+3G→T	Novel
242715	1	5	259	C259Y	776G→A	C259Y	Substitution	P829L	Novel
105005	1	5	325	Y325C	974A→G	Y325C	Substitution		Novel
174632	1	5	381	G381S	1141G→A	G381S	Substitution	R3063C	3X
246620	1	5	381	G381S	1141G→A	G381S	Substitution		3X
300007	2	5	381	G381S	1141G→A	G381S	Substitution		3X
110080	1	6	432	A432V	1295C→T	A432V	Substitution		Novel
110890	1	6	436	C436Y	1307G→A	C436Y	Substitution		Novel
386040	1	7	508	C508R	1522T→C	C508R	Substitution		1X[f,g]
243738	1	9	605	L605R	1814T→C	L605R	Substitution		Novel
306546	1	10	727	L727P	2180T→C	L727P	Substitution		Novel
400001	2	11	910	910delDVV	2729_2737del	910_912del	Deletion		Novel
204555	1	12	967	W967R	2899T→C	W967R	Substitution		1X[f]
248712	1	15	1412	Y1412D	4234T→G	Y1412D	Substitution		Novel
170795	1	15	1503	G1503R	4507G→C	G1503R	Substitution		Novel
113994	1	15	1993	F1992L: T1993del	5976_5978del	F1992L:T1993del	Substitution/deletion		1X[f]
401405	1	17	2370	C2370S	7110T→A	C2370S	Substitution		Novel
374687	1	IVS17	2404	IVS17+4del4	IVS17+4delAGTG	V2356_G2403del[e]	Deletion[e]		Novel

(Table IB continued on next page)

Table IB. *(continued)*

CRISP ID[b]	*n*[c]	Exon/ Intron	Codon	Mutation Designation	cDNA Change	Protein Change	Variant Effect	Other[d] Indeterminate Variants	Previous Description
200009	2	19	2516	R2516C	7546C→T	R2516C	Substitution		2X
295940	1	19	2516	R2516C	7546C→T	R2516C	Substitution		2X
108730	1	23	2771	E2771K	8311G→A	E2771K	Substitution		3X + 4X[f]
136055	1	23	2771	E2771K	8311G→A	E2771K	Substitution		3X + 4X[f]
407593	1	23	2771	E2771K	8311G→A	E2771K	Substitution		3X + 4X[f]
327933	1	23	2809	P2809L	8426C→T	P2809L	Substitution		Novel
196164	1	24	2979	F2979del	8935_8937del	F2979del	Deletion		1X[f]
199540	1	25	3036	T3036I	9107C→T	T3036I	Substitution		Novel
334672	1	27	3168	F3168del	9502_9504del	F3168del	Deletion		Novel
173203	1	27	3186	H3186Y	9556C→T	H3186Y	Substitution		2X
293317	1	27	3186	H3186Y	9556C→T	H3186Y	Substitution		2X
236202	1	29	3287	L3287del	9859_9861del	L3287del	Deletion		Novel
292362	1	39	3753	R3753W	11257C→T	R3753W	Substitution		3X[f]
271662	1	39	3753	R3753Q	11258G→A	R3753Q	Substitution		Novel
200007	2	41	3834	L3834R	11501T→G	L3834R	Substitution		Novel
223635	1	43	3908	L3908R	11723T→G	L3908R	Substitution		Novel
PKD2									
300696	1	4	322	R322W	964C→T	R322W	Substitution		1X[f]
120010	2	4	325	R325Q	974G→A	R325Q	Substitution		Novel
146563	1	IVS8	633	IVS8+5G→A	IVS8+5G→A	L573fs[e]	Truncation[e]		Novel

Table IC. Details of probable mutations in the CRISP cohort: Likely, mutation group C

CRISP ID[b]	*n*[c]	Exon/Intron	Codon	Mutation Designation	cDNA Change	Protein Change	Variant Effect	Other[d] Indeterminate Variants	Previous Description
PKD1									
150534	1	15	1870	N1870H	5608A→C	N1870H	Substitution		Novel
170121	1	15	1999	G1999S	5995G→A	G1999S	Substitution		Novel
294105	1	15	1999	G1999V	5996G→T	G1999V	Substitution	R3965S	Novel
245990	1	15	2278	G2278R	6832G→A	G2278R	Substitution		Novel
320182	1	15	2297	H2297P	6890A→C	H2297P	Substitution		Novel
343233	1	23	2866	L2866P	8597C→T	L2866P	Substitution		Novel
200004	2	25	3007	S3007P	9019T→C	S3007P	Substitution		Novel
380166	1	27	3168	F3168L	9504C→G	F3168L	Substitution		Novel
200003	2	27	3171	A3171P	9511G→C	A3171P	Substitution		Novel
113094	1	36	3604	E3604K	10810G→T	E3604K	Substitution		Novel
117537	1	37	3651	G3651S	10951G→A	G3651S	Substitution		Novel
229428	1	45	4137	L4137P	12410T→C	L4137P	Substitution		Novel
148014	1	45	4139	L4139P	12416C→T	L4139P	Substitution	*R807Q*	Novel
PKD2									
457036	1	5	378	S378del	1131_1133del	S378del	Deletion	R1698W	Novel
419258	1	13	804	S804N	2411G→A	S804N	Substitution		Novel
479334	1	14	878	R878del	2632_2634del	R878del	Deletion	R4276W	Novel

[a]CRISP, Consortium for Radiologic Imaging Study of PKD.
[b]Patient or pedigree number.
[c]Number of patients from family in CRISP study.
[d]Other indeterminate variant (variant group I) found in the proband. *PKD2* change in italics.
[e]Predicted change.
[f]Details of previous description in Human Gene Mutation Database (HGMD) and/or ADPKD Mutation Database (PKDB).
[g]Same family as published.

Table IIA. Classification of probable pathogenic mutations: Missense[a]

Mutation Designation	GD[b]	Orthologs GV[c]	Protein Domain Conservation[d]	Splicing Predictions[e]	Segregation Analysis[f]	Other Description[g]	Other Indeterminate Variant[h]	Variant Score[i]	Mutation Group[j]
PKD1									
C230S	112	0	WSC (I)	N/P	S2	Novel	Y	11	B
C259Y	194	0	WSC (I)	N/P	N/A	Novel	Y	11	B
Y325C	194	0	PKDI (C)	N/P	N/A	Novel	N	12	B
G381S	56	0	N/D	N/P	F9	3X + G381C[k]	2N, 1Y	17	B
A432V	64	0	C-lectin (HC)	ND (0.11)	N/A	Novel	N	11	B
C436Y	194	0	C-lectin (I)	N/P	N/A	Novel	N	14	B
C508R	180	0	C-lectin (HC)	N/P	F6	Novel	N	17	B
L605R	102	15	N/D	N/P	F3	Novel	N	11	B
L727P	98	5	N/D	N/P	F3	Novel	N	11	B
W967R	101	0	PKDII (HC)	N/P	N/A	1X[k]	N	13	B
Y1412D	160	0	PKDVIII (C)	N/P	N/A	Novel	N	11	B
G1503R	125	0	PKDIX (HC)	N/P	N/A	Novel	N	11	B
N1870H	68	0	PKDXIII (HC)	N/P	N/A	Novel	N	10	C
G1999S	56	0	PKDXV (C)	N/P	N/A	Novel, G1999V	N	10	C
G1999V	109	0	PKDXV (C)	N/P	N/A	Novel, G1999S	Y	8	C
G2278R	125	60	REJ (I)	N/P	N/A	Novel	N	9	C
H2297P	77	29	REJ (NC)	N/P	F4	Novel	N	10	C
C2370S	112	0	REJ (I)	N/P	N/A	Novel, C2370R[k]	N	13	B
R2516C	180	0	REJ (NC)	N/P	S2	2X	2N	16	B
E2771K	56	0	REJ (C)	N/P	N/A	3X + 4X[k]	3N	25	B
P2809L	98	0	N/D	N/P	F4	Novel	N	12	B
L2866P	98	0	N/D	N/P	F3	Novel	N	10	C
S3007P	74	80 (E, fi)	N/D	N/P	F2	Novel	N	5	C
T3036I	89	0	GPS (HC)	N/P	N/A	Novel	N	11	B
F3168L	22	0	PLAT (HC)	N/P	F12	Novel, F3168del	N	10	C
A3171P	27	0	PLAT (C)	N/P	S2	Novel	N	10	C
H3186Y	83	0	PLAT (HC)	ND (0.24)	N/A	2X	2N	18	B
E3604K	56	0	TM (C)[m]	N/P	N/A	Novel	N	10	C
G3651S	56	0	N/D	N/P	N/A	Novel	N	7	C
R3753W	101	0	PC-A (HC)	N/P	N/A	3X[k], R3753Q	N	18	B
R3753Q	43	0	PC-A (HC)	N/P	N/A	Novel, R3753W	N	13	B
L3834R	102	22	PC-B (HC)	N/P	S2	Novel	N	12	B
L3908R	102	15	TM (NC,DSS)	N/P	F2	Novel	N	12	B
L4137P	98	0	GPB (DSS)	N/P	N/A	Novel	N	10	C
L4139P	98	15	GPB (DSS)	N/P	N/A	Novel	Y	6	C
PKD2									
R322W	101	0	PC-A (I)	N/P	F3	1X[k]	N	18	B
R325Q	43	26	PC-A (HC)	N/P	F3	Novel	N	13	B
S804N	46	0	ERL (HC)	N/P	N/A	Novel	N	10	C

Table IIB. Classification of probable pathogenic mutations: Atypical splicing changes[a]

Mutation Designation	RT-PCR Change[n]	Splicing Predictions[e]	Segregation Analysis[f]	Previous Description[g]	Other Indeterminate Variant[h]	Variant Score[i]	Mutation Group[j]
PKD1							
IVS15-9G→A	Y	NA, IVS15-7 (<0.1[i])	N/A	Novel	N	27	A
IVS15-10C→A	Y	NA, IVS15-8 (0.98)	S2	Novel	N	34	A
IVS17+4del4	N/A	LD (0.95 to <0.1)	F3	Novel	N	16	B
IVS37-10C→A	Y	LA (0.34 to <0.1)	FP	1X[k]	Y	21	A
IVS43+14del20	Y	Skipping ex43, small IVS[37]	N/A	2X[k]	N	26	A
PKD2							
IVS4+3delAAGT	Y	LD (0.99 to <0.1)	F3	1X[k]	N	25	A
IVS8+5G→A	N/A	ND, IVS8-4 (0.28)	N/A	Novel	N	12	B

Table IIC. Classification of probable pathogenic mutations: Small deletions

Mutation Designation	Conservation Orthologs[l]	Protein Domain Conservation[d]	Segregation Analysis[f]	Other Description[g]	Other Indeterminate Variant[h]	Variant Score[i]	Mutation Group[j]
PKD1							
910delDVV	C	PKDII (C)	F2	Novel	N	14	B
F1992L; T1993del	C	PKDXV (C)	N/A	1X[k]	N	13	B
F2979del	I	N/D	N/A	1X[k]	N	13	B
F3168del	I	PLAT (I)	N/A	Novel	N	15	B
L3287del	I	TM (HC)	N/A	Novel	N	14	B
PKD2							
S378del	HC	N/D	N/A	Novel	Y	7	C
R878del	I	N/D	N/A	Novel	Y	7	C

[a]GD, Grantham distance; GV, Grantham variation; MG, mutation group; RT-PCR, reverse transcription–PCR.
[b]Score of chemical difference between the normal and mutated residue (high score, greater difference).
[c]Score of chemical difference between orthologs (human, mouse, rat, dog, chicken, frog, and fish [fi]; 0 = completed conserved). Residue and species shown if GV>GD.
[d]Domain containing residue (see Concise Methods for details): WSC, PKD repeats, REJ, GPS, PLAT; TM, transmembrane region; PC, polycystin motif A or B; GPB, G-protein binding region; ERL, ER localization signal[51]; N/D, no domain defined. Level of conservation: I, invariant; HC, highly conserved (>80%); C, conserved (80 to 50%); NC, not conserved (<50%); DSS, predicted to disrupt secondary structure.
[e]ND, novel donor; NA, novel acceptor; LD, loss of donor; LA, loss of acceptor; N/P, none predicted; 0.1 to 1, weakly to strongly predicted splice site.
[f]Segregation demonstrated in: S, sibs #; F, family # (at least two generations); N/A, samples not available; FP, segregation in family in previous description.
[g]Previous description of variant as mutation (#X ref.) or recurrence within study (#X); or substitution of same residue.
[h]Other indeterminate variant found associated with probable mutation: Y, yes (see Table I for details); N, no.
[i]Variant score: >11, MG = B; 5 to 10, MG = C. Atypical splicing changes proven by RT-PCR, MG = A.
[j]A, definite; B, highly likely; C, likely.
[k]Details of previous descriptions in HGMD and/or PKDB.
[l]I, invariant; HC, highly conserved (>80%); C, conserved (80 to 50%).
[m]Loss of positive residue after TM.
[n]Abnormal splicing demonstrated by RT-PCR: Y, yes; N/A, sample not available.
[o]Not predicted by Splice Site Prediction Neural Network.

Table IIIA. Classification of indeterminate variants: Missense

Designation	GD[a]	Orthologs GV[b]	Protein Domain Conservation[c]	Splicing Predictions[d]	Segregation Analysis[e]	Other Description[f]	Other Variant[g]	Variant Score[h]
PKD1								
V466L	32	32 (L;fr)	Lectin (NC)	SWA(0.65→0.60)	N/A	Novel, V466M	N	−2
V466M	21	32	Lectin (NC)	N/P	F2	Novel, V466L	N	4
V609G	109	29	N/P	N/P	N/A	Novel	I	3
V690G	109	32	N/P	N/P	N/A	Novel, V690D[i]	I	4
P829L	48	0	N/P	N/P	N/A	Novel	B	2
N970K	94	23	PKDIII (NC)	N/P	N/A	Novel	I	3
H1093D	81	83 (Y, fi)	PKDIV (HC)	N/P	N/A	Novel, H1093Y	N	4
R1340W	101	110	PKDVII (NC)	N/P	N/A	1X + 1X[i]	I	−1
S1352N	46	46 (N;fi)	PKDVII (NC)	N/P	N/A	Novel	N	−4
R1411C	180	112	PKDVIII (NC)	N/P	N/A	Novel	A	−1
R1698W	101	43	PKDXI (NC)	N/P	N/A	Novel	C	1
V2267M	21	0	REJ (NC)	N/P	S2	Novel	A	−1
I2463T	89	0	REJ (NC)	N/P	N/A	Novel	A	1
L2696P	48	145	REJ (NC)	N/P	N/A	Novel	N	3
R3063C	180	110	N/P	N/P	N/A	Novel	B	1
R3965S	110	43	N/P	N/P	N/A	Novel	C	1
S4144T	58	0	GPB (NC)	N/P	N/A	Novel	A	0
R4276W	101	86	N/P	N/P	N/A	2X; 1X[i]	C	−4
PKD2								
R807Q	43	0	ERL (NC)	N/P	N/A	1X; 1X[i]	C	2

Table IIIB. Classification of indeterminate variants: Splicing variants

Designation	RT-PCR Change[j]	Splicing Predictions[d]	Segregation Analysis[e]	Other Description[f]	Other Variant[g]	Variant Score[h]
PKD1						
IVS7+3G→T	N/A	WD (0.56→0.21)	S2	Novel	B	1
IVS35–14C→T	N/A	SSA (0.95→0.96)	N/A	Novel	I	−4

[a]Score of chemical difference between the normal and mutated residue (high score, greater difference).
[b]Score of chemical difference between orthologs (human, mouse, rat, dog, chicken, frog [fr] and fish [fi]; 0 = completed conserved). Species that match the substitution are shown.
[c]Domain containing residue: PKD, C-type lectin, REJ, GPB, ERL; N/P, no domain predicted; HC, highly conserved (>80%); NC, not conserved (<50%).
[d]Predicted splicing changes; N/P, none predicted; SSA, slightly strengthens acceptor; SWA, slightly weakens acceptor; WD, weakens donor; 0.1 to 1, weakly to strongly predicted.
[e]Segregation demonstrated in: S, sibs #; F, family # (at least two generations); N/A, samples not available.
[f]Recurrent within study (#X) or previously described as mutation (#X, ref); substitution of same residue.
[g]Other variant (mutation group, A, B, or C; indeterminate, I; or none, N); see Table I or IV for details.
[h]Composite variant score: 4 to −4 = indeterminate variant.
[i]Details of previous description as a mutation or variant of the same residue in HGMD and/or PKDB.
[j]Aberrant splicing demonstrated by RT-PCR; N/A, not available.

Table IVA. Classification of newly described likely polymorphic changes: Missense

Designation	GD[a]	Orthologs GV[b]	Protein Domain Conservation[c]	Splicing Predictions[d]	Segregation Analysis[e]	Other Description[f]	Other Variant[g]	Variant Score[h]
PKD1								
A140V	64	64 (V;c,fr)	LRR (NC)	N/P	S2	Novel	A	−9
H201Q	24	89	WSC (NC)	N/P	N/A	Novel	A	−9
R324L	102	110	PKDI (NC)	N/P	N/A,S2	2X + 1X[i]	A,B	−6
P556S	74	110	N/P	N/P	N/A	Novel	A	−7
R601W	101	110	N/P	N/P	N/A	Novel	A	−7
V715F	50	50 (F;fr)	N/P	N/P	N/A	Novel	A	−11
P839S	74	93	N/P	N/P	N/A	Novel	A	−7
S903G	56	142 (G;fr)	PKDII (NC)	N/P	N/A	Novel	A	−11
N1034S	46	80	PKDIV (C)	N/P	N/A	2X	A,N	−7
H1093Y	83	83 (Y;fi)	PKDIV (C) (Y,2X)	N/P	N/A	Novel, H1093D	A	−12
L1106V	32	198	PKDIV (NC) (V,3X)	N/P	N/A	2X	2A	−16
V1339M	21	84	PKDVII (NC)	N/P	N/A	4X	A,2B,N	−15
R1351W	101	125	PKDVII (NC)	N/P	N/A	Novel	A	−7
A1422T	58	106 (T;m,c)	PKDVIII (NC)	N/P	N/A	Novel	A	−11
A1516T	58	99 (T;fi)	PKDIX (NC)	N/P	N/A	2X	A,B	−14
H1777P	77	87	PKDXII (NC)	N/P	N/A	Novel	A	−7
A1790V	64	107 (V;m,r,fr)	PKDXII (C) (V, 2X)	N/P	N/A	Novel	A	−13
E1811D	45	45 (D;fr,fi)	PKDXIII (NC) (D,1X)	N/P	N/A	Novel, E1811K[j]	A	−12
A1871T	58	126	PKDXIII (NC) (A,1X)	N/P	N/A	Novel	C	−6
V2153M	21	32	N/P	N/P	N/A	Novel	A	−7
R2200C	180	110	REJ (NC)	N/P	N/A	10X + 1X[i]	4A,4B,2N	−29
V2345A	64	107	REJ (NC)	N/P	S2	Novel	A	−7
R2383H	29	112	REJ (NC)	N/P	N	Novel	B	−27
R2404W	101	180	REJ (NC)	N/P	N/A	Novel	A	−7
R2477C	180	71	REJ (NC)	N/P	N	Novel	C	−14
R2515W	101	110	REJ (NC)	N/P	N/A	Novel	A	−7
V2628M	21	124 (M;r)	REJ (NC)	N/P	N/A	Novel	B	−11
E2742K	56	98	REJ (NC)	N/P	S2	Novel	A	−5
W2882R	101	181	N/P	N/P	N/A	5X	2A,2B,C	−20
V2897I	29	121	N/P	N/P	N	Novel	N	−22
T3207M	81	58	PLAT (NC) (M, PC1L2)	N/P	N/A	2X	2A	−9

(Table IVA continued on next page)

Table IVA. Continued

Designation	GD[a]	Orthologs GV[b]	Protein Domain Conservation[c]	Splicing Predictions[d]	Segregation Analysis[e]	Other Description[f]	Other Variant[g]	Variant Score[h]
A3211T	58	106 (T;r,m)	PLAT (NC)	N/P	N/A	Novel	C	−8
A3391V	64	99	N/P	N/P	N/A	Novel	A	−7
V3409L	32	121	N/P	N/P	N/A	8X	4A,B,C,2I,N	−29
P3412S	74	108	N/P	N/P	N/A	2X	B,N	−5
G3414S	56	80 (S;r,m.c)	N/P	N/P	N/A	Novel	B	−9
S3779R	110	98	N/P	N/P	N/A	Novel	A	−7
E3786Q	29	140 (Q;r,m)	N/P	N/P	N/A	Novel	A	−13
H3840Q	24	89	PC-B (NC)	N/P	N/A	Novel	A	−9
A4098T	58	113	TM (NC)	N/P	N/A	Novel	A	−7
PKD2								
R119H	29	125	N/P	N/P	N/A	Novel	B	−7
V262M	21	32	N/P	N/P	N/A	Novel	A	−7

Table IVB. Classification of newly described likely polymorphic changes: Splicing variants

Designation	RT-PCR Change[i]	Splicing Predictions[d]	Segregation Analysis[e]	Other Description[f]	Other Variant[g]	Variant Score[h]
PKD1						
IVS2-11G→C	N/A	SSA (0.88→0.93)	N/A	Novel	B	−8
IVS4+10G→A	N/A	N/P (0.95)	N/A	Novel	B	−8
IVS12-15C→T	N/A	SSA (0.8→0.84)	N/A	5X	3A,B,C	−25
IVS14-15A→G	N/A	SSA (0.96→0.97)	N/A	2X	2A	−10
PKD2						
1869C→G	N	ND (0.39)	F3	Novel	C	−8

Table IVC. Classification of newly described likely polymorphic changes: Deletion

Designation	Orthologs Conservation[k]	Protein Domain Conservation[c]	Segregation Analysis[e]	Other Description[f]	Other Variant[g]	Variant Score[h]
PKD1						
2894delANS	NC	N/P	N	Novel	B	−23

[a]Score of chemical difference between the normal and mutated residue (high score, greater difference).
[b]Score of chemical difference between orthologs (human, mouse [m], rat [r], dog [d], chicken [c], frog [fr], and fish [fi]; 0 = completely conserved). Species that match the substitution are shown.
[c]Domain containing residue: LRR, WSC, PKD, REJ, PLAT; N/P, no domain predicted; C, conserved (80 to 50%); NC, not conserved (<50%); indicates whether variant residue matches domain sequence (residue #X).
[d]Predicted splicing changes; N/P, none predicted; SSA, slightly strengthens acceptor; ND, normal donor; 0.1 to 1, weakly to strongly predicted.
[e]Segregation demonstrated in: S, sibs #; F, family # (at least two generations); N/A, samples not available.
[f]Recurrent within study (#X) or previously described as mutation (#X, ref); substitution of same residue.
[g]Other variant (mutation score, A, B or C; indeterminate variants, I; or no variant, N); see Table I or V for details.
[h]Composite variant score: ≤−5 = neutral polymorphism.
[i]Aberrant splicing demonstrated by RT-PCR; N/A, not available; N, no.
[j]Details of a previous description as a mutation or a variant of the same residue in HGMD and/or PKDB.

Table V. Details of molecular analysis of mutation negative cases

CRISP ID[a]	*N*[b]	Gene Linkage[c]	Deletion Analysis[d]	Indeterminate Variants or Novel Polymorphisms				Variant Group
				Designation[e]	Splicing Prediction[f]	Segregation Analysis[g]	Other Description	
100002	2	NC	NVD	V2897I	N/P	N	Novel	P
100006	2	*PKD2*	N/A	9795C→T	N/P	N	Novel	P
100009	5	*PKD1*	NA;HEM	N/D				
103227	1	N/P	N/A	V609G	N/P	N/A	Novel	I
				N970K	N/P	N/A	Novel	I
118641	1	N/P	N/A	V690G	N/P	N/A	Novel (V690D)[h]	I
				R1340W	N/P	N/A	1X[h]	I

(Table V continued on next page)

Table V. Continued

CRISP ID[a]	*n*[b]	Gene Linkage[c]	Deletion Analysis[d]	Indeterminate Variants or Novel Polymorphisms				Variant Group
				Designation[e]	Splicing Prediction[f]	Segregation Analysis[g]	Other Description	
157925	1	N/P	NVD	10251G→T	N/P	N/A	Novel	P
160928	1	N/P	NVD	3006G→C	N/P	N/A	Novel	P
188086	1	N/P	N/A	N/D				
213454	1	N/P	N/A	3930C→T	N/P	N/A	Novel	P
235752	1	N/P	NVD	N/D				
244111	1	N/P	PV	N/D				
252086	1	N/P	N/A	V466L	SWA (0.65→0.60)	N/A	Novel (V466M)	I
				IVS35-14C→T	SSA (0.95→0.96)	N/A	Novel	I
259940	1	N/P	N/A;HEM	N/D				
299663	1	N/P	N/A	V466M	N/P	F2	Novel (V466L)	I
				8176C→T	SSA (0.57→0.62)	N	Novel	P
313893	1	N/P	NVD	N/D				
364664	1	N/P	NVD	N/D				
369941	1	N/P	NVD	N/D				
406737	1	N/P	NVD	S1352N	N/P	N/A	Novel	I
				IVS6-46A→G	N/P	N/A	Novel	P
407132	1	N/P	N/A	L2696P	N/P	N/A	Novel	I
464923	1	N/P	NVD	N/D				
476972	1	N/P	N/A	H1093D	N/P	N/A	Novel (H1093Y)	I
493328	1	N/P	PV;HEM	N/D				

[a]Patient or pedigree number.
[b]Number of patients from family in CRISP study.
[c]NC, linkage analysis not conclusive; N/P, not possible.
[d]FIGE analysis for large deletions; NVD, no variant detected; N/A, samples not available for analysis; PV, possible variant but not conclusively characterized; HEM, apparent homozygosity (hemizygosity) for one or more rare variants suggesting an uncharacterized deletion mutation.
[e]Details of amino acid changes shown in Tables III and IV. N/D, no novel variants detected; *PKD2* changes italicized.
[f]Predicted splicing changes; N/P, none predicted; SWA, slightly weakens acceptor; SSA, slightly strengthens acceptor.
[g]N, does not segregate; N/A, samples not available; F#, segregation demonstrated in family (at least two generations).
[h]Previous description of variant as a mutation or change of the same residue, details in HGMD and/or PKDB.

REFERENCES

1. European Polycystic Kidney Disease Consortium: The polycystic kidney disease 1 gene encodes a 14 kb transcript and lies within a duplicated region on chromosome 16. *Cell* 77: 881–894, 1994
2. Hughes J, Ward CJ, Peral B, Aspinwall R, Clark K, San Millan JL, Gamble V, Harris PC: The polycystic kidney disease 1 (*PKD1*) gene encodes a novel protein with multiple cell recognition domains. *Nat Genet* 10: 151–160, 1995
3. International Polycystic Kidney Disease Consortium: Polycystic kidney disease: The complete structure of the *PKD1* gene and its protein. *Cell* 81: 289–298, 1995
4. Mochizuki T, Wu G, Hayashi T, Xenophontos SL, Veldhusien B, Saris JJ, Reynolds DM, Cai Y, Gabow PA, Pierides A, Kimberling WJ, Breuning MH, Deltas CC, Peters DJM, Somlo S: PKD2, a gene for polycystic kidney disease that encodes an integral membrane protein. *Science* 272: 1339–1342, 1996
5. Peters DJM, Sandkuijl LA: Genetic heterogeneity of polycystic kidney disease in Europe. *Contrib Nephrol* 97: 128–139, 1992
6. Torra R, Badenas C, Darnell A, Nicolau C, Volpini V, Revert L, Estivill X: Linkage, clinical features, and prognosis of autosomal dominant polycystic kidney disease types 1 and 2. *J Am Soc Nephrol* 7: 2142–2151, 1996
7. Daoust MC, Reynolds DM, Bichet DG, Somlo S: Evidence for a third genetic locus for autosomal dominant polycystic kidney disease. *Genomics* 25: 733–736, 1995
8. Hateboer N, van Dijk MA, Bogdanova N, Coto E, Saggar-Malik AK, San Millan JL, Torra R, Breuning M, Ravine D: Comparison of phenotypes of polycystic kidney disease types 1 and 2. *Lancet* 353: 103–107, 1999
9. Ong AC, Harris PC: Molecular pathogenesis of ADPKD: The polycystin complex gets complex. *Kidney Int* 67: 1234–1247, 2005
10. Torres VE, Harris PC: Mechanisms of disease: Autosomal dominant and recessive polycystic kidney disease. *Nat Clin Pract Nephrol* 2: 40–55, 2006
11. Yoder BK, Hou X, Guay-Woodford LM: The polycystic kidney disease proteins, polycystin-1, polycystin-2, polaris, and cystin, are co-localized in renal cilia. *J Am Soc Nephrol* 13: 2508–2516, 2002
12. Pazour GJ, San Agustin JT, Follit JA, Rosenbaum JL, Witman GB: Polycystin-2 localizes to kidney cilia and the ciliary level is elevated in orpk mice with polycystic kidney disease. *Curr Biol* 12: R378–R380, 2002
13. Nauli SM, Alenghat FJ, Luo Y, Williams E, Vassilev P, Li X, Elia AE, Lu W, Brown EM, Quinn SJ, Ingber DE, Zhou J: Polycystins 1 and 2 mediate mechanosensation in the primary cilium of kidney cells. *Nat Genet* 33: 129–137, 2003
14. Ravine D, Gibson RN, Walker RG, Sheffield LJ, Kincaid-Smith P, Danks DM: Evaluation of ultrasonographic diagnostic criteria for autosomal dominant polycystic kidney disease 1. *Lancet* 343: 824–827, 1994
15. Grantham JJ, Torres VE, Chapman AB, Guay-Woodford LM, Bae KT, King BF Jr, Wetzel LH, Baumgarten DA, Kenney PJ, Harris PC, Klahr S, Bennett WM, Hirschman GN, Meyers CM, Zhang X, Zhu F, Miller JP: Volume progression in polycystic kidney disease. *N Engl J Med* 354: 2122–2130, 2006
16. Chapman AB, Guay-Woodford LM, Grantham JJ, Torres VE, Bae KT,

Baumgarten DA, Kenney PJ, King BF Jr, Glockner JF, Wetzel LH, Brummer ME, O'Neill WC, Robbin ML, Bennett WM, Klahr S, Hirschman GH, Kimmel PL, Thompson PA, Miller JP: Renal structure in early autosomal-dominant polycystic kidney disease (ADPKD): The Consortium for Radiologic Imaging Studies of Polycystic Kidney Disease (CRISP) cohort. *Kidney Int* 64: 1035–1045, 2003

17. Rossetti S, Chauveau D, Walker D, Saggar-Malik A, Winearls CG, Torres VE, Harris PC: A complete mutation screen of the ADPKD genes by DHPLC. *Kidney Int* 61: 1588–1599, 2002
18. Rossetti S, Strmecki L, Gamble V, Burton S, Sneddon V, Peral B, Roy S, Bakkaloglu A, Komel R, Winearls CG, Harris PC: Mutation analysis of the entire PKD1 gene: Genetic and diagnostic implications. *Am J Hum Genet* 68: 46–63, 2001
19. Veldhuisen B, Saris JJ, de Haij S, Hayashi T, Reynolds DM, Mochizuki T, Elles R, Fossdal R, Bogdanova N, van Dijk MA, Coto E, Ravine D, Norby S, Verellen-Dumoulin C, Breuning MH, Somlo S, Peters DJM: A spectrum of mutations in the second gene for autosomal dominant polycystic kidney disease (PKD2). *Am J Hum Genet* 61: 547–555, 1997
20. Grantham R: Amino acid difference formula to help explain protein evolution. *Science* 185: 862–864, 1974
21. Mathe E, Olivier M, Kato S, Ishioka C, Hainaut P, Tavtigian SV: Computational approaches for predicting the biological effect of p53 missense mutations: A comparison of three sequence analysis based methods. *Nucleic Acids Res* 34: 1317–1325, 2006
22. Tavtigian SV, Deffenbaugh AM, Yin L, Judkins T, Scholl T, Samollow PB, de Silva D, Zharkikh A, Thomas A: Comprehensive statistical study of 452 BRCA1 missense substitutions with classification of eight recurrent substitutions as neutral. *J Med Genet* 43: 295–305, 2006
23. Abkevich V, Zharkikh A, Deffenbaugh AM, Frank D, Chen Y, Shattuck D, Skolnick MH, Gutin A, Tavtigian SV: Analysis of missense variation in human BRCA1 in the context of interspecific sequence variation. *J Med Genet* 41: 492–507, 2004
24. Ng PC, Henikoff S: Predicting the effects of amino acid substitutions on protein function. *Annu Rev Genomics Hum Genet* 7: 61–80, 2006
25. Chenevix-Trench G, Healey S, Lakhani S, Waring P, Cummings M, Brinkworth R, Deffenbaugh AM, Burbidge LA, Pruss D, Judkins T, Scholl T, Bekessy A, Marsh A, Lovelock P, Wong M, Tesoriero A, Renard H, Southey M, Hopper JL, Yannoukakos K, Brown M, Easton D, Tavtigian SV, Goldgar D, Spurdle AB: Genetic and histopathologic evaluation of BRCA1 and BRCA2 DNA sequence variants of unknown clinical significance. *Cancer Res* 66: 2019–2027, 2006
26. Sharp AM, Messiaen LM, Page G, Antignac C, Gubler MC, Onuchic LF, Somlo S, Germino GG, Guay-Woodford LM: Comprehensive genomic analysis of *PKHD1* mutations in ARPKD cohorts. *J Med Genet* 42: 336–349, 2005
27. Harris PC, Bae K, Rossetti S, Torres VE, Grantham JJ, Chapman A, Guay-Woodford L, King BF, Wetzel LH, Baumgarten D, Kenney PJ, Consugar M, Klahr S, Bennett WM, Meyers CM, Zhang Q, Thompson PA, Zhu F, Miller JP, Consortium C: Cyst number but not the rate of cystic growth is associated with the mutated gene in ADPKD. *J Am Soc Nephrol* 17: 3013–3019, 2006
28. Peral B, Gamble V, San Millán JL, Strong C, Sloane-Stanley J, Moreno F, Harris PC: Splicing mutations of the polycystic kidney disease 1 (*PKD1*) gene induced by intronic deletion. *Hum Mol Genet* 4: 569–574, 1995
29. Li A, Tian X, Sung S-W, Somlo S: Identification of two novel polycystic kidney disease-1-like genes in human and mouse genomes. *Genomics* 81: 596–608, 2003
30. Parnell SC, Magenheimer BS, Maser RL, Rankin CA, Smine A, Okamoto T, Calvet JP: The polycystic kidney disease-1 protein, polycystin-1, binds and activates heterotrimeric G-proteins *in vitro*. *Biochem Biophys Res Commun* 251:625–631, 1998
31. Watnick TJ, Piontek KB, Cordal TM, Weber H, Gandolph MA, Qian F, Lens XM, Heumann HPH, Germino GG: An unusual pattern of mutation in the duplicated portion of *PKD1* is revealed by use of a novel strategy for mutation detection. *Hum Mol Genet* 6: 1473–1481, 1997
32. Cooper DN, Krawczak M, Antonorakis SE: The nature and mechanisms of human gene mutation. In: *Metabolic and Molecular Bases of Inherited Disease*, 7th ed, edited by Scriver C, Beaudet AL, Sly WS, Valle D, New York, McGraw-Hill, 1995, pp 259–291
33. Watnick TJ, Gandolph MA, Weber H, Neumann HPH, Germino GG: Gene conversion is a likely cause of mutation in *PKD1*. *Hum Mol Genet* 7: 1239–1243, 1998
34. Nickerson ML, Weirich G, Zbar B, Schmidt LS: Signature-based analysis of MET proto-oncogene mutations using DHPLC. *Hum Mutat* 16: 68–76, 2000
35. Ariyurek Y, Leeuwen IL, Spruit L, Ravine D, Breuning MH, Peters DJ: Large deletions in the polycystic kidney disease 1 (*PKD1*) gene. *Hum Mutat* 23: 99, 2004
36. Sellner LN, Taylor GR: MLPA and MAPH: new techniques for detection of gene deletions. *Hum Mutat* 23: 413–419, 2004
37. Koulen P, Cai Y, Geng L, Maeda Y, Nishimura S, Witzgall R, Ehrlich BE, Somlo S: Polycystin-2 is an intracellular calcium release channel. *Nat Cell Biol* 4: 191–197, 2002
38. Qian F, Boletta A, Bhunia AK, Xu H, Liu L, Ahrabi AK, Watnick TJ, Zhou F, Germino GG: Cleavage of polycystin-1 requires the receptor for egg jelly domain and is disrupted by human autosomal-dominant polycystic kidney disease 1-associated mutations. *Proc Natl Acad Sci U S A* 99: 16981–16986, 2002
39. Pagenstecher C, Wehner M, Friedl W, Rahner N, Aretz S, Friedrichs N, Sengteller M, Henn W, Buettner R, Propping P, Mangold E: Aberrant splicing in MLH1 and MSH2 due to exonic and intronic variants. *Hum Genet* 119: 9–22, 2006
40. Cartegni L, Chew SL, Krainer AR: Listening to silence and understanding nonsense: Exonic mutations that affect splicing. *Nat Rev Genet* 3: 285–298, 2002
41. Cartegni L, Wang J, Zhu Z, Zhang MQ, Krainer AR: ESEfinder: A web resource to identify exonic splicing enhancers. *Nucleic Acids Res* 31: 3568–3571, 2003
42. King K, Flinter FA, Nihalani V, Green PM: Unusual deep intronic mutations in the COL4A5 gene cause X linked Alport syndrome. *Hum Genet* 111: 548–554, 2002
43. Lantinga-van Leeuwen IS, Leonhard WN, Dauwerse H, Baelde HJ, van Oost BA, Breuning MH, Peters DJ: Common regulatory elements in the polycystic kidney disease 1 and 2 promoter regions. *Eur J Hum Genet* 13: 649–659, 2005
44. Cooper DN, Krawczak M: The mutational spectrum of single base-pair substitutions causing human genetic disease: Patterns and predictions. *Hum Genet* 85: 55–74, 1990
45. Matyas G, De Paepe A, Halliday D, Boileau C, Pals G, Steinmann B: Evaluation and application of denaturing HPLC for mutation detection in Marfan syndrome: Identification of 20 novel mutations and two novel polymorphisms in the FBN1 gene. *Hum Mutat* 19: 443–456, 2002
46. Ponting CP, Hofmann K, Bork P: A latrophilin/CL-1-like GPS domain in polycystin-1. *Curr Biol* 9: R585–R588, 1999
47. Bycroft M, Bateman A, Clarke J, Hamill S, Sanford R, Thomas R, Chothia C: The structure of a PKD domain from polycystin-1: Implication for polycystic kidney disease. *EMBO J* 18: 297–305, 1999
48. Moy GW, Mendoza LM, Schulz JR, Swanson WJ, Glabe CG, Vacquier VD: The sea urchin sperm receptor for egg jelly is a modular protein with extensive homology to the human polycystic kidney disease protein, PKD1. *J Cell Biol* 133: 809–817, 1996
49. Bateman A, Sandford R: The PLAT domain: A new piece in the PKD1 puzzle. *Curr Biol* 9: R588–R590, 1999
50. Reese MG, Eeckman FH, Kulp D, Haussler D: Improved splice site detection in Genie. *J Clin Endocrinol Metab* 4: 311–323, 1997
51. Cai Y, Maeda Y, Cedzich A, Torres VE, Wu G, Hayashi T, Mochizuki T, Park JH, Witzgall R, Somlo S: Identification and characterization of polycystin-2, the PKD2 gene product. *J Biol Chem* 274: 28557–28565, 1999