Genomics 104 (2014) 512–519

Contents lists available at ScienceDirect

Genomics

journal homepage: www.elsevier.com/locate/ygeno

Characterization of the polycystic kidney disease 2 gene promoter

Qingsong Wang [a], Gang Han [a], Jianhua Ye [b], Xianjun Gao [a], Hongjing Niu [a,1], Jingwen Zhao [a], Yawen Chai [a], Ning Li [b], HaiFang Yin [a,*]

[a] Research Centre of Basic Medical Science,Department of Cell Biology, Tianjin Medical University, Qixiangtai Road, Heping District, Tianjin 300070, China
[b] State Key Laboratory for Agrobiotechnology, China Agricultural University, Beijing 100193, China

ARTICLE INFO

Article history:
Received 27 March 2014
Accepted 15 August 2014
Available online 27 August 2014

Keywords:
PKD2
G-quadruplex
Sp1
Transcription
G4R1

ABSTRACT

The key regulatory elements for *PKD2* transcription remain unclear. To identify these core elements, we characterized porcine *PKD2* promoter with bioinformatics and molecular tools and found porcine *PKD2* promoter bearing typical features of enriched CpG and less TATA. Further studies demonstrated that the core region was located in fragment −483 to +100. Subsequent biophysical binding assays and mutation experiments revealed that G4 motif and Sp1 are critical regulators for mediating the transcription of porcine *PKD2*. Moreover, the same regulatory pattern was reproduced in human *PKD2* promoter region, indicating the critical role of G4 and Sp1 in regulating *PKD2*.

1. Introduction

Autosomal dominant polycystic kidney disease (ADPKD), one of the most common inherited kidney disorders, is genetically heterogeneous, with a prevalence estimated between 1:400 and 1:1000 in all races worldwide [1]. Approximately 50% of ADPKD patients will develop end-stage renal failure (ESRD) at age 60 [2] and currently there is no effective treatment available except kidney transplantation. Mutations in polycystic kidney disease 2 gene (*PKD2*) account for 15% of ADPKD cases [3,4], therefore it is important to elucidate the role of *PKD2* in the pathogenesis of ADPKD.

To investigate the role of *PKD2* in the development of ADPKD, we previously cloned and characterized porcine *PKD2* cDNA as swine is an emerging and promising animal model for various human diseases e.g. cystic fibrosis [5], xenotransplantation [6], and Huntington's disease [7], due to its physiological and anatomical similarities to humans. Furthermore, we attempted to establish a transgenic pig model of ADPKD by over-expressing porcine *PKD2*. However the transgenic pigs failed to show relevant clinical features during their development though the results were positive at the molecular level [8]. We speculated that one possible explanation is attributed to the chronic nature of ADPKD; another reason is likely due to the lack of endogenous regulatory elements in the transgenic construct, in which a ubiquitous promoter was applied, whereas the expression of *PKD2* is developmentally regulated [9,10]. Therefore we wish to identify the core promoter region and regulatory elements involved in the transcription and expression of *PKD2*, thereby further to improve the design of transgenic pig model of ADPKD.

In this study, we focused on the regulatory elements involved in the expression of porcine *PKD2*. To identify the regulatory elements of porcine *PKD2*, we analyzed sequences in close proximity to transcriptional start site and showed that the upstream region of porcine *PKD2* is highly GC-rich, whereas no typical TATA- and CCAAT-box was revealed. As it was previously reported that G-quadruplex (G4), consisting of guanine runs, could form in the GC-rich regulatory region and interact with the Sp1 transcription factor [11,12]. With the QGRS Mapper software, we were able to predict several putative G4-forming sequences and Sp1 binding sites, which partially overlapped [13,14]. Recent studies have shown that G4 DNA structures in the regulatory regions of several oncogenes play a significant role in regulating gene expression [15–17]. Thus G4 elements associated with Sp1 are likely to be an important regulatory element involved in the transcription of *PKD2*. Subsequent studies verified this assumption by showing that G4 and Sp1 play critical roles in regulating the expression of *PKD2* both in human and swine.

Abbreviations: ADPKD, autosomal dominant polycystic kidney disease; ATP, adenosine triphosphate; BAC, bacterial artificial chromosome; CD, circular dichroism; ChIP, chromatin immunoprecipitation; E2F, E2F-myc activator/cell cycle regulator; EGFR, early growth response factor; ESRD, end-stage renal failure; G4, G-quadruplex; G4R1, G4 nucleic acid resolvase 1; MZF, myeloid zinc finger protein; *PKD2*, polycystic kidney disease 2 gene; RACE, rapid amplification of cDNA ends; TF2B, transcription factor II B; TSS, transcription start site; UTR, untranslated region; ZBP, zinc binding protein.

* Corresponding author. Fax: +86 22 83336537.
E-mail address: haifangyin@tmu.edu.cn (H. Yin).
[1] Present address: Jinhua Municipality Blood Centre, Qingchun Road, Jinhua 321000, China.

http://dx.doi.org/10.1016/j.ygeno.2014.08.006
0888-7543/

2. Results

2.1. Identification of porcine PKD2 transcription start site (TSS)

In order to identify the TSS of porcine *PKD2*, we carried out 5′ RACE and a 455 bp DNA fragment was obtained (Figs. S1a and S1b). The transcription start site 'C' with a putative cap site 'GGGAG' was located 100 bp upstream from the translation start codon (ATG) of *PKD2* (Figs. S1c and 1a). Similar to human *PKD2*, porcine *PKD2* is composed of 15 exons and 14 introns [18], the size of exons and introns is listed in Supplementary Table 1, and the first exon of porcine *PKD2* is 700 bp, covering 5′ UTR and partial coding sequences.

Subsequent sequence analysis of the entire porcine *PKD2* with EMBOSS CpGPlot revealed that the upstream region is highly GC-rich with two putative CpG islands adjacent to TSS (data not shown) and lacks a typical TATA or CAAT box.

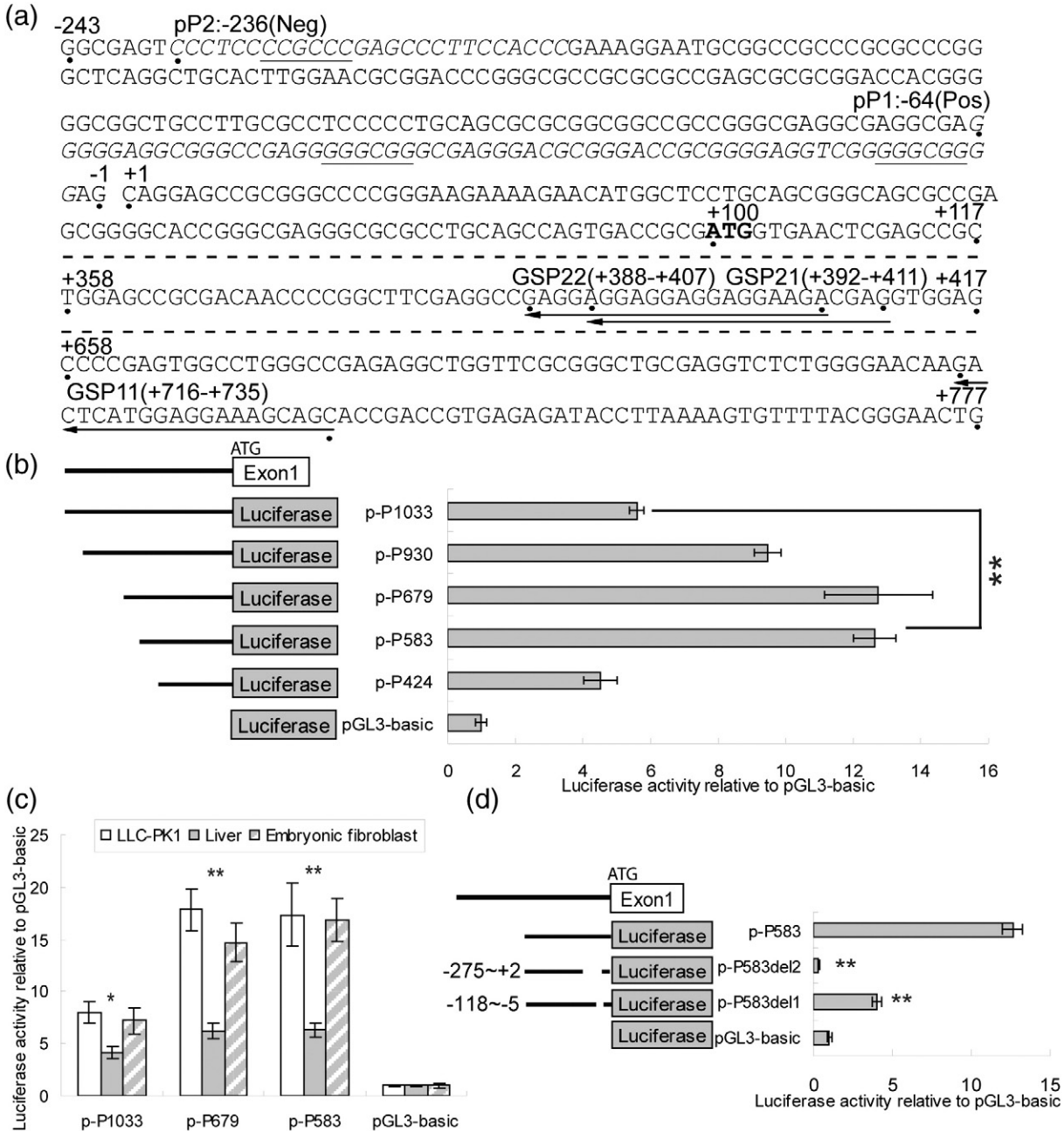

Fig. 1. Nucleotide sequences of the 5′ flanking region and deletion analysis of porcine *PKD2*. a. The transcription start site was designated as +1. The underlined sequence in the *PKD2* promoter region is putative transcription factor Sp1 binding sites. G4 DNA structure candidates (pP1 and pP2) in the *PKD2* promoter are shown in italic letters as negative- or positive-strands. GSP22, GSP21 and GSP11 were denoted as primers used in the 5′ RACE underlined with arrows. The positions of the numbered nucleotides were identified by dots beneath them. The dash lines indicated the omitted sequences. Neg: negative strand; Pos: positive strand. b. Diagram of different deletion constructs for the 5′ flanking region of *PKD2* and assays for the promoter activity of each construct in LLC-PK1 cells. Cells were transiently transfected with each individual construct along with the renilla luciferase-expressing vector (pRL-TK). The luciferase activity was measured 24 h after transfection and normalized to the pGL3-Basic plasmid, which contained no insert. c. Analyses of different porcine *PKD2* promoter constructs in other porcine cells. Porcine *PKD2* promoter p-P583, p-P679 and p-P1033 were tested in LLC-PK1, porcine liver cells and embryonic fibroblasts. d. Two mutant constructs containing different deletions were evaluated in LLC-PK1 cells. The promoter activity of each mutant was determined by luciferase activity after normalization to the pGL3-Basic plasmid. The pRL-TK was used as an internal control for all the transfection experiments (*$p < 0.05$, **$p < 0.005$).