ADPKD occurrence corresponds with *PKD1* orthologs harboring guanine quaduplex sequence motifs

Monica Ellis, Emma Swayze, Aaron Zebolsky, Gloria Alvarado, Greg Vanden Heuvel, and Erik Larson

Western Michigan University Homer Stryker MD School of Medicine, Kalamazoo, MI 49008

INTRODUCTION

Autosomal dominant polycystic kidney disease (ADPKD) is thought to arise predominately from an inherited mutation in the *PKD1* gene followed by a somatic "second hit" mutation in the normal allele, leading to cyst formation. Cysts are clonal in nature and can number in the thousands with each cyst possibly resulting from an independent somatic mutation. *PKD1* appears to be genetically unstable, but the mechanisms of this instability are not understood. Interestingly, the human and feline *PKD1* orthologs are both subject to inactivating mutagenesis, while murine *PKD1* is not.

METHODS

We examined the human, feline, and murine *PKD1* genes for sequence differences that could explain the apparent instability in humans and cats.

RESULTS

We found that while *PKD1* gene organization and coding is conserved, there is an over-representation of guanine triplets in human and feline, but not murine, *PKD1*. Guanine-rich DNA is a hallmark of unstable loci and is found at oncogenic translocation sites and regions of programmed recombination. These guanines fold into four-stranded structures called G-quadruplex (G4) DNA, which can and do interfere with replication and repair. Thus, the presence of guanine rich DNA in *PKD1* may explain why the human and feline orthologs are unstable. Consistent with that model, we find that G4 sequence motifs are distributed throughout human and feline *PKD1*, present on both strands, and in tandem arrays in several introns. Circular Dichroism spectroscopy confirms that representative G4 sequence motifs from human *PKD1* form G4 DNA structures under physiological conditions. The possibility that G4 formation directly contributes to *PKD1* inactivation is supported by the fact that cataloged somatic mutations are clustered near predicted G4 forming motifs, with a pathogenic 20 nucleotide deletion directly overlapping with a G4 DNA structural motif. Previous studies have shown that a long polypurine-polypyrimidine tract in intron 21 of the human *PKD1* gene (IVS21) can adopt triplex structures. Thus, multiple DNA structures may contribute to *PKD1* instability.

CONCLUSIONS

Taken together, the identification of sequences that form G4 DNA in *PKD1* may provide a molecular rationale for ADPKD and identify a DNA structure that can be targeted for diagnosis or disease prevention.