European Journal of Human Genetics (2005) 13, 649–659

www.nature.com/ejhg

ARTICLE

Common regulatory elements in the polycystic kidney disease 1 and 2 promoter regions

Irma S Lantinga-van Leeuwen[1], Wouter N Leonhard[1], Hans Dauwerse[1], Hans J Baelde[2], Bernard A van Oost[3], Martijn H Breuning[1] and Dorien JM Peters*[,1]

[1]*Center for Human and Clinical Genetics, Leiden University Medical Center, Leiden, The Netherlands;* [2]*Department of Pathology, Leiden University Medical Center, Leiden, The Netherlands;* [3]*Department of Laboratory Animal Science, Utrecht University, The Netherlands*

The *PKD1* and *PKD2* genes are mutated in patients with autosomal dominant polycystic kidney disease (ADPKD), a systemic disease, with the formation of renal cysts as main clinical feature. The genes are developmentally regulated and aberrant expression of *PKD1* or *PKD2* leads to cystogenesis. To date, however, the transcription factors regulating expression of these genes have hardly been studied. To identify conserved putative transcription factor-binding sites, we cloned and characterized the 5′-flanking regions of the murine and canine *Pkd1* genes and performed a multispecies comparison by including sequences from the human and *Fugu rubripes* orthologues as well as the *Pkd2* promoters from mouse and human. Sequence analysis revealed a variety of conserved putative binding sites for transcription factors and no TATA-box element. Nine elements were conserved in the mammalian *Pkd1* promoters: AP2, E2F, E-Box, EGRF, ETS, MINI, MZF1, SP1, and ZBP-89. Interestingly, six of these elements were also found in the mammalian *Pkd2* promoters. Deletion studies with the mouse *Pkd1* promoter showed that a ~280 bp fragment is capable of driving luciferase reporter gene expression, whereas reporter constructs containing larger fragments of the *Pkd1* promoter showed a lower activity. Furthermore, mutating a potential E2F-binding site within this 280 bp fragment diminished the reporter construct activity, suggesting a role for E2F in regulating cell cycle-dependent expression of the *Pkd1* gene. Our data define a functional promoter region for *Pkd1* and imply that E2F, EGRF, Ets, MZF1, Sp1, and ZBP-89 are potential key regulators of *PKD1* and *PKD2* in mammals.

European Journal of Human Genetics (2005) **13,** 649–659. doi:10.1038/sj.ejhg.5201392
Published online 16 March 2005

Keywords: *PKD1*; *PKD2*; ADPKD; polycystic kidney disease

Introduction

Two genes, *PKD1* and *PKD2*, are mutated in patients with autosomal dominant polycystic kidney disease (ADPKD). This disease is characterized by the formation of multiple fluid-filled cysts in both kidneys, ultimately leading to renal failure and a need for dialysis or a kidney transplant in most patients. Although the kidney is the most severely affected organ, the disease is systemic and extra-renal manifestations such as cysts formation in liver and pancreas, hypertension, and cerebral aneurysms are frequently observed.[1,2] ADPKD has a prevalence of 1 : 1000 and the majority of patients, ~85%, have a mutation in the *PKD1* gene.[3] This gene encodes a large protein, polycystin-1, which forms multiprotein complexes at the cell membrane and is thought to function in cell–cell/

*Correspondence: Dr DJM Peters, Center for Human and Clinical Genetics, Leiden University Medical Center, Wassenaarseweg 72, 2333AL Leiden, The Netherlands. Tel: +31 71 5276048;
Fax: +31 71 5276075; E-mail: d.j.m.peters@lumc.nl
Received 9 March 2004; revised 24 January 2005; accepted 25 January 2005

cell–matrix interactions, signal transduction, and mechanosensation.[4–7] Its C-terminus has been shown to interact with the *PKD2* gene product, polycystin-2,[8–10] which is mutated in ~15% of the ADPKD patients. Polycystin-2 functions as a cation channel, activated by, or in complex with, polycystin-1.[11–14]

In accordance with the systemic nature of the disease, polycystin-1 and polycystin-2 are expressed in many tissues.[15–17] Within the kidney, polycystin-1 and -2 expression is highest in epithelium lining the distal tubules and collecting ducts. Expression of *PKD1* and *-2* is developmentally regulated, with the highest level in late fetal and early neonatal life.[18,19] Mice with homozygous targeted mutations in *Pkd1* or *Pkd2* are embryonic lethal, showing severe renal and pancreatic cysts and cardiovascular abnormalities.[20–22] The interaction between polycystin-1 and -2 is probably needed for maintenance of the epithelial architecture of the kidneys, ductal structures in other tissues, and vascular integrity.[20,21,23,24] Both decreased and increased expression of polycystin-1 may result in cyst formation as has been shown by studies on loss of heterozygosity (LOH) in cysts, immunohistochemical analyses, and knockout and transgenic overexpression studies (reviewed in Igarashi and Somlo[24]). To date, little is known about the transcription factors regulating *PKD1* expression, although putative TCF-binding sites have been reported for the human gene.[25]

Comparative gene analysis is a powerful approach for obtaining insight into gene regulation and function. We therefore cloned and sequenced the *Pkd1* promoter regions of mouse and dog, and compared these sequences to the proximal promoter regions of the human[25] and the *Fugu rubripes* orthologues.[26] We also included *PKD2* sequences[27,28] in our comparative analysis, since this gene has highly overlapping but not identical expression patterns.[17] Functional promoter analyses using reporter assays revealed the highest expression in renal epithelial cells with the most proximal 280 bp promoter sequence. Disruption of the E2F-binding site within this fragment strongly reduced reporter activity.

Materials and methods

Cloning and sequence analysis of the *Pkd1* 5′-flanking regions

The murine *Pkd1* gene was isolated from a 129/SV mouse strain PAC library (RPCI-21) by PCR screening using exon 15 primers (based on GenBank accession no. U70209). One positive clone (427F19) was identified and PCR analysis showed that it contains exon 1, exon 15, and exon 46 sequences. The clone was further characterized by Southern blot analysis, using a selection of restriction enzymes and exon 1, exons 2–5, exons 11–12, and exon 15 probes. The dog genomic BAC library[29] was screened with mouse-specific oligonucleotide probes, directed to a conserved human/mouse region in exon 6, and canine-specific intron 41 oligonucleotide probes, based on the limited available sequence data of dog *Pkd1* (exons 41–42, GenBank accession no. AF027359). Primer sequences are available upon request. Three positive canine clones (A11, D9, H6) were identified and further characterized by Southern blot analysis using a variety of restriction enzymes and probes.

A 4.3 kb *Bgl*II–*Hin*dIII fragment from murine *Pkd1* was cloned and completely sequenced. The sequence was shown to contain ~2.7 kb upstream sequence, exon 1, and part of the first intron of *Pkd1*. In parallel, a 5.8 kb *Hin*dIII DNA fragment from the canine *Pkd1* was cloned, containing ~3 kb 5′-upstream sequence. Sequence analysis of this GC-rich 5.8 kb *Hin*dIII clone proved difficult, therefore 0.7 and 1.3 kb *Bam*HI subclones were sequenced with vector-specific primers, yielding approximately 1 kb upstream sequence.

Consensus sequences of transcription factor-binding sites were identified using MatInspector 5.3 software (www.genomatix.de) using a matrix similarity score of 0.85. Multiple sequence alignments were performed using ClustalW (www.ebi.ac.uk/clustalw).

Construction of the luciferase reporter plasmids

Fragments of the murine *Pkd1* promoter were cloned into the promoterless luciferase expression vector pGL3-Basic (Promega Benelux, Leiden, The Netherlands). The 1.2 kb-pGL3 construct was generated by inserting a 1.2 kb *Sac*I-*Nae*I *Pkd1*-fragment (−1052 to +157) into the *Sac*I/*Sma*I sites of the multiple cloning site of pGL3-Basic. The 0.7 kb-pGL3 (−531/+157) and 0.3 kb-pGL3 (−127/+157) constructs were derived from the 1.2 kb-pGL3 construct by digesting with *Sac*I/*Spe*I and *Sac*I/*Stu*I, respectively, followed by self-ligation. The 2.3 kb-pGL3 construct (−2141/+157) was obtained by inserting a 1.1 kb *Sac*I fragment (−2141 to −1053) in the same orientation into the *Sac*I site of the 1.2 kb-pGL3 vector. All reporter constructs were verified by restriction enzyme digestion and DNA sequence analysis. Base pair positions in the murine sequence are indicated by assigning the position corresponding to the human transcriptional initiation site at +1 (Figure 1).

Mutagenesis of the putative E2F site in the 0.3 kb-pGL3 construct was performed by PCR using the forward primer 5′-GCTGGGTCTGCAGTGCAG**AT**CGAATGCGCGAGCAG-3′, containing the mutation (shown in bold-face) and an immediate upstream located *Pst*I restriction site (underlined), and the reverse primer 5′-TCCAGCGGTTC-CATCTTCCA-3′ (located in the pGL3 vector). The PCR product was digested with *Pst*I and *Hin*dIII and subsequently cloned into the wild-type 0.3 kb-pGL3 construct cut with *Pst*I and *Hin*dIII. The nucleotide sequence of the mutated 0.3 kb-pGL3 construct, mutE2F-pGL3, was confirmed by sequencing.

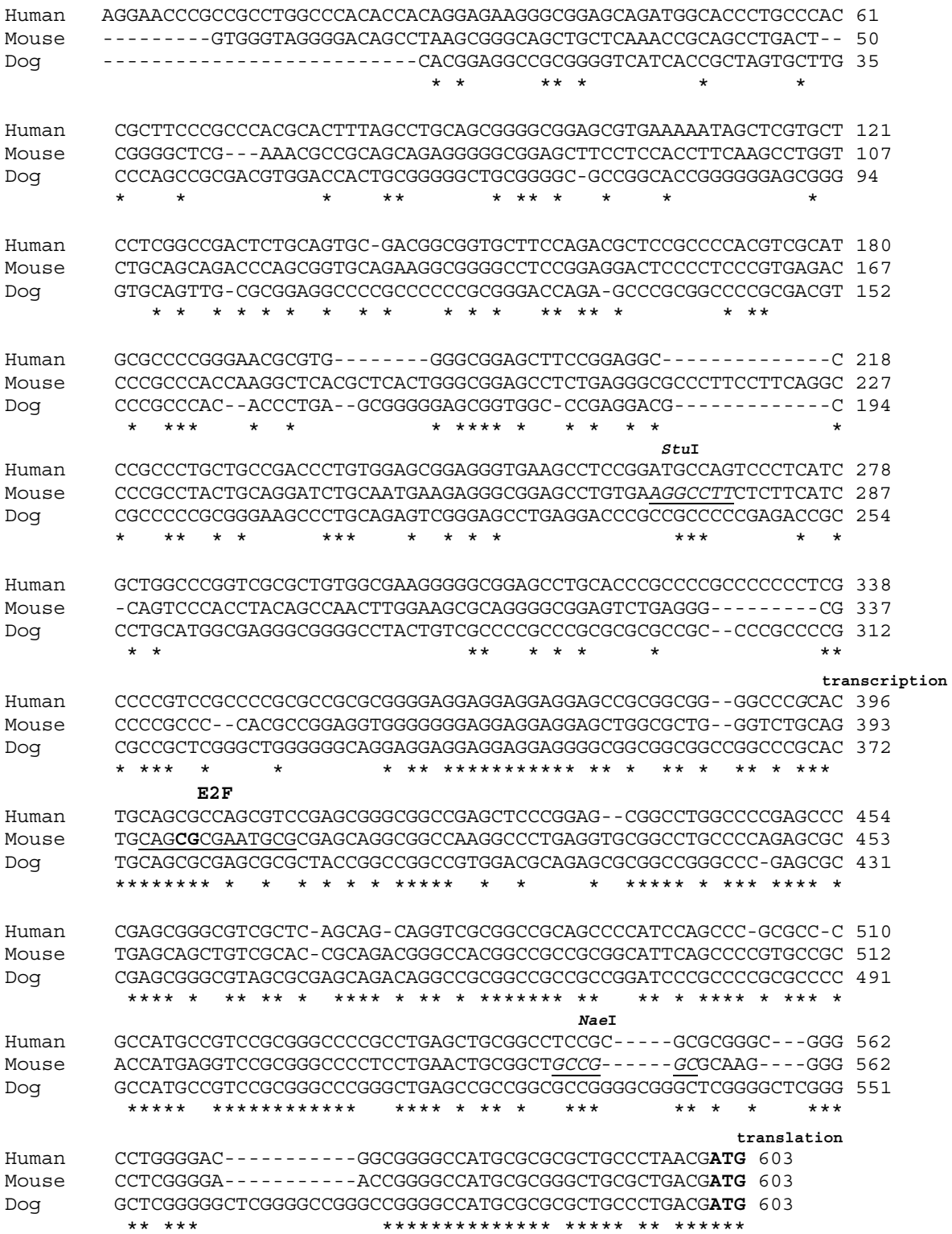

Figure 1 Alignment of the *Pkd1* 5′-flanking sequences of the mouse (Accession no. AY332760), dog (Accession no. AY332761), and human (Accession no. L39891). The transcription start site (+1) for the human *PKD1* is indicated in italics. The translation start sites (ATG start codon) are indicated in bold face. Restriction sites of *Stu*I and *Nae*I, used to generate deletion constructs, are underlined and italicized. Mutations generated in the DNA consensus-binding site for E2F (underlined) are in bold face. Conserved nucleotides between the three sequences are indicated by asterisks.

Cell culture, transfection and luciferase assays

The renal cell lines MDCK (dog, distal tubule), mIMCD3 (mouse, collecting duct), NRK-52E (rat, proximal tubule), LLC-PK1 (pig, proximal tubule) and the human embryonic kidney HEK293, and mouse hepatocyte-like cell line mhAT3 were grown in Dulbecco's modified Eagle's medium F12 (with glutamax) supplemented with 10% fetal bovine serum, 1 mM sodium pyrovate, 10 mM HEPES, and 100 U/ml penicillin/streptomycin in humidified 5% CO_2. The human colon carcinoma cell line HCT-E8 was cultured in RPMI 1640 with similar supplements and 2 mM glutamax. HEK293 medium also contained 4.5 g/l glucose. All media and supplements were obtained from Gibco Invitrogen Corporation (Breda, The Netherlands).

We transfected promoter constructs (500–800 ng) and 5 ng DNA from *renilla reniformis* plasmids using Transfast reagent (Promega) or Fugene (Roche Diagnostics, Almere, The Netherlands; only HEK293), as recommended by the manufacturers. After overnight incubation, cells were lysed in lysis buffer (Promega) and luciferase activity was measured in a luminometer (Lumat LB 9507). The data was normalized for the transfection efficiency by using the Dual Luciferase reporter assay system (Promega) according to the manufacturer's instructions. The activity of the promoterless vector, pGL3-Basic, which contained no insert, was measured to determine background activity. All transfections experiments were performed in triplicate and repeated at least three times, except for the transfections with the mutE2F-pGL3 construct which were performed twice, in independent experiments.

RNA analysis

RNA was isolated from cultured MDCK and mIMCD3 cells using Trizol reagent (Invitrogen Life Technologies, Breda, The Netherlands) and cDNA was synthesized with SuperScript II (RNase H) reverse transcriptase (GibcoBRL, Breda, The Netherlands). Real-time *Pkd1* PCR analysis of four RNA samples from each cell line, obtained in two independent experiments, was performed using TaqMan technology as published previously using the primers and a Taqman probe located in exons 45 and 46 of the *Pkd1* gene (*Pkd1* 3′PCR).[30] The expression of the housekeeping gene hypoxanthine guanine phosphoribosyl transferase (HPRT) served as reference for gene expression.

Results

Cloning of the *Pkd1* promoter regions of the mouse and dog

From a mouse genomic PAC and a canine genomic BAC library, we isolated one murine and three canine clones covering exon 1 and 5′ flanking regions of the *Pkd1* genes. A 4.3 kb *Bgl*II–*Hin*dIII mouse fragment and a 5.8 kb *Hin*dIII canine fragment containing the 5′-upstream regions were subcloned and sequenced. We obtained approximately 2.9 kb sequence upstream of the murine and 1.3 kb sequence upstream of the canine ATG start codon (GenBank Accession Nos. AY332760 and AY332761).

Sequence analysis of the *Pkd1* and *Pkd2* promoter regions

Sequence analysis revealed that both the murine and canine *Pkd1* promoter regions are GC-rich and lack a TATA or CAAT box, as has been reported for human *PKD1*.[25] Alignment of the human, murine and canine sequences (Accession nos. L39891, AY332760, and AY332761, respectively) showed that an approximately 300 bp fragment was considerably conserved between the three species, showing 66% sequence identity (Figure 1). In particular, the $\sim$30 bp sequences in the immediate vicinity of the translational start site are highly conserved (Figure 1). This region of the *Pkd1* promoter contains a putative E-box-binding site. Previous attempts to experimentally identify the murine transcriptional start site were unsuccessful, probably because of the GC-rich composition of the region.[31] Therefore, base pair positions in the murine and canine sequences are indicated with respect to the ATG start codon, or by assigning the position corresponding to the human transcriptional initiation site at $+1$ (Figure 1).

Computer predictions of the 600 bp region upstream of the ATG start codon suggested the presence of multiple binding sites for transcription factors (MatInspector, Genomatix). Comparative analysis of the human, murine, and canine proximal promoter regions revealed conserved putative binding sites for AP2, E2F, E-box factors, early growth response factors (EGRF), and one or more members of the family of Ets factors (Ets1 and GABP in mouse and human, Elk-1 in dog). MINI, MZF1, and ZBP-89 elements, as well as binding sites for the ubiquitous transcription factor Sp1 were also conserved (Figure 2).

Some sites like the most proximal E-box are present at similar positions within the *Pkd1* promoter sequences while others are not, as depicted in Figure 2.

When we included the *Pkd1* orthologue of the pufferfish *Fugu rubripes* (Accession no. AF013614) in our sequence comparison, we found that the three mammals and *Fugu* share potential binding sites for E2F, E-box, and Ets factors, MZF1 and ZBP-89 (Table 1), indicating that they are strong candidate regulators of *Pkd1*. The numerous Sp1 sites present in the mammalian species were not found in the *Fugu Pkd1* promoter.

The almost indistinguishable phenotype resulting from mutations in either *PKD1* or *PKD2* suggests that polycystin-1 and -2 share the same signaling pathway or interact to form a functional unit. Comparison with the previously reported human *PKD2* (Accession no. AF243476) and murine *Pkd2* (Accession no. AF242389) proximal promoter sequences revealed that most of the conserved mammalian/*Fugu* putative *Pkd1*-promoter elements as well as Sp1 sites are also present in *Pkd2* (Table 1). The AP2 site

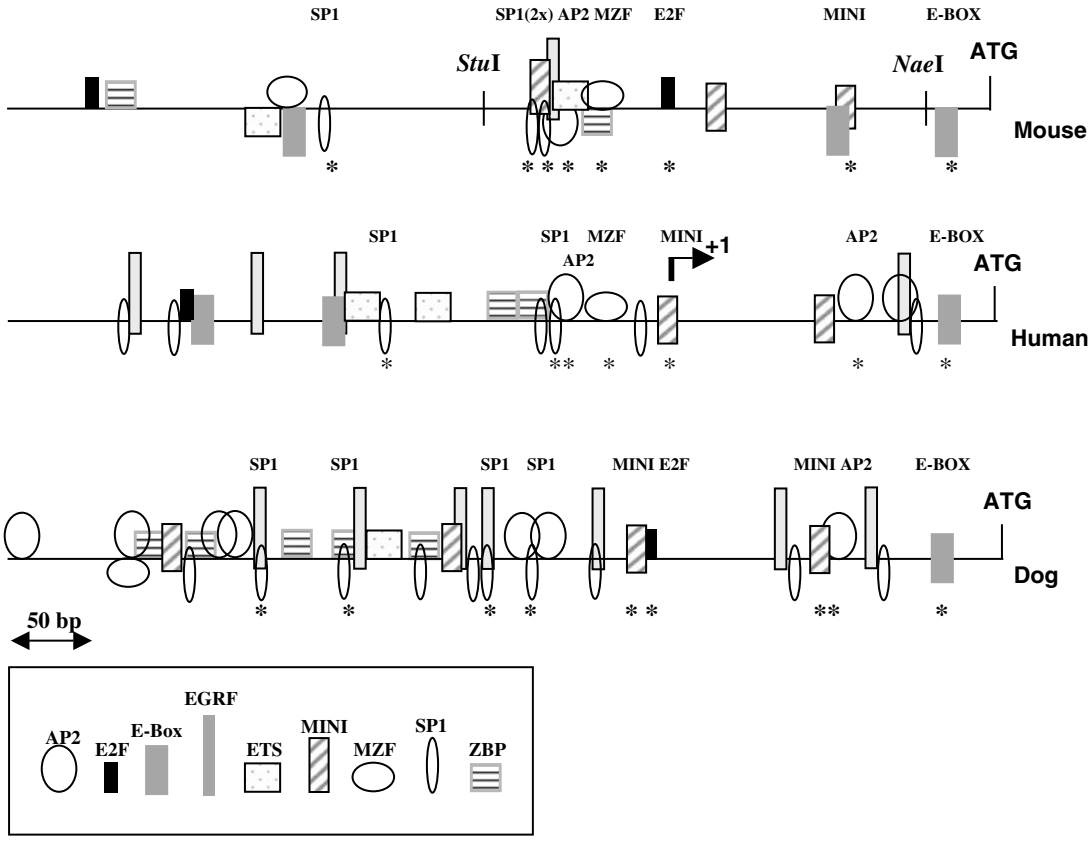

Figure 2 Organization of the murine, human and canine proximal promoter regions. Potential transcription factor binding sites present in the 600-bp region upstream of the translation start site (ATG) are indicated. Sites present at similar positions (according to the sequence-alignment in Figure 1) in at least two of the species are highlighted by asterisks. The transcription start site for human *PKD1* is indicated with an arrow. The restriction sites *Stu*I and *Nae*I used to construct deletion mutants are indicated.

Table 1 Conserved regulatory elements in the proximal promoter of *Pkd1* and *Pkd2*

Putative binding site	*Present in Pkd1: human, mouse, dog*	*+Fugu*	*Present in Pkd2: human, mouse*
AP2	x		
E2F	x	x	x
E-Box	x	x	
EGRF	x		x
ETS	x	x	x
MINI	x		
MZF1	x	x	x
SP1	x		x
ZBP-89	x	x	x

Putative binding sites for transcription factor families conserved in the proximal promoter of the human, mouse and dog *Pkd1* genes are indicated, and their presence or absence in corresponding promoter sequences of *Fugu Pkd1* (Accession no. AF013614) and human/mouse *Pkd2* (Accession nos. AF243476 and AF242389). For detailed information about the promoter analysis and the consensus binding sites, see the website of Genomatix (genomatix.gsf.de). Abbreviations: AP2, activator protein 2; E-Box, E-box-binding factor; EGRF, early growth response factor; ETS, Ets factor; MINI, muscle initiator sequence; MZF1, myeloid zinc-finger 1; SP1, stimulating protein 1, ZBP; zinc-binding protein.

reported previously in the mouse *Pkd2* promoter and not in human *PKD2*[28] was not identified with our search criteria.

Analysis of the distal promoter region of the murine and canine *Pkd1* genes showed no overall homology with human. However, a human 109 bp sequence containing a T-cell factor (TCF) site, and implicated in the β-catenin responsiveness of the human *PKD1* promoter[25] showed a striking homology with a murine sequence at 1.8 kb distance of the murine ATG start codon (Figure 3). This murine sequence showed an imperfect match to the TCF consensus sequence (CTTTGWW). The MatInspector program predicted TCF sites at positions 1.0, 2.8 and 2.9 kb upstream the murine ATG start codon, and one most distantly in the available canine sequence (1300 bp). In the conserved T-rich sequence flanking the TCF-like site in the murine sequence, three putative *fork head* homolog 3 (HFH-3) sites (core sequence TRTTTRT) were present (Figure 3).

Deletion analysis of the murine *Pkd1* promoter

The functional activity of the murine *Pkd1* promoter was analyzed by transient transfection experiments using

```
                                                                  TCF
Human   GATCTCTGGCACATTTTATTTGCTCTGTCTCAC----CACATGGATTTTGTTTTTTTGTTTCTTTGTTTTTTGAGATGGA 76
Mouse   GAGTACTGGCCTGAGAAGCGCATGCTACCATGCATGCTCTGTGAGTTGGTTTTGTTTGTTTGTTTGTTTTTCGAGACAGG 1140
        **   *****             **  *   *        **  **   *** ******* ********* ****  *

Human   GTCTCACTCTTGTTGCCCAGGCTGGAGTGCCAT--------------------------- 109
Mouse   GTTTCTCTGT-GTAGCCCTGGCTTTCCTGGAACTCACTCTGTAGACCAGGCTGGCCTTGGA 1200
        ** ** ** * ** **** ****    **  *
```

Figure 3 Sequence comparison of a human 109-bp PKD1 promoter sequence, mediating β-catenin activation,[25] and the murine *Pkd1* distal promoter. The human 109 bp promoter fragment, located 2.1 kb upstream the human ATG codon shows high homology with a murine sequence located 1.8 kb upstream the ATG codon. (Human *PKD1*: Accession no. L39891, bp 1422–1530; ATG = 3648–3650; mouse *Pkd1*: Accession no. Y332760, bp 1061–1172; ATG = 2969–2971). The putative TCF site in the human sequence (boxed) overlaps with several putative HFH3 sites in the murine sequence (underlined). A 28 bp sequence deletion that was reported to abolish the β-catenin induction of a human *PKD1* promoter construct is presented in bold.

luciferase reporter gene constructs. A series of deletion constructs, containing 0.3, 0.7, 1.2, and 2.3 kb murine promoter fragments cloned into the pGL3-Basic vector upstream of the luciferase reporter gene, were transiently transfected into a variety of cell lines. We used renal epithelial cell lines derived from different nephronic segments, namely proximal tubule (NRK-52E, LLC-PK1), distal tubule (MDCK) and collecting duct (mIMCD), HEK293 (human embryonic kidney) cells, and hepatocyte-like mhAT3 and colon carcinoma HCT-E8 cells (Figure 4).

The 300 bp fragment (-127 to $+157$) stimulated luciferase activity in all cell lines, showing high levels of activity in the renal cell lines and lower activity in the colon and liver cell lines (Figure 4). In contrast, a well-characterized 1.3 kb promoter region of the kidney-specific Ksp-Cadherin gene[32] showed promoter activity only in the renal epithelial cell lines, and not in the colon, liver, and HEK293 cells (not shown). To correlate reporter gene expression to endogenous *Pkd1* expression, we applied real-time PCR to determine *Pkd1* expression in the mouse-derived cell lines mIMCD3 and mhAT3. These two cell lines were selected since the other cell lines transfected with the *Pkd1* reporter constructs were all derived from different species. We used *HPRT* expression as a reference. *Pkd1* expression in the renal cell line mIMCD3 was approximately 2.5-fold higher compared to the hepatocyte-like mhAT3 cells (1.0 *versus* 2.6), showing that in these cell lines the luciferase assay results (63 *versus* 12) are in line with the endogenous *Pkd1* expression (Figure 4c).

A reduced stimulation, up to 50%, was observed following transfection of the 0.7 kb-pGL3 construct into renal epithelial cell lines. This region of the murine *Pkd1* promoter contained one each of the two putative E2F, MZF, ZBP, and Ets1-binding sites, and one of the three E-box elements. A further reduction in expression levels was found with the 1.2 kb-pGL3 construct, suggesting the presence of an additional inhibitory element(s) in the 1.2–0.7 kb fragment. The 2.3 kb-pGL3 construct showed stimulation levels rather similar to the 1.2 kb-pGL3 construct (Figure 4d).

Since a variety of data suggest cell cycle-dependent expression of *Pkd1*, we further analyzed the putative E2F-binding site.[33–35] Therefore, we mutated the E2F3 site by replacing the conserved CG by AT (see also Figure 1, Table 2). This markedly decreased the promoter activity. The transfected MDCK and mIMCD3 cell lines showed 40–50% activity compared to the wild-type 0.3 kb-pGL3 construct (Figure 4e). According to the MatInspector program, this two base pair change did not create a new transcription factor-binding site.

Discussion

Comparative genome analysis offers a powerful method for detecting functional regulatory sequences. Since transcrip-

Figure 4 *In vitro* characterization of the transcriptional activity of the murine *Pkd1* promoter. (**a**) Structure of the *Pkd1* reporter gene fragments. ATG, translational start site. (**b**, **d**) Characterization of promoter activity in cell lines. MDCK, LLC-PK1, NRK-52E, mIMCD-3, HEK293, HCT-E8, and mhAT3 cells were transiently transfected with each of the indicated reporter constructs. Luciferase activity was measured 24 h after transfection. Cells were cotransfected with *renilla reniformis* luciferase to control for transfection efficiency. Normalized luciferase activity is shown relative to the pGL3-Basic plasmid (**b**) or relative to the plasmid containing 0.3 kb of 5′-flanking region (**d**). The results represent the mean $\pm$ SE of three separate transfections, each performed in triplicate. (**c**) Real-time PCR analysis on RNA samples of the renal epithelial cell line mIMCD3 and hepatocyte-like cell line mhAT3. *Pkd1* expression was normalized to expression of the hypoxanthine guanine phosphoribosyl transferase (HPRT) gene that served as reference. Measurements, in duplicate, were performed on four RNA samples, obtained in two independent experiments. The mean *Pkd1* expression of the mIMCD3 cell line has been set to 100%. Error bars represent standard error. (**e**) Site-directed mutagenesis of the mouse *Pkd1* promoter. A mouse *Pkd1* reporter construct (E2Fmut) containing a double nucleotide change in a putative E2F site was transfected into MDCK and mIMCD3 cells. Activity of the wild-type 0.3 kb-pGL3 construct was set at 100%. The results represent the mean $\pm$ SE of three separate transfections, each performed in duplicate.

tion factor-binding sites are usually small in size and highly degenerate, database searches used to identify TF-binding sites in regulatory regions invariably identify a large number of false sites.[36] A multispecies comparison allows the identification of sites that are of functional importance, as these sites are under selective pressure.[37,38]

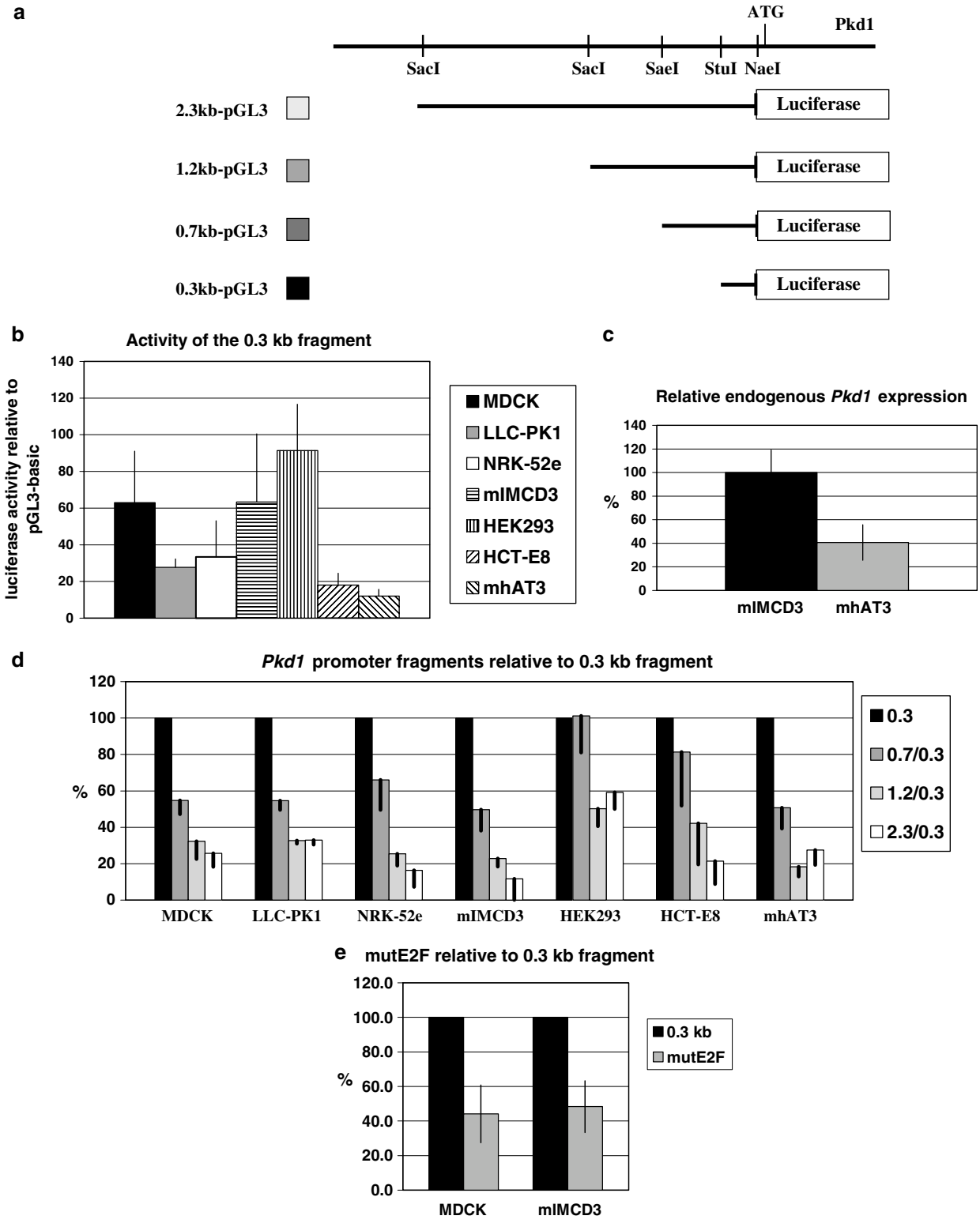

Table 2 Location and sequence motifs of the elements common to the upper (+) or lower (−) strand of the mammalian *Pkd1* promoter sequences

	Human		*Mouse*		*Dog*	
	Position	*Sequence (5′–3′)*	*Position*	*Sequence (5′–3′)*	*Position*	*Sequence (5′–3′)*
AP2	346–357 (+)	cgccccgcgccg	341–352 (+)	cgccccacgccg	7–18 (−)	gaccccgcggcc
	518–529 (−)	ggcccgcggacg			79–90 (−)	ctccccccggtg
	549–560 (−)	cgcccgcgcgcg			118–129 (−)	gtcccgcggggg
					133–144 (+)	agcccgcggccc
					288–299 (+)	cgcccgcgcgcg
					306–317 (+)	cgccccgcgccg
					499–510 (−)	ggcccgcggacg
E2F	101–108 (+)	gcgtgaaa	55–62 (+)	gctcgaaa	375–387 (+)	cagcgcgagcgcg
			396–408 (+)	cagcgcgaatgcg		
EBOX	111–122 (+)	tagctcgtgctc	156–165 (−)	ctcacgggag	578–587 (+)	gccatgcgcg
	191–200 (−)	cccacgcgtt	501–510 (−)	ggcacggggc		
	578–587 (+)	gccatgcgcg	578–587 (+)	gccatgcgcg		
EGRF	68–79 (−)	gtgcgtgggcgg	341–352 (−)	cggcgtggggcg	151–163 (−)	gtgtgggcgggac
	139–150 (+)	gtgcgacggcgg			194–205 (−)	ccgcgggggcgg
	192–203 (+)	acgcgtggggcg			261–272 (+)	tggcgagggcgg
	550–561 (+)	gcgcgcgggcgg			287–298 (−)	gcgcgcgggcgg
					350–361 (+)	gggcggcggcgg
					465–476 (−)	atccggcggcgg
					520–531 (−)	cggcgccggcgg
ETS	60–71 (−)	gcgggaagcggt	138–147 (−)	tccggaggcc	199–212 (+)	cccgcgggaagccc
	205–214 (−)	tccggaagct	349–358 (+)	gccggaggtg		
	258–267 (+)	tccggatgcc				
MINI	385–405 (+)	cggggcccgcactgcagcgcc	337–357 (+)	gccccgccccacgccggagg	86–106 (−)	gcgcaactgcaccccgctccc
	486–506 (+)	ccgcagccccatccagcccgc	424–444 (−)	ggcaggccgcacctcagggc	251–271 (−)	cgccctcgccatgcagggcgg
			507–527 (−)	cgcggacctcatggtgcggc	361–381 (+)	gccggcccgcactgcagcgcg
					486–506 (−)	cgcggacggcatggcggggcg
MZF	358–365 (+)	cgcgggga	152–159 (−)	ggagggga	82–89 (+)	cgggggga
			358–365 (+)	gggggggа		
SP1	64–76 (−)	cgtgggcgggaag	165–177 (−)	ggtgggcggggtc	110–122 (−)	gggggcggggcc
	91–103 (+)	gcggggcggagcg	317–329 (+)	caggggcggagtc	150–162 (−)	tgtgggcgggacg
	215–227 (−)	gcagggcggggcc	336–348 (−)	gtggggcggggcg	190–201 (−)	cgggggcggcgtc
	322–334 (−)	gggggggcggggcg			235–247 (−)	cgggggcggcggg
	341–353 (−)	gcggggcggacgg			264–276 (+)	cgagggcggggcc
	379–391 (+)	ccgcggcggggcc			283–295 (−)	cgcgggcggggcg
	553–565 (+)	cgcgggcgggcct			301–313 (−)	gcggggcggggcg
					347–359 (+)	gaggggcggcggc
					474–486 (−)	gcggggcgggatc
					529–541 (+)	ccggggcgggctc
ZBP	302–313 (−)	ctccgccccctt	73–84 (−)	ctccgccccctc	82–93 (−)	ccgctccccccg
	324–335 (+)	ccccgccccccc	358–369 (−)	ctcctccccccc	112–123 (+)	ccccgccccccg
					169–180 (−)	ccgctccccgc
					192–203 (+)	cgccgccccgc
					237–248 (+)	cgccgccccga

Positions correspond to Figure 1.

Therefore, we analyzed the *Pkd1* promoter regions of several mammalian species. In the proximal 5′-flanking region of approximately 600 bp, we identified nine elements that were conserved in the mammalian *Pkd1* promoters. When we also included the *Pkd1* promoter of the pufferfish *Fugu Rubripes*, five of these elements, E2F, E-Box, ETS, MZF1, and ZBP-89, were found conserved. These factors are therefore likely to be essential regulators of *Pkd1* gene regulation. Binding sites for the transcription factors AP2, EGRF, MINI, and Sp1, which are conserved in human, mouse, and dog *Pkd1* but not in fish, suggest a regulatory role for these factors in functions specific for mammals.

Besides the finding of putative binding sites for a transcription factor in multiple *Pkd1* gene promoters, conserved positions within these promoter sequences makes it even more likely that a site is a functional binding site. However, sites at other positions may also modulate the promoter activity and putative binding sites for different regulatory factors are sometimes overlapping. Further studies are required to determine if indeed sites bind regulatory factors. The dinucleotide that was mutated in this study, and clearly diminished promoter reporter activity, was conserved in mouse, human, and dog and recognized as a putative E2F-binding factor in mouse and dog.

PKD1 and *PKD2* patients have essentially a similar phenotype, although PKD2 patients have a milder progression towards end-stage renal failure.[39] In addition, expression of the proteins polycystin-1 and polycystin-2 is highly coordinated and both proteins can interact.[14] Therefore, we were interested to identify elements common to the promoters of both genes. Four elements were found in the mammalian *Pkd2* promoter, mammalian *Pkd1* promoter, and in the *Pkd1* promoter of the pufferfish *Fugu Rubripes*: E2F, ZBP-89, ETS, and MZF1. These factors are therefore likely to be crucial regulators of *Pkd1* and *Pkd2* expression. Site-directed mutagenesis of the E2F site resulted in a diminished reporter activity, underscoring the relevance of this site. The presence of a putative E2F site suggests a cell cycle-dependent expression, and is supportive for experimental data showing downregulation of polycystin-1 in differentiating cells, and may be correlated to the increased expression of polycystin-1 and -2 upon ischaemic injury in kidneys.[33–35]

Also ZBP-89, a Kruppel-type zinc-finger protein, is implicated in cell proliferation.[40] ZBP-89 binds GC-rich sequences and has been shown to be either a repressor or an activator of gene expression, depending on the promoter. It modulates cell proliferation partly by forming complexes with transcription (co)-factors like Sp1, p300, and p53.[40]

The ETS gene family contains a diverse group of helix-turn-helix transcription factors that control cell proliferation, differentiation, and survival. Ets1 is predominantly expressed in endothelial cells or their precursors and is essential for embryonic blood vessel development and angiogenesis. The finding of Ets1 elements in the promoters of *Pkd1* and *Pkd2* is of particular interest since the *Pkd1* promoter has been shown to direct the expression of a reporter gene in the endothelium of a targeted *Pkd1* mouse model.[22] Moreover, ADPKD patients have an increased risk to develop cerebral aneurisms and several *Pkd1* and *Pkd2* mouse models show cardiovascular abnormalities.[22,41] It is therefore reasonable to speculate that ETS factors play a role in regulating the endothelium-specific expression of the *Pkd1* and *Pkd2* genes.

Like ZBP-89, MZF-1 belongs to the Kruppel family of zinc-finger proteins. The myeloid zinc-finger 1 (MZF-1) is expressed in totipotent hemopoietic cells as well as in myeloid progenitors but not in fully differentiated blood cells.[42] The conserved MZF-1-binding site suggests expression of the *Pkd* genes in hemotopoietic lineages. The role of polycystin-1 and -2 in hemopoietic cells is not well characterized, although one study demonstrated polycystin-1 expression in a human erythroid and macrophage cell line.[33]

Binding sites for AP2, MINI, and E-box elements, identified in all three mammalian *Pkd1* promoters, were not found in human/mouse *Pkd2*. In contrast, GATA-1 sites, present in both human and murine *Pkd2* proximal promoter regions, were not identified in the proximal *Pkd1* promoter sequences, whereas the *Pkd2* conserved H4TF-1 sites and NF-1 sites[28] were only identified in the distal *Pkd1* promoter of the mouse. These transcription factors may govern distinct expression of *PKD1* and *PKD2*, which may underly the exhibited differences in expression patterns for polycystin-1 and polycystin-2.

In TATAA-less promoters, the transcription factor Sp1, for which several binding sites are present in the mammalian *Pkd1* and *Pkd2* gene promoters, is thought to be involved in transcription initiation.[43] Although generally considered to be ubiquitous, Sp1 has also been implicated in regulation of tissue-specific gene expression. Evenmore, Sp1 expression is temporally and spatially regulated during nephrogenesis. Interestingly, Sp1 was reported as a regulator of the Wilms' tumor (WT1) and Von Hippel–Lindau (VHL) tumor suppressor genes[44,45] that are involved in renal cancers.

It has recently become clear that both *PKD* genes are expressed in the primary cilia of cells and that targeted mutations in the *Pkd* genes results in ciliary defects.[7] A key promoter element present in a group of genes expressed in ciliated cells, the X-box,[46] was not evident in our comparative analysis of the proximal promoter. An extended search, however, showed that putative X-box-binding sites were present within 2 kb distance of the ATG start codons in the *Pkd1* genes and also in the distal promoters of the *Pkd2* genes. Likewise, putative binding sites for *Cut* homeobox homologues were found in the distal promoters of the *Pkd1* and *Pkd2* genes. The murine *Cut* homolog, *Cux-1,* is highly expressed in the developing kidney and cyst epithelium of polycystic kidneys from C57BL/6J$^{cpk/cpk}$ mice and regulates cell proliferation.[47]

Reporter constructs revealed that a 280 bp fragment of the murine *Pkd1* promoter was capable of directing reporter gene expression *in vitro*. As expected from the systemic nature of ADPKD, the promoter constructs are active in renal and nonrenal epithelial cell lines, with the highest expression in renal cells. In order to correlate luciferase activity to *PKD1* expression, we performed real-time PCR on two murine cell lines that showed different

levels of luciferase activity. These experiments confirmed that *Pkd1* expression was higher in the renal epithelial mIMCD3 cells compared to the mhAT3 hepatoma cell line. Deletion mutant analysis showed reduced activities of the larger constructs, implying the presence of (multiple) inhibitory elements. Our data are in agreement with the study of the human *PKD1* promoter that showed reduced activity with a 0.95 kb fragment compared to the 0.2 kb proximal promoter. In contrast to the mouse, the human *PKD1* promoter contains several large Alu repeats, located upstream 1 kb distance of the start codon, making it difficult to compare the human and mouse distal promoter analyses. Nevertheless, we identified TCF sites in the murine and canine distal promoter sequences (not shown), supporting the suggestion by Rodova *et al*[25] that *PKD1* may be a target of the β-catenin/TCF pathway. In addition, we found that the human 109 bp sequence, mediating the β-catenin induction of their human 2.0 kb-PKD1 promoter construct, was highly homologous to a murine sequence. The homology was most prominent in a 28 bp stretch where deletion was needed to abolish the induction, whereas deletion of the TCF site only was not effective. Three putative hepatocyte nuclear factor-3 (HNF-3)/HFH sites were in this fragment of the murine *Pkd1* sequence but not identified in the human promoter. However, these data suggest that HFH-3 or another factor binding to this highly conserved T-rich sequence may partly mediate the β-catenin responsiveness of the *PKD1* promoter. HFH-3 is expressed in the distal tubules of embryonic and adult mouse kidney, and binding sites for this transcriptional activator have been identified in promoters of genes expressed in the epithelium of the renal distal tubules. These include anion exchanger AE1, E-Cadherin, Vasopressin type 2 receptor, and Ksp-cadherin gene promoter.[48]

In conclusion, we isolated and characterized the 5′-flanking region of the murine and canine *Pkd1* genes. A multispecies comparison allowed us to identify conserved binding sites for several transcription factors that are likely to mediate the developmentally and cell-type-specific expression of *PKD1* and *PKD2*, and may be of help in understanding the pathogenesis of ADPKD.

Acknowledgements

We thank Lia Spruit and Serge Versteeg for technical assistance and Stefan White for reading the manuscript. This work was supported by a grant from the Dutch Kidney Foundation (98.1745).

References

1 Ravine D, Walker RG, Gibson RN *et al*: Phenotype and genotype heterogeneity in autosomal dominant polycystic kidney disease. *Lancet* 1992; **340**: 1330–1333.
2 Gabow PA: Autosomal dominant polycystic kidney disease: more than a renal disease. *Am J Kidney Dis* 1990; **16**: 403–413.
3 Peters DJM, Sandkuijl LA: Genetic heterogeneity of polycystic kidney disease in Europe; In: Breuning MH, Devoto M, Romeo G (eds) Contributions to nephrology, Vol. 97: Polycystic kidney disease Basel: Karger, 1992, pp 128–139.
4 Arnould T, Kim E, Tsiokas L *et al*: The polycystic kidney disease 1 gene product mediates protein kinase C alpha-dependent and c-Jun N-terminal kinase-dependent activation of the transcription factor AP-1. *J Biol Chem* 1998; **273**: 6013–6018.
5 Parnell SC, Magenheimer BS, Maser RL, Zien CA, Frischauf AM, Calvet JP: Polycystin-1 activation of c-Jun-N-terminal kinase and AP-1 is mediated by heterotrimeric G proteins. *J Biol Chem* 2002; **277**: 19566–19572.
6 Bhunia AK, Piontek K, Boletta A *et al*: PKD1 induces p21(waf1) and regulation of the cell cycle via direct activation of the JAK-STAT signaling pathway in a process requiring PKD2. *Cell* 2002; **109**: 157–168.
7 Nauli SM, Alenghat FJ, Luo Y *et al*: Polycystins 1 and 2 mediate mechanosensation in the primary cilium of kidney cells. *Nat Genet* 2003; **33**: 129–137.
8 Tsiokas L, Kim E, Arnould T, Sukhatme VP, Walz G: Homo- and heterodimeric interactions between the gene products of PKD1 and PKD2. *Proc Natl Acad Sci USA* 1997; **94**: 6965–6970.
9 Qian F, Germino FJ, Cai Y, Zhang X, Somlo S, Germino GG: PKD1 interacts with PKD2 through a probable coiled-coil domain. *Nat Genet* 1997; **16**: 179–183.
10 Newby LJ, Streets AJ, Zhao Y, Harris PC, Ward CJ, Ong AC: Identification, characterization, and localization of a novel kidney polycystin-1-polycystin-2 complex. *J Biol Chem* 2002; **277**: 20763–20773.
11 Gonzalez-Perrett S, Kim K, Ibarra C *et al*: Polycystin-2, the protein mutated in autosomal dominant polycystic kidney disease (ADPKD), is a Ca^{2+}-permeable nonselective cation channel. *Proc Natl Acad Sci USA* 2000; **98**: 1182–1187.
12 Vassilev PM, Guo L, Chen XZ *et al*: Polycystin-2 is a novel cation channel implicated in defective intracellular Ca(2+) homeostasis in polycystic kidney disease. *Biochem Biophys Res Commun* 2001; **282**: 341–350.
13 Koulen P, Cai Y, Geng L *et al*: Polycystin-2 is an intracellular calcium release channel. *Nat Cell Biol* 2002; **4**: 191–197.
14 Hanaoka K, Qian F, Boletta A *et al*: Co-assembly of polycystin-1 and -2 produces unique cation-permeable currents. *Nature* 2000; **408**: 990–994.
15 Peters DJM, van de WA, Spruit L *et al*: Cellular localization and tissue distribution of polycystin-1. *J Pathol* 1999; **188**: 439–446.
16 Ong AC, Harris PC, Biddolph S, Bowker C, Ward CJ: Characterisation and expression of the PKD-1 protein, polycystin, in renal and extrarenal tissues. *Kidney Int* 1999; **55**: 2091–2116.
17 Foggensteiner L, Bevan AP, Thomas R *et al*: Cellular and subcellular distribution of polycystin-2, the protein product of the PKD2 gene. *J Am Soc Nephrol* 2000; **11**: 814–827.
18 Geng L, Segal Y, Pavlova A *et al*: Distribution and developmentally regulated expression of murine polycystin. *Am J Physiol (Renal physiol)* 1997; **272**: F451–F459.
19 Guillaume R, Trudel M: Distinct and common developmental expression patterns of the murine Pkd2 and Pkd1 genes. *Mech Dev* 2000; **93**: 179–183.
20 Lu W, Peissel B, Babakhanlou H *et al*: Perinatal lethality with kidney and pancreas defects, in mice with a targetted pkd1 mutation. *Nat Genet* 1997; **17**: 179–181.
21 Wu G, D'Agati V, Cai Y *et al*: Somatic inactivation of Pkd2 results in polycystic kidney disease. *Cell* 1998; **93**: 177–188.
22 Boulter C, Mulroy S, Webb S, Fleming S, Brindle K, Sandford R: Cardiovascular, skeletal, and renal defects in mice with a targeted disruption of the Pkd1 gene. *Proc Natl Acad Sci USA* 2001; **98**: 12174–12179.
23 Peters DJM, Breuning MH: Autosomal dominant polycystic kidney disease: modification of disease progression. *Lancet* 2001; **358**: 1439–1444.
24 Igarashi P, Somlo S: Genetics and pathogenesis of polycystic kidney disease. *J Am Soc Nephrol* 2002; **13**: 2384–2398.

25 Rodova M, Islam MR, Maser RL, Calvet JP: The polycystic kidney disease-1 promoter is a target of the beta-catenin/T-cell factor pathway. *J Biol Chem* 2002; **277**: 29577–29583.
26 Sandford R, Sgotto B, Aparicio S *et al*: Comparative analysis of the polycystic kidney disease 1 (PKD1) gene reveals an integrated membrane glycoprotein with multiple evolutionary conserved domains. *Hum Mol Genet* 1997; **6**: 1483–1489.
27 Mochizuki T, Wu G, Hayashi T *et al*: PKD2, a gene for polycystic kidney disease that encodes an integral membrane protein. *Science* 1996; **272**: 1339–1342.
28 Park JH, Li L, Cai Y *et al*: Cloning and characterization of the murine pkd2 promoter. *Genomics* 2000; **66**: 305–312.
29 Li R, Mignot E, Faraco J *et al*: Construction and characterization of an eightfold redundant dog genomic bacterial artificial chromosome library. *Genomics* 1999; **58**: 9–17.
30 Lantinga-van Leeuwen IS, Dauwerse JG, Baelde HJ *et al*: Lowering of *Pkd1* expression is sufficient to cause polycystic kidney disease. *Hum Mol Genet* 2004; **13**: 3069–3077.
31 Lohning C, Nowicka U, Frischauf A-M: The mouse homolog of PKD1: sequence analysis and alternative splicing. *Mamm Genome* 1997; **8**: 307–311.
32 Whyte DA, Li C, Thomson RB *et al*: Ksp-cadherin gene promoter. I. Characterization and renal epithelial cell-specific activity. *Am J Physiol* 1999; **277**: F587–F598.
33 Aguiari G, Piva R, Manzati E *et al*: K562 erythroid and HL60 macrophage differentiation downregulates polycystin, a large membrane-associated protein. *Exp Cell Res* 1998; **244**: 259–267.
34 Zhao Y, Haylor JL, Ong AC: Polycystin-2 expression is increased following experimental ischaemic renal injury. *Nephrol Dial Transplant* 2002; **17**: 2138–2144.
35 Ibraghimov-Beskrovnaya O, Dackowski WR, Foggensteiner L *et al*: Polycystin: *in vitro* synthesis, *in vivo* tissue expression, and subcellular localization identifies a large membrane-associated protein. *Proc Natl Acad Sci USA* 1997; **94**: 6397–6402.
36 Murphy K: Database searches for binding sites. *Science* 2000; **288**: 2319.
37 Loots GG, Ovcharenko I, Pachter L, Dubchak I, Rubin EM: rVista for comparative sequence-based discovery of functional transcription factor binding sites. *Genome Res* 2002; **12**: 832–839.
38 Mrowka R, Steinhage K, Patzak A, Persson PB: An evolutionary approach for identifying potential transcription factor binding sites: the renin gene as an example. *Am J Physiol Regul Integr Comp Physiol* 2003; **284**: R1147–R1150.
39 Hateboer N, Dijk MA, Coto E *et al*: European PKD1-PKD2 Study Group: comparison of phenotypes of polycystic kidney disease types 1 and 2. *Lancet* 1999; **353**: 103–107.
40 Bai L, Merchant JL: ZBP-89 promotes growth arrest through stabilization of p53. *Mol Cell Biol* 2001; **21**: 4670–4683.
41 Wu G, Markowitz GS, Li L *et al*: Cardiac defects and renal failure in mice with targeted mutations in Pkd2. *Nat Genet* 2000; **24**: 75–78.
42 Gaboli M, Kotsi PA, Gurrieri C *et al*: Mzf1 controls cell proliferation and tumorigenesis. *Genes Dev* 2001; **15**: 1625–1630.
43 Blake MC, Jambou RC, Swick AG, Kahn JW, Azizkhan JC: Transcriptional initiation is controlled by upstream GC-box interactions in a TATAA-less promoter. *Mol Cell Biol* 1990; **10**: 6632–6641.
44 Cohen HT, Bossone SA, Zhu G, McDonald GA, Sukhatme VP: Sp1 is a critical regulator of the Wilms' tumor-1 gene. *J Biol Chem* 1997; **272**: 2901–2913.
45 Zatyka M, Morrissey C, Kuzmin I *et al*: Genetic and functional analysis of the von Hippel–Lindau (VHL) tumour suppressor gene promoter. *J Med Genet* 2002; **39**: 463–472.
46 Swoboda P, Adler HT, Thomas JH: The RFX-type transcription factor DAF-19 regulates sensory neuron cilium formation in *C. elegans*. *Mol Cell* 2000; **5**: 411–421.
47 Ledford AW, Brantley JG, Kemeny G *et al*: Deregulated expression of the homeobox gene Cux-1 in transgenic mice results in downregulation of p27(kip1) expression during nephrogenesis, glomerular abnormalities, and multiorgan hyperplasia. *Dev Biol* 2002; **245**: 157–171.
48 Overdier DG, Ye H, Peterson RS, Clevidence DE, Costa RH: The winged helix transcriptional activator HFH-3 is expressed in the distal tubules of embryonic and adult mouse kidney. *J Biol Chem* 1997; **272**: 13725–13730.