

Germline mutations in *PRKCSH* are associated with autosomal dominant polycystic liver disease

Published online 10 February 2003; doi:10.1038/ng1104

Polycystic liver disease (PCLD, OMIM 174050) is a dominantly inherited condition characterized by the presence of multiple liver cysts of biliary epithelial origin. Fine mapping established linkage to marker *D19S581* ($Z_{\max} = 9.65$; $\theta = 0.01$) in four large Dutch families with PCLD. We identified a splice-acceptor site mutation (1138-2A→G) in *PRKCSH* in three families, and a splice-donor site mutation (292+1G→C) in *PRKCSH* segregated completely with PCLD in another family. The protein encoded by *PRKCSH*, here named hepatocystin, is predicted to localize to the endoplasmic reticulum. These findings establish germline mutations in *PRKCSH* as the probable cause of PCLD.

We collected samples from 108 individuals in 4 separate Dutch families affected with PCLD; 32 individuals fulfilled the ultrasonographic criteria for PCLD^{1,2} (Fig. 1). All clinical investigations were done by the same physicians (J.D., R.S. and J.J.) and the study was approved by the local Medical Ethical Committee (CMO Regio Arnhem-Nijmegen, the Netherlands). All affected individuals from these families had polycystic livers, but females who had had multiple pregnancies had a more severe phenotype with gross hepatomegaly. This condition is distinct from autosomal dominant polycystic kidney disease type 1 (ADPKD-1, OMIM 173900) and type 2 (ADPKD-2, OMIM 173910), and none of the affected individuals showed evidence of renal disease or cerebral aneurysm as in ADPKD.

A previous study established linkage of PCLD to a 12.5-cM interval on chromosome 19p (ref. 3). Fine mapping in these individuals identified a maximum multipoint lod score of 10.96 for the interval between *D19S583* and *D19S581* (Fig. 1 and see Supplementary Table 1 online). Critical recombination events for individuals II-3 and II-4 (of family 3) indicated that marker *D19S586* defined the telomeric end and marker *D19S906* the centromeric boundary, limiting the maximum genetic region to 2.1 cM (Fig. 1). Affected individuals from families 1–3 shared common alleles for markers *D19S583*, *D19S581* and *D19S584*, suggesting a common founder, but the risk-associated haplotype for family 4 was completely different. Several individuals from the affected families carried the risk haplotype but were clinically unaffected. This might be explained by the relatively late onset of the disease, mostly after 40 years of age, but non-penetrance also has a role. For example, individual

I-1 (of family 3) did not have PCLD at 79 years of age, even though she carries the risk haplotype and has a clinically affected brother as well as affected children (Fig. 1).

The genomic interval contained 78 genes and expressed-sequence tag clusters. In the absence of a clear candidate gene, we decided to carry out exon screening using flanking intronic primers. After sequencing 677 exons representing 94% of exons retained in the genetic region, we detected a heterozygous mutation of adenine to guanine (1138-2A→G) in an affected individual of family 1 at the splice-acceptor site of exon 16 of *PRKCSH* (Fig. 2). *PRKCSH* consists of 18 exons in a 15-kb interval with an open reading frame of 1581 bp. Sequence analysis showed a perfect segregation of this mutation with the disease in all affected members of families 1–3, which was corroborated by *BanI* restriction endonuclease digestion (Fig. 2). A heterozygous change from guanine to cytosine (292+1G→C) at the splice-donor site of intron 4 segregated with all affected individuals from family 4. This was confirmed with *DdeI* digestion (Fig. 2). The mutations were not present in the unaffected family members who did not have the risk haplotype, in the (unrelated) spouses or in a set of 400 ethnically appropriate control alleles, indicating that the sequence changes observed are not common polymorphisms.

RT-PCR analysis showed that the 1138-2A→G mutation results in the inclusion of the first base of intron 15 in the sequence and the introduction of a premature stop codon. The 292+1G→C mutation abolishes the splice-donor site, causing the retention of the first five bases in the sequence and the appearance of a premature stop codon. Both mutations probably result in a shortened and abnormal protein (Fig. 2).

Analysis of northern blots containing mRNA from multiple human tissues showed ubiquitous expression of *PRKCSH* in a variety of organs including liver, heart and pancreas (see Supplementary Fig. 1 online). The gene encodes an acidic protein of 527 amino acids and 59 kDa, which we call hepatocystin. Computational analysis indicated that hepatocystin contains a signal sequence for translocation across the endoplasmic reticulum (ER) membrane, a low-density lipoprotein-receptor domain, two putative Ca²⁺-binding EF domains and a prominent glutamic-acid repeat around the middle of the protein. A His-Asp-Glu-Leu (HDEL) sequence at the C terminus suggests a possible ER localization; this sequence is deleted in mutant hepatocystin (see Supplementary Fig. 2 online). BLASTP analysis indicated a relatively high degree of conservation of hepatocystin in other species, including *Mus musculus* (74% identity, 79% similarity), *Bos taurus* (71% identity, 74% similarity), *Caenorhabditis elegans* (38% identity, 54% similarity) and *Drosophila melanogaster* (32% identity, 43% similarity; see Supplementary Fig. 3 online).

The discovery of a gene associated with PCLD follows discovery of the molecular basis of other cystic disorders, such as ADPKD type 1 (ref. 4) and type 2 (ref. 5) and, recently, of autosomal recessive PKD^{6,7,8}. The product of *PRKCSH* was originally identified in a search for protein-kinase C (PKC) substrate proteins⁹. Previous studies suggested that hepatocystin is not only a substrate for phosphorylation by PKC¹⁰, but that it also interacts with fibroblast-growth-factor receptors as part of a complex with the adaptor protein Grb2 and the Ras guanine nucleotide exchange factor Sos¹¹. Hepatocystin has also been implicated in the cellular response to advance glycation endproducts (AGE) associated with aging and hyperglycemia¹². Binding of AGE to their receptors has been shown to induce various responses, including chemotaxis, cell activation and secretion of cytokines and growth factors¹³. These observations are consistent with a possible role for hepatocystin in the control of cell growth or differentiation. Hepatocystin has also been identified as the non-catalytic β subunit of glucosidase-II, an ER resident enzyme that trims newly synthesized asparagine-linked oligosaccharide chains and participates in ER quality-control processes¹⁴. A role for hepatocystin in the ER is supported by the presence of an N-terminal signal peptide and a C-terminal ER localization signal in the protein. Thus, expression of an

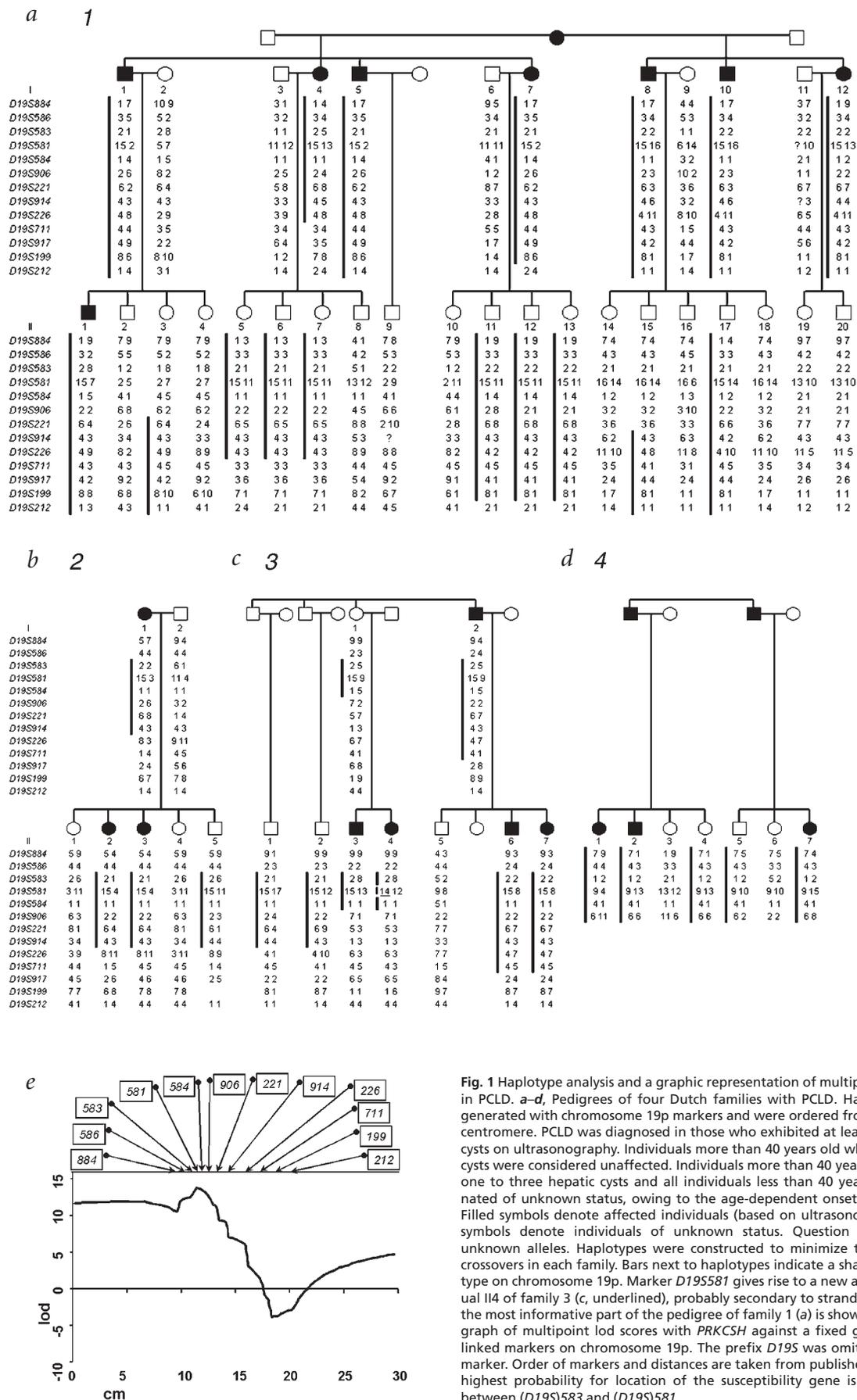
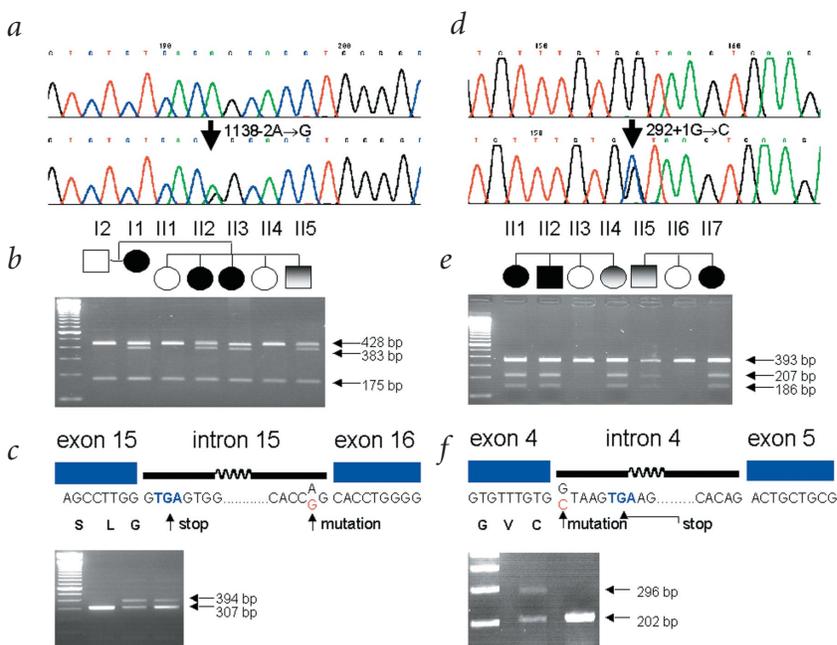


Fig. 1 Haplotype analysis and a graphic representation of multipoint lod scores in PCLD. **a–d**, Pedigrees of four Dutch families with PCLD. Haplotypes were generated with chromosome 19p markers and were ordered from telomere to centromere. PCLD was diagnosed in those who exhibited at least four hepatic cysts on ultrasonography. Individuals more than 40 years old who had no liver cysts were considered unaffected. Individuals more than 40 years old who had one to three hepatic cysts and all individuals less than 40 years were designated of unknown status, owing to the age-dependent onset of liver cysts³. Filled symbols denote affected individuals (based on ultrasonography); open symbols denote individuals of unknown status. Question marks denote unknown alleles. Haplotypes were constructed to minimize the number of crossovers in each family. Bars next to haplotypes indicate a shared risk haplotype on chromosome 19p. Marker *D19S581* gives rise to a new allele in individual III4 of family 3 (**c**, underlined), probably secondary to strand slippage. Only the most informative part of the pedigree of family 1 (**a**) is shown. **e**, FASTMAP graph of multipoint lod scores with *PRKCSH* against a fixed genetic map of linked markers on chromosome 19p. The prefix *D19S* was omitted from each marker. Order of markers and distances are taken from published sources. The highest probability for location of the susceptibility gene is in the region between (*D19S*)583 and (*D19S*)581.

Fig. 2 Analysis of *PRKCSH* mutations associated with PCLD. **a**, Electrophoretogram of heterozygous A→G change in the intron 15 splice-acceptor site compared with the wild-type sequence. **b**, The pedigree structure shows the segregation of the 1138-2A→G mutation in family 2. Digestion with *Ban*I produced fragments of 428, 383, 175 and 45 bp in carriers of the 1138-2A→G mutation. In family 2, individuals I2, I1, I11 and I12 carry the mutated allele. Individual I15 is phenotypically unaffected at age of 34 years, despite carrying the risk haplotype. **c**, RT-PCR analysis of *PRKCSH* RNA of the index individual. Experiments were done with RNA from liver cyst wall (lane 3) and from whole blood leukocytes (lane 4) of the 1138-2A→G carrier and compared with control liver RNA (lane 2). RT-PCR of *PRKCSH* exons 15–16 yielded a 307-bp product in control liver (lane 2), and the mutant generated an extra 394-bp fragment in liver cyst wall (lane 3) and leukocytes (lane 4). The relative expression of the 394-bp fragment is greatest in liver cyst wall. Sequence analysis of the 394-bp RT-PCR fragment showed that the 1138-2A→G mutation retains the first G of intron 15 in the sequence and introduces a premature stop codon immediately thereafter. The amino-acid sequence is given below the genomic sequence. **d**, Electrophoretogram of heterozygous G→C change at the intron 4 splice-donor site compared with the wild-type sequence. **e**, The 292+1G→C mutation creates a new *Dde*I restriction site, and digestion produced fragments of 393, 207 and 186 bp in carriers of the 292+1G→C in family 4. Individual II4 and II5 of this family are examples of non-penetrance, as they are unaffected at ages 35 and 44 years, respectively, but carry the mutation. **f**, RT-PCR of exons 4–5 yielded a 202-bp product in control liver (lane 3), and an extra 296-bp fragment was seen in liver from the 292+1G→C carrier (lane 2). Sequence analysis showed that the 292+1G→C mutation retains the first five bases of intron 5 in the sequence, followed by a stop codon. The corresponding amino-acid sequence is given below the genomic sequence.



abnormal hepatocystin in PCLD could alter the processing of oligosaccharide chains of various secretory and transmembrane proteins, perhaps including receptors that control cell proliferation. Future studies will have to distinguish between these possibilities.

GenBank accession number. *PRKCSH* gene sequence, NM_002743.

Note: Supplementary information is available on the Nature Genetics website.

Acknowledgments

The authors thank the family members for their support and cooperation, Pie Medical for free lending of the mobile ultrasound scanner and H.G. Brunner, M. Boehm and J.C. Zenklusen for advice and inspiring discussions. J.P.H.D. is an Investigator of the Royal Netherlands Academy of

Arts and Sciences and recipient of a grant from the Niels Stensen Foundation and of a travel grant from the Netherlands Organisation of Scientific Research. This project was supported in part by an unrestricted educational grant from Schering-Plough Corporation.

Competing interests statement

The authors declare that they have no competing financial interests.

**Joost P.H. Drenth^{1,2},
Rene H.M. te Morsche¹, Renate Smink¹,
Juan S. Bonifacino² & Jan B.M.J. Jansen¹**

¹The Department of Medicine, Division of Gastroenterology and Hepatology, University Medical Center, St. Radboud, Nijmegen, The Netherlands. ²Cell Biology and Metabolism Branch, National Institute of Child Health and Human Development, Building 18T, Room 101, National Institutes of Health, Bethesda,

Maryland 20892, USA. Correspondence should be addressed to J.P.H.D. (e-mail: JoostPHDrenth@CS.com).

Received 25 November 2002; accepted 21 January 2003.

1. Qian, Q. *et al. Hepatology* **37**, 164–171 (2003).
2. Tahvanainen, P. *et al. J. Hepatol.* **38**, 39–43 (2003).
3. Reynolds, D.M. *et al. Am. J. Hum. Genet.* **67**, 1598–1604 (2000).
4. The European Polycystic Kidney Disease Consortium. *Cell* **77**, 881–894 (1994).
5. Mochizuki, T. *et al. Science* **272**, 1339–1342 (1996).
6. Ward, C.J. *et al. Nat. Genet.* **30**, 259–269 (2002).
7. Xiong, H. *et al. Genomics* **80**, 96–104 (2002).
8. Onuchic, L.F. *et al. Am. J. Hum. Genet.* **70**, 1305–1317 (2002).
9. Hirai, M. & Shimizu, N. *Biochem. J.* **270**, 583–589 (1990).
10. Sakai, K. *et al. Genomics* **5**, 309–315 (1989).
11. Curto, M., Frankel, P., Carrero, A. & Foster, D.A. *Biochem. Biophys. Res. Commun.* **243**, 555–560 (1998).
12. Li, Y.M. *et al. Proc. Natl. Acad. Sci. USA* **93**, 11047–11052 (1996).
13. Thornalley, P.J. *Cell Mol. Biol. (Noisy-le-grand)* **44**, 1013–1023 (1998).
14. Trombetta, E.S., Fleming, K.G. & Helenius, A. *Biochemistry* **40**, 10717–10722 (2001).

