Latest news on the major gene for polycystic kidney disease, *PKD1*

M. H. Breuning and D. J. M. Peters

Department of Genetics, University of Leiden, The Netherlands

The gene responsible for the most common form of polycystic kidney disease, the *PKD1* gene on chromosome 16, was identified a year ago [1], giving a focus to the research into this important cause of renal failure. The first thing to do was to elucidate the primary structure of the gene. With a messenger RNA of an estimated 14000 nucleotides, spread over a piece of chromosome of approximately 50000 base pairs, this was an area where only the large research groups, or strong combinations of groups could compete. The focus on the *PKD1* gene led to the appearance of three papers, each containing important and interesting pieces of the puzzle [2-4].

First, the American PKD Consortium reported the genomic sequence of the *PKD1* gene, and deduced coding and non-coding parts (exons and introns) by comparing the sequence with other known sequences using specialized computer programs. Thus leucine-rich repeats were found in the protein coded by the *PKD1* gene [2]. Such leucine-rich motifs are found outside the cell and engage in protein–protein interactions.

Second, the International PKD Consortium took this analysis further and compared the genomic sequence with that of *PKD1* cDNA. The entire exon–intron structure of the *PKD1* gene was elucidated, revealing 46 exons in a 14.5 kb transcript that codes for a 4304 amino acid protein [3]. This group identified a novel repeated domain characteristic for the *PKD1* gene.

Third, the Oxford group, led by Dr Peter Harris, reported a set of overlapping cDNAs covering the complete transcript of 14148 base pairs [4]. The protein coded by the *PKD1* gene, which was given the name polycystin, was proposed to be a glycoprotein with several transmembrane domains, and a cytoplasmic C-tail. The repeated domains reported by the International Consortium show homology to immunoglobulin domains (Figure 1). All groups agree that polycystin is an integral membrane protein involved in cell–cell and cell–matrix interactions.

The results have also shown that Murphy’s law applies to DNA sequencing. The first reported *PKD1* sequence was found to contain an error. Due to compression in a guanine/cytosine rich stretch, two ‘C’s were missed. The sequence on which all three groups now agree predicts a shorter protein. Antibodies have been described against peptides coded by the very end of the gene first reported [5]. These data have to be regarded with caution since the frameshift resulting from the two extra ‘C’s has removed the peptide used from the predicted protein [6].

*PKD* research is now entering a very exciting phase, the functional analysis of the the protein polycystin. The first results show that one needs to be extremely careful when switching from the relative certainties of cloning and gene mapping to the capriciousness of proteins and antibodies.

**Key**

- signal sequence
- amino flanking region
- LRR
- carboxy flanking region
- putative hinge
- C-type lectin domain
- Ig-like repeat
- FNIII-related domain
- putative TM region

Fig. 1. Proposed model of the *PKD1* protein, polycystin (see key for details). Reprinted with the permission of Nature Genetics, and Dr PC Harris, see reference 4.
References


