RESEARCH

Generation of a Transcriptional Map for a 700-kb Region Surrounding the Polycystic Kidney Disease Type 1 (*PKD1*) and Tuberous Sclerosis Type 2 (*TSC2*) Disease Genes on Human Chromosome 16p13.3

Timothy C. Burn,[1] Timothy D. Connors, Terence J. Van Raay, William R. Dackowski, John M. Millholland, Katherine W. Klinger, and Gregory M. Landes

Department of Human Genetics, Integrated Genetics, Framingham, Massachusetts 01701

A 700-kb region of DNA in human chromosome 16p13.3 has been shown to contain the polycystic kidney disease 1 (*PKD1*) and the tuberous sclerosis type 2 (*TSC2*) disease genes. An estimated 20 genes are present in this region of chromosome 16. We have initiated studies to identify transcribed sequences in this region using a bacteriophage P1 contig containing 700 kb of DNA surrounding the *PKD1* and *TSC2* genes. We have isolated 96 unique exon traps from this interval, with 23 of the trapped exons containing sequences from five genes known to be in the region. Thirty exon traps have been mapped to additional transcription units based on data base homologies, Northern analysis, or their presence in cDNA or reverse transcriptase (RT)–PCR products. We have mapped the human *RNPS* gene to the cloned interval. We have obtained cDNAs or RT–PCR products from eight novel genes, with sequences from seven of these genes having homology to sequences in the data bases. Two of the newly identified genes represent human homologs for rat and murine genes identified previously. We have isolated three exon traps with homology to sequences in the data bases but have been unable to confirm the presence of these exon traps in expressed sequences. In addition, we have isolated 43 exon traps that do not map to our existing cDNAs or PCR products and have no homology to sequences in the data bases. In this report we present a transcriptional map for the 700 kb of DNA surrounding the *PKD1* and *TSC2* genes.

A 700-kb CpG-rich region in band 16p13.3 has been shown to contain the disease gene for 90% of the cases of autosomal dominant polycystic kidney disease (*PKD1*) (Germino et al. 1992; Somlo et al. 1992; European Polycystic Kidney Disease Consortium 1994) as well as the tuburin gene (*TSC2*), responsible for one form of tuberous sclerosis (European Chromosome 16 Tuberous Sclerosis Consortium 1993). An estimated 20 genes are present in this region of chromosome 16 (Germino et al. 1993). In addition to the *PKD1* and *TSC2* genes, four additional genes from this interval have been characterized previously and include the *ATP6C* proton pump gene (Gillespie et al. 1991), the *CCNF* gene encoding cyclin F (Kraus et al. 1994), the *sazD* gene encoding a transducin-like protein (Weinstat-Saslow et al. 1993). and the *hERV1* gene (Lisowsky et al. 1995).

Studies of the region surrounding *PKD1* in 16p13.3 are complicated by the duplication of a portion of the genomic interval more proximally at 16p13.1 (European Polycystic Kidney Disease Consortium 1994). This duplication makes the assembly of a transcriptional map particularly challenging. For these reasons, we have chosen to use exon trapping to identify transcribed sequences in the 700 kb of DNA surrounding the *PKD1* gene. Several exon trapping methodologies and vectors have been described for the rapid and efficient isolation of coding regions from genomic DNA (Auch and Reth 1990; Duyk et al. 1990; Buckler et al. 1991; Church et al. 1994). The major advantage of exon trapping is that the

[1]**Corresponding author.**
E-MAIL tburn@world.std.com; FAX (508) 620-1203.

expression of cloned genomic DNAs (cosmid, P1, or YAC) is driven by a heterologous promoter in tissue culture cells. This allows for coding sequences to be identified without prior knowledge of their tissue distribution or developmental stage of expression. A second advantage of exon trapping, and particularly important for our application, is that exon trapping allows for the identification of coding sequences from only the cloned template of interest, which eliminates the risk of characterizing highly conserved transcripts from duplicated loci. This is not the case for either cDNA selection or direct library screening. Exon trapping has been used successfully to identify transcribed sequences in the Huntington's disease locus (Ambrose et al. 1992; Taylor et al. 1992; Duyao et al. 1993) and *BRCA1* locus (Brody et al. 1995; Brown et al. 1995). In addition, a number of disease-causing genes have been identified using exon trapping, including the genes for Huntington's disease (Huntington's Disease Collaborative Research Group 1993), neurofibromatosis type 2 (Trofatter et al. 1993), Menkes disease (Vulpe et al. 1993), Batten Disease (International Batten Disease Consortium 1995), and the gene responsible for the majority of Long-QT syndrome cases (Wang et al. 1996).

In preliminary studies by our group, the 700-kb interval around the *PKD1* locus was cloned as a series of 14 overlapping bacteriophage P1 clones, with a minimal tiling series being defined by 11 P1 clones (Dackowksi et al., this issue). Here, we describe the isolation and characterization of 96 individual exon traps from the genomic P1 contig, with the exon traps being used as a framework for the assembly of a transcriptional map.

RESULTS

The Identification of Coding Sequences Using Exon Trapping

The starting reagent for exon trapping experiments was a 14-clone bacteriophage P1 contig containing 700 kb of genomic DNA from the region surrounding the *PKD1* gene in chromosome 16p13.3 (Dackowski et al., this issue). Exon trapping was performed using an improved trapping vector (Burn et al. 1995), with the resulting exon traps being characterized by DNA sequence analysis. In a number of cases, exons were trapped multiple times from overlapping P1 clones, whereas in other cases, exons were trapped both as individual exons and in combination with one or more flanking exons. The resulting unique exon traps were individually labeled and hybridized to the dot-blotted P1 contig to authenticate the exon trap as well as obtain additional positional information. In >98% of the cases examined, hybridization to the parental P1 clone was observed with additional hybridization to one or more overlapping P1 clones being seen in many cases (summarized in Fig. 1).

Comparison of Trapped Exons to Human Genes Reported Previously

To determine the relative efficiency of the exon trapping procedure, exon traps were compared with the cDNA sequences for those genes known to be in the interval around the *PKD1* gene (Fig. 1). Single exon traps were obtained from the human homolog of the *ERV1* (Lisowsky et al. 1995) and the *ATP6*C proton pump genes (Gillespie et al. 1991). In contrast, 8 individual exon traps were isolated from the *TSC2* gene and 10 from the *CCNF* gene (European Chromosome 16 Tuberous Sclerosis Consortium 1993; Kraus et al. 1994). We obtained trapped sequences from three of the exons present in the *PKD1* gene (American *PKD1* Consortium 1995; Hughes et al. 1995; International Polycystic Kidney Disease Consortium 1995). We were unable to identify exon traps containing sequences from the *sazD* gene, despite obtaining an exon trap from the adjacent *hERV1* gene as well as 16 additional exon traps from the 109.8C and 47.2H P1 clones, where the *sazD* gene is located. Based solely on these data, our overall efficiency for obtaining exon traps from genes in the region of *PKD1* could be estimated at 83%. This is in close agreement to previous studies using exon trapping, which reported the isolation of at least one exon trap from seven of eight genes known to be present in a chromosome 6 cosmid contig (North et al. 1993).

Sequences present in two exon traps (Genome Sequence Database L75926 and L75927), localizing to the region of overlap between the 96.4B and 64.12C P1 clones, were shown to contain sequences from the human homolog to the murine *RNPS* gene described previously (GenBank accession no. L37368), encoding an S-phase-prevalent DNA/RNA-binding protein (Schmidt and Werner 1993). A comparison of these exon traps to the dbEST data base indicated that they were also contained in cDNA 52161 from the I.M.A.G.E. Consortium (Lennon et al.

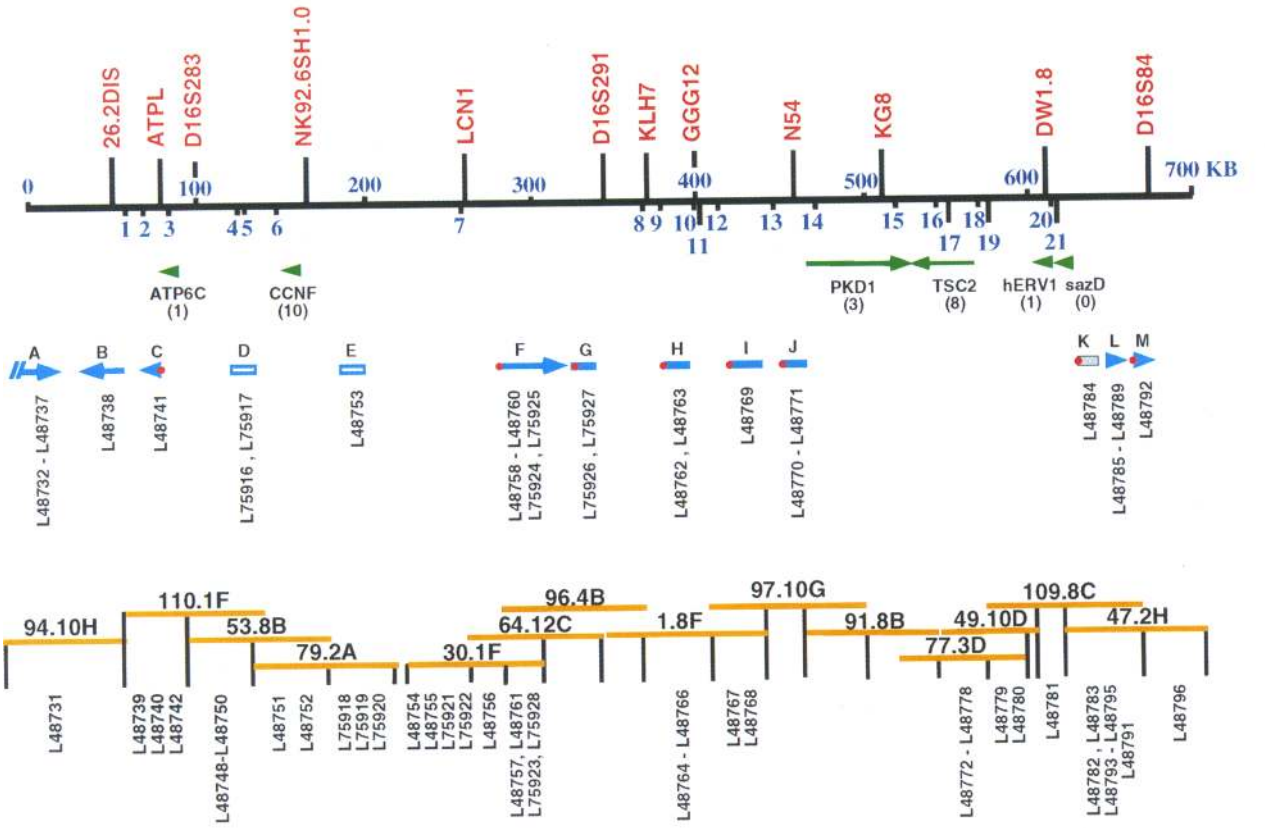

Figure 1 Schematic diagram of the P1 contig and trapped exons. The horizontal line at the *top* shows the position of relevant DNA markers (in red) with the scale (in kilobases) being indicated in navy blue. The position of *Not*I sites is shown in navy blue below the horizontal line. The position and orientation of the known genes is indicated by green arrows with the number of exon traps obtained from each gene shown in parentheses. The position of the transcription units described in this report (*A–M*) are shown below the known genes. Authenticated transcription units are shown in light blue, with the orientation being denoted by an arrow, if known. Transcriptional units that have not been authenticated are denoted by open light blue boxes (transcripts *D* and *E*), whereas the gray box denotes the LLRep3 sequences encoding a potential pseudogene (transcript *K*). Transcriptional units with EST hits are denoted by solid red circles. The accession numbers of corresponding exon traps are shown below each transcriptional unit. P1 clones are indicated by orange lines with the name of the clone above the line. The position of trapped exons that did not map to characterized transcripts are shown below the P1 contig. Vertical lines denote the interval within the P1 clones detected by the exon traps in hybridization studies.

1996). On the basis of these data, the h*RNPS* gene can be mapped to 16p13.3 near DNA marker D16S291 (transcript G in Fig.1).

Identification of Human Homologs to Two Rat Genes and a Murine Gene

At least three human homologs to rodent genes reported previously are present in the 700 kb examined based on comparison of the 96 unique exon traps with the data bases. Two exon traps from the 1.8F P1 clone were found to have a high level of homology to the murine ΦAP3 transcription factor described previously (Fognani et al. 1993). The two exon traps were linked by PCR, with the resulting 1.1-kb PCR product being 85% identical at the nucleotide level to the murine Φ*AP3* cDNA (Fognani et al. 1993). Hybridization of the Φ*AP3*-like exon traps to the dot-blotted P1 contig indicated that the gene lies in the nonoverlapping region of the 1.8F P1, between the DNA markers KLH7 and GGG12 (transcript H in Fig. 1). Significant homology was also seen between two exon traps from the 97.10G P1 and the rat *Rab26 ras*-related GTPase (Wagner et al. 1995). The *Rab26*-like exon traps were linked by

reverse transcriptase (RT)–PCR (transcript J in Fig. 1) with the encoded sequences being 94% (83/88) identical at the protein level to Rab26 (Fig. 2D). We also have obtained an exon trap that is 86% identical (170/197) at the nucleotide level to the rat augmentor of liver regeneration described previously, ALR (Hagiya et al. 1994). The *ALR*-like exon trap was shown to contain sequences from the recently described *hERV1* gene, which encodes a functional homolog to yeast ERV1 (Lisowsky et al. 1995). Data base searches by Lisowsky and co-workers (1995) revealed homology to only ERV1; however, a comparison of hERV1 and rat ALR indicates that the two proteins are 89% (111/125) identical.

Exon Traps Containing Sequences from Novel Genes

To correlate exon traps with individual transcripts, cDNA library screening and PCR-based approaches were used to clone transcribed sequences containing selected exon traps. RT–PCR was used to link individual exon traps together in cases where the two exon traps had homology to similar sequences in the data bases. In cases where only single exon traps were available, we relied on 3′ rapid amplification of cDNA ends (RACE) or cDNA library screening to obtain additional sequences. Using this strategy, we have obtained six partial cDNAs, three RT–PCR products, and a 3′ RACE product (Fig. 1; Table 1). Sequences from the exon traps and cloned products where used to map the position, and when possible, the orientation of the corresponding transcription units.

Six unique exon traps, containing sequences from at least eight exons, were shown to be from a transcriptional unit in our centromeric most P1 clone, 94.10H (transcript A in Fig. 1). A 2-kb cDNA linking the six exon traps was isolated and shown to hybridize to an 8-kb transcript (Fig. 3, below). Additional hybridization studies indicated that the gene was oriented centromeric to telomeric, with at least 6 kb of the transcript originating from sequences centromeric of our P1 contig. Extensive homology was observed between the translated cDNA and a variety of protein kinases; however, the presence of the conserved HRDLKPEN motif encoded in exon trap L48734, as well as the partial cDNA, suggests that it encodes a serine/threonine kinase (van-der-Geer et al. 1994). We have also isolated cDNAs using sequences from a separate 94.10H exon trap (L48738) and determined the position and orientation of the corresponding transcription unit (transcript B in Fig. 1). Two cDNA species were obtained using exon trap L48738 as a probe, with the only homology between the two species arising from the 109 bases contained in the exon trap. Using oligonucleotide probes, the transcription unit was mapped to a position near the 26-6DIS DNA marker, in a telomeric to centromeric orientation; however, we were only able to map one of the cDNA species to our P1 contig (transcript B in Fig. 1). On the basis of these data, we hypothesized that the second cDNA species originated from a region outside of our P1 contig, possibly from the duplicated 26-6PROX marker located farther centromeric in 16p13.3 (Gillespie et al. 1990).

In addition to the *ATP6*C gene, the 110.1F P1 clone contains at least two additional genes. Using BLASTX to search the protein data bases, we observed significant homology between sequences encoded by exon trap L48741 and the *N*-acetylglucosamine-6-phosphate deacetylase (nagA) proteins from *Caenorhabditis elegans* (Wilson et al. 1994), *Escherichia coli* (Plumbridge 1989), and *Haemophilus* (Fleischmann et al. 1995). An alignment of the nagA proteins to the translated exon trap revealed the presence of multiple conserved regions (Fig. 2A), suggesting that the exon trap contains sequences from the human *nagA* gene. We have cloned additional sequences from the *nagA*-like transcript using 3′ RACE and mapped the transcription unit to a region between *Not*I sites 2 and 3 in Figure 1. The gene is oriented telomeric to centromeric with *Not*I site 2 being present in the 3′ untranslated region (UTR) of the RACE clone (transcript C in Fig.1). Two additional exon traps (L75916 and L75917), mapping to the region of overlap between the 110.1F and 53.8B P1 clones (transcript D in Fig. 1), were shown to have homology with the chicken netrins (Kennedy et al. 1994; Serafini et al. 1994) and the *C. elegans* UNC-6 protein (Ishii et al. 1992) (Fig. 2). The netrins define a family of diffusable factors involved in axon outgrowth, whereas UNC-6 has been shown to have a role in the migration of mesodermal cells and axons (Ishii et al. 1992). Sequences encoded by exon trap L75917 were shown to have significant homology with the carboxy-terminal most epidermal growth factor (EGF) repeat found in the netrin and UNC-6 proteins (Fig. 2B). The second netrin-like trap (L75916) encodes sequences from the more divergent carboxy-terminal domain of

Table 1. Data Base Homologies

Gene[a]	Independent Exon Traps[b]	Clone[c]	Transcript Size	Database Homology[d]	Accession Number of Best Hit[e]	P value[f]
A	6	2 kb (cDNA)	8 kb	Probable protein kinase [*S. cerevisiae*]	Z48149	6.3e-83
B	1	1.3 kb (cDNA)	2.5	No Significant homology		
C	1	0.55 kb (Exon Trap) 0.6 kb (3' RACE)	1.4 kb	N-acetylglucosamine-6-phosphate deacetylase [*C. elegans*]	P34480	7.4e-73
D	2	Exon trap (159 bp) Exon trap (196 bp)	- -	Netrin-2 [*G. gallus*] Netrin-2 [*G. gallus*]	B54665 B54665	3.7e-11 6.1e-33
E	1	Exon trap (100 bp)	-	ABC1 gene product [*M. musculus*]	P41233	0.0047
F	3 2	1.1 kb (RT-PCR) 2.8 kb (cDNA)	7 kb 7 kb	ABC2 gene product [*M. musculus*] ABC1 gene product [*M. musculus*]	P41234 P41233	3.0e-28 7.1e-65
G	2	1.8 kb (cDNA)	2.5 kb	RNA-Binding protein [*Homo sapiens*]	L37368	2.6e-176
H	2	1.2 kb (RT-PCR)	2.5 kb	phi AP3 [*M. musculus*]	S41688	2.9e-169
I	1	0.45 kb (Exon Trap)	3.0 + 4.5 kb	No significant homologies		
J	2	0.24kb (RT-PCR)	2 kb	Rab26 [*R. norvegicus*]	U18771	3.6e-56
K	1	Exon trap (219 bp)	§	40S Ribosomal protein S4 [*Homo sapiens*]	P15880	7.3e-18
L	5	1.7 kb (cDNA)	1.6 kb	60s Ribosomal protein L3 [*Homo sapiens*]	S34195	6.7e-233
M	1	0.7 kb (cDNA)	1.3 kb	Hypothetical 17.2 Kd protein [*C. elegans*]	P34436	6.2e-10

[a]Gene as denoted in Fig. 1.
[b]Number of the trapped exon present in cloned cDNA or PCR product.
[c]Size of clone with type of clone indicated in parentheses.
[d]Significant homology in data bases as determined by BLASTX.
[e]Accession number of best hit.
[f]Smallest sum probability for the best data base match.
(-) Northern analysis was not performed because of the small size of the exon traps.
[§]Up to 200 copies of LLREP3 are present in the genome.

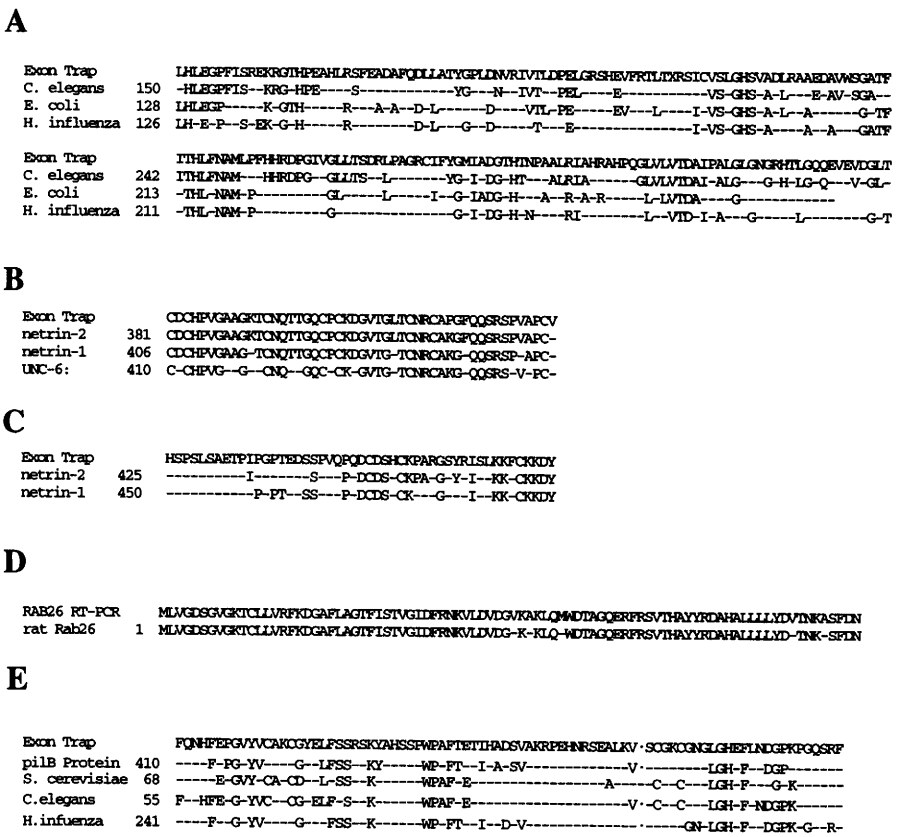

Figure 2 Alignment of selected exon traps with sequences in the data bases. (*A*) An alignment of sequences encoded by exon trap L48741 and nagA from *C. elegans, E. coli,* and *Haemophilus.* In all cases, identical residues are shown below the translated exon trap, whereas mismatches are denoted by the hyphens. The numbers at *left* designate the position of the amino acids from each protein. (*B*) The EGF repeat from netrin-1, netrin-2, and UNC-6 are shown aligned to one of the translated netrin-like exon traps (L75917). (*C*) An alignment of sequences from the second netrin-like exon trap (L75916) and netrin-1 and netrin-2 is shown. (*D*) An alignment of the translated Rab26-like RT–PCR product and rat Rab26. (*E*) Sequences encoded by exon trap L48792 are shown aligned to sequences from the pilB transcriptional repressor from *N. gonorrhoeae,* sequences predicted by computer analysis to be encoded by cosmid F44E2.6 from *C. elegans,* the YCL33C gene product from yeast (GenBank accession no. P25566), and a transcriptional repressor from *Haemophilus.* (·)Positions where gaps were inserted in the protein sequence to maintain alignment.

the netrins (Fig. 2C). This region is the least conserved between UNC-6 and the netrins, with sequences being 63% conserved between netrin-1 and -2 and 29% conserved between netrin-2 and UNC-6 (Serafini et al. 1994). We have been unable to confirm the presence of the netrin-like exon traps in transcribed sequences based on RT–PCR or cDNA library screening. These data suggest that the exon traps originate from a gene with a restricted tissue or developmental expression pattern, or alternatively from an untranscribed pseudogene.

Four exons were shown to have a high level of homology to the murine ABC1 and ABC2 proteins (Luciani et al. 1994), which represent a novel subclass of mammalian ABC (ATPase binding cassette) transporters. Three of the ABC-like exons (L48758–L48760) map to the region of overlap between the 30.1F, 64.12C, and 96.4B P1 clones (transcript F in Fig. 1), whereas the fourth (L48753) maps to the 79.2A P1 exclusively (transcript E in Fig. 1). We have been unsuccessful in identifying tissues that express sequences containing exon trap L48753 using both RT–PCR and PCR screening of cDNA libraries. Again, it is unclear whether the exon trap originated from an untranscribed pseudogene or a gene with a restricted pattern of expression. We have, however, obtained a 1.1-kb RT–PCR product that links the three exon traps from transcript F, with the RT–PCR product detecting a 7-kb message on Northern blots (data not shown). We have also obtained a cDNA from this region based on a search of the dbEST data base, with sequences from exon traps L75924 and L75925 being contained in cDNA 49233 from the I.M.A.G.E. Consortium (Lennon et al. 1996). Because data base searches indicated that the cDNA also encodes sequences with homology to murine ABC1 (Table 1), we hypothesized that the 1.1-kb RT–PCR product and the cDNA originate from the same transcription unit. The presence of both cloned reagents in the same transcription unit has been confirmed using RT–PCR (T.D. Connors, T.J. VanRaay, K.W. Klinger, G.M. Landes, and T.C. Burn, in prep.). These results, coupled with hybridization studies, indicate that a novel

gene encoding an ABC-transporter is located between the LCN1 and D16S291 markers in a centromeric to telomeric orientation.

Five trapped exons from P1s 109.8C and 47.2H were shown to contain sequences with homology to the human ribosomal protein L3 cDNA, with hybridization studies indicating that the L3-like gene is oriented centromeric to telomeric (transcript L in Fig. 1). The cumulative percent identity between the trapped exons and the reported human ribosomal protein L3 cDNA was 74% (537/724) at the nucleotide level. These data have been confirmed by analysis of a full-length cDNA, with the comparison suggesting that we have identified a novel ribosomal protein L3 subtype (T.J. Van Raay, T.D. Connors, K.W. Klinger, G.M. Landes, T.C. Burn, in prep.). In addition, we have obtained an exon trap (L48792) from a gene that is located telomeric of the L3-like gene (transcript M in Fig. 1). Exon trap L48792 was shown to have homology to a computer-predicted 17.2-kD protein encoded by cosmid F44E2.6 from *C. elegans* (Wilson et al. 1994). Using sequences from exon trap L48792, we have isolated a 600-bp partial cDNA and determined that the corresponding gene is oriented centromeric to telomeric. A 1.3-kb message is detected by the cDNA on Northern blots (Fig. 3). Sequences conserved between the partial cDNA and the hypothetical 17.2-kD protein were also conserved in the pilB protein from *Neisseria gonorrhoeae* (Taha et al. 1988), a hypothetical 19.3-kD protein from yeast (GenBank accession no. P25566), and a fimbrial transcription regulation repressor from *Haemophilus* (Fleischmann et al. 1995) (Fig. 2E). The pilB protein has homology to histidine kinase sensors and has been shown to have a role in the repression of pilin production in *N. gonorrhoeae* (Taha et al. 1988, 1991). An additional exon trap from region of overlap between the 109.8C and 47.2H P1 clones was shown to contain human LLRep3 sequences (Slynn et al. 1990). Hybridization studies indicated that the LLRep3 sequences (transcript K in Fig. 1) were located between the *sazD* and L3-like genes. However, because there are nearly 200 copies of LLRep3 in the murine and human genomes (Heller et al. 1988), with the majority of the copies representing pseudogenes, no additional characterization was performed on this exon trap.

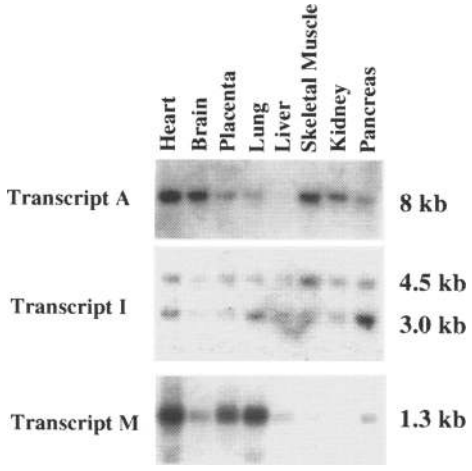

Figure 3 Expression pattern of selected transcription units. Hybridization of cDNAs (transcripts A and M) or an exon trap (transcript I) to poly(A)$^+$ RNA. Multiple-tissue Northern blots were obtained from Clontech. The sizes of the transcripts are shown at *right*. Exposure times were 1 (transcripts A and M) or 4 days (transcript I).

Forty-three of the trapped exons did not have significant homology to sequences in the protein or DNA data bases, nor were expressed sequence tags (ESTs) containing sequences from the exon traps observed in dbEST. The absence of ESTs containing sequences from these novel exon traps is not surprising because one of the criterion for selecting exon traps for further analysis was the presence of an EST in the data base. These trapped exons are likely to represent bona fide products, because in many cases they were trapped multiple times from different P1 clones and in combination with flanking exons. An example of this is given by exon trap L48769, which maps to the region of overlap between the 1.8F and 97.10G P1 clones. Sequences contained in the L48769 exon trap were obtained from both the 1.8F and 97.10G P1 clones (transcript I in Fig. 1), with a comparison of the exon traps from this region indicating that at least four individual exons are present in trap L48769 (data not shown). There were no sequences in the protein data bases with homology to exon trap L48769; however, an EST containing sequences from the exon trap was identified. The presence of the exon trap in transcribed sequences was also confirmed by Northern analysis, with the exon trap hybridizing to a 3.0- and 4.5-kb message (Fig. 3).

Comparison of the Exon Traps and cDNAs to the EST Data Base

A search of the dbEST data base using exon traps, cDNAs, and RT–PCR products revealed the pres-

ence of ESTs for 61.5% (8/13) of the novel genes identified in our studies (indicated by solid red circles in Fig. 1). In addition to the EST from transcript I described above, we found three ESTs containing sequences from the *nagA* gene (transcript C). Two ESTs from the ABC transporter gene (transcript F) were identified, whereas a total of five ESTs were found to contain sequences from the RNA-binding protein gene (transcript G). Single ESTs were identified for the human homologs to the murine Φ*AP3* (transcript H) and rat *Rab26* genes (transcript J). ESTs containing sequences from transcript M, encoding a potential transcriptional repressor, and LLRep3 (transcript K) were also identified.

DISCUSSION

We have used exon trapping to identify transcribed sequences in a P1 contig containing ~700 kb of DNA surrounding the *PKD1* and *TSC2* genes. Exon traps have been obtained from a minimum of 18 genes in this interval. The 18 genes identified include 5 genes reported previously from the interval and 1 previously characterized gene whose location was unknown. Additional exon traps have been mapped to genes based on their presence in cDNAs, RT–PCR products, or their hybridization to distinct mRNA species on Northern blots. Transcripts D and E have been tentatively labeled as genes based on the presence of homologous sequences in the protein data bases; however, we have been unable to confirm their presence in transcribed sequences. We have included the LLRep3 transcriptional unit (transcript K) in the 18 identified genes. However, the LLRep3 transcriptional unit may be a pseudogene, as there are 200 LLRep3 copies in the genome with the majority being pseudogenes (Heller et al. 1988). We have also identified 43 exon traps that have not been mapped to transcription units and lack significant homology to sequences in the data bases; some of these novel exon traps are likely to contain sequences from additional genes. We have not, however, tested these novel exon traps in "zoo-blots" to confirm their cross-species conservation, nor have we attempted to determined the number of genes that these exons may represent.

Exon traps representing a number of biologically interesting genes have been identified in the 700-kb region. Exons traps containing sequences from a novel protein kinase gene were observed in the centromeric-most P1. The kinase-like exon traps detected an 8-kb transcript by Northern analysis (Fig. 3), with sequences of one of the exon traps predicting a motif characteristic of a protein–serine/threonine kinase (van-der-Geer et al. 1994). Because the kinase gene is oriented centromeric to telomeric, with the 5′ 6 kb of the transcript lying outside of our P1 contig, it possible that the gene extends into the familial Mediterranean fever (FMF) critical region, which has a telomeric boundary mapping to the VK5 DNA marker (Aksentijevich et al. 1993), <100 kb centromeric of the 94.10H P1 clone.

We have also identified exon traps from a gene with homology to the netrin family of proteins. Analysis of the full-length cDNA and predicted peptide will be required to examine more completely the conservation of functional domains between the netrin-like gene encoded in the *PKD1* locus and the chicken netrins. However, the low level of homology between the netrin-like trap and the C-domain of the netrins may indicate that the exon traps originate from a gene encoding a closely related protein and not a homolog to either of the chicken netrins. To date, we have been unable to identify a cDNA library containing sequences from the netrin-like exon traps. Given that netrins function as target-derived neural chemoattractants, it is likely that the netrin-like gene has a restricted spatial and temporal pattern of expression. The identification of two exon traps from a gene with an apparently restricted expression pattern highlights one of the major advantages of exon trapping. That is the ability to identify transcribed sequences independent of their tissue-specific or developmental-specific patterns of expression.

Our exon trapping results indicate that there is at least one novel ABC transporter in the *PKD1* locus. Exon traps with homology to the murine *ABC1* and *ABC2* genes (Luciani et al. 1994) were obtained from the 79.2A P1 clone as well as from the region of overlap between the 30.1F, 64.12C, and 96.4B P1 clones. However, only the latter exon traps have been successfully mapped to a transcript, suggesting that the 79.2A ABC-like exon traps contain sequences from a gene with a restricted pattern of expression or an untranscribed pseudogene. Exon traps from the ABC transporter encoded by transcript F encode sequences with homology to the R-domain of the murine *ABC1* and *ABC2* genes (Luciani et al. 1994). The R-domain is believed to have a regulatory role based on the comparison to a conserved region in CFTR (Luciani et al. 1994). To

date, only ABC1, ABC2, and CFTR have been shown to contain an R-domain (Luciani et al. 1994). The cloning and characterization of a full-length cDNA for this novel ABC transporter gene will allow for a more comprehensive analysis (T.D. Connors, T.J. VanRaay, K.W. Klinger, G.M. Landes, and T.C. Burn, in prep.).

We report the cloning and characterization of sequences from human homologs to two previously characterized rat genes and a murine gene. We have obtained sequences from the human homolog to the murine Φ*AP3* gene, encoding a zinc finger-containing transcription factor (Fognani et al. 1993). The mΦAP3 protein is believed to function as a negative regulator for genes encoding proteins responsible for the inhibition of cell cycling (Fognani et al. 1993). In addition, we describe the identification of sequences from the human homolog to the rat *Rab26* gene encoding a GTP-binding protein involved in the regulation of vesicular transport (Nuoffer and Balch 1994; Wagner et al. 1995). Finally, we note that the human homolog to yeast ERV1 is also the human homolog to the recently cloned rat augmenter of liver regeneration (ALR), which is capable of augmenting hepatocytic regeneration following hepatectomy (Hagiya et al. 1994).

Sequences encoded by transcript M were shown to have homology to pilB from *N. gonorrhoeae* (Taha et al. 1988) as well as to sequences from *C. elegans* (Wilson et al. 1994), yeast, and *Haemophilus* (Fleischmann et al. 1995). The pilB protein, which acts to repress pilin production in *N. gonorrhoeae,* has been shown to have homology to histidine kinase sensors (Taha et al. 1991). However, residues conserved between pilB, transcript M, and the *C. elegans,* yeast, and *Haemophilus* sequences do not include the conserved histidine kinase domains from pilB (Taha et al. 1991). These findings suggest that the conserved region in transcript M has a function that is independent of the proposed histidine kinase sensor activity of pilB. It is unclear at present what the functional role of this conserved motif is.

The region of highest gene density appears to be at the telomeric end of the cloned interval, particularly the region between *TSC2* and D16S84, with a minimum of five genes mapping to this region (transcription units K, L, and M, *sazD* and *hERV1*). We have also mapped 17 novel exon traps to this region, with many of these likely to contain sequences from additional genes. Despite the identification of numerous genes from this region, we were unable to obtain an exon trap from the *sazD* gene. Similar gaps in transcriptional maps have been observed when independently derived maps from the *BRCA1* locus have been compared (Brody et al. 1995; Friedman et al. 1995; Osborne-Lawrence et al. 1995; Rommens et al. 1995). These findings, coupled with our results suggest that no transcriptional mapping methodology is 100% efficient, and a combination of approaches may be warranted. However, even using a combination of five methods, Brody and co-workers (1995) were unable to identify clones for at least two known genes in a 600-kb region.

In summary, we have used exon trapping to identify transcribed sequences from a 700-kb region in 16p13.3. We describe 96 exon traps from this region containing sequences from at least 18 genes, with 43 of the exon traps currently being unassigned to transcripts. Further study of this region will be aided greatly by the exon traps and cloned reagents described in this report. The continued characterization of this region and the genes encoded in it will provide valuable information about genome organization and gene expression.

METHODS

Tissue Culture

COS-7 cells (ATCC CRL-1651) were maintained in Dulbecco's modified Eagle medium (DMEM; Life Technologies, Inc., Gaithersberg, MD) supplemented with 10% fetal bovine serum (Life Technologies, Inc.), 2 mM L-glutamine (Life Technologies, Inc.) and penicillin/streptomycin (Life Technologies, Inc.).

Exon Trapping

Genomic P1 clones were prepared for exon trapping experiments by digestion with *Pst*I, double digestion with *Bam*HI–*Bgl*II, or partial digestion with limiting amounts of *Sau*3AI. *Sau*3AI digestions were performed by incubating 1 μg of P1 DNA with increasing amounts of *Sau*3AI (0.1–1.0 units) for 30 min at 37°C followed by phenol/chloroform/isoamyl alcohol (24:24:1) extraction and ethanol precipitation. For the *Sau*3AI partial digests, ligations were performed using those reactions yielding 2- to 6-kb products. The *Bam*HI–*Bgl*II-digested and *Sau*3AI-digested P1 DNAs were ligated to *Bam*HI-cut and dephosphorylated pSPL3B, whereas *Pst*I-digested P1 DNA was subcloned into *Pst*I-cut dephosphorylated pSPL3B. A modification present in the pSPL3B vector results in improved exon trapping efficiency, with the HIV *tat*-derived background being reduced to <1% of the resulting exon trapping products (Burn et al. 1995). Ligations were performed in triplicate using 50 ng of vector DNA and 1, 3, or 6 mass equivalents

of digested P1 DNA. Transformations were performed following an overnight 16°C incubation, with one-tenth and one-half of the transformation being plated on LB + AMP plates. After overnight growth at 37°C, colonies were scraped off those plates having the highest tranformation efficiency (based on a comparison with "no insert" ligation controls) and miniprepped using the alkaline lysis method. To examine the proportion of the pSPL3B containing insert, a small portion of the miniprep was digested with *Hin*dIII, which cuts pSPL3B on each side of the multiple cloning site. Approximately 10 μg of the remaining miniprep DNA was ethanol precipitated, resuspended in 100 μl of sterile PBS, and electroporated into ~2×10^6 COS-7 cells (in 0.7 ml of ice-cold PBS) using a Bio-Rad GenePulser electroporator (1.2 kV, 25 μF, and 200 Ω). The electroporated cells were incubated for 10 min on ice before their addition to a 100-mm tissue culture dish containing 10 ml of prewarmed complete DMEM.

Cytoplasmic RNA was isolated 48 hr post-transfection. The transfected COS-7 cells were removed from tissue culture dishes using 0.25% trypsin/1 mM EDTA (Life Technologies, Inc.). Trypsinized cells were washed in DMEM/10% FCS and resuspended in 400 μl of ice-cold TKM (10 mM Tris-HCl at pH 7.5, 10 mM KCl, 1 mM $MgCl_2$) supplemented with 1 μl of RNAsin (Promega, Madison, WI). After adding 20 μl of 10% Triton X-100, the cells were incubated for 5 min on ice. The nuclei were removed by centrifugation at 1200 rpm for 5 min at 4°C. Thirty microliters of 5% SDS was added to the supernatant, with the cytoplasmic RNA being purified further by three rounds of extraction using phenol/chloroform/isoamyl alcohol (24:24:1). The cytoplasmic RNA was ethanol-precipitated and resuspended in 50 μl of H_2O. RT–PCR was performed as described (Church et al. 1994), using commercially available exon trapping oligonucleotides (Life Technologies, Inc.). The resulting CUA-tailed products were shotgun-subcloned into pAMP10 as recommended by the manufacturer (Life Technologies, Inc.). Random clones from each ligation were analyzed by colony PCR using the secondary PCR primers (Life Technology, Inc.). Products containing trapped exons, based on comparison with the 177-bp "vector-only" product, were selected for sequencing.

Characterization of Exon Traps

Miniprep DNA containing the pAMP10/exon traps was prepared from overnight cultures by alkaline lysis using the EasyPrep manifold or a QIAwell 8 system as described by the manufacturers (Pharmacia, Piscataway, NJ and Qiagen Inc., Chatsworth, CA, respectively). DNA was sequenced with fluoroscein-labeled M13 universal and reverse primers using the Pharmacia Autoread Sequencing kit and Pharmacia ALF automated DNA sequencers. Sequences were assembled and analyzed using the Sequencher sequence analysis software (GeneCodes, Ann Arbor, MI). Sequences from the exon traps were used to search the nonredundant nucleotide and dbEST (on Dec. 16, 1995) data bases using the BLASTN program (Altschul et al. 1990), whereas protein data bases were searched using BLASTX (Altschul et al. 1990; Gish and States 1993). In searching dbEST, only identical or near identical sequences were examined in detail.

cDNA Library Screening and RT–PCR

cDNA library screening was performed using the Genetrapper cDNA Positive Selection System with adult kidney, liver, or heart libraries as described by the manufacturer (Life Technologies, Inc.). For Genetrapper experiments, two oligonucleotides (Table 2) were designed from each exon trap using Oligo 4.0 (National Biosciences, Inc., Plymouth, MN). The first oligonucleotide was biotinylated and used for selection, whereas the second oligonucleotide was used in the repair reaction as described by the manufacturer (Life Technologies, Inc.). RT–PCR was performed essentially as described using poly(A)$^+$ RNA from adult

Table 2. Oligonucleotides Used to Clone Additional Sequences

Gene[a]	Method[b]	Oligonucleotide 1[c]	Oligonucleotide 2[d]	Clone Size[e]
A	Genetrapper	TGACGCCGTGCCCATCCAGT	CAGCGTGGTGTTATGTTCCT	2.0 kb
B	Genetrapper	TTGGGCCTGTGCTGAACTAC	CGGCAAGCTGGTGATTAACA	1.3 kb
C	3' RACE	CGGCAGAGGATGCTGTGT	GCGGAGCCACCTTCATCA	0.6 kb
F	RT-PCR	GACGCTGGTGAAGGAGC	TCGCTGACCGCCAGGAT	1.1 kb
H	RT-PCR	CTGTCGGGAAGGTCTCACTG	GTTCACCGCCTTGGAGGATT	1.1 kb
J	RT-PCR	GTGTGGGGAAGACCTGTCTG	AGGAGGCCTTGTTGGTGACA	0.24 kb
L	Genetrapper	ACGGACACCTGGGCTTC	AAACGGGAGGAGGTGGA	1.7 kb
M	Genetrapper	TGTGGCTATGAGCTGTTCTC	GCAGTCCCGATTCTGAATAT	0.7 kb

[a]Gene as denoted in Fig. 1.
[b]Method used to clone additional sequences. Lifetechnologies Genetrapper system, 3' RACE, and RT–PCR were performed as described in Methods.
[c]Sequence of oligonucleotides used to obtain additional sequences. For the Genetrapper system, this oligonucleotide was used in the direct selection step. In the case of 3' RACE experiments, this oligonucleotide was the external primer. While in RT–PCR experiments, the designated oligonucleotide was used as a sense primer.
[d]Sequence of oligonucleotides. In the Genetrapper experiments, this oligonucleotide was used in the repair step. For 3' RACE experiments, this was the internal primer, for RT–PCR experiments, this was the antisense primer.
[e]Size of clone obtained using the primer pair.

brain and placenta (Kawasaki 1990). Sense and anti-sense RT–PCR primers (Table 2) were designed from different exons to minimize the risk of characterizing a product amplified from genomic DNA. To clone sequences 3′ to selected exon traps, RACE was performed as described (Frohman 1994). In cases where RNA from a desired source was unavailable, PCR was performed using commercially available cDNA libraries as template. All PCR products were cloned using the pGEM-T vector as described by the manufacturer (Promega).

ACKNOWLEDGMENTS

We thank Linda Petry for expert technical assistance and Norman Doggett (Los Alamos National Laboratory) for cosmids used for the initial phases of this work. This work was supported by a grant from the National Institutes of Health to K.W.K. (DK44853). The sequences of the exon traps have been submitted to the Genome Sequence Database (accession nos. L48731–L48742, L48747–L48796, and L75916–L75928) and to dbEST.

The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 USC section 1734 solely to indicate this fact.

REFERENCES

Aksentijevich, I., E. Pras, L. Gruberg, Y. Shen, K. Holman, S. Helling, L. Prosen, G.R. Sutherland, R.I. Richards, M. Ramsburg, et al. 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., W. Gish, W. Miller, E.W. Myers, and D.J. Lipman. 1990. Basic local alignment search tool. *J. Mol. Biol.* **215:** 403–410.

Ambrose, C., M. James, G. Barnes, C. Lin, G. Bates, M. Altherr, M. Duyao, N. Groot, D. Church, J.J. Wasmuth, et al. 1992. A novel G protein-coupled receptor kinase gene cloned from 4p16.3. *Hum. Mol. Genet.* **1:** 697–703.

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. Genet.* **4:** 575–582.

Auch, D. and M. Reth. 1990. Exon trap cloning: Using PCR to rapidly detect and clone exons from genomic DNA fragments. *Nucleic Acids Res.* **18:** 6743–6744.

Brody, L.C., K.J. Abel, L.H. Castilla, F.J. Couch, D.R. Mckinley, G.Y. Yin, P.P. Ho, S. Merajver, S.C. Chandrasekharappa, J.Z. Xu, J.L. Cole, J.P. Struewing, et al. 1995. Construction of a transcription map surrounding the *BRCA1* locus of human chromosome 17. *Genomics* **25:** 238–247.

Brown, M.A., K.A. Jones, H. Nicolai, M. Bonjardim, D. Black, R. Mcfarlane, P. Dejong, J.P. Quirk, H. Lehrach, and E. Solomon. 1995. Physical mapping, cloning, and identification of genes within a 500-kb region containing *BRCA1*. *Proc. Natl. Acad. Sci.* **92:** 4362–4366.

Buckler, A.J., D.D. Chang, S.L. Graw, J.D. Brook, D.A. Haber, P.A. Sharp, and D.E. Housman. 1991. Exon amplification: A strategy to isolate mammalian genes based on RNA splicing. *Proc. Natl. Acad. Sci.* **88:** 4005–4009.

Burn, T.C., T.D. Connors, K.W. Klinger, and G.M. Landes. 1995. Increased exon-trapping efficiency through modifications to the pSPL3 splicing vector. *Gene* **161:** 183–187.

Church, D.M., C.J. Stotler, J.L. Rutter, J.R. Murrell, J.A. Trofatter, and A.J. Buckler. 1994. Isolation of genes from complex sources of mammalian genomic DNA using exon amplification. *Nature Genet.* **6:** 98–105.

Dackowski, W.R., T.D. Connors, A.E. Bowe, V. Stanton Jr., D. Housman, N.A. Doggett, G.M. Landes, and K.W. Klinger. 1996. The region surrounding the *PKD1* gene: A 700-kb P1 contig from a YAC-deficient interval. *Genome Res.* (this issue).

Duyao, M.P., S.A. Taylor, A.J. Buckler, C.M. Ambrose, C. Lin, N. Groot, D. Church, G. Barnes, J.J. Wasmuth, D.E. Housman, et al. 1993. A gene from chromosome 4p16.3 with similarity to a superfamily of transporter proteins. *Hum. Mol. Genet.* **2:** 673–676.

Duyk, G.M., S.W. Kim, R.M. Myers, and D.R. Cox. 1990. Exon trapping: A genetic screen to identify candidate transcribed sequences in cloned mammalian genomic DNA. *Proc. Natl. Acad. Sci.* **87:** 8995–8999.

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

European Polycystic Kidney Disease Consortium. 1994. The polycystic kidney disease 1 gene encodes a 14 kb transcript and lies within a duplicated region on chromosome 16. *Cell* **77:** 881–894.

Fleischmann, R.D., M.D. Adams, O. White, R.A. Clayton, E.F. Kirkness, A.R. Kerlavage, C.J. Bult, J.F. Tomb, B.A. Dougherty, J.M. Merrick, K. McKenney, G. Sutton, et al. 1995. Whole-genome random sequencing and assembly of Haemophilus influenzae Rd. *Science* **269:** 496–512.

Fognani, C., G. Della-Valle, and L.E. Babiss. 1993. Repression of adenovirus E1A enhancer activity by a novel zinc finger-containing DNA-binding protein related to the GLI-Kruppel protein. *EMBO J.* **12:** 4985–4992.

Friedman, L.S., E.A. Ostermeyer, E.D. Lynch, P. Welcsh, C.I. Szabo, J.E. Meza, L.A. Anderson, P. Dowd, M.K. Lee, S.E. Rowell, J. Ellison, J. Boyd, et al. 1995. 22 genes from chromosome 17q21: Cloning, sequencing, and

characterization of mutations in breast cancer families and tumors. *Genomics* **25:** 256–263.

Frohman, M.A. 1994. On beyond classic RACE (rapid amplification of cDNA ends). *PCR Methods Applic.* **4:** S40–S58.

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

Germino, G.G., S. Somlo, D. Weinstat-Saslow, and S.T. Reeders. 1993. Positional cloning approach to the dominant polycystic kidney disease gene, *PKD1*. *Kidney Int. Suppl.* **39:** S20–S25.

Gillespie, G.A., G.G. Germino, S. Somlo, D. Weinstat-Saslow, M.H. Breuning, and S.T. Reeders. 1990. Cosmid walking and chromosome jumping in the region of *PKD1* reveal a locus duplication and three CpG islands. *Nucleic Acids Res.* **18:** 7071–7075.

Gillespie, G.A., S. Somlo, G.G. Germino, D. Weinstat-Saslow, and S.T. Reeders. 1991. CpG island in the region of an autosomal dominant polycystic kidney disease locus defines the 5′ end of a gene encoding a putative proton channel. *Proc. Natl. Acad. Sci.* **88:** 4289–4293.

Gish, W. and D.J. States. 1993. Identification of protein coding regions by database similarity search. *Nature Genet.* **3:** 266–272.

Hagiya, M., A. Francavilla, L. Polimeno, I. Ihara, H. Sakai, T. Seki, M. Shimonishi, K.A. Porter, and T.E. Starzl. 1994. Cloning and sequence analysis of the rat augmenter of liver regeneration (ALR) gene: Expression of biologically active recombinant ALR and demonstration of tissue distribution. *Proc. Natl. Acad. Sci.* **91:** 8142–8146.

Heller, D.L., K.M. Gianola, and L.A. Leinwand. 1988. A highly conserved mouse gene with a propensity to form pseudogenes in mammals. *Mol. Cell. Biol.* **8:** 2797–2803.

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

Huntington's Disease Collaborative Research Group. 1993. A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes. *Cell* **72:** 971–983.

International Batten Disease Consortium. 1995. Isolation of a novel gene underlying Batten disease, CLN3. *Cell* **82:** 949–957.

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

Ishii, N., W.G. Wadsworth, B.D. Stern, J.G. Culotti, and E.M. Hedgecock. 1992. UNC-6, a laminin-related protein, guides cell and pioneer axon migrations in C. elegans. *Neuron* **9:** 873–881.

Kawasaki, E.S. 1990. Amplification of RNA. In *PCR Protocols: A guide to methods and applications* (ed. M.A. Innis, D.H. Gelfand, J.J. Sninsky, and T.J. White), pp. 21–27. Academic Press, San Diego, CA.

Kennedy, T.E., T. Serafini, J.R. de-la-Torre, and M. Tessier-Lavigne. 1994. Netrins are diffusible chemotropic factors for commissural axons in the embryonic spinal cord. *Cell* **78:** 425–435.

Kraus, B., M. Pohlschmidt, A.L. Leung, G.G. Germino, A. Snarey, M.C. Schneider, S.T. Reeders, and A.M. Frischauf. 1994. A novel cyclin gene (*CCNF*) in the region of the polycystic kidney disease gene (*PKD1*). *Genomics* **24:** 27–33.

Lennon, G.G., C. Auffray, N. Polymeropoulos, and M.B. Soares. 1996. The I.M.A.G.E. consortium: An integrated molecular analysis of genomes and their expression. *Genomics* **33:** 151–152.

Lisowsky, T., D.L. Weinstatsaslow, N. Barton, S.T. Reeders, and M.C. Schneider. 1995. A new human gene located in the *PKD1* region of chromosome 16 is a functional homologue to ERV1 of yeast. *Genomics* **29:** 690–697.

Luciani, M.F., F. Denizot, S. Savary, M.G. Mattei, and G. Chimini. 1994. Cloning of two novel ABC transporters mapping on human chromosome 9. *Genomics* **21:** 150–159.

North, M.A., P. Sanseau, A.J. Buckler, D. Church, A. Jackson, K. Patel, J. Trowsdale, and H. Lehrach. 1993. Efficiency and specificity of gene isolation by exon amplification. *Mamm. Genome* **4:** 466–474.

Nuoffer, C. and W.E. Balch. 1994. GTPases: Multifunctional molecular switches regulating vesicular traffic. *Annu. Rev. Biochem.* **63:** 949–990.

Osborne-Lawrence, S., P.L. Welcsh, M. Spillman, S.C. Chandrasekharappa, T.D. Gallardo, M. Lovett, and A.M. Bowcock. 1995. Direct selection of expressed sequences within a 1-Mb region flanking *BRCA1* on human chromosome 17q21. *Genomics* **25:** 248–255.

Plumbridge, J.A. 1989. Sequence of the nagBACD operon in Escherichia coli K12 and pattern of transcription within the nag regulon. *Mol. Microbiol.* **3:** 505–515.

Rommens, J.M., F. Durocher, J. Mcarthur, P. Tonin, J.F. Leblanc, J. Allen, C. Samson, L. Ferri, S. Narod, K. Morgan, and J. Simard. 1995. Generation of a transcription map at the HSD17B locus centromeric to *BRCA1* at 17q21. *Genomics* **28:** 530–542.

Schmidt, G. and D. Werner. 1993. Sequence of a complete murine cDNA reflecting an S phase-prevalent transcript encoding a protein with two types of nucleic acid binding motifs. *Biochim. Biophys. Acta* **1216:** 317–320.

Serafini, T., T.E. Kennedy, M.J. Galko, C. Mirzayan, T.M. Jessell, and M. Tessier-Lavigne. 1994. The netrins define a family of axon outgrowth-promoting proteins homologous to C. elegans UNC-6. *Cell* **78:** 409–424.

Slynn, G., D. Jenner, W. Potts, P. Elvin, J.E. Morten, and A.F. Markham. 1990. Human cDNA sequence homologous to the mouse LLRep3 gene family. *Nucleic Acids Res.* **18:** 681.

Somlo, S., B. Wirth, G.G. Germino, D. Weinstat-Saslow, G.A. Gillespie, H. Himmelbauer, L. Steevens, P. Coucke, P. Willems, L. Bachner, et al. 1992. Fine genetic localization of the gene for autosomal dominant polycystic kidney disease (*PKD1*) with respect to physically mapped markers. *Genomics* **13:** 152–158.

Taha, M.K., M. So, H.S. Seifert, E. Billyard, and C. Marchal. 1988. Pilin expression in Neisseria gonorrhoeae is under both positive and negative transcriptional control. *EMBO J.* **7:** 4367–4378.

Taha, M.K., B. Dupuy, W. Saurin, M. So, and C. Marchal. 1991. Control of pilus expression in Neisseria gonorrhoeae as an original system in the family of two-component regulators. *Mol. Microbiol.* **5:** 137–148.

Taylor, S.A., R.G. Snell, A. Buckler, C. Ambrose, M. Duyao, D. Church, C.S. Lin, M. Altherr, G.P. Bates, N. Groot, et al. 1992. Cloning of the alpha-adducin gene from the Huntington's disease candidate region of chromosome 4 by exon amplification. *Nature Genet.* **2:** 223–227.

Trofatter, J.A., M.M. MacCollin, J.L. Rutter, J.R. Murrell, M.P. Duyao, D.M. Parry, R. Eldridge, N. Kley, A.G. Menon, K. Pulaski, et al. 1993. A novel moesin-, ezrin-, radixin-like gene is a candidate for the neurofibromatosis 2 tumor suppressor. *Cell* **72:** 791–800.

van-der-Geer, P., T. Hunter, and R.A. Lindberg. 1994. Receptor protein-tyrosine kinases and their signal transduction pathways. *Annu. Rev. Cell Biol.* **10:** 251–337.

Vulpe, C., B. Levinson, S. Whitney, S. Packman, and J. Gitschier. 1993. Isolation of a candidate gene for Menkes disease and evidence that it encodes a copper-transporting ATPase. *Nature Genet.* **3:** 7–13.

Wagner, A.C., M.Z. Strowski, B. Goke, and J.A. Williams. 1995. Molecular cloning of a new member of the Rab protein family, Rab 26, from rat pancreas. *Biochem. Biophys. Res. Commun.* **207:** 950–956.

Wang, Q., M.E. Curran, I. Splawski, T.C. Burn, J.M. Millholland, T.J. VanRaay, J. Shen, K.W. Timothy, G.M. Vincent, T. deJager, P.J. Schwartz, J.A. Towbin, et al. 1996. Positional cloning of a novel potassium channel gene: KVLQT1 mutations cause cardiac arrhythmias. *Nature Genet.* **12:** 17–23.

Weinstat-Saslow, D.L., G.G. Germino, S. Somlo, and S.T. Reeders. 1993. A transducin-like gene maps to the autosomal dominant polycystic kidney disease gene region. *Genomics* **18:** 709–711.

Wilson, R., R. Ainscough, K. Anderson, C. Baynes, M. Berks, J. Bonfield, J. Burton, M. Connell, T. Copsey, J. Cooper, A. Coulson, M. Craxton, et al. 1994. 2.2 Mb of contiguous nucleotide sequence from chromosome III of C. elegans. *Nature* **368:** 32–38.

Received February 21, 1996; accepted in revised form April 26, 1996.

Generation of a transcriptional map for a 700-kb region surrounding the polycystic kidney disease type 1 (PKD1) and tuberous sclerosis type 2 (TSC2) disease genes on human chromosome 16p3.3.

T C Burn, T D Connors, T J Van Raay, et al.

Genome Res. 1996 6: 525-537
Access the most recent version at doi:10.1101/gr.6.6.525