Cell, Vol. 87, 979–987, December 13, 1996, Copyright ©1996 by Cell Press

The Molecular Basis of Focal Cyst Formation in Human Autosomal Dominant Polycystic Kidney Disease Type I

Feng Qian,* Terry J. Watnick,* Luiz F. Onuchic, and Gregory G. Germino
Department of Medicine, Division of Nephrology
The Johns Hopkins University School of Medicine
720 Rutland Street
Baltimore, Maryland 21205

Summary

Autosomal dominant polycystic kidney disease (ADPKD) is a common disease and an important cause of renal failure. It is characterized by considerable intrafamilial phenotypic variation and focal cyst formation. To elucidate the molecular basis for these observations, we have developed a novel method for isolating renal cystic epithelia from single cysts and have used it to show that individual renal cysts in ADPKD are monoclonal. Loss of heterozygosity was discovered within a subset of cysts for two closely linked polymorphic markers located within the *PKD1* gene. Genetic analysis revealed that it was the normal haplotype that was lost. This study provides a molecular explanation for the focal nature of cyst formation and a probable mechanism whereby mutations cause disease. The high rate at which "second hits" must occur to account for the large number of cysts observed suggests that unique structural features of the *PKD1* gene may be responsible for its mutability.

Introduction

Autosomal dominant polycystic kidney disease (ADPKD) is one of the most common inherited diseases of humans; it is estimated to affect 1 in 1000 of the population (Gabow, 1993). Renal cysts are the major clinical features of the disease and appear to increase in size and number throughout the lifetime of an individual. This process results in renal failure in approximately half of affected individuals by age 50 years. Although the renal lesion is the most prominent feature, ADPKD is a systemic disorder with a variety of other manifestations including liver cysts, cerebral aneurysms, and a variety of cardiac valvular abnormalities.

ADPKD exhibits considerable variability with respect to both its renal and extrarenal manifestations. Genetic heterogeneity is likely to account for at least some of these differences. Linkage studies have determined that there are at least three forms of ADPKD. PKD1, which is most common and accounts for 85%–95% of all cases, maps to chromosome 16p13.3 (Reeders et al., 1985). The second type, PKD2, which affects most of the remaining families, maps to chromosome 4q13–23 (Kimberling et al., 1993; Peters et al., 1993). A small number of families have another form, which has not yet been mapped (Bogdanova et al., 1995; Daoust et al., 1995; de Almeida et al., 1995). Although all types of ADPKD present with an identical profile of extrarenal manifestations (including liver cysts and aneurysms), PKD1 is the most severe, with a lower median survival and a higher risk of progressing to end-stage renal disease (Ravine et al., 1992).

While unique mutations in disease-causing genes may account for interfamilial variation, they cannot explain why related individuals who presumably share a common mutation have different clinical presentations. Severe childhood cases of ADPKD born into families that exhibit the classic adult presentation of the disease are dramatic examples of this phenomenon (Blyth and Ockenden, 1971; Kaariainen, 1987; Fick et al., 1993; Zerres et al., 1993). Fick et al. (1994) argued that these families are examples of anticipation and proposed that an unstable trinucleotide repeat might be responsible for these findings.

Since these initial clinical observations were made, the genes for PKD1 and PKD2 have been identified. *PKD1* encodes a 14 kb mRNA that is derived from 46 exons that extend over ~50 kb of genomic DNA (The European Polycystic Kidney Disease Consortium, 1994; The American PKD1 Consortium, 1995; Hughes et al., 1995; The International Polycystic Kidney Disease Consortium, 1995). An unusual feature of the gene is that ~70% of its length is replicated in at least three copies clustered together on 16p13.1. The other loci, which also are transcribed, have a sequence nearly identical to that of *PKD1* in their shared segments. The *PKD1* gene product, polycystin, is 4302 amino acids in length and is likely to be an integral membrane glycoprotein that regulates cell–cell or cell–matrix interactions. *PKD2* encodes a 5.4 kb mRNA within ~68 kb of genomic DNA (Mochizuki et al., 1996). The predicted protein of 968 amino acids has significant homology to the voltage-activated Ca^{2+} channel α_{1E} as well as to a portion of polycystin. It is thought to encode an integral membrane protein that may function as an ion channel or pore whose activity is regulated by polycystin. Neither gene has any trinucleotide repeats within its mRNA that are >5 units in length, nor are any found within the complete gene sequence of *PKD1*.

Mutation analysis of *PKD1* has been greatly hindered by high sequence similarity to its replicated loci. Nonetheless, the mutations reported to date for both *PKD1* and *PKD2* have been stable nucleotide substitutions, deletions, or insertions and do not explain why family members with the same germline mutation exhibit dramatic phenotypic variability (Peral et al., 1995, 1996a, 1996b; Mochizuki et al., 1996). In fact, Peral et al. (1996b) recently reported a set of fraternal twins with the same germline nonsense mutation in *PKD1* yet with remarkably different phenotypes. One child had a severe infantile-onset form of the disease, whereas the other had no cysts at the age of 5 years. The authors suggested that the difference in clinical presentation was most likely due to the effect of a small number of genetic modifying factors.

Although breeding experiments in mice with recessive

*These authors contributed equally to this work and are listed in alphabetical order.

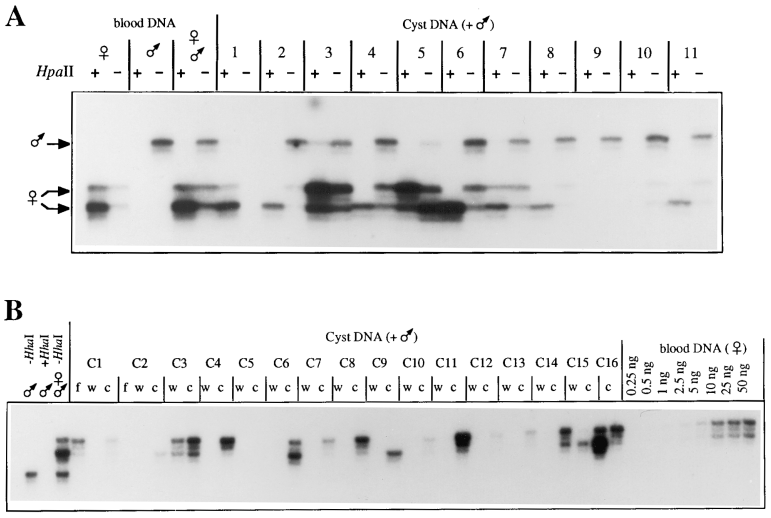

Figure 1. Clonality Assay for Single Cysts Derived from Kidney Donors JHU93 and JHU273

(A) Cystic epithelia were isolated from 11 intact cysts (1–11) from donor JHU93. Male genomic DNA (♂) was added to an aliquot of each cyst-derived DNA sample as well as DNA isolated from the donor's nucleated blood cells. One half of each mix was digested with HpaII and then used as template for PCR using the androgen receptor primers AR1 and AR2. The amplification products are presented in lanes labeled +. The other undigested half of each mix was used as template for a control amplification done in parallel using identical reaction conditions. The results are presented in lanes labeled −. The symbols to the left of the figure identify the position of the AR alleles of the donor (♀) and male control. Three bands are present in the − lanes of all samples except for that which had only male DNA as template (♂) and cyst 9, which yielded no products. All samples derived from cysts yielded a single band after digestion if the reaction was complete (cysts 2, 4, 5, 6, 8, and 11). In contrast, DNA from blood of the donor yielded two bands despite complete digestion with HpaII prior to PCR. The clonal status of samples 1, 3, and 7 were deemed uninterpretable because of incomplete digestion.

(B) Sixteen intact cysts (C1–16) from donor JHU273 were analyzed using a protocol similar to that described above. Each mix of male- and cyst-derived DNA was digested with HhaI rather than HpaII prior to amplification using AR1 and AR2. Both the final rinse (w) and fluid containing the cyst-lining epithelial cells (c) were evaluated for each sample. For C1 and C2, the original cyst fluid (f) was similarly examined. A dilution series of DNA prepared from nucleated blood cells of JHU273 was analyzed using an identical procedure (right). These data suggest that a minimum of 2.5 ng of cyst DNA was used as template for each amplification. Assuming that the average cell contains ~10 pg of DNA and one tenth of the total yield of cyst DNA was used for each reaction, each sample was estimated to contain between 2500 and 500,000 cells.

cystic disease support a role for the effect of other loci in disease expression, genetic background probably does not explain all aspects of phenotypic variability in human ADPKD (Iakoubova et al., 1995). Even within a single kidney, fewer than 1% of nephrons actually develop cysts, suggesting that cystogenesis is a focal process. Histopathologic studies of ADPKD kidneys have confirmed that cyst formation begins with the localized outgrowth of a tubule in any nephron segment, including segments with different embryologic origins (Evan and McAteer, 1992; Schäfer et al., 1994). These observations suggest that cyst formation is a two-step process and that an inherited mutation at one of the ADPKD gene loci is necessary but insufficient for cystogenesis.

We have hypothesized that the focal nature of cyst formation in ADPKD probably holds clues to an understanding of the pathogenesis of this disorder. To study this process, we have developed a novel method for isolating epithelial cells from single renal cysts that minimizes contamination by other cells and have used it to show that renal cysts in ADPKD are monoclonal. We have demonstrated loss of heterozygosity (LOH) within individual cysts for two closely linked polymorphic markers located within the *PKD1* gene. Genetic analysis has confirmed that it is the normal haplotype that is lost in somatic tissues. This study provides a molecular explanation for the focal nature of cyst formation and a probable mechanism whereby mutations cause disease.

Results

Renal Cysts in ADPKD Are Monoclonal

X-chromosome inactivation has been commonly used to demonstrate the clonal nature of tumors (Fearon et al., 1987). Recently, a polymerase chain reaction (PCR)–based clonality assay has been developed that can be used for small quantities of template (Gilliland et al., 1991). This method exploits methylation-sensitive restriction sites (HpaII and HhaI) close to the highly polymorphic $(CAG)_n$ repeat in the X-linked androgen receptor gene. The methylation status of these sites correlates with X-inactivation (Allen et al., 1992). Willman et al. (1994) used this approach to show the monoclonal nature of histiocystosis X. We have modified this technique to determine the clonality of single cysts in ADPKD kidneys.

Initial studies using DNA prepared from large specimens (>5 g) containing multiple cysts failed to detect monoclonality (data not shown). We then attempted to isolate the epithelial cells lining a single cyst using microdissection. Again, most samples were polyclonal, but histologic analysis revealed that the microdissected specimens were in fact multilayered structures with multiple nonepithelial cellular elements (data not shown). To circumvent these difficulties we adapted a method previously used to study toad bladder epithelia (Handler et al., 1979) and applied it to the isolation of cyst-lining cells. In this method, the cyst remains intact during the entire procedure. EDTA injected into the cyst lumen is used to dissociate the epithelial cells from the underlying basement membrane, which serves as a natural barrier to contaminating cells.

Results of the clonality assay of cysts isolated from two individuals with ADPKD are presented in Figure 1 as representative examples. Of the 11 samples that were prepared from individual cysts of donor JHU93 (Figure 1A), two failed to yield sufficient product for analysis (samples 9 and 10). The results of three other samples

were deemed uninterpretable because of incomplete digestion (samples 1, 3, and 7). The remainder (samples 2, 4, 5, 6, 8, and 11) yielded a single product after digestion. Samples prepared from 16 individual cysts of donor JHU273 were similarly analyzed (Figure 1B). The clonal status of three samples could not be determined because their final rinse solutions also yielded PCR products (C3 and C15) or because no PCR product was detected (C5). Of the remaining 13 cysts, 11 (85%) yielded a single PCR product after HhaI digestion. Only 2 of the 13 cysts (C6 and C14) yielded two products, and in each instance the bands were of unequal intensity, suggesting a clonal bias.

A total of 76 renal cysts derived from 8 affected females were analyzed for clonality, and 62 cysts (82%) were found to be monoclonal (Table 1). It is very likely that the actual percentage of monoclonal cysts is >82%. In some samples, rinse solutions were not tested for luminal contamination (Table 1). Later analyses revealed that the rinse solution of approximately one half of the apparently polyclonal cysts yielded PCR products, suggesting probable contamination (Figure 1). The small number of cysts that are truly polyclonal may have resulted from the fusion of two or more formerly neighboring monoclonal cysts.

Allelic Loss of a *PKD1* Intragenic Microsatellite, *KG8*

It is possible that cysts may develop from a cluster of cells sharing a common X-chromosome inactivation status rather than from clonal expansion of a single progenitor cell. To prove that renal cysts are truly monoclonal and to investigate the molecular basis of clonality, we tested for LOH for *PKD1* in individual cysts using the microsatellite *KG8* (Snarey et al., 1994). This marker lies within the 3′ untranslated region (3′UTR) of the *PKD1* mRNA (Figure 2). Two alleles were equally amplified in most samples. In a subset, however, only a single band was amplified. We included primers for the androgen receptor as an internal control. Two bands corresponding to the androgen receptor alleles were equally amplified in all samples, including those that had a single band for *KG8*. Figure 3A is a representative example of data using these markers. Two cysts from this kidney (C8 and C14) were found to have LOH. A total of 46 cysts from four donors were evaluated with *KG8*, and 17% were found to be hemizygous for this marker (Table 1). At least two cyst preparations per donor had loss of a *KG8* allele, and in each case it was the same allele.

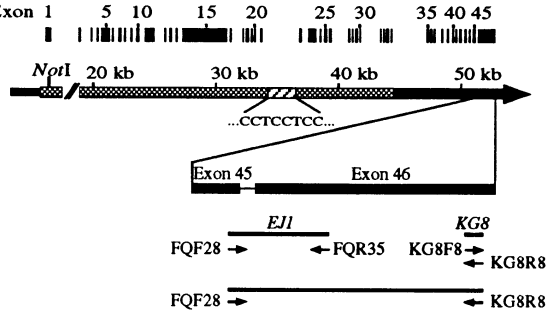

Figure 2. Position of Intragenic *PKD1* Polymorphisms Used to Assay LOH

The *PKD1* gene contains 46 exons and is bisected by a polypyrimidine tract of ~2.5 kb (hatched box). The replicated portion of the gene begins with the first exon and ends in intron 34 (stippled bar). The microsatellite *KG8* lies within the 3′UTR of the gene. *EJ1* is a polymorphic locus located in exons 45–46. The primers used to generate these markers and their relative positions are indicated. The PCR product that contains *KG8* and *EJ1* is approximately 1.6 kb in length and is generated using FQF28 and KG8R8.

Loss of a Second Closely Linked Marker, *EJ1*

We sought to confirm the *KG8* results using a second marker that was either intragenic or immediately proximal to *PKD1*, but the closest known highly polymorphic marker was at least 70 kb away. In the course of performing mutation analyses, we discovered a polymorphic locus, *EJ1*, in exons 45–46 that is located approximately 1.6 kb proximal to *KG8* (Figure 2). It has at least three allelic variants (A1–A3) that can be detected using heteroduplex analysis. We used this marker to test for

Table 1. Summary of Clonality Assay and LOH for Single Renal Cysts from ADPKD Patients

Patients	Number of Cysts Evaluated[a]	Monoclonal Cysts		Cysts with LOH	*KG8* Allele Lost	*EJ1* Allele Lost
		Number	%			
JHU93	6	6	100	ND	ND	ND
JHU188	9	7	78	ND	NI	NI
JHU244	12	10	83	ND	NI	NI
JHU245	10	5	50	C3	110 nt	A1
				C4	110 nt	A1
JHU246	13	12	92	C4	106 nt	A2
				C14	106 nt	A2
JHU288[b]	10	8	80	C8	104 nt	NI
				C14	104 nt	NI
JHU273[b]	13	11	85	C8	106 nt	A2
				C12	No loss	A2→A4
				C14	106 nt	A2
JHU304	3	3	100	ND	ND	NI
Total	76	62	82			

[a] Number of cysts from which androgen receptor-specific PCR products could be amplified after HhaI or HpaII digestion.
[b] Samples for which the final rinses were tested for contamination.
ND, not determined; NI, not informative, nt, length in nucleotides.

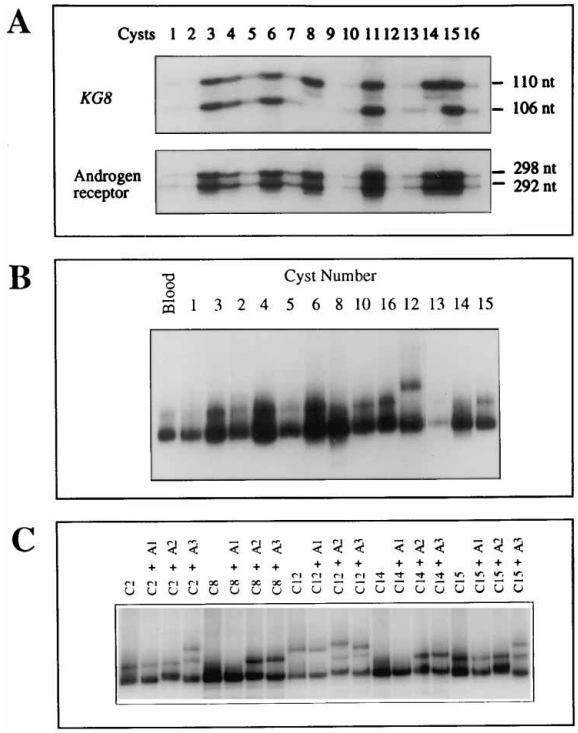

Figure 3. Loss of Heterozygosity for *PKD1* Markers in Cysts from Kidney Donor JHU273

(A) DNA isolated from cyst samples was used as template for PCR using *KG8* and the androgen receptor primers AR1 and AR2 (without prior HhaI digestion). Sample numbers in this figure correspond to those used in Figure 1. The numbers on the right indicate the lengths of the *KG8* and androgen receptor alleles. In two cysts (C8 and C14), only one *KG8* allele, 110 bp long, was amplified (top), whereas two bands were equally amplified by AR1 and AR2 (bottom). nt, length in nucleotides.

(B) Heteroduplex analysis of *EJ1* amplified from the blood of JHU273 (lane 1) shows that the patient is informative at this locus. Sample numbers in this figure correspond to those used in Figure 1. The majority of renal cysts assayed have the same pattern as lane 1. C8 and C14, however, demonstrate LOH at this locus since heteroduplex formation is no longer present. C12 shows a novel heteroduplex pattern.

(C) Mixing studies were performed to determine which *EJ1* allele remained in cyst samples with apparent LOH (C8 and C14) or a novel heteroduplex pattern (C12). Samples with patterns representative of blood (C2 and C15) were included as controls. Unlabeled *EJ1* alleles A1, A2, and A3 were mixed separately with radiolabeled *EJ1* that had been amplified from each cyst and then analyzed for heteroduplex formation. The addition of A1 did not alter the preexisting heteroduplex patterns of any of the cysts, whereas the addition of A3 resulted in a novel pattern in cysts without LOH. Although mixing with A2 did not change the patterns for C2 and C15, it restored the control heteroduplex pattern in C8 and C14 and yielded a unique band in C12. These results suggest that A1 is present in all cysts, whereas A2 has been lost from C8 and C14 and replaced by a novel allele in C12. A3 is missing from all cysts because this allele is not present in JHU273.

LOH in cysts from donors known to be polymorphic at this locus. There was complete concordance between the studies using *KG8* and *EJ1*, with one exception (discussed below) (Figure 3B). All samples that lacked a *KG8* allele also lacked one for *EJ1*. Mixing studies (see Experimental Procedures) were used to demonstrate

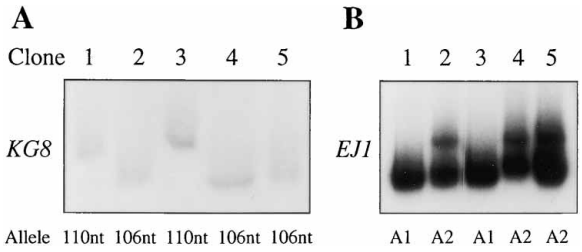

Figure 4. Haplotype Analysis of Donor JHU273

A 1.6 kb PCR fragment containing *KG8* and *EJ1* was amplifed from blood and then cloned (Figure 2). Individual clones were assayed for their respective *KG8* (A) and *EJ1* alleles (B). Clones were assigned an *EJ1* allele based on mixing studies with unlabeled A1 (B) and A2 (data not shown). All clones contained either *KG8* allele 110 bp and *EJ1* allele A1 (clones 1 and 3) or *KG8* allele 106 bp and *EJ1* A2 allele (clones 2, 4, and 5). nt, length in nucleotides.

that cyst samples isolated from the same donor had loss of an identical allele (Figure 3C). A total of 30 cysts from three donors were evaluated with this marker, and 7 were found to have either a deletion or novel somatic mutation involving this locus (Table 1).

Loss of the Unaffected Haplotype

The concordance between the *KG8* and *EJ1* data strongly suggested that the lost alleles belong to the same haplotype. We confirmed this hypothesis using a method that exploited the physical proximity of the loci. A genomic fragment containing *EJ1* and *KG8* was amplified from the peripheral blood of each donor showing LOH for both markers using primers FQF28 and KG8R8 (Figure 2). The products were cloned and the haplotype of each was determined. A representative example of this analysis for JHU273 is shown in Figure 4. For each donor, analysis of cloned genomic segments confirmed that the lost *KG8* and *EJ1* alleles indeed belonged to the same haplotye (Table 1). No haplotype common to all donors was found to be lost in the samples.

We next sought to determine whether it was the normal or affected haplotype that was lost in cystic tissue. Linkage analysis using *KG8* and *3′HVR* was performed on the only two families available for testing. *3′HVR* was linked to the disease in these families, with maximum lod scores of 1.34 (pedigree 1, Figure 5A) and 2.18 (pedigree 2, Figure 5B) at a recombination fraction of 0.0. A positive lod score (1.93) was also obtained with *KG8* in pedigree 2. While linkage of ADPKD to *KG8* was only weakly positive in pedigree 1 (maximum lod score of 0.32), two-point analysis of *3′HVR* and *KG8* favored the haplotype shown in Figure 5A (maximum lod score of 0.99 at a recombination fraction of 0.0). These results strongly sugest that it was the chromosome 16 haplotype not associated with disease that was lost in cystic tissue (Figure 5).

A Novel Somatic Mutation

One cyst sample was discovered to have discordant results with *KG8* and *EJ1*. This sample (C12) did not have loss of a *KG8* allele but did have a unique heteroduplex pattern with *EJ1* (Figure 3). Mixing studies determined that the A2 allele of *EJ1*, which was lost in the other two

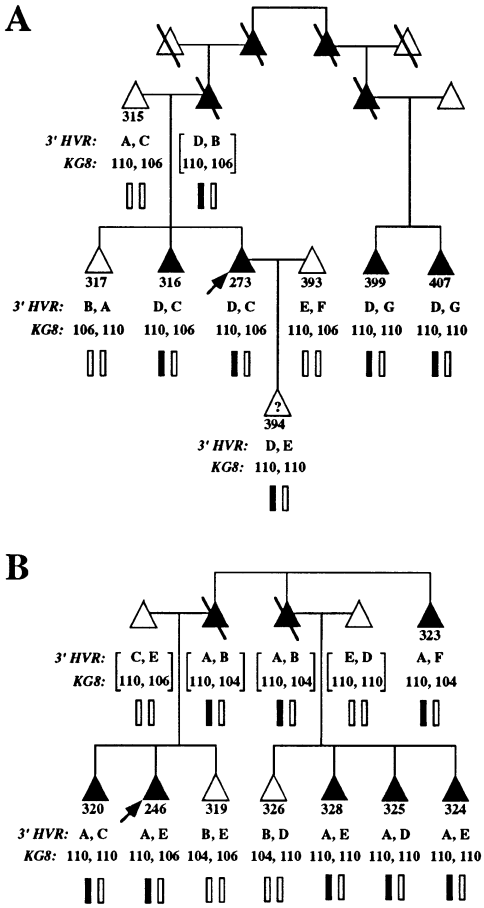

Figure 5. Linkage Analysis of Two Polymorphic Markers Linked to PKD1 in the Families of Donors JHU273 (A) and JHU246 (B)

PKD1 is linked to 3′HVR and KG8 with maximum lod scores of 1.34 and 0.32, respectively, at a recombination fraction of 0.0 for both markers in pedigree 1 (A). The posterior probability of linkage of the disease in this family to PKD1 is >95%. Two-point analysis of 3′HVR and KG8 favors the haplotype shown in (A). In pedigree 2 (B), the maximum lod scores for 3′HVR and KG8 are 2.18 and 1.93, respectively, at a recombination fraction of 0.0. The genotypes D,110 and A,110 segregate with ADPKD in pedigree 1 (A) and pedigree 2 (B), respectively. In each kidney donor (arrows), it is the *KG8* allele that does not segregate with disease which is lost in cysts demonstrating LOH. Affected genotypes are indicated by closed bars; deduced genotypes are shown in brackets. The pedigrees as shown do not include all family members. Triangles are used to protect the identity of participants.

cysts (C8 and C14) from this donor, had been replaced by a novel allele (A4) in C12. We cloned both the A1 and A4 alleles of the *EJ1* locus from this sample and determined the sequence of the novel allele (Figure 6A). Comparison of the sequence of A4 to that of A1 revealed two differences. The first is a 2 bp deletion (ΔC12) at positions 12,694–12,695 of the cDNA sequence (HUMPKD1A, GenBank accession number L33243). The ΔC12 deletion is predicted to cause a reading frameshift resulting in a truncated protein. This mutation also creates a new MwoI endonuclease restriction site not present in the sequence of either the A1 or A2 alleles of *EJ1* (Figure 6B). This site was used to confirm the presence of the deletion in the original C12 DNA sample (Figure

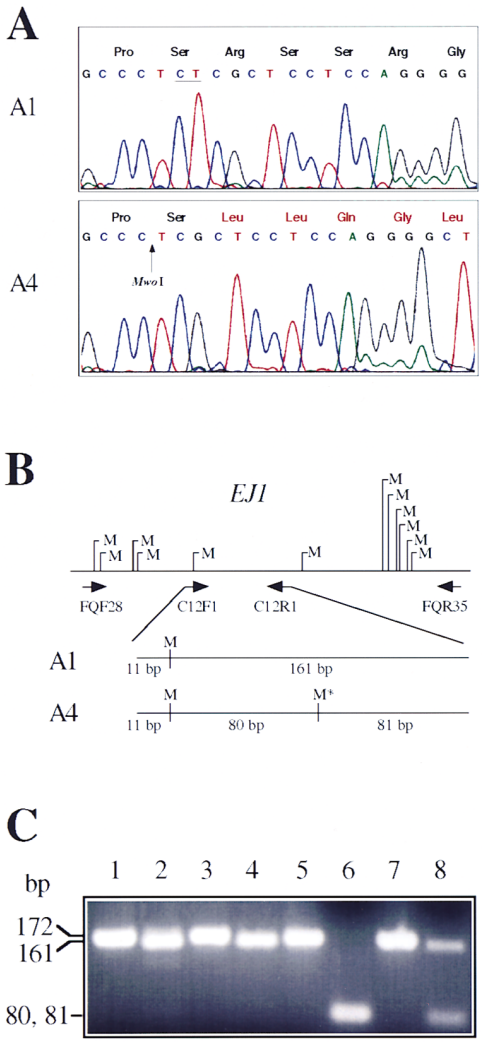

Figure 6. A 2 bp Deletion in C12 (ΔC12) Creates a Novel MwoI Site

(A) *EJ1* was amplified from C12, and clones of two alleles (A1 and A4) were identified. The sequence of the novel allele, A4, is shown along with the corresponding sequence of A1. A4 contains a 2 bp deletion (underlined in A1) that creates a new MwoI site (arrow). The deletion was confirmed by analysis of two clones sequenced in both orientations.

(B) (Top) The MwoI restriction map of *EJ1* A1 (M). An internal PCR product of 172 bp was used to test for the presence of the novel MwoI site (M*). (Bottom) Comparison of the MwoI maps of this subfragment of A1 and A4. Cleavage at the new site is predicted to result in two fragments of nearly identical length in A4.

(C) To confirm that the DC12 deletion was a somatic mutation, the 172 bp subfragment was amplified from the following templates: DNA isolated from peripheral blood of JHU273 (lanes 1 and 2) and cloned alleles A1 (lanes 3 and 4) and A4 (lanes 5 and 6) as well as from the original *EJ1* PCR product (lanes 7 and 8). Undigested products (lanes 1, 3, 5, and 7) and products digested with MwoI (lanes 2, 4, 6, and 8) were analyzed on a 4% Nusieve gel. As predicted, two indistinguishable bands of 80 and 81 bp are present in lanes 6 and 8 and absent in lanes 2 and 4. A 161 bp fragment is also observed in the original *EJ1* PCR product (lane 8). The 11 bp restriction fragment common to all could not be detected. The results were confirmed using an independently derived C12 *EJ1* PCR product (data not shown).

6C). The donor's blood DNA lacks the restriction site, proving that the mutation was acquired and not present in the donor's germline. The second difference in the A4 sequence is a C-to-T transition at position 12,617, which has no effect on the protein sequence. This is the same sequence variant associated with the A2 allele, suggesting that the ΔC12 deletion occurred on this background.

Discussion

A notable but unexplained clinical feature of all forms of ADPKD is the phenotypic variability exhibited by family members who share a common mutation. The focal nature of renal cysts is an extreme example of this variability since all renal tubular epithelial cells within a kidney have an identical mutation. These features suggest that other factors are required for cyst formation. Identification of these other components may explain the pathogenesis of the disease and possibly provide new avenues for developing therapies.

To evaluate the molecular mechanisms responsible for cystogenesis, we developed a novel method for isolating epithelial cells from individual cysts. This technique minimizes contamination from other cell types and avoids biases resulting from the use of cultured cells. We have used this approach to analyze multiple kidneys from females and have shown, using two independent methods, that renal cysts in ADPKD are clonal in origin. Our data suggest that somatic mutation of the previously normal *PKD1* allele is the second, rate-limiting step in cystogenesis and is likely to account for the focal formation of renal cysts. The rate at which one acquires these "second hits" probably plays a major role in determining the course of disease and probably explains much of its phenotypic variability. It may be speculated that a similar two-hit mechanism may also be responsible for causing the extrarenal manifestations of PKD1.

Loss of the normal allele of *PKD1* in cystic tissue suggests a molecular recessive mechanism of disease similar to that seen with numerous tumor suppressor genes. If our model is correct, however, the frequency of "second hits" in cystic epithelia must be extremely high to account for the thousands of cysts that are observed. This is in contrast to the small number of tumors found in other inherited disorders of the kidney in which somatic mutation of the normal allele is required for disease. Wilms' tumor (Huff et al., 1991) and von Hippel Lindau's disease (Foster et al., 1994) are representative examples.

Our data suggest that there is indeed a high rate of somatic mutation in *PKD1*. This is best illustrated by sample JHU273 (Table 1). Of the 13 cysts that were shown to be monoclonal, 2 had loss of alleles for *KG8* and *EJ1*; 1 had a unique 2 bp deletion; and 8 had no detectable mutation of the normal allele. The last group did not have a common X-chromosome inactivation pattern, suggesting that they had arisen from at least two independent progenitor cells (Figure 1). Assuming that the somatic mutations arose after X-chromosome inactivation was completed, a minimum of four independent mutations must have occurred in the 13 cysts that were examined. We cannot determine whether the two cysts with LOH for *KG8* had identical mutations or were independent events since they had identical X-chromosome inactivation patterns. Cysts from other kidneys, however, that had a similar pattern of LOH for *PKD1* markers had different X chromosomes inactivated, suggesting that independent events had resulted in LOH (data not shown). Formal proof of our hypothesis must await the development of methods that can be used for mutation analysis of the full length of the *PKD1* gene.

The very high rate of somatic mutation predicted by our model is surprising since no known dynamic elements have been identified within the genomic sequence of *PKD1* and since the adult kidney is thought to have a relatively low mitotic index. However, we have previously reported an extremely unusual 2.5 kb polypyrimidine tract within intron 21 that may be responsible for the gene's increased rate of mutation (The American PKD1 Consortium, 1995). Similar but much shorter elements present within other genes have been shown to undergo triple-helix formation both in vitro (Young et al., 1991) and in vivo (Rao et al., 1988). Wang et al. (1996) recently have shown that triplex formation induces mutagenesis in a mammalian cell culture system and demonstrated a requirement for excision and transcription-coupled repair in this process. The authors hypothesized that formation of the triple helix causes a stall in transcription that triggers gratuitous and potentially error-prone repair. They proposed that naturally occurring triple helices may similarly trigger repair and mutagenesis and thus constitute endogenous sources of genetic instability.

We postulate that the polypyrimidine tract within *PKD1* may cause ongoing errors in its transcription-coupled repair that result in a high frequency of somatic mutation. This model can explain the multiplicity of second hits as well as the apparent development of new cysts throughout the lifetime of an individual. Likewise, the genetic instability possibly associated with this unusual genomic structure may be responsible for the high incidence of PKD1 within the population. If this element is proven to be responsible for the gene's mutability, PKD1 will be the first example of a disease that results from this novel mechanism of somatic mutation.

The indistinguishable clinical presentation of patients with PKD2 suggests a two-step process in this disease as well. Whereas the initial step is certainly a germline mutation of *PKD2*, the nature of the second event is not yet known. Presently it cannot be excluded that there may be an uncharacterized unstable element hiding within the genomic structure of *PKD2* that results in frequent somatic inactivation in a manner analogous to that of *PKD1.* The relative infrequency of PKD2 argues against this hypothesis, however. The phenotypic similarity of the disorders may offer an alternative hypothesis. Investigators have postulated that the gene products of PKD1 and PKD2 may be interacting partners of a common pathway. It has been suggested that the function of PKD1 may be to regulate the activity of PKD2 (Mochizuki et al., 1996). If this is correct, somatic inactivation of *PKD1* may be the second step that leads to clonal expansion in PKD2 and possibly other forms of ADPKD. This model predicts that the frequency with

which cysts form in PKD1, PKD2, and PKD3 will be determined by the rate of somatic mutation of *PKD1*. The identical number of renal cysts observed in the three disorders is consistent with this hypothesis.

In this study, we have determined that random inactivation of the normal *PKD1* allele in somatic tissue is the likely molecular explanation for the observed clinical variability and the focal formation of renal cysts in the most common form of ADPKD. Both the high incidence of the disease in the population and the large number of second hits that must occur to account for the number of cysts observed suggest that unique features of the *PKD1* gene structure may be responsible for its mutability. Our data suggest a molecular recessive mechanism of disease since both alleles are mutated in renal cysts. These findings have important implications for investigators interested in developing models systems and suggest that treatment strategies directed at replacement of polycystin may prevent cyst formation and end-stage renal disease.

Experimental Procedures

Preparation of Cystic Epithelial Cells from a Single Cyst

Cystic kidneys were processed within 24 hours of removal from the patient and were maintained at 4°C. The surface of the cyst was first rinsed with PBS and then its contents were drained by needle and syringe. The needle was left inserted in the cyst for the duration of the washing and incubation steps. The cavity of intact cysts was rinsed a minimum of three times with Ca^{2+}- and Mg^{2+}-free PBS. The last rinse from some cysts was collected and stored on ice. PBS containing 2 mM EDTA was then injected into the lumen of the intact cyst. After a 20 min incubation, the cyst was massaged several times to assist in detachment of the epithelial cells from the basement membrane. Cysts that maintained the extraction solution (PBS/EDTA) in their lumina for the duration of the incubation period were considered intact. Only the epithelial cells of intact cysts were harvested by drainage. The cystic epithelial cells in the extraction solution (PBS/EDTA) and the last wash solutions were centrifuged at 1500 rpm (Beckman) for 15 min. The pellet was used for DNA preparation using the Puregene DNA extraction kit (Gentra) according to the manufacturer's protocol. To assist in DNA precipitation, 10 μg of glycogen was added to each sample.

Clonality Assay

The DNA samples were digested with 10 units of HpaII or HhaI in a volume of 10 μl overnight at 37°C. The enzyme was inactivated by heating at 95°C for 10 min and then used directly as template for PCR amplification. A set of control amplifications without prior restriction digestion was done in parallel using identical reaction conditions for samples from donors JHU93, JHU188, JHU244, and JHU304. In all cases, two alleles were amplified from each sample. PCR was performed for 28 cycles of 94°C for 20 sec, 65°C for 20 sec, and 72°C for 20 sec after an initial denaturation at 94° for 5 minutes. The PCR mixture (40 μl) contained a final concentration of 10 mM of each androgen receptor primer (AR1 and AR2), 5 nM ^{32}P–end-labeled AR1, 2 units of Taq DNA polymerase (Bio-Rad Laboratories), and 2 mM $MgCl_2$. The sequences of the primers were obtained from sequences reported by Tilley et al. (1989): AR1, 5′-GCTGTGAAGGTTGCTGTTCCTCAT-3′ (bp 485–508; HUMARA, GenBank accession number M21748), and AR2, 5′-TCCAGAATCTG TTCCAGAGCGTGC-3′ (bp 230–253). The PCR products were diluted with 5× volume of sequencing loading buffer (95% formamide, 10 mM EDTA [pH 8.0], 0.1% bromophenol blue, and 0.1% xylene cyanol FF) and separated in a 6% denaturing polyacrylamide gel. The dried gel was examined by autoradiography using X-Omat XAR film (Kodak) at −80°C overnight with an intensifying screen. An aliquot of DNA from a male with a different number of CAG repeats at the androgen receptor locus (so that the alleles could be distinguished) was added to each cyst DNA sample prior to PCR to control for the completeness of HpaII or HhaI digestion. In most cyst samples, the male-specific product was not amplified after HpaII or HhaI digestion, as expected.

Loss of Heterozogosity Assay Using *KG8*

The microsatellite *KG8* that is present in the 3′UTR of *PKD1* mRNA was used to distinguish the mutant and normal alleles (Snarey et al., 1994). Cystic DNA samples served as template for PCR amplification. PCR amplification was performed for 28 cycles using an identical protocol as that used for the androgen receptor except that the reaction included a final concentration of 1 μM of each *KG8* primer (KG8F8, 5′-CTCCCAGGGTGGAGGAAGGTG-3′ [bp 13,925–13,945; HUMPKD1A, GenBank accession number L33243] and KG8R8, 5′-GCAGGCACAGCCAGCTCCGAG-3′ [bp 14,014–14,034]), 5 nM ^{32}P–end-labeled KG8F8, and 1.5 mM $MgCl_2$. The PCR products were diluted with 5× volume of sequencing loading buffer and separated in a 6% denaturing polyacrylamide gel. The dried gel was examined by autoradiography using X-Omat XAR film (Kodak) at −80°C overnight with an intensifying screen.

Heteroduplex Analysis

Either 200 ng of genomic DNA (isolated from whole blood using the Puregene kit) or 5 μl of cyst DNA was used as template for amplification of a 540 bp product (*EJ1*) using primers FQF28 (5′-CACGCCTTGCGTGGAGAG-3′, 12,539–12,556; HUMPKD1A) and FQR35 (5′-ATGGGCCACGGGAAGATCC-3′, 12,977–12,995). PCR was performed as follows: denaturation at 94°C for 5 min; 35 cycles of 94°C for 30 sec, 62°C for 30 sec, and 72°C for 30 sec; and a final extension of 72°C for 10 min. The total PCR volume was 30 μl using 2 units of Taq DNA polymerase (Boehringer Mannheim), 0.2 μl dCTP, and a final $MgCl_2$ concentration of 1.5 mM.

Heteroduplex analysis was performed using Hydrolink Mutation Detection Enhancement gels (MDE, AT Biochem) following the manufacturer's protocol. Urea was added to the gel to a final concentration of 15% to minimize band broadening. The radiolabeled PCR products were initially denatured by heating at 95°C for 5 minutes and then allowed to cool to room temperature gradually over 1–2 hr before loading. Gels were run at 700 V for 14–16 hr, dried, and placed on X-Omat XAR film (Kodak) at room temperature or on a Phosphoimager cassette (Molecular Dynamics).

Mixing Studies

The polymorphic locus *EJ1* was identified by heteroduplex analysis using the primers FQF28–FQR35 as described above. This PCR product spans part of exons 45 and 46 as well as the 90 bp intron between them and contains the coiled-coil domain of polycystin. Three heteroduplex patterns were detected in a screen of 20 normal and 45 affected individuals. The PCR products that gave each pattern were cloned into pCRII (Invitrogen) and then sequenced to determine the molecular basis for the polymorphism. A1 was found to contain the published sequence (HUMPKD1A, GenBank accession number L33243), whereas A2 had a conservative base pair change (C-to-T) at position 12,617, which did not alter the amino acid sequence (leucine, amino acid 4136). A3 had eight consecutive guanines in intron 45 instead of the seven contained in the published genomic sequence (51,325–51,331; HUMPKD1GEN, GenBank accession number L39891). All three alleles were used for mixing studies.

EJ1 was amplified from clones of each of the alleles using the PCR conditions described above. Four microliters of radiolabeled *EJ1* amplifed using cyst DNA as template was mixed in separate tubes with 4 μl of cold product derived from each of the three cloned alleles. Heteroduplex analysis was performed as described above. The products obtained using sample C12 as template were cloned into pCRII and assayed using A1–A2 in mixing studies. Two independent clones containing the mutant allele were sequenced using an automated sequencer (Applied Biosystems, Inc.) and standard protocols.

The 2 bp deletion identified in the C12 mutant clones created a new MwoI site. To test for the presence of this site, a nested 172 bp PCR product was amplified using primers C12F1 (5′-CTC

TGCCCAGGGTGCAGC-3′, genomic position 51,350–51,367; HUMPKD1GEN) and C12R1 (5′-GAGGTGGAGGGGTGCGAG-3′, cDNA position 12,760–12,777; HUMPKD1A) and the following templates: peripheral blood leukocyte DNA of the same donor, plasmid DNA containing the *EJ1* alleles cloned from C12, and the original *EJ1* products obtained from C12 in two independent reactions. PCR products were digested with MwoI (New England Bio Labs) according to the manufacturer's specfications and analyzed on a 4% Nusieve gel (FMC Bioproducts).

Haplotype and Family Analysis

The allelic genomic fragments that include both the *KG8* and *EJ1* loci were amplified from each individual using the primers FQF28 and KG8R8 and 200 ng of peripheral blood leukocyte DNA as template. The 50 μl reaction included 4 units of rTth DNA polymerase, XL (Perkin Elmer), and a final concentration of 1.1 mM magnesium acetate and was performed using the following conditions: 95°C for 3 min; 35 cycles of 95°C for 20 sec and 68°C for 2 min; and a final extension of 68°C for 10 min. The 1.6 kb products were cloned into pCRII, and then individual clones were analyzed for the variants they contained at each locus using the microsatellite assay (*KG8*) and mixing studies (*EJ1*) as described above.

Family members of patients donating kidneys were recruited to participate after receiving permission from the donors. Blood samples were obtained after receiving informed consent and according to institutional guidelines. DNA was prepared from whole blood using the Puregene kit. The *KG8* microsatellite assay was performed as described above. The marker *3′HVR* was also used to confirm linkage of ADPKD to chromosome 16 markers. Five micrograms of genomic DNA was digested with PvuII (New England Bio Labs), electrophoresed overnight through a 1% agarose gel, and then transferred to nylon membrane (Schleicher and Schuel) using standard Southern blot protocols. *3′HVR* was labeled using Rediprime (Amersham) and hybridized to Southern blots in hybridization buffer (0.5 M Na_2HPO_4 [pH 7.2], 1 mM EDTA, 1% BSA, and 7% SDS) at 66°C. Blots were washed two or three times at 65°C with 0.1% SSC, 0.1% SDS and imaged using a Phosphoimager screen (Molecular Dyanamics).

Two-point linkage analyses were performed using the MLINK program of the LINKAGE 5.1 package. The disease gene frequency was set at 0.001 and penetrance set at 0.98 for individuals >30 years old. Sex-averaged recombination fractions were used (0.05 for *3′HVR* and 0.0 for *KG8*). The posterior probability of linkage to PKD1 was determined for pedigree 1 (Figure 5A) using an estimate of the prior probability of linkage (α) of 0.85 (Narod, 1991; Peters and Sandkuijl, 1992; Ravine et al., 1992; Peters et al., 1993; Mochizuki et al., 1996).

Acknowledgments

Correspondence should be addressed to G. G. G. We are grateful to all ADPKD family members for their invaluable participation and the Polycystic Kidney Research Foundation for assistance in the collection of tissue samples used in the study. We thank Xiangbin Zhang, Edward Lee, and Martin Daoust for their help in cyst preparation; Dr. Klaus Piontek and Sidney McGaughey for their help in the preparation of the manuscript; Dr. Irene Maumenee and Thomas Mitchell of the Johns Hopkins Hereditary Eye Center for assistance in the linkage calculations; and Dr. Joseph Handler for helpful discussions. This work was supported by the Polycystic Kidney Research Foundation (grant 95004), the National Insitutes of Health (grant DK48006), and the McKutcheon Foundation. G. G. G. is the Irving Blum Scholar of The Johns Hopkins University School of Medicine.

Received July 19, 1996; revised September 30, 1996.

References

Allen, R.C., Zoghbi, H.Y., Moseley, A.B., Rosenblatt, H.M., and Belmont, J. W. (1992). Methylation of HpaII and HhaI sites near the polymorphic CAG repeat in the human androgen-receptor gene correlates with X chromosome inactivation. Am. J. Hum. Genet. *51*, 1229–1239.

The American PKD1 Consortium (1995). Analysis of the genomic sequence for the autosomal dominant polycystic kidney disease (PKD1) gene predicts the presence of a leucine-rich repeat. Hum. Mol. Gen. *4*, 575–582.

Blyth, H., and Ockenden, B.G. (1971). Polycystic disease of kidney and liver presenting in childhood. J. Med. Genet. *8*, 257–284.

Bogdanova, N., Dworniczak, B., Dragova, D., Todorov, V., Dimitrakov, D., Kalinov, K., Hallmayer, J., Horst, J., and Kalaydjieva, L. (1995). Genetic heterogeneity of polycystic kidney disease in Bulgaria. Hum. Genet. *95*, 645–650.

Daoust, M.C., Reynolds, D.M., Bichet, D.G., and Somlo, S. (1995). Evidence for a third genetic locus for autosomal dominant polycystic kidney disease. Genomics *25*, 733–736.

de Almeida, S., de Almeida, E., Peters, D., Pinto, J.R., Tavora, I., Lavinha, J., Breuning, M., and Prata, M.M. (1995). Autosomal dominant polycystic kidney disease-evidence for the existence of a third locus in a Portuguese family. Hum. Genet. *96*, 83–88.

The European Polycystic Kidney Disease Consortium (1994). The polycystic kidney disease gene encodes a 14 kb transcript and lies within a duplicated region on chromosome 16. Cell *77*, 881–894.

Evan, A.P., and McAteer, J.A. (1992). Cyst cells and cyst walls. In The Cystic Kidney, K.D. Gardner and J. Bernstein, eds. (London: Kluwer Academic Publishers), pp. 21–41.

Fearon, E.R., Hamilton, S.R., and Vogelstein, B. (1987). Clonal analysis of human colorectal tumors. Science *238*, 193–197.

Fick, 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.

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

Foster, K., Prowse, A., van den Berg, A., Fleming, S., Hulsbeek, M. M., Crossey, P.A., Richards, F.M., Cairns, P., Affara, N.A., Ferguson-Smith, M.A., et al. (1994). Somatic mutations of the von Hippel-Lindau disease tumour suppressor gene in non-familial clear cell renal carcinoma. Hum. Mol. Genet. *3*, 2169–2173.

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

Gilliland, D.G., Blanchard, K.L., Levy, J., Perrin, S., and Bunn, H.F. (1991). Clonality in myeloproliferative disorders: analysis by means of the polymerase chain reaction. Proc. Natl. Acad. Sci. USA *88*, 6848–6852.

Handler, J.S., Steele, R.E., Sahib, M.K., Wade, J.B., Preston, A.S., Lawson, N.L., and Johnson, J.P. (1979). Toad urinary bladder epithelial cells in culture: maintenance of epithelial structure, sodium transport and reponse to hormones. Proc. Natl. Acad. Sci. USA *76*, 4151–4155.

Huff, V., Miwa, H., Haber, D.A., Call, K.M., Housman, D., Strong, L. C., and Saunders, G. F. (1991). Evidence for WT1 as a Wilms tumor (WT) gene: intragenic germinal deletion in bilateral WT. Am. J. Hum. Genet. *48*, 997–1003.

Hughes, J., Ward, C.J., Peral, B., Aspinwall, R., Clark, K., San Millán, J.L., Gamble, V., and Harris, P.C. (1995). The polycystic kidney disease 1 (PKD1) gene encodes a novel protein with multiple cell recognition domains. Nature Genet. *10*, 151–160.

Iakoubova, O.A., Dushkin, H., and Beier, D.R. (1995). Localization of a murine recessive polycystic kidney disease mutation and modifying loci that affect disease severity. Genomics *26*, 107–114.

The International Polycystic Kidney Disease Consortium (1995). Polycystic kidney disease: the complete structure of the *PKD1* gene and its protein. Cell *81*, 289–298.

Kaariainen, H. (1987). Polycystic kidney disease in children: a genetic and epidemiological study of 82 Finnish patients. J. Med. Genet. *24*, 474–481.

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.

Mochizuki, T., Wu, G., Hayashi, T., Xenophontos, S.L., Veldhuisen, B., Saris, J.J., Reynolds, D.M., Cai, Y., Gabow, P.A., Pierides, A., et al. (1996). *PKD2*, a gene for polycystic kidney disease that encodes an integral membrane protein. Science *272*, 1339–1342.

Narod, S. (1991). Counselling under genetic heterogeneity: a practical approach. Clin. Genet. *39*, 125–131.

Peral, B., Gamble, V., San Millán, J.L., Strong, C.Sloane-Stanley, J., Moreno, F., and Harris, P.C. (1995). Splicing mutations of the polycystic kidney disease 1 (PKD1) gene induced by intronic deletion. Hum. Mol. Gen. *4*, 569–574.

Peral, B., San Millán, J.L., Ong, A.C., Gamble, V., Ward, C.J., Strong, C., and Harris, P.C. (1996a). Screening the 3′ region of the polycystic kidney disease 1 (PKD1) gene reveals six novel mutations. Am. J. Hum. Genet. *58*, 86–96.

Peral, B., Ong, A.C.M., San Millán, J.L., Gamble, V., Rees, L., and Harris, P.C. (1996b). A stable, nonsense mutation associated with a case of infantile onset polycystic kidney disease 1 (PKD1). Hum. Mol. Gen. *5*, 539–542.

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

Peters, D.J.M., Spruit, L., Saris, J J., Ravine, D., Dandkuijl, L.A., Fossdal, R., Boersma, J., van Eijk, R., Borby, S., Constantinou-Deltas, C.D., et al. (1993). Chromosome 4 localization of a second gene for autosomal dominant polycystic kidney disease. Nature Genet. *5*, 359–362.

Rao, B.S., Manor, H., and Martin, R.G. (1988). Pausing in simian virus 40 DNA replication by a sequence containing (dG-dA)27·(dT-dC)27. Nucleic Acids Res. *16*, 8077–8094.

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

Schäfer, K., Gretz, N., Bader, M., Obermüller, I., Eckardt, K.-U., Kriz, W., and Bachmann, S. (1994). Characterization of the Han:SPRD rat model for hereditary polycystic kidney disease. Kidney Int. *46*, 134–152.

Snarey, A., Thomas, S., Schneider, M.C., Pound, S.E., Barton, N., Wright, A.F., Somlo, S., Germino, G.G., Harris, P.C., and Reeders, S.T. (1994). Linkage disequilibrium in the region of the autosomal dominant polycystic kidney disease gene (PKD1). Am. J. Hum. Genet. *55*, 365–371.

Tilley, W.D. Marcelli, M., Wilson, J.D., and McPhaul, M.J. (1989). Characterization and expression of a cDNA encoding the human androgen receptor. Proc. Natl. Acad. Sci. USA *86*, 327–331.

Wang, G., Seidman, M.M., and Glazer, P.M. (1996). Mutagenesis in mammalian cells induced by triple helix formation and transcription-coupled repair. Science *271*, 802–805.

Willman, C.L., Busque, L., Griffith, B.B., Favara, B.E., McClain, K.L., Duncan, M.H., and Gilliland, D.G. (1994). Langerhans'-cell histiocytosis (histiocytosis X): a clonal proferative disease. N. Engl. J. Med. *331*, 154–160.

Young, SD. L., Krawczyk, S.H., Matteucci, M.D., and Toole, J.J. (1991). Triple helix formation inhibits transcription elongation in vitro. Proc. Natl. Acad. Sci. USA *88,* 10023–10026.

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