Cell, Vol. 77, 881–894, June 17, 1994, Copyright © 1994 by Cell Press

The Polycystic Kidney Disease 1 Gene Encodes a 14 kb Transcript and Lies within a Duplicated Region on Chromosome 16

The European Polycystic Kidney Disease Consortium*

Summary

Autosomal dominant polycystic kidney disease (ADPKD) is a common genetic disorder that frequently results in renal failure due to progressive cyst development. The major locus, *PKD1*, maps to 16p13.3. We identified a chromosome translocation associated with ADPKD that disrupts a gene (*PBP*) encoding a 14 kb transcript in the *PKD1* candidate region. Further mutations of the *PBP* gene were found in *PKD1* patients, two deletions (one a de novo event) and a splicing defect, confirming that *PBP* is the *PKD1* gene. This gene is located adjacent to the *TSC2* locus in a genomic region that is reiterated more proximally on 16p. The duplicate area encodes three transcripts substantially homologous to the *PKD1* transcript. Partial sequence analysis of the *PKD1* transcript shows that it encodes a novel protein whose function is at present unknown.

Introduction

A landmark study by Dalgaard (1957) showed that autosomal dominant polycystic kidney disease (ADPKD), also termed adult polycystic kidney disease, is one of the most common genetic diseases in humans (approximately 1 in 1000 individuals affected). The major feature of this disease is the development of cystic kidneys that commonly leads to renal failure in adult life. This simple description, however, belies the diverse clinical phenotype of ADPKD, which is now considered a systemic disorder (reviewed by Gabow, 1990) and one that occasionally presents in childhood (Fink et al., 1993; Zerres et al., 1993). Extrarenal manifestations include liver cysts (Milutinovic et al., 1980) and (more rarely) cysts of the pancreas (Gabow, 1993) and other organs. Intracranial aneurysms occur in approximately 5% of patients and are a significant cause of morbidity and mortality due to subarachnoid hemorrhage (Chapman et al., 1992). More recently, an increased prevalence of cardiac valve defects (Hossack et al., 1988), herniae (Gabow, 1990), and colonic diverticulae (Scheff et al., 1980) has been reported.

The major cause of morbidity in ADPKD, however, is progressive renal disease characterized by the formation and enlargement of fluid-filled cysts, resulting in grossly enlarged kidneys. Renal function deteriorates as normal tissue is compromised by cystic growth, resulting in end stage renal disease (ESRD) in more than 50% of patients by the age of 60 years (Gabow et al., 1992): ADPKD accounts for 8%–10% of all renal transplantation and dialysis patients in Europe and the United States (Gabow, 1993). Biochemical studies have suggested several potential causes of cyst formation and development, including abnormal epithelial cell growth, alterations to the extracel-

The European Polycystic Kidney Disease Consortium is comprised of the following groups:

Group 1:
Christopher J. Ward, Belén Peral, Jim Hughes, Sandra Thomas, Vicki Gamble, Angela B. MacCarthy, Jackie Sloane-Stanley, Veronica J. Buckle, Lyndal Kearney, Douglas R. Higgs, Peter J. Ratcliffe, and Peter C. Harris†
Medical Research Council Molecular Haematology Unit
Institute of Molecular Medicine
John Radcliffe Hospital
Headington, Oxford OX3 9DU
England

Group 2:
Jeroen H. Roelfsema, Lia Spruit, Jasper J. Saris, Hans G. Dauwerse, Dorien J. M. Peters, and Martijn H. Breuning
Department of Human Genetics
Leiden University
2333 AL Leiden
The Netherlands

†*Correspondence should be addressed to Peter C. Harris.*

Group 3:
Mark Nellist,[1] Phillip T. Brook-Carter,[1] Magitha M. Maheshwar,[1] Isabel Cordeiro,[2] Heloisa Santos,[2] Pedro Cabral,[3] and Julian R. Sampson[1]
[1]Institute of Medical Genetics
University of Wales College of Medicine
Cardiff CF4 4XN
Wales
[2]Genetics Unit
Hospital Santa Maria
[3]Department of Neurology
Hospital D. Estefânia
1699 Lisbon
Portugal

Group 4:
Bart Janssen, Arjenne L. W. Hesseling-Janssen, Ans M. W. van den Ouweland, Bert Eussen, Senno Verhoef, Dick Lindhout, and Dicky J. J. Halley
Department of Clinical Genetics
Erasmus University
and University Hospital
3015 GE Rotterdam
The Netherlands

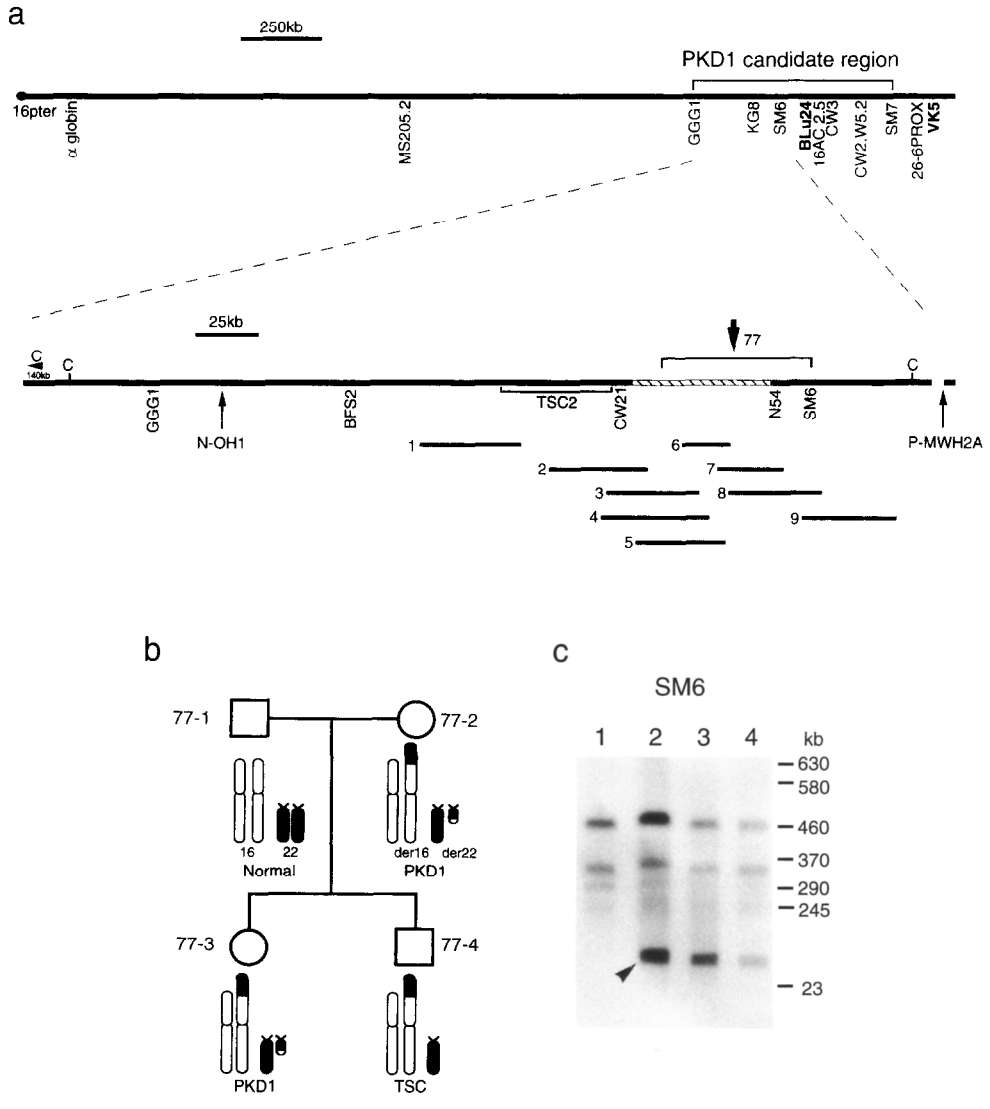

Figure 1. A Chromosome Translocation Associated with *PKD1*

(a) (Top) A long-range map of the terminal region of the short arm of chromosome 16 showing the *PKD1* candidate region defined by genetic linkage analysis. The positions of selected DNA probes and microsatellites used for haplotype, linkage, or heterozygosity analysis are indicated. Markers previously described in linkage disequilibrium studies are shown in bold (adapted from Harris et al., 1990, 1991; Germino et al., 1992; Somlo et al., 1992; Peral et al., 1994; Snarey et al., 1994). (Bottom) A detailed map of the distal part of the *PKD1* candidate region showing the area of 16p13.3 duplicated in 16p13.1 (hatched); ClaI restriction sites (C); the breakpoints in the somatic cell hybrids *N-OH1* and *P-MWH2A*; DNA probes; and the *TSC2* gene. The limits of the position of the translocation breakpoint found in family 77 (see [b]), determined by evidence of heterozygosity (in 77-4) and by PFGE (see [c] and text) is also indicated. The contig covering the family 77 breakpoint region consists of the following cosmids: 1, *CW9D*; 2, *ZDS5*; 3, *JH2A*; 4, *REP59*; 5, *JC10.2B*; 6, *CW10III*; 7, *SM25A*; 8, *SMII*; 9, *NM17*.

(b) Pedigree of family 77, which segregates a 16;22 translocation, showing the chromosomal composition of each subject. Individuals 77-2 and 77-3 have the balanced products of the exchange and have *PKD1*; 77-4 is monosomic for 16p13.3→16pter and 22q11.21→22pter and has TSC.

(c) PFGE of DNA from members of family 77: 77-1 (1), 77-2 (2), 77-3 (3), and 77-4 (4), digested with ClaI and hybridized with *SM6*. In addition to the normal fragments of 340 kb and partially digested fragment of 480 kb, a proximal breakpoint fragment of approximately 100 kb (shown by arrow) is seen in individuals 77-2, 77-3, and 77-4, concordant with segregation of the der(16) chromosome.

lular matrix, and changes in cellular polarity and secretion (reviewed by Gabow, 1991; Wilson and Sherwood, 1991). The primary defect in ADPKD, however, remains unclear, and considerable effort has therefore been applied to identifying the defective gene(s) in this disorder by genetic approaches.

The first step toward positional cloning of an ADPKD gene was the demonstration of linkage of one locus, now designated as the polycystic kidney disease 1 (*PKD1*) locus, to the α-globin cluster on the short arm of chromosome 16 (Reeders et al., 1985). Subsequently, families with ADPKD unlinked to markers of 16p were described (Kimberling et al., 1988; Romeo et al., 1988), and a second ADPKD locus (*PKD2*) has recently been assigned to chromosome region 4q13-q23 (Kimberling et al., 1993; Peters et al., 1993). It is estimated that approximately 85% of

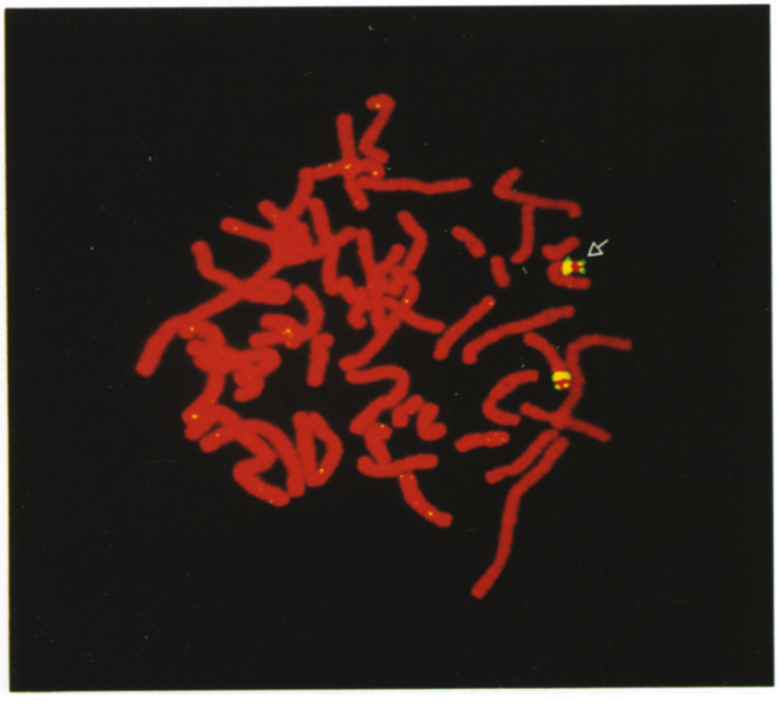

Figure 2. FISH of Cosmid from the Duplicated Area to a Normal Male Metaphase

Duplication of the locus *CW10III* (cosmid 6 in Figure 1a) is illustrated with two sites of hybridization on 16p; the distal site (the *PKD1* region) is shown by an arrow. The signal from the proximal site (16p13.1) is stronger than that from the distal, indicating that sequences homologous to *CW10III* are reiterated in 16p13.1.

ADPKD is due to *PKD1* (Peters and Sandkuijl, 1992), with *PKD2* accounting for most of the remainder. *PKD2* appears to be a milder condition with a later age of onset and ESRD (Parfrey et al., 1990; Gabow et al., 1992; Ravine et al., 1992).

The position of the *PKD1* locus was refined to chromosome band 16p13.3, and many markers were isolated from that region (Breuning et al., 1987, 1990; Reeders et al., 1988; Germino et al., 1990; Hyland et al., 1990; Himmelbauer et al., 1991). Their order and the position of the *PKD1* locus have been determined by extensive linkage analysis in normal and *PKD1* families and by the use of a panel of somatic cell hybrids (Reeders et al., 1988; Breuning et al., 1990; Germino et al., 1990). An accurate long range restriction map (Harris et al., 1990; Germino et al., 1992) has located the *PKD1* locus in an interval of approximately 600 kb between the markers *GGG1* and *SM7* (Harris et al., 1991; Somlo et al., 1992) (see Figure 1a). The density of CpG islands and identification of many mRNA transcripts indicated that this area is rich in gene sequences. Germino et al. (1992) estimated that the candidate region contains approximately 20 genes.

Identification of the *PKD1* gene from within this area has thus proved difficult, and other means to pinpoint the disease gene were sought. Linkage disequilibrium has been demonstrated in a Scottish population between *PKD1* and the proximal marker *VK5* (Pound et al. , 1992) and in a Spanish population between *PKD1* and *BLu24* (see Figure 1a) (Peral et al., 1994). Studies with additional markers have shown evidence of a common ancestor in a proportion of each population (Peral et al., 1994; Snarey et al., 1994), but the association has not precisely positioned the *PKD1* locus.

Disease-associated genomic rearrangements, detected by cytogenetics or pulsed-field gel electrophoresis (PFGE), have been instrumental in the identification of the genes associated with many genetic disorders. In the case of the tuberous sclerosis (TSC) locus *TSC2*, which lies within 16p13.3, deletions were detected by PFGE, within the interval thought to contain the *PKD1* gene, and their characterization was a significant step toward the rapid identification of the *TSC2* gene (European Chromosome 16 Tuberous Sclerosis Consortium, 1993).

We have now identified a pedigree in which the two distinct phenotypes, typical ADPKD or TSC, are seen in different members. In this family the two individuals with ADPKD are carriers of a balanced chromosome translocation with a breakpoint within 16p13.3. We have located the chromosome 16 translocation breakpoint, and a gene disrupted by this rearrangement has been defined; the discovery of additional mutations of that gene in other *PKD1* patients shows that we have identified the *PKD1* gene.

Results

A Translocation Associated with ADPKD

A major pointer to the identity of the *PKD1* gene was provided by a Portuguese pedigree (family 77) with both ADPKD and TSC (Figure 1b). Cytogenetic analysis showed that the mother, 77-2, has a balanced translocation, 46XX t(16;22)(p13.3;q11.21), that was inherited by her daughter, 77-3. The son, 77-4, has the unbalanced karyotype 45XY/ – 16-22+der(16)(16qter→16p13.3::22q11.21→22qter) and consequently is monosomic for 16p13.3→16pter as well as for 22q11.21→22pter. This individual has the clinical phenotype of TSC (see Experimental Procedures); the

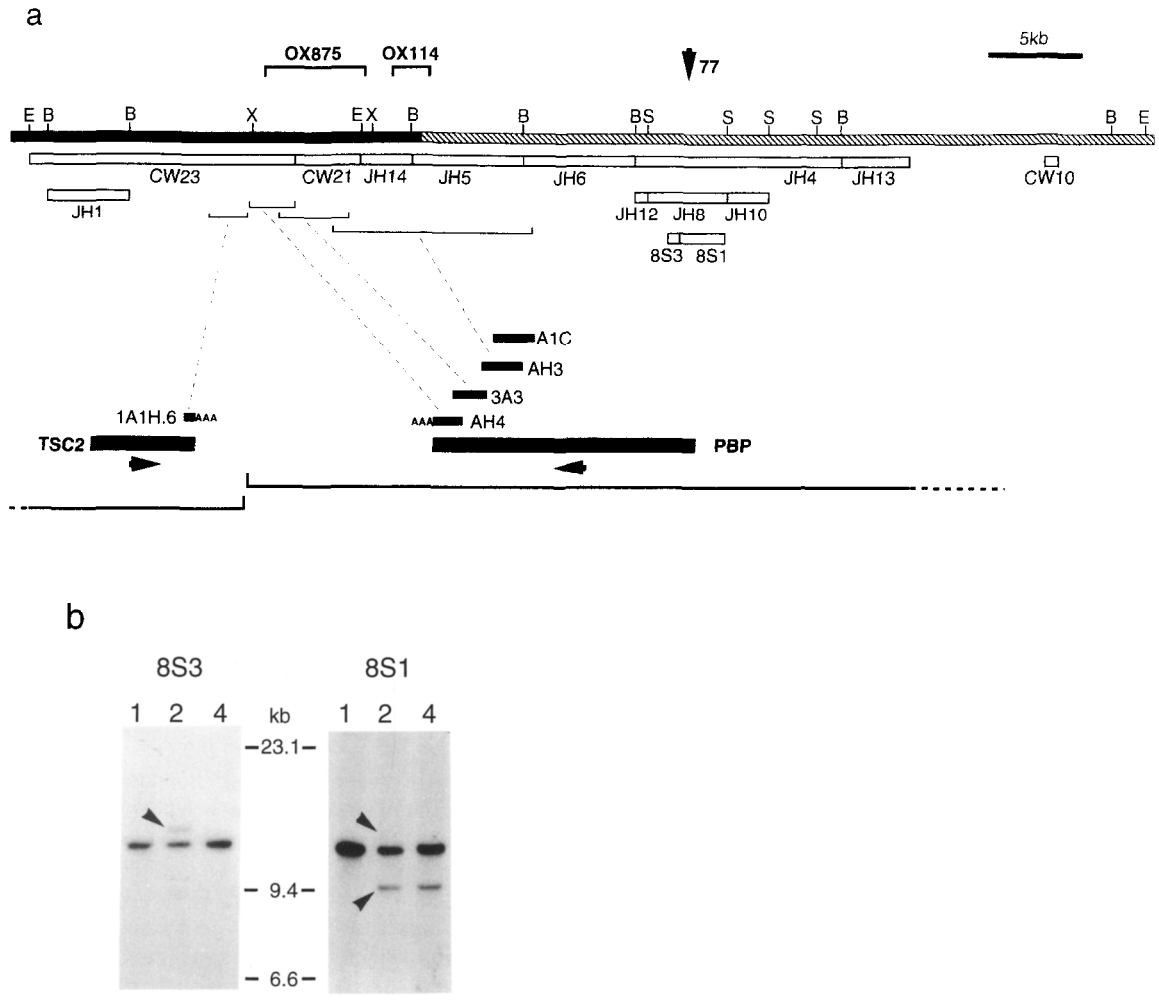

Figure 3. A Detailed Map of the *PBP* Gene Region Showing the Family 77 Translocation Breakpoint

(a) A detailed map of the translocation region showing the precise localization of the family 77 breakpoint and the region that is duplicated in 16p13.1 (hatched). DNA probes (open boxes), the transcripts *PBP* and *TSC2* (closed boxes; with direction of transcription indicated by an arrow), and cDNAs (stippled boxes) are shown below the genomic map. The known genomic extent of each gene is indicated at the bottom of the diagram, and the approximate genomic locations of each cDNA are indicated under the genomic map. The positions of genomic deletions found in *PKD1* patients OX875 and OX114 are also indicated. Restriction sites for EcoRI (E) and incomplete maps for BamHI (B), SacI (S), and XbaI (X) are shown.

(b) Southern blots of BamHI-digested DNA from individuals 77-1 (1), 77-2 (2), and 77-4 (4) hybridized with *8S3* (left) and *8S1* (right) (see [a]). *8S3* detects a novel fragment on the telomeric side of the breakpoint (12 kb; shown by arrow), associated with the der(22) chromosome in 77-2, but not 77-4; *8S1* identifies a novel fragment on the centromeric side of the breakpoint (9 kb; shown by arrow), associated with the der(16) chromosome in 77-2 and 77-4. The telomeric breakpoint fragment is also seen weakly with *8S1* (shown by arrow), indicating that the breakpoint lies in the distal part of *8S1*. The *8S3* and *8S1* loci are both duplicated; the normal BamHI fragment detected at the 16p13.3 site by these probes is 11 kb (see [a]), but a similarly sized fragment is also detected at the 16p13.1 site. Consequently, the breakpoint fragments are much fainter than the normal (16p13.1 plus 16p13.3) band.

most likely explanation was that the *TSC2* locus located within 16p13.3 was deleted in the unbalanced karyotype.

Further analysis revealed that the mother (77-2) and the daughter (77-3) with the balanced translocation have the clinical features of ADPKD (see Experimental Procedures), while the parents of 77-2 (aged 67 and 82 years) were cytogenetically normal, with no clinical features of TSC and no renal cysts on ultrasound examination. Although kidney cysts can be a feature of TSC, no other clinical signs of TSC were identified in 77-2 or 77-3, making it unlikely that the polycystic kidneys were due to TSC. We therefore investigated the possibility that the translocation disrupted the *PKD1* locus in 16p13.3 and proceeded to identify and clone the region containing the breakpoint.

Family 77 was analyzed with polymorphic markers from 16p13.3. Individual 77-4 was hemizygous for *MS205.2* and *GGG1*, but heterozygous for *SM6* and more proximal markers, locating the translocation breakpoint between *GGG1* and *SM6* (see Figure 1a). Fluorescence in situ hybridization (FISH) of a cosmid from the *TSC2* region, *CW9D* (cosmid 1 in Figure 1a), to metaphase spreads showed that it hybridized to the der(22) chromosome of 77-2, placing the breakpoint proximal to *CW9D* and indicating that 77-4 was hemizygous for this region, consistent with his TSC

phenotype. DNA from members of family 77 was digested with ClaI, separated by PFGE, and hybridized with *SM6*, revealing a breakpoint fragment of ~100 kb in individuals with the der(16) chromosome (Figure 1c). The small size of this novel fragment enabled the breakpoint to be localized distal to *SM6* in a region of just 60 kb (Figure 1a). A cosmid contig covering this region was therefore constructed (see Experimental Procedures for details).

The Translocation Breakpoint Lies within a Region Duplicated Elsewhere on Chromosome 16 (16p13.1)

It was previously noted that the region between *CW21* and *N54* (Figure 1a) was duplicated at a more proximal site on the short arm of chromosome 16 (Germino et al., 1992; European Chromosome 16 Tuberous Sclerosis Consortium, 1993). Figure 2 shows that a cosmid, *CW10III*, from the duplicated region hybridizes to two points on 16p: the distal *PKD1* region and a proximal site positioned in 16p13.1. The structure of the duplicated area is complex, with each fragment present once in 16p13.3 reiterated two to four times in 16p13.1 (see Figure 2). Cosmids spanning the duplicated area in 16p13.3 were subcloned (see Figure 3a; see Experimental Procedures for details), and a restriction map was generated. A genomic map of the *PKD1* region was constructed using a radiation hybrid, *Hy145.19*, that contains the distal portion of 16p but not the duplicate site in 16p13.1.

To localize the family 77 translocation breakpoint, subclones from the target region were hybridized to DNA from patient 77-2, digested with ClaI, and separated by PFGE. Once probes mapping across the breakpoint were identified, they were hybridized to conventional Southern blots of family 77 DNA. Figure 3b shows that novel BamHI fragments were detected from the centromeric and telomeric side of the breakpoint, which was localized to the distal part of the probe *8S1* (Figure 3a). Hence, the balanced translocation was not associated with a substantial deletion, and the breakpoint was located more than 20 kb proximal to the *TSC2* locus (Figure 3a). These results supported the hypothesis that ADPKD in individuals with the balanced translocation (77-2 and 77-3) was not due to disruption of the *TSC2* gene, but indicated that a separate gene mapping just proximal to *TSC2* was likely to be the *PKD1* gene.

The Polycystic Breakpoint Gene Is Disrupted by the Translocation

Localization of the family 77 breakpoint identified a precise region in which to look for a candidate for the *PKD1* gene. During the search for the *TSC2* gene, we identified other transcripts not associated with TSC, including a large transcript (~14 kb) partially represented in the cDNAs *3A3* and *AH4* that mapped to the genomic fragments *CW23* and *CW21* (Figure 3a). The orientation of the gene encoding this transcript had been determined by the identification of a poly(A) tract in the cDNA, *AH4*: the 3′ end of this gene lies very close to the *TSC* gene, in a tail-to-tail orientation (European Chromosome 16 Tuberous Sclerosis Consortium, 1993). To determine whether this gene crossed the translocation breakpoint, genomic probes from within the duplicated area and flanking the breakpoint were hybridized to Northern blots. Probes from both sides of the breakpoint, between *JH5* and *JH13*, identified the 14 kb transcript (Figure 3a; see below for details). Therefore, this gene (previously called *3A3*, but now designated the polycystic breakpoint [*PBP*] gene) extended over the family 77 breakpoint and consequently was a candidate for the *PKD1* gene. A walk was initiated to increase the extent of the *PBP* cDNA contig, and several novel cDNAs were identified using probes from the single copy (nonduplicated) region (see Experimental Procedures for details). A cDNA contig was constructed that extended ~5.7 kb, including ~2 kb into the area that is duplicated (Figure 3a).

Expression of the *PBP* Gene

Initial studies of the expression pattern of the *PBP* gene were undertaken with cDNAs that map entirely within the single copy region (e.g., *AH4* and *3A3*). Figure 4a shows that the ~14 kb transcript was identified by *3A3* in various tissue-specific cell lines. From this and other Northern blots, we concluded that the *PBP* gene was expressed in all of the cell lines tested, although often at a low level. The two cell lines that showed the highest level of expression were fibroblasts and a cell line derived from an astrocytoma, G-CCM. Significant levels of expression were also obtained in cell lines derived from kidney (G401) and liver (Hep3B). Measuring the expression of the *PBP* gene in tissue samples by Northern blotting proved difficult because such a large transcript is susceptible to minor RNA degradation. However, initial results with an RNAase protection assay, using a region of the gene located in the single copy area (see Experimental Procedures), showed a moderate level of expression of the *PBP* gene in tissue obtained from normal and polycystic kidney (data not shown). The widespread expression of the *PBP* gene is consistent with the systemic nature of ADPKD.

Identification of Transcripts That Are Partially Homologous to the *PBP* Transcript

Novel cDNAs were identified with the genomic fragments, *JH4* and *JH8*, that map to the duplicated region (see Figure 3a; see Experimental Procedures). However, when these cDNAs were hybridized to Northern blots, a more complex pattern than that seen with *3A3* was observed. As well as the ~14 kb *PBP* transcript, three other partially homologous transcripts were identified, designated homologous gene A (*HG-A*) (~21 kb), *HG-B* (~17 kb), and *HG-C* (8.5 kb) (Figure 4b). There were two possible explanations for these results: either the *HG* transcripts were alternatively spliced forms of the *PBP* gene, or the *HG* transcripts were encoded by genes located in 16p13.1. To determine the genomic location of the *HG* loci, a fragment from the 3′ end of one *HG* cDNA (*HG-4/1.1*) was isolated. *HG-4/1.1* hybridized to all three *HG* transcripts, but not to the *PBP* transcript, and on a hybrid panel it mapped to 16p13.1 (not the *PKD1* area). These results show that all the *HG* transcripts are related to each other outside the region of homology with the *PBP* transcript and that the *HG* loci map to the proximal site (16p13.1).

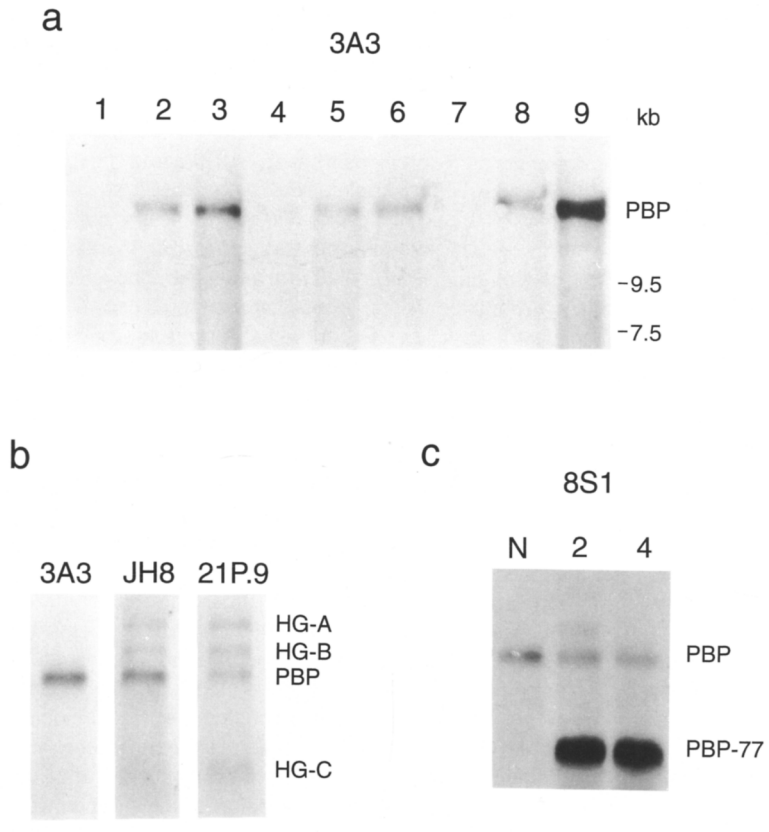

Figure 4. Northern Blot Analysis of the *PBP* Transcript

(a) *PBP* cDNA, *3A3*, hybridized to a Northern blot containing ~1 μg of poly(A)-selected mRNA per lane of the tissue-specific cell lines: lane 1, MJ, Epstein–Barr virus–transformed lymphocytes; lane 2, K562, erythroleukemia; lane 3, FS1, normal fibroblasts; lane 4, HeLa, cervical carcinoma; lane 5, G401, renal Wilm's tumor; lane 6, Hep3B, hepatoma; lane 7, HT29, colonic adenocarcinoma; lane 8, SW13, adrenal carcinoma; lane 9, G-CCM, astrocytoma. A single transcript of approximately 14 kb is seen; the highest level of expression is in fibroblasts and in the astrocytoma cell line G-CCM. Although in this comparative experiment little expression is seen in lanes 1, 4, and 7, we have demonstrated at least a low level of expression in these cell lines on other Northern blots and by RT–PCR (see later).

(b) A Northern blot containing ~20 μg of total RNA from cell line G-CCM hybridized with cDNAs or a genomic probe that each identify different parts of the *PBP* gene. (Left) A single ~14 kb transcript is seen with a cDNA from the single copy area, *3A3*. (Right) A cDNA, *21P.9*, that is homologous to parts of the region that is duplicated (*JH12*, *JH8*, and *JH10*; see Figure 3a) hybridizes to the *PBP* transcript and three novel transcripts: *HG-A* (~21 kb), *HG-B* (~17 kb), and *HG-C* (8.5 kb). A similar pattern of transcripts is seen with cDNAs and genomic fragments that hybridize to the area between *JH5* and *JH13*, with the exception of the *JH8* area. (Middle) *JH8* hybridizes to the transcripts *PBP*, *HG-A*, and *HG-B* but not to *HG-C*.

(c) A Northern blot of 20 μg of total fibroblast RNA from normal control (N), 77-2 (2), and 77-4 (4) hybridized with *8S1*, which contains the 16;22 translocation breakpoint (see Figure 3). A transcript of ~9 kb (*PBP-77*) is identified in the two patients with this translocation, but not in the normal control. *PBP-77* is a chimeric *PBP* transcript formed owing to the translocation and is not seen in RNA from 77-2 or 77-4 with probes that map distal to the breakpoint.

An Abnormal Transcript Associated with the Family 77 Translocation

As the *PBP* gene was transcribed across the region disrupted by the family 77 translocation breakpoint, in a proximal to distal direction on the chromosome (see Figure 3a), it was possible that a novel transcript originating from the *PBP* promotor would be found in this family. Figure 4c shows that using a probe to the *PBP* transcript that mapped mainly proximal to the breakpoint, a novel transcript of approximately 9 kb (*PBP-77*) derived from the der(16) product of the translocation was detected. Interestingly, the *PBP-77* transcript appears to be expressed at a higher level than the normal product of *PBP*. These results confirmed that the family 77 translocation disrupts the *PBP* gene and supports the hypothesis that this is the *PKD1* gene.

Mutations of the *PBP* Gene in Other ADPKD Patients

To prove that the *PBP* gene is the defective gene at the *PKD1* locus, we analyzed this region for mutations in patients with typical ADPKD. The 3′ end of the *PBP* gene was most accessible to study as it maps outside the duplicated area. To screen this region, BamHI digests of DNA from 282 apparently unrelated ADPKD patients were hybridized with the probe *1A1H.6* (see Figure 3a). In addition, a large EcoRI fragment (41 kb) that contains a significant proportion of the *PBP* gene was assayed by field inversion gel electrophoresis (FIGE) in 167 ADPKD patients, using the probe *CW10*. Two genomic rearrangements were identified in ADPKD patients by these procedures, each identified by both methods.

The first rearrangement was identified in patient OX875 (see Experimental Procedures for clinical details), who was shown to have a 5.5 kb genomic deletion within the 3′ end of the *PBP* gene (for details see Figures 5a and 5b and Figure 3a). This genomic deletion results in an ~3 kb internal deletion of the transcript (*PBP-875*), with the ~500 bp adjacent to the poly(A) tail intact. In this family, linkage of ADPKD to chromosome 16 could not be proven because although OX875 has a positive family history of ADPKD, there were no living, affected relatives. However, paraffin-embedded tissue from her affected father (now deceased) was available. We demonstrated, using the polymerase chain reaction (PCR), amplification of a 220 bp fragment spanning the deletion (data not shown), and hence that this individual had the same rearrangement as

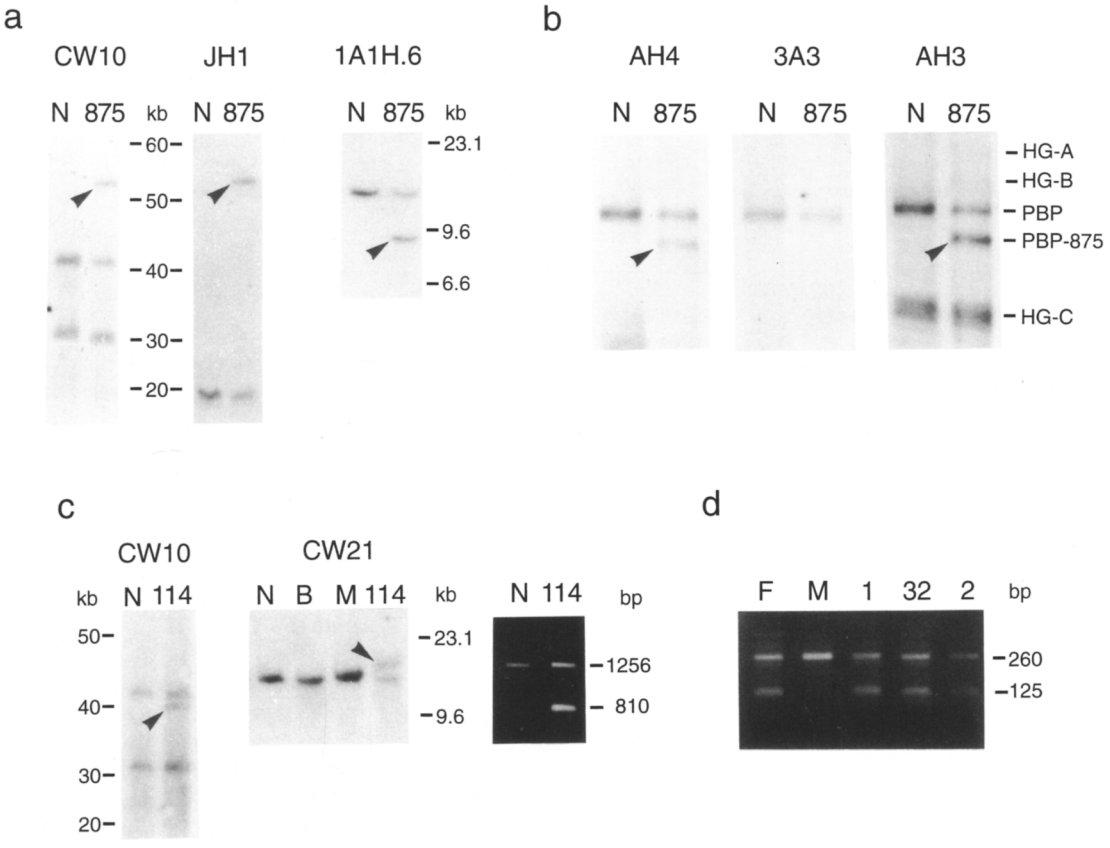

Figure 5. Analysis of Intragenic Mutations of the *PBP* Gene

(a) FIGE of DNA from normal (N) and ADPKD patient OX875 (875), digested with EcoRI and hybridized with *CW10* (left) and *JH1* (middle). Normal fragments of 41 kb (plus a 31 kb fragment from the 16p13.1 site) (*CW10*) and 18 kb (*JH1*) are identified with these probes; OX875 has an additional 53 kb band (shown by arrow). The EcoRI site separating these two fragments is removed by the deletion (see Figure 3a). The right panel shows a Southern blot of BamHI-digested DNA (as above) hybridized with *1A1H.6*. A novel fragment of 9.5 kb is seen in DNA from OX875, as well as the normal 15 kb fragment. These results indicate that OX875 has a 5.5 kb deletion; its position was determined more precisely by mapping relative to two XbaI sites that flank the deletion (see Figure 3a).

(b) Northern blot of total fibroblast RNA, as in (a), hybridized with the cDNAs *AH4*, *3A3*, and *AH3*. A novel transcript (*PBP-875*) of ~11 kb is seen with *AH4* (the band is reduced in intensity because the probe is partly deleted) and *AH3* (shown by arrow), which flank the deletion, but not *3A3*, which is entirely deleted (see Figure 3a). The transcripts *HG-A*, *HG-B*, and *HG-C*, from the duplicated area, are seen with *AH3* (see Figure 4b).

(c) (Left) FIGE of DNA from normal (N) and ADPKD patient OX114 (114), digested with EcoRI and hybridized with *CW10*; a novel fragment of 39 kb (shown by arrow) is seen in OX114. (Middle) DNA, as above, plus the normal mother (M) and brother (B) of OX114, digested with BamHI and hybridized with *CW21*. A larger than normal fragment of 19 kb (shown by arrow) was detected in OX114, but not other family members, owing to deletion of a BamHI site; together these results are consistent with a 2 kb deletion (see Figure 3a). (Right) RT–PCR of RNA, as above, with primers flanking the OX114 deletion (see Experimental Procedures). A novel fragment of 810 bp is seen in OX114, indicating a deletion of 446 bp in the *PBP* transcript.

(d) RT–PCR of RNA from ADPKD patient OX32 (32) plus the probands, normal mother (M) and affected father (F), and sibs (1) and (2), using the *C* primer pair from *3A3* (see Experimental Procedures). A novel fragment of 125 bp is detected in each of the affected individuals.

OX875. This result and analysis of two unaffected sibs of OX875 who did not have the deletion showed that this mutation was transmitted with ADPKD.

The second rearrangement detected by hybridization was a 2 kb genomic deletion within the *PBP* gene, found in ADPKD patient OX114 (see Experimental Procedures for clinical details; see Figures 5c and 3a). No abnormal *PBP* transcript was identified by Northern blot analysis, but using primers flanking the deletion (see Experimental Procedures), a shortened product was detected by reverse transcription (RT)–PCR (Figure 5c). This was cloned and sequenced and shown to have a frameshift deletion of 446 bp (between base pairs 1746 and 2192 of the sequence shown in Figure 7). OX114 is the only member of the family with ADPKD (she has no children), and ultrasound analysis of her parents at age 78 (father) and age 73 (mother) showed no evidence of renal cysts. Somatic cell hybrids were produced from OX114, and the deleted chromosome was found by haplotype analysis to be of paternal origin. The father of OX114 is now deceased, but analysis of DNA from the brother of OX114 (OX984) with seven microsatellite markers from the *PKD1* region (see Experimental Procedures) showed that he shares the same paternal chromosome, in the *PKD1* region, as OX114. Renal ultrasound revealed no cysts in OX984 at age 53, and no deletion was detected by DNA analysis (Figure 5c). Hence, the

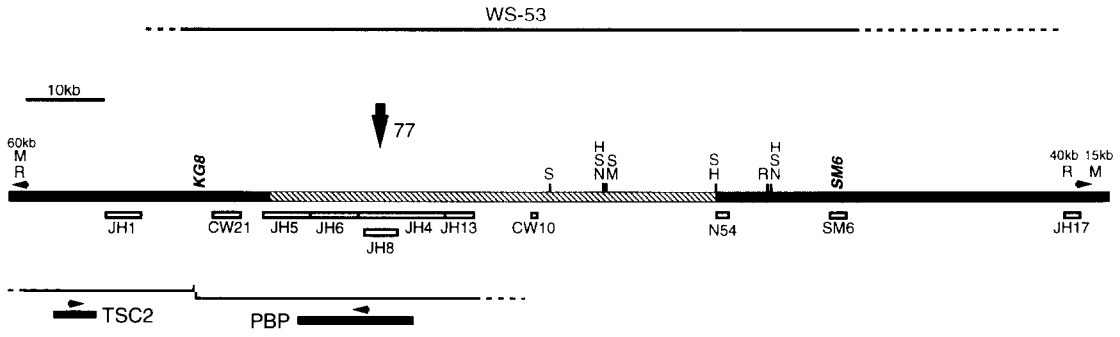

Figure 6. Map of the Region Containing the *TSC2* and *PBP* Genes Showing the Area Deleted in Patient WS-53

Localization of the distal end of the WS-53 deletion was previously described (European Chromosome 16 Tuberous Sclerosis Consortium, 1993), and we have now localized the proximal end between *SM6* and *JHI7*. The size of the aberrant MluI fragment in WS-53, detected by *JH1* and *JH17*, is 90 kb, and these probes lie on adjacent MluI fragments of 120 kb and 70 kb, respectively. Therefore, the WS-53 deletion is ~100 kb. Restriction sites for MluI (M), NruI (R), and NotI (N) and partial maps for SacII (S) and BssHII (H) are shown. DNA probes (open boxes) and the *TSC2* and *PBP* transcripts (closed boxes) are indicated below the line with their known genomic extents (brackets). The locations of the microsatellites *KG8* and *SM6* and the family 77 breakpoint are also indicated.

deletion in OX114 is a de novo event associated with the development of ADPKD. Although it is not possible to show that the ADPKD is chromosome 16 linked, the location of the *PBP* gene indicates that this is a de novo *PKD1* mutation.

To identify more *PKD1*-associated mutations, we analyzed single copy regions of the *PBP* gene by RT–PCR using RNA isolated from lymphoblastoid cell lines established from ADPKD patients. cDNA from 48 unrelated patients was amplified with the primer pair *3A3 C* (see Experimental Procedures), and the product of 260 bp was analyzed on an agarose gel. In one patient, OX32, an additional smaller product (125 bp) was identified, consistent with a deletion or splicing mutation. OX32 comes from a large family in which the disease can be traced through three generations. Analysis of RNA from two affected sibs of OX32 and his parents showed that the abnormal transcript segregates with *PKD1* (Figure 5d).

Amplification of normal genomic DNA with the *3A3 C* primers generates a product of 418 bp; sequencing showed that this region contains two small introns (5′, 75 bp; 3′, 83 bp) flanking a 135 bp exon. The product amplified from genomic DNA from OX32 was normal in size, excluding a genomic deletion. However, heteroduplex analysis of that DNA revealed larger heteroduplex bands, consistent with a mutation within that genomic interval. The abnormal RT–PCR product from OX32 was cloned and sequenced: this demonstrated that, although present in genomic DNA, the 135 bp exon was missing from the abnormal transcript. Sequencing of OX32 genomic DNA demonstrated a G→C transition at +1 of the splice donor site following the 135 bp exon. This mutation was confirmed in all available affected family members by digesting amplified genomic DNA with the enzyme BstNI: a site is destroyed by the base substitution. The splicing defect results in an in-frame deletion of 135 bp from the *PBP* transcript (3696–3831 bp of the sequence shown in Figure 7). Together, the three intragenic mutations confirm that the *PBP* gene is the defective gene at the *PKD1* locus.

A Deletion That Disrupts the *TSC2* and the *PKD1* Gene

We previously identified a deletion (WS-53) that disrupts the *TSC2* gene and the *PKD1* gene (European Chromosome 16 Tuberous Sclerosis Consortium, 1993), although its full proximal extent was not determined. Further study has shown that the deletion extends ~100 kb (see Figure 6 for details) and deletes most, if not all, of the *PKD1* gene. This patient has TSC but also has unusually severe polycystic disease of the kidneys (full clinical details will be provided in a later paper). Other patients with a similar phenotype are presently under investigation.

Characterization of the *PKD1* Gene

To characterize the *PKD1* gene further, evolutionary conservation was analyzed by "zoo blotting." Using probes from the single copy 3′ region (*3A3*) and from the duplicated area (*JH4* and *JH8*), the *PKD1* gene was conserved in other mammalian species, including horse, dog, pig, and rodents (data not shown). No evidence of related sequences were seen by hybridization at normal stringency in chicken, frog, or fruit fly. The degree of conservation was similar when probes from the single copy or the duplicated region were employed.

The full genomic extent of the *PKD1* gene is not yet known, although results obtained by hybridization to Northern blots show that it extends from at least as far as *JH13*. Several CpG islands have been localized 5′ of the known extent of the *PKD1* gene (Figure 6), although there is no direct evidence that any of these are associated with this gene.

The cDNA contig extending 5631 bp to the 3′ end of the *PKD1* transcript was sequenced; when possible, more than one cDNA was analyzed, and in all regions both strands were sequenced (Figure 7). We estimate that this accounts for ~40% of the *PKD1* transcript. An open reading frame was detected that runs from the 5′ end of the region sequenced and spans 4842 bp, leaving a 3′ untranslated region of 789 bp that contains the previously de-

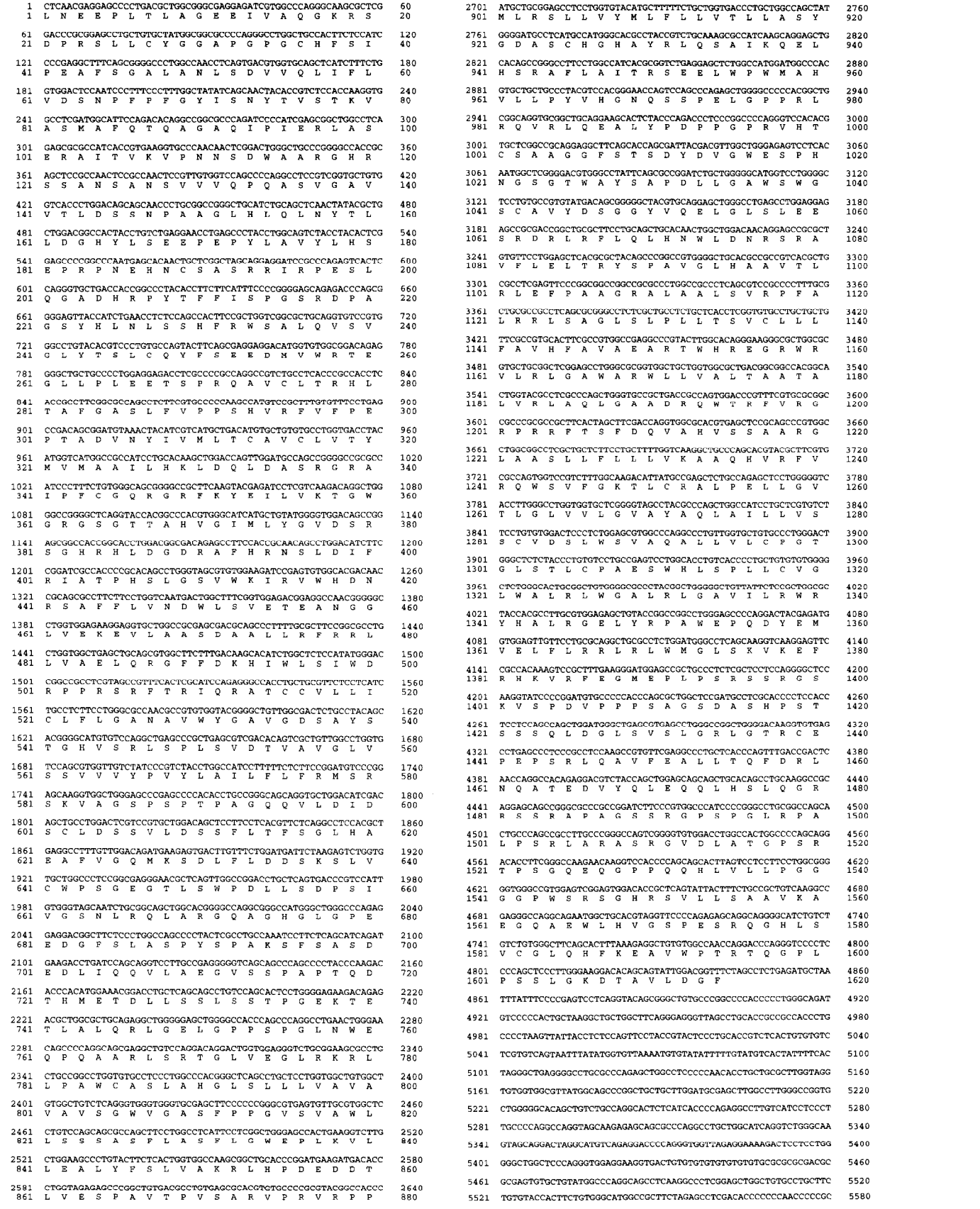

Figure 7. The Partial Nucleotide Sequence of the *PKD1* Transcript Extending 5631 bp to the 3′ End of the Gene

The corresponding predicted protein is shown below the sequence and extends from the start of the nucleotide sequence. The GT repeat *KG8* is in the 3′ untranslated region between 5430 and 5448 bp.

scribed microsatellite *KG8* (Peral et al., 1994; Snarey et al., 1994). A polyadenylation signal is present at nucleotides 5598–5603, and a poly(A) tail was detected in two independent cDNAs (*AH4* and *AH6*) at position 5620. Comparison with the cDNAs *HG-4* and *11BHS21*, which are encoded by genes in the duplicate 16p13.1 region, show that 1866 bp at the 5′ end of the currently known *PKD1* sequence lies within the duplicated area. The predicted amino acid sequence from the available open reading frame extends 1614 residues and is shown in Figure 7. A search of the SwissProt and NBRF data bases with the available protein sequence using the BLAST program (Altschul et al., 1990) identified only short regions of similarity (notably, between amino acids 690–770 and 1390–1530) to a diverse group of proteins; no highly significant areas of homology were recognized. The importance of the short regions of similarity is unclear, as the search for protein motifs with the ProSite program did not identify any recognized functional protein domains within the *PKD1* gene.

Discussion

We have presented compelling evidence that mutations of the *PBP* gene give rise to the typical phenotype of ADPKD. The location of this gene within the *PKD1* candidate region and the available genetic evidence from the families with mutations show that this is the *PKD1* gene. Four *PKD1*-associated mutations have been described: a de novo translocation, which was subsequently transmitted with the phenotype; two intragenic deletions (one a de novo event); and a splicing defect. The task of identifying and characterizing the *PKD1* gene has been more difficult than for other disorders because more than three quarters of the gene is embedded in a region of DNA that is duplicated elsewhere on chromosome 16. This segment of 40–50 kb of DNA, present as a single copy in the *PKD1* area (16p13.3), is reiterated as several divergent copies in the more proximal region, 16p13.1. This proximal site contains three gene loci (*HG-A*, *HG-B*, and *HG-C*) that each produce polyadenylated mRNAs and share substantial homology to the *PKD1* gene; it is not known whether these partially homologous transcripts are translated into functional proteins.

Although gene amplification is known as a major mechanism for creating protein diversity during evolution, the discovery of a human disease locus embedded within an area duplicated relatively recently is a novel observation. In this case, because of the recent nature of the reiteration, the whole duplicated genomic region retains a high level of homology, not just the exons. The sequence of events leading to the duplication and which sequence represents the original gene locus are not yet clear. However, preliminary evidence of homology of the 3′ ends of the three *HG* transcripts, which are different from the 3′ end of the *PKD1* gene, indicate that the loci in 16p13.1 have probably arisen by further reiteration of sequences at this site, after it separated from the distal locus.

The practical implications of this duplication for isolating and characterizing a full-length transcript and for the detection of *PKD1* mutations are considerable. All of the cDNA libraries screened contain sequences of the *PKD1* transcript but also of the three similar *HG* transcripts; careful comparison of these to the genomic sequence from the *PKD1* region will be required to construct a full-length cDNA contig that faithfully represents the entire *PKD1* gene. Screening for mutations within the area duplicated is also difficult, whether Southern blotting or a PCR-based assay for base pair changes (e.g., single strand conformation polymorphism or heteroduplex analysis) is employed. In each case the degree of similarity between the duplicate regions means that loci from both areas are detected simultaneously, even if an RT–PCR approach is adopted. In such circumstances slight sequence variations between the duplicate loci are likely to mask variability due to mutation.

Identification of a disease gene does not necessarily provide immediate insight into the pathophysiology of a disorder. Analysis of the available sequence has not highlighted significant homologies to known proteins, so the role of the PKD1 protein remains unclear. It is possible that characterization of the full protein sequence will shed more light on its function. Even the mechanism by which the detected mutations cause PKD1 is not clear. It has previously been argued that PKD1 could be recessive at the cellular level, with a second somatic mutation required to give rise to cystic epithelium (Reeders, 1992). This "two hit" process is thought to be the mutational mechanism giving rise to several dominant diseases, such as neurofibromatosis (Legius et al., 1993) and TSC (Green et al., 1994), that result from a defect in the control of cellular growth. If this were the case, however, we might expect that a proportion of constitutional *PKD1* mutations would be inactivating deletions as seen in these other disorders.

Four *PKD1* mutations have been characterized, and in each case an abnormal transcript has been identified by Northern analysis or by RT–PCR. Three of these mutations are effectively internal deletions in the 3′ end of the *PKD1* gene, one of which is a small in-frame deletion. The location of these mutations may, however, reflect some ascertainment bias as it is this single copy area that has been screened most intensively for mutations. Nevertheless, no additional deletions were detected when a large part of the gene was screened by FIGE, and studies by PFGE showed no large deletions of this area in 75 *PKD1* patients. It is possible that the mutations detected so far result in the production of an abnormal protein that causes disease through a gain of function. However, it is also possible that these mutations eliminate the production of functional protein from this chromosome and result in the *PKD1* phenotype by haploinsufficiency or only after loss of the second *PKD1* homolog by somatic mutation. Clearly, a larger number of mutations need to be studied and the 5′ end of the gene located and characterized before the full spectrum of *PKD1* mutations can be accessed.

One mutation that probably deletes the entire *PKD1* gene has been identified (WS-53), but in this case it also disrupts the adjacent *TSC2* gene, and the resulting phenotype is of TSC with severe cystic kidney disease. Renal cysts are common in TSC so that the phenotypic significance of deletion of the *PKD1* gene in this case is difficult

to assess. It is clear that not all cases of renal cystic disease in TSC are due to disruption of the *PKD1* gene; chromosome 9–linked *TSC* (*TSC1*) families also manifest cystic kidneys, and we have analyzed many *TSC2* patients with kidney cysts who do not have deletion of the *PKD1* gene. A number of patients with TSC and severe cystic kidney disease are presently being analyzed in detail to see whether they also involve disruption of the *PKD1* gene.

Familial studies indicate that de novo mutations probably account for only a small minority of all ADPKD cases; a recent study detected five possible spontaneous mutations in 209 families (Davies et al., 1991). However, in our study 1 of 3 intragenic mutations detected was a novel mutation, and the *PKD1*-associated translocation was also a de novo event. Furthermore, the mutations detected in the two familial cases do not account for a significant proportion of the local *PKD1*. The OX875 deletion was only detected in 1 of 282 unrelated cases, and the splicing defect was seen in only 1 of 48 unrelated cases. Nevertheless, studies of linkage disequilibrium have found evidence of common haplotypes associated with *PKD1* in a proportion of some populations (Peral et al., 1994; Snarey et al., 1994), suggesting that common mutations will be identified.

Once a larger range of mutations has been characterized, it will be possible to evaluate whether the type and location of mutation determines disease severity and whether there is a correlation between mutation and extrarenal manifestations. Previous studies have provided some evidence that the risk of cerebral aneurysms "runs true" in families (Huston et al., 1993) and that some *PKD1* families exhibit a consistently mild phenotype (Ryynanen et al., 1987). A recent study has concluded that there is evidence of anticipation in ADPKD families, especially if the disease is transmitted through the mother (Fink et al., 1994). Furthermore, analysis of families with early manifestation of ADPKD show that there is a significant intrafamilial recurrence risk and that childhood cases are most often transmitted maternally (Fink et al., 1993; Zerres et al., 1993). This pattern of inheritence is reminiscent of that seen in diseases in which an expanded trinucleotide repeat was found to be the mutational mechanism (reviewed by Mandel, 1993). However, no evidence for an expanding repeat correlating with *PKD1* has been found in this region, although such a sequence cannot be excluded.

There is ample evidence that early presymptomatic diagnosis of *PKD1* is helpful because it allows complications such as hypertension and urinary tract infections to be monitored and treated quickly (Ravine et al., 1991). The identification of mutations within a family will allow rapid screening of that and other families with the same mutation. Given the difficulties of identifying mutations in this region, however, genetic linkage analysis is likely to remain important for presymptomatic diagnosis. The accuracy and ease of linkage-based diagnosis will be improved by the identification of the *PKD1* gene: a microsatellite lies in the 3′ untranslated region of this gene (*KG8*), and several CA repeats are located 5′ of the gene (see Figures 1a and 6; Peral et al., 1994; Snarey et al., 1994).

The identification of the *PKD1* gene is a first step toward understanding the molecular pathology of this complex disorder. Characterization of a full-length transcript and definition of the full range of mutations that cause *PKD1* are now required, progress that is likely to be impeded because this area is duplicated. The information now available will, however, allow study of the PKD1 protein to determine both its normal role in the cell and the mechanism by which mutation of the *PKD1* gene causes the major renal cystic disease in humans.

Experimental Procedures

Clinical Details of Patients

Family 77

Patients 77-2 and 77-3 are 48 and 17 years old, respectively, and have typical ADPKD. Both have bilateral polycystic kidneys, and 77-2 has impaired renal function. Neither patient manifests any signs of TSC (apart from cystic kidneys) on clinical and ophthalmological examination or by a computerized tomography scan of the brain.

Patient 77-4 is 13 years old and severely mentally retarded and has multiple signs of TSC, including adenoma sebaceum, depigmented macules, and periventricular calcification on computerized tomography scan. Renal ultrasound reveals a small number of bilateral renal cysts.

ADPKD Patients

Patient OX875, age 46, developed ESRD from ADPKD. Progressive decline in renal function has been observed over 17 years; ultrasound examinations documented enlarging polycystic kidneys with less extensive hepatic cystic disease. Both kidneys were removed after renal transplantation, and pathological examination showed typical advanced cystic disease in kidneys weighing 1920 g and 3450 g (normal average, 120 g).

Patient OX114, age 54, developed ESRD from ADPKD: diagnosis was made by radiological investigation during an episode of abdominal pain at age 25. A progressive decline in renal function and the development of hypertension were subsequently observed. Ultrasonic examination demonstrated enlarged kidneys with typical cystic disease, with less severe hepatic involvement.

Patient OX32 is a member of a large kindred affected by typical ADPKD in which several members have developed ESRD. The patient himself has been observed for 12 years, with progressive renal failure and hypertension following ultrasonic demonstration of polycystic kidneys.

No signs of TSC were observed on clinical examination of any of the ADPKD patients.

DNA Electrophoresis and Hybridization

DNA extraction, restriction digests, electrophoresis, Southern blotting, hybridization, and washing were performed by standard methods or as previously described (Harris et al., 1990). FIGE was performed with the Bio-Rad FIGE mapper using program 5 to separate fragments from 25–50 kb. High molecular weight DNA for PFGE was isolated in agarose blocks and separated on the Bio-Rad CHEF DRII apparatus using appropriate conditions.

Genomic DNA Probes and Somatic Cell Hybrids

Many of the DNA probes used in this study have been described previously: *MS205.2* (*D16S309*; Royle et al., 1992); *GGG1* (*D16S259*; Germino et al., 1990); *N54* (*D16S139*; Himmelbauer et al., 1991); and *SM6* (*D16S665*), *CW23*, *CW21*, and *JH1* (European Chromosome 16 Tuberous Sclerosis Consortium, 1993). Microsatellite probes for haplotype analysis were *KG8* and *W5.2* (Snarey et al., 1994); *SM6*, *CW3*, and *CW2* (Peral et al., 1994); *16AC2.5* (Thompson et al., 1992); *SM7* (Harris et al., 1991); and *VK5AC* (Aksentijevich et al., 1993).

Probes isolated during this study were *JH4*, *JH5*, and *JH6* (11 kb, 6 kb, and 6 kb BamHI fragments, respectively) and *JH13* and *JH14* (4 kb and 2.8 kb BamHI–EcoRI fragments, respectively), all from the cosmid *JH2A*; *JH8* and *JH10* are 4.5 kb and 2 kb SacI fragments, respectively, and *JH12* is a 0.6 SacI–BamHI fragment, all from *JH4*; *8S1* and *8S3* are 2.4 kb and 0.6 kb SacII fragments, respectively, from

JH8; *CW10* is a 0.5 kb NotI–MluI fragment of *SM25A*; *JH17* is a 2 kb EcoRI fragment of *NM17*.

The somatic cell hybrids *N-OH1* (Germino et al., 1990), *P-MWH2A* (European Chromosome 16 Tuberous Sclerosis Consortium, 1993), and *Hy145.19* (Himmelbauer et al., 1991) have previously been described. Somatic cell hybrids containing the paternally derived (*BP2-10*) and maternally derived (*BP2-9*) chromosomes from OX114 were produced by the method of Deisseroth and Hendrick (1979).

Constructing a Cosmid Contig

Cosmids were isolated from chromosome 16–specific and total genomic libraries, and a contig was constructed using the methods and libraries previously described (European Chromosome 16 Tuberous Sclerosis Consortium, 1993). To ensure that cosmids were derived from the 16p13.3 region (not the duplicate 16p13.1 area), initially probes from the single copy area were used to screen libraries (e.g., *CW21* and *N54*). Two cosmids mapped entirely within the area duplicated, *CW10III* and *JC10.2B*. To establish that these were from the *PKD1* area, they were restriction mapped and hybridized with the probe *CW10*. The fragment sizes detected were compared with results obtained with hybrids containing only the 16p13.3 area (*Hy145.19*) or only the 16p13.1 region (*P-MWH2A*).

FISH

FISH was performed essentially as previously described (Buckle and Rack, 1993). The hybridization mixture contained 100 ng of biotin II–dUTP-labeled cosmid DNA and 2.5 μg of human *Cot-1* DNA (Bethesda Research Laboratories), which was denatured and annealed at 37°C for 15 min prior to hybridization at 42°C overnight. After stringent washes, the site of hybridization was detected with successive layers of fluorescein-conjugated avidin (5 μg/ml) and biotinylated anti-avidin (5 μg/ml) (Vector Laboratories). Slides were mounted in Vectashield (Vector Laboratories) containing 1 μg/ml propidium iodide and 1 μg/ml 4′,6′-diamidino-2-phenylindole (DAPI) to allow concurrent G-banded analysis under UV light. Results were analyzed and images captured using a Bio-Rad MRC 600 confocal laser scanning microscope.

cDNA Screening and Characterization

Fetal brain cDNA libraries in λ phage (Clonetech; Stratagene) were screened by standard methods with genomic fragments in the single copy area (equivalent to *CW23* and *CW21*) or with a 0.8 kb PvuII–EcoRI single copy fragment of *AH3*. Six *PBP* cDNAs were characterized, including two previously described cDNAs, *AH4* (1.7 kb) and *3A3* (2.0 kb) (European Chromosome 16 Tuberous Sclerosis Consortium, 1993), and four novel cDNAs, *AH3* (2.2 kb), *AH6* (2.0 kb), *A1C* (2.2 kb), and *B1E* (2.9 kb). A Striatum library (Stratagene) was screened with *JH4*, and an *HG-C* cDNA, *11BHS21* (3.8 kb), was isolated; *21P.9* is a 0.9 kb PvuII–EcoRI subclone of this cDNA. A *HG-A* or *HG-B* cDNA, *HG-4* (7 kb), was also isolated by screening the fetal brain library (Stratagene) with *JH8*. *HG-4/1.1* is a 1.1 kb PvuII–EcoRI fragment from the 3′ end of *HG-4*. *1A1H.6* is a 0.6 kb HindIII–EcoRI subclone of a *TSC2* cDNA, *1A-1* (1.7 kb), which was isolated from the Clonetech library. Each cDNA was subcloned into Bluescript and sequenced utilizing a combination of sequential truncation and oligonucleotide primers using DyeDeoxy′ Terminators (Applied Biosystems) and an ABI 373A DNA Sequencer (Applied Biosystems) or by hand with Sequenase T7 DNA polymerase (U. S. Biochemicals).

RNA Procedures

Total RNA was isolated from cell lines and tissues by the method of Chomczynski and Sacchi (1987), and enrichment for mRNA was made using the poly(AT) tract mRNA isolation system (Promega). For RNA electrophoresis, 0.5% agarose denaturing formaldehyde gels were used that were Northern blotted, hybridized, and washed using standard procedures. The 0.24–9.5 kb RNA (GIBCO BRL) size standard was used, and hybridization of the probe (*1-9B3*) to the 13 kb utrophin transcript (Love et al., 1989) in total fibroblast RNA was used as a size marker for the large transcripts.

RT–PCR was performed with 2.5 μg of total RNA by the method of Brown et al. (1990) with random hexamer primers, except that AMV reverse transcriptase (Life Sciences) was employed. To characterize the deletion of the *PBP* transcript in OX114, we used the following primers: *AH3F9*, 5′-TTT GAC AAG CAC ATC TGG CTC TC-3′; *AH3B7*, 5′-TAC ACC AGG AGG CTC CGC AG-3′. These were emplyed in a DMSO-containing PCR buffer (Dodé et al., 1990) with 0.5 mM $MgCl_2$ and 36 cycles of 94°C for 1 min, 61°C for 1 min, and 72°C for 2 min plus a final extension of 10 min.

The *3A3 C* primers used to amplify the cDNA from OX32 and DNA were these: *3A3 C1*, 5′-CGC CGC TTC ACT AGC TTC GAC-3′; *3A3 C2*, 5′-ACG CTC CAG AGG GAG TCC AC-3′. These were employed in a PCR buffer and cycle previously described (Harris et al., 1991) with 1 mM $MgCl_2$ and an annealing temperature of 61°C.

PCR products for sequencing were amplified with Pful (Stratagene) and ligated into the SrfI site in PCR-Script (Stratagene) in the presence of SrfI.

RNAase Protection

Tissues from normal and end stage polycystic kidneys were immediately homogenized in guanidinium thiocyanate. RNA was purified on a cesium chloride gradient, and 30 μg of total RNA was assayed by RNAase protection by the method of Melton et al. (1984) using a genomic template generated with the *3A3 C* primers.

Heteroduplex Analysis

Heteroduplex analysis was performed essentially as described by Keen et al. (1991). Samples were amplified from genomic DNA with the *3A3 C* primers, heated at 95°C for 5 min, and incubated at room temperature for at least 30 min before loading on a Hydrolink gel (AT Biochem). Hydrolink gels were run for 12–18 hr at 250 V and fragments observed after staining with ethidium bromide.

Extraction and Amplification of Paraffin-Embedded DNA

DNA from formalin-fixed, paraffin wax–embedded kidney tissue was prepared by the method of Wright and Manos (1990), except that after proteinase K digestion overnight at 55°C, the DNA was extracted with phenol plus chloroform before ethanol precipitation. Approximately 50 ng of DNA was used for PCR with 1.5 mM $MgCl_2$ and 40 cycles of 94°C for 1 min, 59°C for 1 min, and 72°C for 40 s, plus a 10 min extension at 72°C.

The oligonucleotide primers designed to amplify across the genomic deletion of OX875 were as follows: *AH4F2*, 5′-GGG CAA GGG AGG ATG ACA AG-3′; *JH14B3*, 5′-GGG TTT ATC AGC AGC AAG CGG-3′. These produced a product of ~220 bp in individuals with the OX875 deletion.

Acknowledgments

We particularly thank Dr. C. G. Winearls, C. Strong, and K. Clark for their important contributions to the project and Prof. Sir D. J. Weatherall, Prof. G. J. B. van Ommen, Prof. P. Harper, and Prof. H. Galjaard for continued support. R. Snell, D. Shaw, A. O. M. Wilkie, K. Fleming, and D. Ravine are thanked for helpful discussions; D. R. M. Lindsell, A. Cahill, J. Fawcett, and W. Landells for assistance with collection of clinical material; P. Thompson for cytogenetics; and V. Chowdhury and R. Boone for technical advice. D. J. Blake and I. Ceccherini contributed important reagents, and L. Rose is thanked for administrative assistance. We also thank the patients, families, and doctors who have contributed to this project and without whom this study would not have been possible. This work was supported financially by the Medical Research Council, the Wellcome Trust, the Oxford Kidney Unit Trust Fund, the Dutch Kidney Foundation, the Prevention Fund (The Netherlands), the Tuberous Sclerosis Association, Action Research, the University of Wales, and The Netherlands Organization for Scientific Research.

Received May 6, 1994; revised May 20, 1994.

References

Aksentijevich, I., Pras, E., Gruberg, L., Shen, Y., Holman, K., Helling, S., Prosen, L., Sutherland, G. R., Richards, R. I., Ramsburg, M., Dean, M., Pras, M., Amos, C. I., and Kastner, D. L. (1993). Refined mapping of the gene causing familial Mediterranean fever, by linkage and homozygosity studies. Am. J. Hum. Genet. *53*, 451–461.

Altschul, S. F., Warren, G., Miller, W., Myers, E. W., and Lipman, D. J. (1990). Basic alignment search tool. J. Mol. Biol. *215*, 403–410.

Breuning, M. H., Reeders, S. T., Brunner, H., Ijdo, J. W., Saris, J. J., Verwest, A., van Ommen, G. J. B., and Pearson, P. L. (1987). Improved early diagnosis of adult polycystic kidney disease with flanking DNA markers. Lancet *ii*, 1359–1361.

Breuning, M. H., Snijdewint, F. G. M., Brunner, H., Verwest, A., Ijdo, J. W., Saris, J. J., Dauwerse, J. G., Blonden, L., Keith, T., Callen, D. F., Hyland, V. J., Xiao, G. H., Scherer, G., Higgs, D. R., Harris, P., Bachner, L., Reeders, S. T., Germino, G., Pearson, P. L., and van Ommen, G. J. B. (1990). Map of 16 polymorphic loci on the short arm of chromosome 16 close to the polycystic kidney disease gene (*PKD1*). J. Med. Genet. *27*, 603–613.

Brown, C. J., Flenniken, A. M., Williams, B. R. G., and Willard, H. F. (1990). X chromosome inactivation of the human *TIMP* gene. Nucl. Acids Res. *18*, 4191–4195.

Buckle, V. J., and Rack, K. (1993). Fluorescent in situ hybridisation. In Human Genetic Disease Analysis: A Practical Approach, Volume 2, K. E. Davies, ed. (Oxford: IRL Press), pp. 59–82.

Chapman, A. B., Rubinstein, D., Hughes, R., Stears, J. C., Earnest, M. P., Johnson, A. M., Gabow, P. A., and Kaehny, W. D. (1992). Intracranial aneurysms in autosomal dominant polycystic kidney disease. N. Engl. J. Med. *327*, 916–920.

Chomczynski, P., and Sacchi, N. (1987). Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. Anal. Biochem. *162*, 156–159.

Dalgaard, O. Z. (1957). Bilateral polycystic disease of the kidneys: a follow-up of two hundred and eighty-four patients and their families. Acta Med. Scand. (Suppl.) *328*, 1–251.

Davies, F., Coles, G. A., Harper, P. S., Williams, A. J., Evans, C., and Cochlin, D. (1991). Polycystic kidney disease re-evaluated: a population-based study. Quart. J. Med. *79*, 477–485.

Deisseroth, A., and Hendrick, D. (1979). Activation of phenotypic expression of human globin genes from non-erythroid cells by chromosome-dependent transfer to tetraploid mouse erythroleukemia cells. Proc. Natl. Acad. Sci. USA *76*, 2185–2189.

Dodé, C., Rochette, J., and Krishnamoorthy, R. (1990). Locus assignment of human α globin mutations by selective amplification and direct sequencing. Br. J. Haemat. *76*, 275–281.

European Chromosome 16 Tuberous Sclerosis Consortium (1993). Identification and characterization of the tuberous sclerosis gene on chromosome 16. Cell *75*, 1305–1315.

Fink, G. M., Johnson, A. M., Strain, J. D., Kimberling, W. J., Kumar, S., Manco-Johnson, M. L., Duley, I. T., and Gabow, P. A. (1993). Characteristics of very early onset autosomal dominant polycystic kidney disease. J. Am. Soc. Nephrol. *3*, 1863–1870.

Fink, G. M., Johnson, A. M., and Gabow, P. A. (1994). Is there evidence for anticipation in autosomal-dominant polycystic disease? Kidney Int. *45*, 1153–1162.

Gabow, P. A. (1990). Autosomal dominant polycystic kidney disease: more than a renal disease. Am. J. Kidney Dis. *16*, 403–413.

Gabow, P. A. (1991). Polycystic kidney disease: clues to pathogenesis. Kidney Int. *40*, 989–996.

Gabow, P. A. (1993). Autosomal dominant polycystic kidney disease. N. Engl. J. Med. *329*, 332–342.

Gabow, P. A., Johnson, A. M., Kaehny, W. D., Kimberling, W. J., Lezotte, D. C., Duley, I. T., and Jones, R. H. (1992). Factors affecting the progression of renal disease in autosomal-dominant polycystic kidney disease. Kidney Int. *41*, 1311–1319.

Germino, G. G., Barton, N. J., Lamb, J., Higgs, D. R., Harris, P., Xiao, G. H., Scherer, G., Nakamura, Y., and Reeders, S. T. (1990). Identification of a locus which shows no genetic recombination with the autosomal dominant polycystic kidney disease gene on chromosome 16. Am. J. Hum. Genet. *46*, 925–933.

Germino, G. G., Weinstat-Saslow, D., Himmelbauer, H., Gillespie, G. A. J., Somlo, S., Wirth, B., Barton, N., Harris, K. L., Frischauf, A.-M., and Reeders, S. T. (1992). The gene for autosomal dominant polycystic kidney disease lies in a 750-kb CpG-rich region. Genomics *13*, 144–151.

Green, A. J., Smith, M., and Yates, J. R. W. (1994). Loss of heterozygosity on chromosome 16p13.3 in hamartomas from tuberous sclerosis patients. Nature Genet. *6*, 193–196.

Harris, P. C., Barton, N. J., Higgs, D. R., Reeders, S. T., and Wilkie, A. O. M. (1990). A long-range restriction map between the α-globin complex and a marker closely linked to the polycystic kidney disease 1 (*PKD1*) locus. Genomics *7*, 195–206.

Harris, P. C., Thomas, S., Ratcliffe, P. J., Breuning, M. H., Coto, E., and Lopez-Larrea, C. (1991). Rapid genetic analysis of families with polycystic kidney disease 1 by means of a microsatellite marker. Lancet *338*, 1484–1487.

Himmelbauer, H., Germino, G. G., Ceccherini, I., Romeo, G., Reeders, S. T., and Frischauf, A.-M. (1991). Saturating the region of the polycystic kidney disease gene with *Not*I linking clones. Am. J. Hum. Genet. *48*, 325–334.

Hossack, K. F., Leddy, C. L., Johnson, A. M., Schrier, R. W., and Gabow, P. A. (1988). Echocardiographic findings in autosomal dominant polycystic kidney disease. N. Engl. J. Med. *319*, 907–912.

Huston, J., Torres, V. E., Sulivan, P. P., Offord, K. P., and Wiebers, D. O. (1993). Value of magnetic resonance angiography for detection of intracranial aneurysm in autosomal dominant polycystic kidney disease. J. Am. Soc. Nephrol. *3*, 1871–1877.

Hyland, V. J., Suthers, G. K., Friend, K., MacKinnon, R. N., Callen, D. F., Breuning, M. H., Keith, T., Brown, V. A., Phipps, P., and Sutherland, G. R. (1990). Probe, *VK5B*, is located in the same interval as the autosomal dominant adult polycystic kidney disease locus, *PKD1*. Hum. Genet. *84*, 286–288.

Keen, J., Lester, D., Inglehearn, C., Curtis, A., and Bhattacharya, S. (1991). Rapid detection of single base mismatches as heteroduplexes on Hydrolink gels. Trends Genet. *7*, 5.

Kimberling, W. J., Fain, P. R., Kenyon, J. B., Goldgar, D., Sujansky, E., and Gabow, P. A. (1988). Linkage heterogeneity of autosomal dominant polycystic kidney disease. N. Engl. J. Med. *319*, 913–918.

Kimberling, W. J., Kumar, S., Gabow, P. A., Kenyon, J. B., Connolly, C. J., and Somlo, S. (1993). Autosomal dominant polycystic kidney disease: localization of the second gene to chromosome 4q13-q23. Genomics *18*, 467–472.

Legius, E., Marchuk, D. A., Collins, F. S., and Glover, T. W. (1993). Somatic deletion of the neurofibromatosis type 1 gene in a neurofibrosarcoma supports a tumour suppressor gene hypothesis. Nature Genet. *3*, 122–126.

Love, D. R., Hill, D. F., Dickson, G., Spurr, N. K., Byth, B. C., Marsden, R. F., Walsh, F. S., Edwards, Y. H., and Davies, K. E. (1989). An autosomal transcript in skeletal muscle with homology to dystrophin. Nature *339*, 55–58.

Mandel, J.-L. (1993). Questions of expansion. Nature Genet. *4*, 8–9.

Melton, D. A., Kreig, P. A., Rebagliati, M. R., Maniatis, T., Zinn, K., and Green, M. R. (1984). Efficient *in vitro* synthesis of biological active RNA and RNA hybridization probes from plasmids containing a bacteriophage SP6 promoter. Nucl. Acid Res. *12*, 7035–7056.

Milutinovic, J., Fialkow, P. J., Rudd, T. G., Agodoa, L. Y., Phillips, L. A., and Bryant, J. I. (1980). Liver cysts in patients with autosomal dominant polycystic kidney disease. Am. J. Med. *68*, 741–744.

Parfrey, P. S., Bear, J. C., Morgan, J., Cramer, B. C., McManamon, P. J., Gault, M. H., Churchill, D. N., Singh, M., Hewitt, R., Somlo, S., and Reeders, S. T. (1990). The diagnosis and prognosis of autosomal dominant polycystic kidney disease. N. Engl. J. Med. *323*, 1085–1090.

Peral, B., Ward, C. J., San Millán, J. L., Thomas, S., Stallings, R. L., Moreno, F., and Harris, P. C. (1994). Evidence of linkage disequilibrium in the Spanish polycystic kidney disease 1 population. Am. J. Hum. Genet. *54*, 899–908.

Peters, D. J. M., and Sandkuijl, L. A. (1992). Genetic heterogeneity of polycystic kidney disease in Europe. Contributions Nephrol. *97*, 128–139.

Peters, D. J. M., Spruit, L., Saris, J. J., Ravine, D., Sandkuijl, L. A., Fossdal, R., Boersma, J., van Eijk, R., Nørby, S., Constantinou-Deltas, C. D., Pierides, A., Brissenden, J. E., Frants, R. R., van Ommen, G.-J. B., and Breuning, M. H. (1993). Chromosome 4 localization of a second gene for autosomal dominant polycystic kidney disease. Nature Genet. *5*, 359–362.

Pound, S. E., Carothers, A. D., Pignatelli, P. M., Macnicol, A. M.,

Watson, M. L., and Wright, A. F. (1992). Evidence for linkage disequilibrium between *D16S94* and the adult onset polycystic kidney disease (*PKD1*) gene. J. Med. Genet. *29*, 247–248.

Ravine, D., Walker, R. G., Gibson, R. N., Sheffield, L. J., Kincaid-Smith, P., and Danks, D. M. (1991). Treatable complications in undiagnosed cases of autosomal dominant polycystic kidney disease. Lancet *337*, 127–129.

Ravine, D., Walker, R. G., Gibson, R. N., Forrest, S. M., Richards, R. I., Friend, K., Sheffield, L. J., Kincaid-Smith, P., and Danks, D. M. (1992). Phenotype and genotype heterogeneity in autosomal dominant polycystic kidney disease. Lancet *340*, 1330–1333.

Reeders, S. T. (1992). Multilocus polycystic disease. Nature Genet. *1*, 235–237.

Reeders, S. T., Breuning, M. H., Davies, K. E., Nicholls, R. D., Jarman, A. P., Higgs, D. R., Pearson, P. L., and Weatherall, D. J. (1985). A highly polymorphic DNA marker linked to adult polycystic kidney disease on chromosome 16. Nature *317*, 542–544.

Reeders, S. T., Keith, T., Green, P., Germino, G. G., Barton, N. J., Lehmann, O. J., Brown, V. A., Phipps, P., Morgan, J., Bear, J. C., and Parfrey, P. (1988). Regional localization of the autosomal dominant polycystic kidney disease locus. Genomics *3*, 150–155.

Romeo, G., Costa, G., Catizone, L., Germino, G. G., Weatherall, D. J., Devoto, M., Roncuzzi, L., Zucchelli, P., Keith, T., and Reeders, S. T. (1988). A second genetic locus for autosomal dominant polycystic kidney disease. Lancet *ii*, 8–10.

Royle, N. J., Armour, J. A. L., Webb, M., Thomas, A., and Jeffreys, A. J. (1992). A hypervariable locus *D16S309* located at the distal end of 16p. Nucl. Acids Res. *20*, 1164.

Ryynanen, M., Dolata, M. M., Lampainen, E., and Reeders, S. T. (1987). Localisation of a mutation producing autosomal dominant polycystic kidney disease without renal failure. J. Med. Genet. *24*, 462–465.

Scheff, R. T., Zuckerman, G., Harter, H. Delmez, J., and Koehler, R. (1980). Diverticular disease in patients with chronic renal failure due to polycystic kidney disease. Ann. Int. Med. *92*, 202–204.

Snarey, A., Thomas, S., Schneider, M. C., Pound, S. E., Barton, N., Wright, A. F., Harris, P. C., Reeders, S. T., and Frischauf, A.-M. (1994). Linkage disequilibrium in the region of the autosomal dominant polycystic kidney disease gene (*PKD1*). Am. J. Hum. Genet., in press.

Somlo, S., Wirth, B., Germino, G. G., Weinstat-Saslow, D., Gillespie, G. A. J., Himmelbauer, H., Steevens, L., Coucke, P., Willems, P., Bachner, L., Coto, E., Lopez-Larrea, C., Peral, B., San Millán, J. L., Saris, J. J., Breuning, M. H., Frischauf, A.-M., and Reeders, S.T. (1992). Fine genetic localization of the gene for autosomal dominant polycystic kidney disease (*PKD1*) with respect to physically mapped markers. Genomics *13*, 152–158.

Thompson, A. D., Shen, Y., Holman, K., Sutherland, G. R., Callen, D. F., and Richards, R. I. (1992). Isolation and characterization of $(AC)_n$ microsatellite genetic markers from human chromosome 16. Genomics *13*, 402–408.

Wilson, P. D., and Sherwood, A. C. (1991). Tubulocystic epithelium. Kidney Int. *39*, 450–463.

Wright, D. K., and Manos, M. M. (1990). Sample preparation from paraffin-embedded tissues. In PCR Protocols: A Guide to Methods and Applications, M. A. Innis, D. H. Gelfland, J. J. Sniksky, and T. J. White, eds. (San Diego, California: Academic Press), pp. 153–166.

Zerres, K., Rudnik-Schöneborn, S., Deget, F., and German Working Group on Paediatric Nephrology (1993). Childhood onset autosomal dominant polycystic kidney disease in sibs: clinical picture and recurrence risk. J. Med. Genet. *30*, 583–588.

GenBank Accession Number

The accession number for the sequence reported in this paper is L 33243.