ISSN: 0975-8585

Research Journal of Pharmaceutical, Biological and Chemical Sciences

Identification of Functional Single Nucleotide Polymorphism in Polycystic Kidney Disease using Computational Tools.

Latika Jaiswal[1], and Ravi Kant Singh[2]*.

[1]Department of Biotechnology, Institute of Engineering & Technology, Lucknow, U.P., India
[2]Department of Biotechnology, Raipur Institute of Technology, Raipur, C.G., India

ABSTRACT

Polycystic kidney disease is one of the most common genetic diseases affecting 600,000 people overall in the U.S, and is a major cause of End stage renal disease (ESRD) in children and adults. The main goal of this work is to understand the mutations that are casually causing the disease, and can be used as a biological marker which will be helpful in development of much advanced and valuable drug. A total of 832 coding SNPs present in both the two genes of ADPKD (PKD1 and 2 genes), whereas in case of Autosomal Recessive PKD, 539 coding SNPs in PKHD1 gene were found to be involved at dbSNP. In this study, those SNPs that can alter the expression and function in PKD genes were evaluated through various computational tools. We have also applied an evolutionary approach by using a sequence-homology base tool SIFT (Sorting Intolerant from Tolerant) to screen out the SNPs, suggested that 23 nsSNPs (17.96%) were intolerant. A structural-based, Polyphen 2 server has classified 288 nsSNPs (36.59%) known to be damaging protein function. I-MUTANT 3.0 suite predicted overall 105 nsSNPs, that were found causing missense mutations further causing instability in protein structure. FASTSNP predicted aboutn130 nsSNPs, found influencing splicing regulation since they are present in splicing modifiers (enhancers and silencers) binding sites. Whereas one SNP was found influencing binding of micro RNA in PKD1 gene by DBSMR. The analysis predicted rs58598099 with a mutation of cysteine to arginine amino acid change at position 508 in the gene PKD1, involved in causing ADPKD was found most damaging by all the functional nsSNP prediction severs used and which could be the main target mutation, results of the study provides useful information about SNPs that affect the polymorphism on phenotype in Polycystic kidney disease genes.

Keywords: Polycystic kidney disease, Biological markers, Computational tools, PKD gene

Corresponding author

ISSN: 0975-8585

INTRODUCTION

Autosomal dominant polycystic kidney disease (ADPKD; MIM173900; 173910) occurs in 1:400 to 1:1000 individuals and is caused by mutations in one of two genes, PKD1 or PKD2 (1, 2). Autosomal recessive PKD (ARPKD; MIM 263200) is less frequent (1:20,000 live births) and occurs as a result of defects in a single gene, PKHD1. Single nucleotide polymorphism plays a vital role in inferring the genetic basis of many complex human diseases. And, the genetics of human phenotypic variation can be well understood by knowing the function of these SNPs. However, it is still one of the major challenges to identify the functional SNPs in a disease-related gene. SNPs account for the more common form of human genetic variations. In the context of human genome, about 500,000 SNPs fall in the coding regions (3). Among these, the nonsynonymous SNPs (nsSNPs) cause changes in the amino acid residues. These are some of the important factors which contribute to the functional diversity of the encoded proteins in the human population (4). Generally the nsSNPs affect gene regulation by altering DNA and transcriptional binding factors (5) and the maintenance of the structural integrity of cells and tissues (6). Also, the non synonymous SNPs affects the functional expression of proteins in the signal transduction of visual, hormonal, and other stimulants (7, 8). SNPs comprising most abundant type of genetic variation, hence, they are now the primal matter underlying most genetic studies and databases. Although other types of variation, including copy number variants, indels, microsatellites, and epigenetic markers remain important to consider and can impact disease. SNPs are largely the easiest to be determined and the most useful and widely applied markers in genetic studies in this modern era.

The acquired polycystic kidney diseases, which are due to germ-line mutations in single genes, inherited as Mendelian traits, include autosomal dominant and autosomal recessive polycystic kidney disease. The age at onset, the severity of symptoms, and the rates of progression to end-stage renal failure or death vary widely in this group of diseases. The two types: type I is caused by mutations in the PKD1 gene and PKD2 gene which leads to Autosomal Dominant type and the other type II caused by mutation in PKHD1 gene leading to Autosomal recessive PKD (9). These 3 genes codes for polycystin1 (PC-1), polycystin2 (PC-2) and fibrocystin respectively. PC-1 is a large, transmembrane protein that interacts with PC-2, a transient receptor potential channel that regulates intracellular calcium. Both proteins localize to the kidney primary cilium, and may act as a flow-dependent mechanosensor regulating the differentiation and proliferation of tubular epithelial cells (10). Polycystic kidney disease (PKD) is passed down through families (inherited), usually as an autosomal dominant trait. If one parent carries the gene, the children have a 50% chance of developing the disorder (11). Tremendous cystic enlargement of both kidneys is characteristic of autosomal dominant polycystic kidney disease. Patients often present with hypertension, hematuria, polyuria, and flank pain and are prone to recurrent urinary tract infections and renal stones. Mutations in the PKHD1 gene cause autosomal recessive polycystic kidney disease. Fibrocystin spans the cell membrane of kidney cells, so that one end of the protein remains inside the cell and the other end projects from the outer surface of the cell. Based on its structure, fibrocystin may act as a receptor, interacting with molecules outside the cell and receiving signals that help the cell respond to its environment. This protein also may be involved in connecting cells together (adhesion), keeping cells apart (repulsion), and promoting the growth and division of cells (proliferation). Fibrocystin is also found in cell structures called primary cilia. Primary cilia are tiny, finger like projections that line the small tubes where urine is formed (renal tubules). The final stage of chronic kidney disease is called ESRD. About one-half of people with the most common type of PKD progress to kidney failure, also called ESRD. The kidneys no longer function and the patient needs dialysis or a kidney transplant. Chronic kidney disease and ESRD affect more than 2 out of every 1,000 people in the United States.

According to the studies conducted by various groups on the PKD1 gene in different regions of Europe given in OMIM (173900) where they have found that the PKD1 mutation being primarily a renal problem, also affects several other organs in humans. Out of 321 individuals evidence of kidney cysts was found in 21% of individuals (12). In another study it was found that out of 397 individuals 48% were found affected by multiple liver cysts (13, 14). Further it was found that 16% of the 529 individuals studied by various groups were affected by multiple problems of the Cerebrovascular and Cardiovascular systems (15, 16, 17). All the test subjects had confirmed PKD1 mutation and the occurrence of multiple problems suggested that the mutation not only affects the kidney but also causes multiple problems in different other organs (Figure-1). According to OMIM entry (173910) the PKD2 gene mutation which is involved in cause of ADPKD can be used as biomarkers for the successful diagnosis of the disease in future. Nine SNPs (out of which SNP IDs present in dbsnp; rs121918039, rs121918040, rs121918041, rs121918042, rs121918043) have been found to be promising biomarkers. (2, 18, 19, 20, 21, 22, 23).

ISSN: 0975-8585

According to OMIM entry (606702) the PKHD1 gene mutation which is involved in cause of ARPKD can be used as biomarkers for the successful diagnosis of the disease in future. One SNP (rs28937907) have been found to be promising biomarkers.

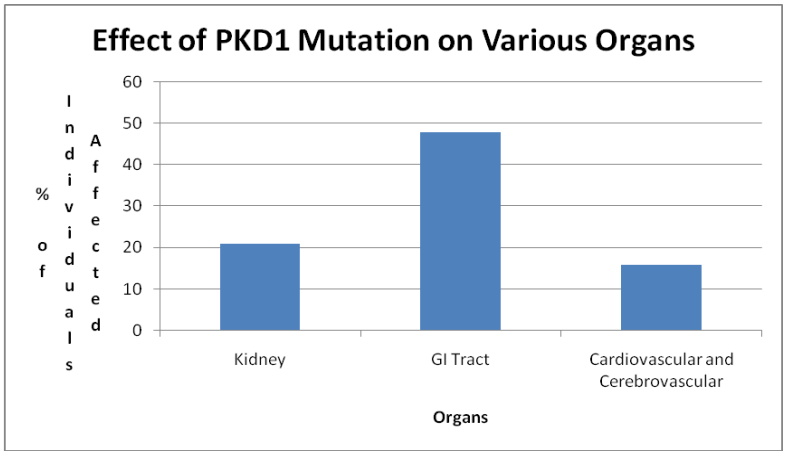

Figure 1: Graph showing the Effect of PKD1 Mutation in Various Organs

MATERIAL AND METHODS

Datasets

The SNPs and their related protein sequence for the Polycystic kidney disease genes were obtained from the dbSNP (http://www.ncbi.nlm.nih.gov/SNP/) build 137 for our computational analysis (24).

Analysis of functional consequences of coding nsSNPs by sequence-homology-based method (SIFT)

We used the program SIFT (25) available at http://blocks.fhcrc.org/sift/SIFT.html to detect the deleterious coding nonsynonymous SNPs. SIFT is a sequence- homology-based tool that presumes that important amino acids will be conserved in the protein family. Hence, changes at well-conserved positions tend to be predicted as deleterious (Ng, 2003). We submitted the query in the form of SNP IDs or as protein sequences. The underlying principle of this program is that SIFT takes a query sequence and uses multiple alignment information to predict tolerated and deleterious substitutions for every position of the query sequence. SIFT is a multistep procedure that, given a protein sequence, (a) searches for similar sequences, (b) chooses closely related sequences that may share similar functions, (c) obtains the multiple alignment of the chosen sequences, and (d) calculates normalized probabilities for all possible substitutions at each position from the alignment. Substitutions at each position with normalized probabilities less than a chosen cutoff are predicted to be deleterious and those greater than or equal to the cutoff are predicted to be tolerated (26). The cutoff value in the SIFT program is a tolerance index of ≥0.05. The higher the tolerance index, the less functional impact a particular amino acid substitution is likely to have.

Simulation for functional change in coding nsSNPs by structure-homology-based method (PolyPhen)

Analyzing the damaged coding nonsynonymous SNPs at the structural level is considered to be very important to understand the functional activity of the protein of concern. We used the server PolyPhen (27), which is available at http://coot.embl.de/PolyPhen/, for this purpose. Input options for the PolyPhen server are protein sequence or SWALL database ID or accession number together with sequence position with two amino acid variants. We submitted the query in the form of protein sequence with mutational position and two amino acid variants. Sequence-based characterization of the substitution site, profile analysis of homologous sequences, and mapping of substitution site to a known protein three-dimensional structure are the parameters taken into account by the PolyPhen server to calculate the score. It calculates PSIC scores for each of the two variants and then computes the PSIC score difference between them. The higher the PSIC score difference is, the higher is the functional impact a particular amino acid substitution is likely to have.

Functional significance of noncoding SNPs in regulatory untranslated regions

Recent studies show that SNPs have functional effects on protein structure by a single change in the amino acid (28, 29) and on transcriptional regulation (29, 30). We used the Web server FastSNP (31) available at http://fastsnp.ibms.sinica.edu.tw for predicting the functional significance of the PKD genes gene and predict the impact of SNP if present in splicing modifier binding sites. The score is generated based on the level of risk, with a ranking of 0, 1,2,3,4 or 5 signifying no, very low, low, medium, high, and very high effect respectively.

Non synonymous SNPs causing protein stability change

The web based version of computational tool which have been used is I-Mutant 3.0 (32) server available at http:/gpcr.biocomp.unibo.it/ to identify the missense mutations causing instability to protein structure. The I-Mutant DDG tool is trained on a DDGMut dataset derived from ProTherm (33) which evaluates the stability change upon single site mutation starting from the protein structure or from the protein sequence. In this, the method allows to predict if a mutation can largely destabilize the protein (ΔΔG<-0.5 Kcal/mol) or largely stabilize (ΔΔG>0.5 Kcal/mol) or have a weak effect (-0.5<=ΔΔG<=0.5 Kcal/mol).

Single Nucleotide Polymorphisms present in micro-RNA binding site

DbSMR (10) is a database of all miRNA binding sites within 200 nucleotide of SNPs which may affect miRNA accessibility to the target site, thereby altering the regulation. Dbsmr is available at http://miracle.igib.res.in/polyreg/, using a consensus approach of three software - miRanda, RNAHybrid and TargetScan. The respective output gives the following details: (The target gene name, the miRNA name, the Target Transcript, the Target Location (Start and End) with respect to the 3' UTR of the target transcript, The miRNA-target binding pattern, Distance of separation of SNP from target site, the SNP ID (as in NCBI dbSNP), number of nucleotides changing mode of interaction compared to wild and poly sequences; The varying shades of red indicate the degree of change, effect of change pertains to effect of SNP in terms of probable gain or loss of the miRNA binding.

The rough estimate of the number of bases changing its intramolecular binding at the target site. It is the ratio of number of bases that change their binding pattern to the total length of the target site.

[Degree = n (Bases with change in pairing pattern) / Total length of target site]

RESULTS AND DISCUSSION

Dataset

The SNPs and their related protein sequence for the genes involved in Polycystic kidney disease were obtained from the dbSNP (http://www.ncbi.nlm.nih.gov/SNP/) build 137 (24) for our computational analysis.

Identification of Deleterious nsSNP found by the SIFT program

The deleterious coding nonsynonymous SNPs were determined by using a sequence homology-based tool, SIFT (25) available at http://blocks.fhcrc.org/sift/SIFT.html. The protein sequences of 788 nsSNPs were submitted independently to the SIFT program to check its tolerance index. The higher the tolerance index, the less functional impact a particular amino acid substitution is likely to have, and vice versa. Among the 128 nsSNPs, 23 were found to be deleterious, having a tolerance index score of ≤0.05 in all three genes of polycystic kidney disease. The results are shown in Table 1.

We observed that, of 23 deleterious nsSNPs, 8 showed a highly deleterious tolerance index score of 0.00, and 6 showed a tolerance index score of 0.01, 5 showed a tolerance index score of 0.02 and 2 nsSNPs with a tolerance index of 0.03, and 0.04, respectively. One nsSNP showed a nucleotide change of A→G, 3 nsSNPs G→A, 5 nsSNPs G→T, 3 nsSNPs G→C, 6 nsSNPs C→T, 1 nsSNP C→A, and 2 nsSNPs T→G. C → T and G → T nucleotide changes occurred the maximum number of times and C → A, T → C and A → G nucleotide changes occurred a minimum number of times, as can be seen from Table 1. The nucleotide change G → T

accounted for the highest number of deleterious nsSNPs, with a SIFT tolerance index of 0.00. This was closely followed by the nucleotide change C → T AND G → A, which showed a tolerance index of 0.02. Also, of the 7 nsSNPs that showed a SIFT tolerance index of 0.00, 3 of them changed to an aromatic amino acid in the mutant type from a non aromatic amino acid in the native protein.

Damaged nsSNP found by the PolyPhen server

To analyze the damaged coding nonsynonymous SNPs at the structural level we used Polyphen server. Seven eighty seven protein sequences of nsSNPs investigated in this work were submitted as input to the PolyPhen server and the results are shown in Table 2. A position-specific independent count (PSIC) score difference of 0.95 and above is considered to be most damaging or probably damaging. It can be seen that, of 788 nsSNPs, 289 were considered to be damaging. All damaging nsSNPs exhibited a PSIC score difference in the range 0.95 to 1.

Nineteen nsSNPs that were observed to be deleterious by the SIFT program also were damaging according to PolyPhen. Hence, we could infer that the results obtained on the basis of sequence details (SIFT) were in good correlation with the results obtained for structural details (PolyPhen), as can be seen from Tables 1 and 2. It can be seen from Tables 1 and 2 that 6 nsSNPs (rs115538130, rs117896488, rs113969111, rs58598099, rs1131408, rs114237522) had a SIFT tolerance index of 0.00 and maximum PSIC score of 1. Hence the mutations occurring with these 6 nsSNPs would be of prime importance in the identification of Polycystic kidney disease caused by the PKD1, PKD2 and PKHD1 genes, according to SIFT and PolyPhen results. It was observed that 19 nsSNPs were predicted to be probably damaging by Polyphen server were also predicted deleterious by SIFT server.

Functional SNPs in Splicing Modifier Binding site

FASTSNP predicted hundred and eleven SNPs to be influencing splicing regulation by their presence in splicing modifiers (enhancers and silencers) binding site (Table 3) (34). 91 SNPs predicted to be influencing splicing regulation had a risk in range of 2–3 (low to medium) and remaining 20 SNPs with a risk in range of 3–4 (medium to high). None of the SNPs in UTR were reported to be present in splicing modifier binding sites.

NSSNPs causing protein stability change using I-MUTANT 3.0 suite

The missense mutations causing instability to protein structure is been identified by I-MUTANT 3.0 server. In order to identify functional stability of nsSNPs of ADPKD and ARPKD genes, we applied I-Mutant 3.0 suite which predicts automatically protein stability changes upon single-site mutations starting from protein sequence alone or protein structure when available. The protein sequence of PKD1 and PKD2 were submitted separately to I-Mutant 3.0 suite as an input file. As a result out of the 2 ADPKD genes, we obtained sixty five nsSNPs shows increase stability and three eighty six nsSNPs shows decrease stability in PKD1 And PKD2 genes of Autosomal Dominant PKD whereas when protein sequence of PKHD1 gene of Autosomal Recessive PKD was submitted to I-Mutant 3.0 we obtained forty nsSNPs shows increase stability and two hundred ninety six nsSNPs shows decrease stability respectively (Table-4). Fourteen nsSNPs shows increase stability with reliability index (RI) 0 and 3, 27 nsSNPs shows increase stability with reliability index 1, 2 and 7, nine nsSNPs shows increase stability with reliability index 4, two nsSNPs shows increase stability with RI 5. Five nsSNPs shows increase stability with reliability index 6. One nsSNP (rs17135779) shows increase stability with reliability index 8 present in PKD1 gene. In PKD2 gene, three nsSNPs shows increase stability with reliability index 0, 6 and 7, four nsSNPs shows increase stability with reliability index 1 respectively.
Whereas five nsSNPs (rs4715227, rs71570543, rs140996978, rs200733734, rs190315828) shows increase stability with reliability index 0 PKHD1 gene. Eight nsSNPs shows increase stability with reliability index 1. Nine nsSNPs shows increase stability with reliability index 2 and 3, three nsSNPs shows increase stability with reliability index 4 present in the gene. Two nsSNPs shows increase stability with reliability index 5 and 6 present in PKD1 gene. One nsSNP shows increase stability with reliability index 7 present in PKHD1 gene respectively.

ISSN: 0975-8585

Identifying Single Nucleotide Polymorphisms in micro-RNA binding site using Dbsmr

In order to identifying those SNPs which may affect miRNA accessibility to the target site, thereby altering the regulation process. One nsSNP (rs1063404) were reported to be influencing binding of micro RNA only in gene PKD1 by dbsmr. Since, none of the SNPs were reported to be influencing binding of micro RNA in gene PKD2.

The columns give the information as target gene name, miRNA target the transcript, the transcript ID, the location of miRNA binding with respect to the 3' UTR of the transcript, the binding modality, the location of SNP with respect to the target site, the SNP ID, the number of bases changing intramolecular conformation, the effect of the change (as gain or loss). None of the SNPs by dbSMR were reported to be influencing binding of micro RNA in gene PKHD1 (gene for ARPKD) as well.

Table 1: List of nsSNPs that were predicted to have functional significance by SIFT and Polyphen2

Gene	dbSNPid	Allele	A.A. Subs	SIFT Score	PSIC	Heterozygosity	Validation
PKD1	rs114251396	C/T	R4276W	**0.01**	**0.999**	N.D.	2
	rs115538130	C/T	R4154C	**0.00**	**1.000**	N.D.	
	rs36221080	G/T	R4150L	**0.01**	**1.000**	N.D.	
	rs2854584	G/C	G3809A	**0.01**	**1.000**	N.D.	
	rs114523369	G/T	G3525V	**0.02**	**0.998**	N.D.	
	rs79648977	G/A	R3183Q	**0.01**	**1.000**	0.003	2,6
	rs117896488	G/A	V3008M	**0.00**	**1.000**	N.D.	2
	rs3952944	C/T	T2496M	**0.01**	**0.994**	N.D.	
	rs2854618	G/A	V1976M	0.02	0.313	N.D.	7
	rs113969111	G/C	W1839C	**0.00**	**1.000**	N.D.	
	rs2575314	G/A	V1566M	**0.02**	**1.000**	N.D.	
	rs55840049	C/T	R1351W	**0.02**	**0.998**	N.D.	2,7
	rs111686944	G/C	D962H	**0.04**	**0.999**	N.D.	
	rs2854624	C/G	P691A	**0.03**	**1.000**	N.D.	
	rs58598099	T/C	C508R	**0.00**	**1.000**	N.D.	
	rs4018169	C/T	T334M	**0.03**	**0.975**	N.D.	
PKD2	rs1131408	G/T	G449V	**0.00**	**1.000**	N.D	
	rs75762896	T/G	F482C	0.02	0.519	N.D.	2
PKHD1	rs4715227	A/G	Q3899R	0.04	0.295	N.D.	2
	rs2784228	C/T	S3821F	0.00	0.441	N.D.	
	rs114237522	G/T	D3668Y	**0.00**	**1.000**	N.D	2
	rs78624439	T/G	I1128S	**0.00**	**0.981**	0.5	
	rs45500692	C/T	T579M	**0.01**	**1.000**	N.D.	2

Table 2: List of nsSNPs that were predicted to be functionally significant by PolyPhen

Gene	dbSNPid	Allele	A.A. Subs	PSIC	Heterozygosity	Validation
PKD1	rs114251396	C/T	R4276W	**0.999**	N.D.	2
	rs144648696	C/T	R4273C	**0.999**	N.D.	

	rs146614757	A/C	E4220A	**0.977**	0.001	
	rs143945578	C/A	P4212T	**0.996**	0.001	
	rs150405871	T/A	S4190T	**0.997**	0.001	
	rs138192074	A/T	T4189S	**0.957**	0.001	
	rs199878927	G/A	R4164H	**1.000**	N.D.	
	rs182827985	C/T	P4162L	**1.000**	N.D.	
	rs115538130	C/T	R4154C	**1.000**	N.D.	
	rs36221080	G/T	R4150L	**1.000**	N.D.	
	rs148478410	G/A	V4146I	**0.994**	N.D.	2
	rs138149588	C/G	D4126E	**1.000**	0.001	
	rs143724907	C/T	L4106F	**1.000**	0.001	
	rs186160250	C/G	S4050C	**1.000**	N.D.	
	rs36224193	T/G	L4046V	**0.984**	N.D.	2
	rs12927338	C/A	L4036I	**0.999**	N.D.	
	rs1804176	G/A	G4028R	**0.999**	N.D.	
	rs201771763	C/T	A4022V	**0.992**	N.D.	
	rs151044657	G/C	K4017N	**1.000**	N.D.	
	rs2855365	C/G	L3983V	**0.998**	N.D.	
	rs2854585	G/C	Q3982H	**0.998**	N.D.	
	rs2854584	G/C	G3809A	**1.000**	N.D.	
	rs138629093	A/T	Y3797F	**1.000**	0	
	rs150071625	T/G	L3744R	**1.000**	0	
	rs200044889	C/T	R3712W	**0.998**	N.D.	
	rs143119214	C/T	H3695Y	**0.993**	0	
	rs36221081	T/G	V3683G	**1.000**	N.D.	
	rs201220835	G/A	R3672Q	**0.959**	N.D.	
	rs140389000	C/T	R3672W	**0.999**	N.D.	
	rs112970540	G/A	V3664I	**0.999**	0.5	
	rs181927900	C/T	R3647W	**1.000**	N.D.	
	rs138089695	G/A	V3631I	**0.999**	N.D.	
	rs140757877	G/C	D3625H	**0.999**	0	
	rs115883514	C/T	A3619V	**0.957**	N.D.	
	rs2854581	G/C	S3595T	**0.999**	N.D.	
	rs79000340	G/A	G3560R	**1.000**	N.D.	2
	rs62624465	G/A	R3548H	**1.000**	N.D.	
	rs2859787	C/G	R3548G	**0.999**	N.D.	
	rs201409107	G/C	G3540A	**0.996**	N.D.	
	rs114523369	G/T	G3525V	**0.998**	N.D.	
	rs62621452	C/A	L3477I	**0.998**	N.D.	
	rs146507511	G/A	D3451N	**0.993**	N.D.	2
	rs137986581	C/T	S3413F	**1.000**	0.001	
	rs149605181	C/T	P3412S	**0.962**	N.D.	2
	rs142799331	C/T	L3381F	**0.999**	0	

	rs139792535	T/C	L3376P	**1.000**	0.001	2
	rs146494724	G/A	R3348Q	**1.000**	N.D.	2
	rs143528690	A/C	N3295T	**1.000**	0.001	
	rs148812376	C/T	R3277C	**1.000**	N.D.	
	rs143453080	C/T	R3272C	**1.000**	0.001	
	rs140791671	G/A	R3247H	**1.000**	0.001	
	rs146732153	C/T	R3244C	**1.000**	N.D.	
	rs144246138	C/T	S3220L	**1.000**	0.001	
	rs200214266	C/T	T3207M	**0.998**	N.D.	
	rs79648977	G/A	R3183Q	**1.000**	0.003	2,6
	rs199738488	G/T	S3178I	**0.993**	N.D.	
	rs139945204	A/T	I3167F	**0.992**	N.D.	2
	rs139210607	C/T	R3152W	**1.000**	0	
	rs149872222	C/T	H3151Y	**1.000**	N.D.	2
	rs145906459	C/T	R3063C	**0.993**	0.001	
	rs200360336	G/C	A3053P	**1.000**	N.D.	
	rs117896488	G/A	V3008M	**1.000**	N.D.	2
	rs150189496	G/A	D2972N	**1.000**	N.D.	2
	rs142039487	G/T	G2970V	**1.000**	0	
	rs139631668	C/T	R2964C	**0.993**	0	
	rs146250848	C/G	S2948W	**0.999**	0	
	rs139438157	G/A	G2919R	**0.99**	N.D.	
	rs200763681	G/A	R2870H	**0.998**	N.D.	
	rs113135029	G/A	G2858D	**1.000**	N.D.	
	rs138294453	C/G	S2845C	**1.000**	0.001	
	rs142888788	G/A	V2844I	**1.000**	0.001	2
	rs2575317	C/T	P2835L	**1.000**	N.D.	
	rs138058871	C/T	L2825F	**0.998**	N.D.	2
	rs201289693	G/A	V2822M	**1.000**	N.D.	2
	rs149151043	G/A	G2814R	**0.997**	N.D.	2
	rs151089809	G/A	V2782M	**0.988**	N.D.	2
	rs200491536	G/A	E2779K	**0.996**	N.D.	
	rs186826765	C/T	T2775M	**1.000**	N.D.	
	rs144979397	C/T	R2765C	**0.999**	N.D.	2
	rs145629362	G/A	R2761H	**0.998**	0.001	
	rs1800569	G/C	R2746P	**0.974**	N.D.	
	rs201050321	C/T	R2746W	**0.992**	N.D.	
	rs199700485	C/A	T2710N	**1.000**	N.D.	
	rs147350387	C/T	T2708M	**1.000**	N.D.	2
	rs148642998	C/T	T2687M	**0.98**	N.D.	2
	rs151308544	C/T	S2682L	**0.959**	0.001	
	rs146636744	G/T	V2655F	**0.993**	0.001	

	rs200646755	C/T	S2612L	**1.000**	N.D.	
	rs138212822	T/C	W2587R	**1.000**	0.001	
	rs149281650	G/A	V2538M	**1.000**	0.001	
	rs146625599	G/A	E2523K	**0.999**	0	
	rs202162543	A/G	H2521R	**0.984**	N.D.	
	rs3952944	C/T	T2496M	**0.994**	N.D.	
	rs4018147	G/A	R2464P	**0.998**	N.D.	
	rs142261601	G/A	A2461T	**0.999**	0.001	
	rs151257298	C/T	R2434W	**1.000**	0.001	
	rs201267349	C/T	T2414M	**1.000**	N.D.	
	rs200159561	C/G	S2385C	**1.000**	N.D.	
	rs146988549	C/T	A2375V	**1.000**	0.001	
	rs202238566	G/T	G2361C	**1.000**	N.D.	
	rs200433577	C/T	R2329W	**1.000**	N.D.	
	rs184394342	C/T	R2327W	**0.999**	N.D.	
	rs143021666	C/T	S2303L	**0.982**	0.001	
	rs148069029	G/A	V2267M	**1.000**	N.D.	
	rs199803853	C/T	R2255C	**1.000**	N.D.	
	rs139971481	C/T	T2250M	**1.000**	N.D.	
	rs145357945	C/G	R2220G	**0.998**	0.001	
	rs200399034	T/G	V2217G	**1.000**	N.D.	
	rs140869992	C/T	R2200C	**0.999**	N.D.	
	rs146096401	C/T	R2166C	**1.000**	N.D.	2
	rs145217118	G/A	R2163Q	**1.000**	N.D.	2
	rs140880305	C/T	R2147W	**0.999**	0.001	
	rs150154235	T/G	F2132C	**1.000**	0.001	
	rs142493877	C/T	A2077V	**0.983**	0.001	
	rs150894453	G/C	D2059H	**1.000**	N.D.	2
	rs186511379	C/T	T2029M	**0.997**	N.D.	6
	rs199943712	C/T	R2024C	**1.000**	N.D.	
	rs144768249	C/T	S2017L	**1.000**	N.D.	
	rs145762096	C/T	S2009L	**0.999**	0.001	
	rs200139724	C/T	T1986M	**0.998**	N.D.	
	rs142174466	G/A	E1981K	**0.991**	0.001	
	rs140434415	G/A	V1971M	**0.997**	0.001	2
	rs191717700	C/T	R1951W	**0.994**	N.D.	6
	rs200001471	G/A	G1944R	**1.000**	N.D.	
	rs185746648	C/T	R1920C	**0.998**	N.D.	6
	rs140902597	C/T	A1900V	**0.969**	0.001	
	rs144137200	G/A	A1871T	**0.999**	N.D.	2
	rs150841859	C/T	R1848C	**0.987**	0.004	2
	rs113969111	G/C	W1839C	**1.000**	N.D.	

	rs149713462	G/A	V1817M	**1.000**	0.001	
	rs149713462	G/A	R1698W	**1.000**	0.001	
	rs139372870	T/C	V1617A	**0.998**	0	
	rs191685707	G/A	V1604M	**1.000**	N.D.	
	rs142589124	C/T	T1589M	**1.000**	0	
	rs2575314	G/A	V1566M	**1.000**	N.D.	
	rs150797875	C/T	R1557C	**1.000**	0	
	rs147596454	C/T	R1546W	**1.000**	N.D.	
	rs138882156	G/A	D1504N	**1.000**	N.D.	2
	rs149295739	A/G	Y1435C	**1.000**	0	
	rs147131259	C/T	R1411C	**1.000**	0.001	
	rs147131259	C/T	T1384I	**1.000**	0.001	
	rs145421882	G/T	E1379D	**1.000**	0	
	rs149210297	A/G	E1379G	**1.000**	N.D.	2
	rs145001864	G/C	V1376L	**1.000**	0	
	rs138789204	A/G	H1369R	**0.999**	0	
	rs201122748	C/T	T1354M	**0.982**	N.D.	
	rs141274774	G/A	S1352N	**0.954**	N.D.	2
	rs55840049	C/T	R1351W	**0.998**	N.D.	2,7
	rs185685883	C/G	T1344M	**0.960**	N.D.	
	rs138096771	C/T	P1343L	**1.000**	0	
	rs147141131	G/A	V1339M	**0.991**	N.D.	2
	rs143953714	C/T	R1285W	**0.998**	N.D.	2
	rs149472570	G/A	V1217M	**1.000**	N.D.	
	rs144338515	C/G	L1215V	**1.000**	0.001	
	rs144211349	C/G	D1162E	**0.981**	N.D.	2
	rs200726460	C/A	P1119H	**1.000**	N.D.	
	rs146352591	C/T	H1093Y	**1.000**	N.D.	2
	rs144389710	C/A	P1080H	**1.000**	0.002	2
	rs182938045	G/A	V1049M	**1.000**	N.D.	6
	rs147910505	C/T	S1047L	**1.000**	N.D.	2
	rs28681051	A/G	N1034S	**1.000**	N.D.	2
	rs28681051	A/T	N1034I	**0.990**	N.D.	2
	rs2369066	G/A	V1011I	**0.996**	0.001	1,2
	rs149041162	A/G	N1010S	**1.000**	0.001	2
	rs111686944	G/C	D962H	**0.999**	N.D.	
	rs147447715	G/A	V911M	**0.994**	0	
	rs151016310	G/A	V882M	**0.994**	N.D.	2
	rs189095400	C/T	R869C	**0.992**	N.D.	6
	rs144610067	C/T	T856M	**1.000**	0.001	2
	rs181738771	T/C	S843P	**0.998**	N.D.	6
	rs140233194	T/C	V821A	**0.977**	0.001	

	rs138575342	C/T	P694L	**1.000**	N.D.	
	rs2854624	C/G	P691A	**1.000**	N.D.	
	rs113000106	C/G	P642R	**0.993**	N.D.	
	rs58598099	T/C	C508R	**1.000**	N.D.	
	rs2855341	G/A	V466M	**1.000**	N.D.	7
	rs147457735	G/A	V424M	**0.969**	0.001	
	rs2519261	G/T	G381V	**1.000**	N.D.	
	rs4018169	C/T	T334M	**0.975**	N.D.	
PKD2	rs145877597	G/A	R322Q	**0.999**	0	
	rs138476749	T/G	L656W	**0.992**	0	
	rs148761063	G/A	G749S	**1.000**	0	
	rs142360839	G/A	D781N	**0.999**	N.D.	
	rs150838063	C/T	R798C	**1.000**	0	
	rs145343957	G/A	S804N	**0.997**	N.D.	2
	rs147654263	G/A	R807Q	**1.000**	N.D.	2
	rs149728995	T/C	S831P	**0.999**	0	
	rs148869484	G/A	V834I	**0.999**	0.001	2
	rs201598917	G/A	R848G	**0.997**	N.D.	
	rs143500495	G/A	R896H	**1.000**	0.002	2
	rs140848008	G/A	V909I	**0.994**	N.D.	
	rs144590958	G/C	R910P	**0.999**	0.001	2
	rs200394279	C/T	R945C	**1.000**	N.D.	
PKHD1	rs146680689	C/T	R3957C	**0.994**	0.001	2
	rs138360112	C/T	S3870L	**0.976**	0	
	rs76572975	G/A	R3842Q	**1.000**	N.D	2
	rs76572975	G/T	R3842L	**0.992**	N.D	2
	rs41273722	C/T	P3780S	**0.999**	N.D	2
	rs185174569	C/T	P3670S	**0.952**	N.D.	
	rs114237522	G/T	D3668Y	**1.000**	N.D	2
	rs201692162	G/C	G3667A	**1.000**	N.D.	
	rs147843070	C/A	P3639H	**0.974**	0	
	rs141349745	C/T	R3637C	**0.999**	N.D	
	rs141340895	G/A	V3567M	**0.972**	N.D.	
	rs200104032	A/T	S3490C	**0.973**	N.D.	
	rs148617572	C/T	R3482C	**1.000**	0	
	rs141132368	T/G	F3473V	**1.000**	0	
	rs149798764	T/C	C3346R	**0.999**	N.D	2
	rs199924729	T/C	F3325L	**0.999**	N.D.	
	rs145141656	C/T	P3221L	**0.999**	N.D.	
	rs141081295	C/G	S3210C	**1.000**	N.D.	
	rs147351244	C/T	H3160Y	**0.996**	N.D	2
	rs202212207	G/C	G3103R	**1.000**	N.D.	

	rs148669411	T/A	V3095D	**0.991**	0	
	rs142146981	A/G	I3081V	**0.999**	N.D	
	rs201333886	C/T	A3079V	**0.997**	N.D.	
	rs146550270	G/A	G3026R	**1.000**	0	
	rs141169758	C/T	S2983L	**0.997**	0	
	rs148990124	C/T	R2891C	**0.995**	0	
	rs142522748	C/A	T2869K	**0.982**	N.D	2
	rs146946292	G/A	G2808D	**1.000**	0	
	rs139555370	G/A	G2648S	**1.000**	N.D	
	rs7766366	A/G	T2641A	**0.999**	N.D	2
	rs201514497	G/T	G2612C	**1.000**	N.D.	
	rs200204857	C/T	P2582S	**0.981**	N.D.	2
	rs142487082	G/A	D2528N	**0.999**	N.D	2
	rs138859228	C/T	P2517S	**0.995**	N.D.	
	rs185747375	G/A	G2497R	**0.999**	N.D.	
	rs137972951	A/G	T2472A	**0.999**	0	
	rs201173272	T/G	I2459M	**0.997**	N.D.	
	rs201881567	T/G	C2422G	**0.998**	N.D.	
	rs199738484	T/G	F2395V	**1.000**	N.D.	
	rs199863048	A/C	T2369P	**0.999**	N.D.	
	rs141360909	C/T	P2356L	**1.000**	0	2
	rs138049936	T/A	I2322N	**0.995**	0	
	rs142976025	C/G	A2212G	**0.996**	0.001	2
	rs151108325	G/A	E2195K	**0.976**	0	
	rs138384224	G/A	G2073R	**0.999**	0	
	rs199589074	G/A	G2041S	**1.000**	N.D.	
	rs200047706	A/G	I1988V	**0.997**	N.D.	
	rs180675584	G/A	G1971D	**1.000**	N.D.	
	rs149553146	C/T	R1911C	**0.97**	N.D	2
	rs202016058	T/G	V1875G	**1.000**	N.D.	
	rs201105958	C/T	S1833L	**0.984**	N.D.	
	rs141103838	G/A	G1712R	**1.000**	N.D	2
	rs138671884	C/T	P1710L	**0.996**	0	
	rs45517932	C/T	L1709F	**1.000**	N.D	2
	rs200391019	C/T	R1624W	**0.992**	N.D.	
	rs150032027	G/C	C1561S	**0.997**	0	
	rs147098776	G/C	D1514H	**0.987**	N.D	
	rs201435075	C/G	S1498C	**0.999**	N.D	
	rs140331370	T/C	L1468P	**0.991**	0	
	rs138242579	G/C	S1435L	**0.998**	N.D	
	rs199552642	C/A	P1428T	**0.98**	N.D.	
	rs144042993	C/T	R1280C	**0.99**	N.D.	2

ISSN: 0975-8585

	rs139820610	G/A	V1269M	**0.964**	0	
	rs141395361	T/C	C1249R	**1.000**	N.D.	
	rs199873951	G/T	G1159V	**0.959**	N.D.	
	rs41273726	A/G	Y1136C	**0.997**	N.D	2
	rs78624439	T/G	I1128S	**0.981**	0.5	
	rs142107837	G/A	G1123S	**0.999**	0	
	rs201721915	T/C	S1105P	**0.998**	N.D.	
	rs150715879	A/G	Y1055C	**0.924**	0	
	rs144387673	A/G	T930A	**0.983**	0	
	rs151037130	A/C	N869T	**0.997**	0	
	rs199568593	T/C	V836A	**0.991**	N.D.	
	rs141645484	T/A	F792I	**1.000**	0	
	rs190315828	C/T	T764I	**0.991**	N.D.	
	rs140608845	C/T	T721M	**0.971**	N.D.	2
	rs201016555	T/G	F635C	**0.988**	N.D.	
	rs145855006	C/G	A616G	**0.999**	0	
	rs199846197	A/G	Y610C	**1.000**	N.D.	
	rs201824369	C/G	P591R	**0.993**	N.D.	
	rs201117261	T/G	F586V	**0.988**	N.D.	2
	rs45500692	C/T	T579M	**1.000**	N.D.	2
	rs143867809	G/T	G573W	**1.000**	N.D.	2
	rs142896856	G/A	R559Q	**0.976**	0.001	
	rs151070471	G/A	R494Q	**0.996**	0	
	rs139770251	C/T	R488W	**0.990**	N.D.	2
	rs201989004	T/C	Y452H	**1.000**	N.D.	
	rs149781976	G/C	G448R	**0.999**	N.D.	
	rs140458350	G/C	G420R	**1.000**	0	
	rs143341567	T/C	F380S	**1.000**	0	
	rs140781915	G/A	R312Q	**0.992**	N.D.	
	rs150102357	A/G	T311A	**1.000**	0	
	rs148070358	G/T	G55V	**1.000**	0	

Table 3: SNPs present in Splicing modifier binding sites

Gene	dbSNP rs# id	Possible Functional effect	RISK
PKD1	rs1804176	Missense (non-conservative); Splicing regulation	(3-4)
	rs2854586	Missense (non-conservative); Splicing regulation	(3-4)
	rs2854585	Missense (non-conservative); Splicing regulation	(3-4)
	rs2854584	Missense (non-conservative); Splicing regulation	(3-4)
	rs2854581	Missense (non-conservative); Splicing regulation	(3-4)
	rs2859787	Missense (non-conservative); Splicing regulation	(3-4)
	rs9925969	Missense (non-conservative); Splicing regulation	(3-4)
	rs9936785	Missense (non-conservative); Splicing regulation	(3-4)
	rs3952944	splicing site	(3-4)

	rs241572	Missense (non-conservative); Splicing regulation	(3-4)
	rs2369063	Missense (non-conservative); Splicing regulation	(3-4)
	rs3952942	Missense (non-conservative); Splicing regulation	(3-4)
	rs2519261	Missense (non-conservative); Splicing regulation	(3-4)
	rs2855335	Missense (non-conservative); Splicing regulation	(3-4)
	rs62038811	Sense/synonymous; Splicing regulation	(2-3)
	rs7203729	Sense/synonymous; Splicing regulation	(2-3)
	rs36221080	Missense (conservative); Splicing regulation	(2-3)
	rs3209986	Missense (conservative); Splicing regulation	(2-3)
	rs12927338	Missense (conservative); Splicing regulation	(2-3)
	rs2855367	Missense (conservative); Splicing regulation	(2-3)
	rs2855365	Missense (conservative); Splicing regulation	(2-3)
	rs36221082	Missense (conservative); Splicing regulation	(2-3)
	rs62624465	Missense (conservative); Splicing regulation	(2-3)
	rs34197769	Missense (conservative); Splicing regulation	(2-3)
	rs45478794	Missense (conservative); Splicing regulation	(2-3)
	rs7194935	Sense/synonymous; Splicing regulation	(2-3)
	rs61747420	Missense (conservative); Splicing regulation	(2-3)
	rs3865207	Missense (conservative); Splicing regulation	(2-3)
	rs12102740	Sense/synonymous; Splicing regulation	(2-3)
	rs1063405	Missense (conservative); Splicing regulation	(2-3)
	rs4787115	Missense (conservative); Splicing regulation	(2-3)
	rs55663903	Sense/synonymous; Splicing regulation	(2-3)
	rs17135779	Missense (conservative); Splicing regulation	(2-3)
	rs1063401	Missense (conservative); Splicing regulation	(2-3)
	rs1063400	Sense/synonymous; Splicing regulation	(2-3)
	rs1063399	Sense/synonymous; Splicing regulation	(2-3)
	rs3204450	Sense/synonymous; Splicing regulation	(2-3)
	rs13333553	Missense (conservative); Splicing regulation	(2-3)
	rs28575767	Sense/synonymous; Splicing regulation	(2-3)
	rs28369051	Missense (conservative); Splicing regulation	(2-3)
	rs3874655	Missense (conservative); Splicing regulation	(2-3)
	rs4990021	Sense/synonymous; Splicing regulation	(2-3)
	rs4018147	Missense (conservative); Splicing regulation	(2-3)
	rs62038819	Missense (conservative); Splicing regulation	(2-3)
	rs2575311	Sense/synonymous; Splicing regulation	(2-3)
	rs56309739	Sense/synonymous; Splicing regulation	(2-3)
	rs4018162	Sense/synonymous; Splicing regulation	(2-3)
	rs62625018	Sense/synonymous; Splicing regulation	(2-3)
	rs2854618	Missense (conservative); Splicing regulation	(2-3)
	rs2575313	Sense/synonymous; Splicing regulation	(2-3)
	rs241571	Missense (conservative); Splicing regulation	(2-3)
	rs71385733	Sense/synonymous; Splicing regulation	(2-3)
	rs2575314	Missense (conservative); Splicing regulation	(2-3)
	rs71385734	Sense/synonymous; Splicing regulation	(2-3)
	rs2239671	Missense (conservative); Splicing regulation	(2-3)
	rs55840049	Missense (conservative); Splicing regulation	(2-3)
	rs2575315	Missense (conservative); Splicing regulation	(2-3)
	rs241573	Missense (conservative); Splicing regulation	(2-3)
	rs2549677	Missense (conservative); Splicing regulation	(2-3)
	rs2575316	Sense/synonymous; Splicing regulation	(2-3)
	rs2369075	Sense/synonymous; Splicing regulation	(2-3)
	rs2855350	Missense (conservative); Splicing regulation	(2-3)
	rs28681051	Missense (conservative); Splicing regulation	(2-3)

	rs2369068	Sense/synonymous; Splicing regulation	(2-3)
	rs2369067	Sense/synonymous; Splicing regulation	(2-3)
	rs2369066	Missense (conservative); Splicing regulation	(2-3)
	rs4018173	Sense/synonymous; Splicing regulation	(2-3)
	rs4018172	Missense (conservative); Splicing regulation	(2-3)
	rs35667726	Sense/synonymous; Splicing regulation	(2-3)
	rs2549676	Missense (conservative); Splicing regulation	(2-3)
	rs2549675	Missense (conservative); Splicing regulation	(2-3)
	rs2855348	Missense (conservative); Splicing regulation	(2-3)
	rs2855347	Missense (conservative); Splicing regulation	(2-3)
	rs58598099	Missense (conservative); Splicing regulation	(2-3)
	rs2855341	Missense (conservative); Splicing regulation	(2-3)
	rs2519260	Sense/synonymous; Splicing regulation	(2-3)
	rs4018169	Missense (conservative); Splicing regulation	(2-3)
	rs2855336	Missense (conservative); Splicing regulation	(2-3)
	rs2855334	Sense/synonymous; Splicing regulation	(2-3)
	rs4999038	Missense (conservative); Splicing regulation	(2-3)
PKD2	rs1805044	Missense (non-conservative); Splicing regulation	(3-4)
	rs1131408	Missense (non-conservative); Splicing regulation	(3-4)
	rs3209753	splicing site	(3-4)
	rs2728118	Sense/synonymous; Splicing regulation	(2-3)
	rs73841280	Sense/synonymous; Splicing regulation	(2-3)
PKHD1	rs7766366	Missense (non-conservative); Splicing regulation	3-4
	rs28937907	Missense (non-conservative); Splicing regulation	3-4
	rs9370096	splicing site	3-4
	rs45627337	Missense (conservative); Splicing regulation	2-3
	rs35445653	Missense (conservative); Splicing regulation	2-3
	rs17752991	Sense/synonymous; Splicing regulation	2-3
	rs45600034	Missense (conservative); Splicing regulation	2-3
	rs45503297	Missense (conservative); Splicing regulation	2-3
	rs765525	Sense/synonymous; Splicing regulation	2-3
	rs9349603	Sense/synonymous; Splicing regulation	2-3
	rs12210295	Sense/synonymous; Splicing regulation	2-3
	rs34796823	Sense/synonymous; Splicing regulation	2-3
	rs970881	Missense (conservative); Splicing regulation	2-3
	rs2435322	Missense (conservative); Splicing regulation	2-3
	rs28939099	Missense (conservative); Splicing regulation	2-3
	rs45517932	Missense (conservative); Splicing regulation	2-3
	rs9689306	Sense/synonymous; Splicing regulation	2-3
	rs41273726	Missense (conservative); Splicing regulation	2-3
	rs62406032	Missense (conservative); Splicing regulation	2-3
	rs4715271	Sense/synonymous; Splicing regulation	2-3
	rs45500692	Missense (conservative); Splicing regulation	2-3
	rs62406036	Sense/synonymous; Splicing regulation	2-3
	rs1896976	Sense/synonymous; Splicing regulation	2-3
	rs9474143	Sense/synonymous; Splicing regulation	2-3
	rs6901799	Sense/synonymous; Splicing regulation	2-3
	rs28939383	Missense (conservative); Splicing regulation	2-3

Table 4: Evaluation of Changes in Protein Stability upon Point Mutation

GENE	dbSNPid	Position	WT	New	DDG Value	RI	Stability	pH	T
PKD1	rs200685883	T4224N	T	N	-0.53	0 Kcal/mol	Increase	7	25
	rs182827985	P4162L	P	L	-0.54	0 Kcal/mol	Increase	7	25

	rs186487281	H4081Y	H	Y	0.43	7 Kcal/mol	Increase	7	25
	rs150724260	T4073I	T	I	-0.01	2 Kcal/mol	Increase	7	25
	rs142768096	G4032V	G	V	-0.07	1 Kcal/mol	Increase	7	25
	rs201771763	A4022V	A	V	0.4	6 Kcal/mol	Increase	7	25
	rs151044657	K4017N	K	N	-0.21	3 Kcal/mol	Increase	7	25
	rs116835405	S3779R	S	R	-0.18	1 Kcal/mol	Increase	7	25
	rs114656915	T3778N	T	N	-0.78	2 Kcal/mol	Increase	7	25
	rs143119214	H3695Y	H	Y	0.25	6 Kcal/mol	Increase	7	25
	rs115883514	A3619V	A	V	-0.02	2 Kcal/mol	Increase	7	25
	rs201209708	A3578V	A	V	-0.1	0 Kcal/mol	Increase	7	25
	rs201409107	G3540A	G	A	-0.52	0 Kcal/mol	Increase	7	25
	rs34197769	A3512V	A	V	0.46	6 Kcal/mol	Increase	7	25
	rs144593342	A3442V	A	V	0.22	0 Kcal/mol	Increase	7	25
	rs137986581	S3413F	S	F	-0.01	1 Kcal/mol	Increase	7	25
	rs139324800	A3391V	A	V	0.12	1 Kcal/mol	Increase	7	25
	rs199738488	S3178I	S	I	0.35	3 Kcal/mol	Increase	7	25
	rs145997957	H3174L	H	L	0.23	3 Kcal/mol	Increase	7	25
	rs149872222	H3151Y	H	Y	-0.13	6 Kcal/mol	Increase	7	25
	rs4787115	F3111L	L	F	-0.99	3 Kcal/mol	Increase	7	25
	rs200360336	A3053P	A	P	0.01	7 Kcal/mol	Increase	7	25
	rs200522524	R3039S	R	S	-0.9	7 Kcal/mol	Increase	7	25
	rs1063401	Q3005E	Q	E	-0.12	7 Kcal/mol	Increase	7	25
	rs200520583	Q2969E	Q	E	-0.42	3 Kcal/mol	Increase	7	25
	rs146250848	S2948W	S	W	-0.04	2 Kcal/mol	Increase	7	25
	rs147788838	V2905I	V	I	-0.59	7 Kcal/mol	Increase	7	25
	rs142888788	V2844I	V	I	-0.77	7 Kcal/mol	Increase	7	25
	rs141717814	Q2735L	Q	L	0.1	3 Kcal/mol	Increase	7	25
	rs147350387	T2708M	T	M	0.27	2 Kcal/mol	Increase	7	25
	rs200509641	A2704V	A	V	0.5	4 Kcal/mol	Increase	7	25
	rs148642998	T2687M	T	M	0.00	0 Kcal/mol	Increase	7	25
	rs145747362	H2638Y	H	Y	0.3	4 Kcal/mol	Increase	7	25
	rs150516444	A2581V	A	V	0.41	4 Kcal/mol	Increase	7	25
	rs200037070	H2546Y	H	Y	0.29	7 Kcal/mol	Increase	7	25
	rs202162543	H2521R	H	R	-0.02	4 Kcal/mol	Increase	7	25
	rs140494005	A2427V	A	V	0.21	1 Kcal/mol	Increase	7	25
	rs143021666	S2303L	S	L	0.5	4 Kcal/mol	Increase	7	25
	rs148496347	A2222V	A	V	0.16	4 Kcal/mol	Increase	7	25
	rs146515679	S1810R	S	R	-0.36	1 Kcal/mol	Increase	7	25
	rs140162759	T1773I	T	I	0.19	5 Kcal/mol	Increase	7	25
	rs143990449	A1743P	A	P	0.11	4 Kcal/mol	Increase	7	25
	rs148538206	D1713E	D	E	0.2	7 Kcal/mol	Increase	7	25
	rs139520275	S1684L	S	L	0.28	2 Kcal/mol	Increase	7	25
	rs146723506	S1619F	S	F	0.26	5 Kcal/mol	Increase	7	25
	rs150031440	T1384I	T	I	-0.09	0 Kcal/mol	Increase	7	25
	rs141274774	S1352N	S	N	-0.35	2 Kcal/mol	Increase	7	25
	rs143624531	T1244I	T	I	-0.01	1 Kcal/mol	Increase	7	25
	rs138910964	S1232T	S	T	-0.07	1 Kcal/mol	Increase	7	25
	rs144211349	D1162E	D	E	-0.27	4 Kcal/mol	Increase	7	25
	rs146352591	H1093Y	H	Y	0.37	6 Kcal/mol	Increase	7	25
	rs2855350	T1056N	T	N	-0.61	2 Kcal/mol	Increase	7	25
	rs147910505	S1047L	S	L	0.26	1 Kcal/mol	Increase	7	25
	rs28681051	N1034S	N	S	0.87	4 Kcal/mol	Increase	7	25
	rs151262575	A898V	A	V	0.17	3 Kcal/mol	Increase	7	25
	rs58598099	C508R	C	R	0.03	1 Kcal/mol	Increase	7	25

	rs200463151	S104P	S	P	-0.4	1 Kcal/mol	Increase	7	25
PKD2	rs201632217	N269S	N	S	-0.15	0 Kcal/mol	Increase	7	25
	rs200759509	T283I	T	I	-0.12	1 Kcal/mol	Increase	7	25
	rs148386297	I408S	I	S	0.6	6 Kcal/mol	Increase	7	25
	rs150314615	T635S	T	S	-0.43	1 Kcal/mol	Increase	7	25
	rs142360839	D781N	D	N	-0.15	7 Kcal/mol	Increase	7	25
	rs145343957	S804N	S	N	-0.73	1 Kcal/mol	Increase	7	25
	rs140848008	V909I	V	I	-0.18	1 Kcal/mol	Increase	7	25
PKHD1	rs199900211	L4037P	L	P	0.2	5 Kcal/mol	Increase	7	25
	rs201210830	R4036W	R	W	0.00	2 Kcal/mol	Increase	7	25
	rs145124509	Q3945E	Q	E	-0.38	2 Kcal/mol	Increase	7	25
	rs4715227	Q3899R	Q	R	0.05	0 Kcal/mol	Increase	7	25
	rs139442289	H3647P	H	P	0.24	7 Kcal/mol	Increase	7	25
	rs145219220	H3630Y	H	Y	0.03	4 Kcal/mol	Increase	7	25
	rs200244173	S3505N	S	N	-0.34	2 Kcal/mol	Increase	7	25
	rs71570543	N3435S	N	S	-0.66	0 Kcal/mol	Increase	7	25
	rs147351244	H3160Y	H	Y	0.02	6 Kcal/mol	Increase	7	25
	rs45503297	D3139Y	D	Y	-0.29	1 Kcal/mol	Increase	7	25
	rs144300382	S2635Y	S	Y	-0.11	5 Kcal/mol	Increase	7	25
	rs144300382	S2635F	S	F	0.21	5 Kcal/mol	Increase	7	25
	rs78361537	A2579V	A	V	0.07	2 Kcal/mol	Increase	7	25
	rs147851214	T2436I	T	I	-0.31	2 Kcal/mol	Increase	7	25
	rs150838488	Q1923L	Q	L	-0.13	4 Kcal/mol	Increase	7	25
	rs140996978	T1896M	T	M	0.17	0 Kcal/mol	Increase	7	25
	rs147933501	S1862L	S	L	0.21	3 Kcal/mol	Increase	7	25
	rs200928412	A1758V	A	V	0.05	3 Kcal/mol	Increase	7	25
	rs185937478	H1699L	H	L	0.29	1 Kcal/mol	Increase	7	25
	rs112357604	H1568Y	H	Y	0.2	6 Kcal/mol	Increase	7	25
	rs191201723	S1400L	S	L	0.32	2 Kcal/mol	Increase	7	25
	rs200733734	H1318L	H	L	-0.01	0 Kcal/mol	Increase	7	25
	rs199643308	A1273P	A	P	-0.35	1 Kcal/mol	Increase	7	25
	rs9296669	A1262V	A	V	-0.1	3 Kcal/mol	Increase	7	25
	rs112182862	A1260V	A	V	0.02	3 Kcal/mol	Increase	7	25
	rs140815495	A1245V	A	V	-0.02	1 Kcal/mol	Increase	7	25
	rs201721915	S1105P	S	P	-0.34	3 Kcal/mol	Increase	7	25
	rs201102541	N1084T	N	T	-0.27	1 Kcal/mol	Increase	7	25
	rs145789167	Q815R	Q	R	-0.09	3 Kcal/mol	Increase	7	25
	rs190315828	T764I	T	I	-0.04	0 Kcal/mol	Increase	7	25
	rs9370096	R760C	R	C	-0.5	3 Kcal/mol	Increase	7	25
	rs141622697	A748V	A	V	0.33	3 Kcal/mol	Increase	7	25
	rs144455663	V712L	V	L	-0.54	3 Kcal/mol	Increase	7	25
	rs149147435	D659N	D	N	-0.32	4 Kcal/mol	Increase	7	25
	rs149781976	G448R	G	R	-0.21	1 Kcal/mol	Increase	7	25
	rs141093030	Y143C	Y	C	-0.88	2 Kcal/mol	Increase	7	25
	rs202133636	Q141R	Q	R	-0.06	1 Kcal/mol	Increase	7	25
	rs149841071	G112R	G	R	-0.47	2 Kcal/mol	Increase	7	25
	rs113399906	N54D	N	D	-0.2	2 Kcal/mol	Increase	7	25
	rs141790557	N53S	N	S	-0.31	1 Kcal/mol	Increase	7	25

DISCUSSION

In recent time, Computational analyses of SNPs in human genes are focused in identification of deleterious or damaging nsSNPs. Hence in the present scenario it has become one of the major interests in human genetics to distinguish those particular mutations that are functionally neutral from those that contribute to causing the disease. Presently, the amino acid substitutions account for approximately half of the

known gene lesions responsible for human inherited diseases (Cooper, 1998). Therefore, the identification of nsSNPs that affect protein functions and relate to disease is an important task. Evaluation of non-neutral SNPs is mainly based on Phylogenetic information extended to a certain degree with structural approaches (PolyPhen). However, there is increasing evidence that many human disease genes are the result of exonic or noncoding mutations affecting regulatory regions (35, 36). Much attention has been focused on modeling the possible phenotypic effect of SNPs that cause amino acid changes using different computational methods, and recently the workers has focused on functional SNPs affecting regulatory regions or the splicing process. Moreover, because of their widespread distribution on the species genome, SNPs become particularly important and valuable as genetic makers in the research for the diseases and corresponding drug against that disease. Currently, millions of human SNPs have reported by high-throughput methods used in wet laboratories. The vast number of SNPs causes a challenge for biologists and bioinformaticians although they provide lot information about the relationships between individuals. Besides numerous ongoing efforts to identify millions of these SNPs, there is now also a focus on studying associations between disease risk and these genetic variations using a molecular epidemiological approach. This SNPs points out a major difficulty faced by scientists in planning costly population-based genotyping, which is to choose target SNPs that are most likely to affect phenotypic functions and ultimately contribute to disease development. It is becoming clear that application of the molecular evolutionary approach may be a powerful tool for prioritizing SNPs to be genotyped in future molecular epidemiological studies. Therefore, our analysis will provide useful information in selecting SNPs that are likely to have potential functional impact and ultimately contribute to an individual's disease susceptibility.

Because autosomal dominant polycystic kidney disease is slowly progressive, there is a window of opportunity to treat the disease by retarding cystic expansion. However, new methods of assessing renal decline, such as monitoring the reduction in the numbers and size of cysts, must be developed and accepted by regulatory agencies.

CONCLUSION

In this work genes involved in Polycystic kidney disease i.e. PKD1, PKD2, PKHD1 was investigated by evaluating the influence of functional SNPs through computation methods. Of a total of 1371 SNPs in the three genes of polycystic kidney disease, 788 were found to be nonsynonymous. Of 788 nsSNPs, 23 were found to be deleterious by SIFT and 289 were damaging as per the PolyPhen server. Nineteen nsSNPs were found to be common in both the SIFT and the PolyPhen server. We conclude that rs58598099 with a mutation of cysteine to arginine at position 508 in the gene PKD1 involved in causing ADPKD was found damaging by all the functional nsSNP prediction severs such that- SIFT, Polyphen 2, I-Mutant 3.0 suite, Fastsnp, Dbsmr and which could be the main target mutation for the Autosomal Dominant Polycystic kidney disease caused by the mutation in PKD1 gene. Further the work in my thesis can be validated since among all those functional SNPs which were found to be damaging or affecting in PKD genes, 7 SNPs (rs34197769, rs40433, rs28939383, rs28937907, rs28939099, rs2435322) were also found damaging associated with the similar genes in Snpedia database (a wiki investigating human genetics database that share information about the effects of variations in DNA, citing peer-reviewed scientific publications) available at (http://www.SNPedia.com). Moreover, this study builds a bridge between evolutionary biology to molecular epidemiology, which may further our understanding of disease-related SNPs and ultimately facilitate SNP genotyping in future studies. By this work we can conclude that by this computational analysis used in this study may also help in analysis of other human genes to identify potentially functional SNPs in it. Moreover, this cystic disease is not only affecting the kidney rather the cerebrovascular, Gastro-intestinal tract and cardiovascular systems were also get severely affected during this disease which had been observed yet in various region. However, this study will be required for further in-vitro validation process and can be helpful in development of much advanced drug for the treatment to reduce the cases of renal failure all over.

REFERENCES

[1] Burn TC, Connors TD, Dackowski WR, Petry LR, Van Raay TJ, Millholland JM, Venet M, (1995). The American PKD1 Consortium: Analysis of the genomic sequence for the autosomal dominant polycystic kidney disease (PKD1) gene predicts the presence of a leucine-rich repeat. Hum Mol Genet 4: 575–582.

[2] Mochizuki T, Wu G, Hayashi T, Xenophontos S, Gabow P (1996): PKD2, a gene for polycystic kidney disease that encodes an integral membrane protein. Science 272: 1339–1342.
[3] Collins F.S, L.D. Brooks (1998), A DNA polymorphism discovery resource for research on human genetic variations, Genomic Res. 8 1229-1231.
[4] Lander ES, (1996). The new genomics: global views of biology, Science 274:536-539.
[5] Barroso JM. Gurnell VE, Croeley M., Agostini JW, Schwabe MA (1999). 'Dominant negative mutations in human PPAR gamma associated with severe insulin resistance diabetes mellitus and hypertension'. Nature 402: 880–883.
[6] Thomas R, McConnell R., Whittacker J., Kirkpatrick P., Bradley J., Stanford R., (1999). Identification of mutations in the repeated part of the autosomal dominant polycystic kidney disease type 1 gene PKD1 by long range PCR, Am. J. Hum. Genet. 65:39-49.
[7] Dryja TP, Mcgee TL, Halu LB, Conley GS, Olsson JE, Reichel E (1990).Mutations within the rhodopsin gene in patients with autosomal dominant retinitis pigmentosa, N. Engl. J. Med. 323:1302–1307.
[8] Smith EP, Boyd J., Frank GR, Takahashi M, Cohen RM., Specker B. (1994). Estrogen resistance caused by a mutation in the estrogen-receptor gene in a man, N. Engl. J. Med. 331: 1056–1061.
[9] Grantham J., (2008). 'Autosomal Dominant Polycystic Kidney Disease'. New Eng. J. Med, 85: 359-1477.
[10] Manoj H., Scaria V., Brahmachari SK, (2009), dbSMR: a novel resource of genome-wide SNPs affecting microRNA mediated regulation'. BMC Bioinformatics, 10(1): 108, 1-6.
[11] Patricia WD, (2004). Mechanisms of disease Polycystic Kidney Disease. N Engl J Med, 350:151-64.
[12] Ravine D, Walker RG, Gibson RN, Sheffield, LJ, Kincaid-Smith P, Danks, DM (1991) Treatable complications in undiagnosed cases of autosomal dominant polycystic kidney disease'. Lancet 337: 127-129.
[13] Dalgaard, OZ (1963). Bilateral polycystic disease of the kidneys. In: Strauss, M. B.; Welt, L. G: Diseases of the Kidney. Boston, Mass.: Little, Brown and Co. (pub.). Pp. 907-910.
[14] Poinso, R., Monges, H., Payan, H. La maladie kystique du foie (1954). Expansion Scientifique Francaise (pub.).
[15] Chapman AB, Johnson A, Gabow PA, Schrier, RW, (1990) The renin-angiotensin-aldosterone system and autosomal dominant polycystic kidney disease.' New Eng. J. Med. 323: 1091-1096.
[16] Chapman JR, Hilson AJW (1980) Polycystic kidneys and abdominal aortic aneurysms'. Lancet 315: 646-647.
[17] Hossack KF, Leddy CL, Johnson AM, Schrier, RW, Gabow PA (1988) Echocardiographic findings in autosomal dominant polycystic kidney disease'. New Eng. J. Med. 319: 907-912.
[18] Xenophontos S., Constantinides R., Hayashi T., Mochizuki T. (1997) A transition frameshift mutation induced by a cytosine insertion in the polycystic kidney disease 2 gene (PKD2)'. Hum. Molec. Genet. 6: 949-952.
[19] Viribay M., Hayashi D., Mochizuki T., Reynolds MF, Harris PC (1997) Novel stop and frameshifting mutations in the autosomal dominant polycystic kidney disease 2 (PKD2) gene'. Hum. Genet. 101: 229-234.
[20] Pei Y., Wang K (1998). 'A novel frameshift mutation induced by an adenosine insertion in the polycystic kidney disease 2 (PKD2) gene'. Kidney Int. 53: 1127-1132.
[21] Koptides M, Hadjimichael C, Koupepidou P, Pierides A, Deltas CC (1999) Germinal and somatic mutations in the PKD2 gene of renal cysts in autosomal dominant polycystic kidney disease'. Hum. Molec. Genet. 8: 509-513.
[22] Reynolds (1999), Aberrant splicing in the PKD2 gene as a cause of polycystic kidney disease'. J. Am. Soc. Nephrol. 10: 2342-2351.
[23] Bergmann C., Bruchle NO, Frank V, Rehder H, Zerres K (2008), Perinatal deaths in a family with autosomal dominant polycystic kidney disease and a PKD2 mutation'. New Eng. J. Med. 359: 318-319.
[24] Sherry ST, Ward MH, Kholodov M, Baker J, Phan L, Smigielski M, Sirotkin K (2001), DbSNP: the NCBI database of genetic variation, Nucleic Acids Res. 29:308–311.
[25] Ng CP, Henikoff S, (2003), SIFT: predicting amino acid changes that affect protein function, Nucleic Acids Res. 31:3812–3814.
[26] Ng CP, Henikoff S (2001), Predicting deleterious amino acid substitutions. Genome Res. 11: 863–874.
[27] Ramensky V, Bork P, Sunyaev S, (2002), Human non-synonymous SNPs: server and survey, Nucleic Acids Res. 30:3894–3900.
[28] Sunyaev S, Ramensky V, Bork P, (2000), Towards a structural basis of human non-synonymous single nucleotide polymorphisms. Trends Genet. 16:198–200.

[29] Prokunina L, Alarcn-Riquelme ME, (2004) Regulatory SNPs in complex diseases: their identification and functional validation. Expert Rev. Mol. Med. 1–15.
[30] Prokunina L, Castillejo-Lopez C (2002), A regulatory polymorphism in PDCD1 is associated with susceptibility to systemic lupus erythematosus in humans. Nat. Genet. 32:666–669.
[31] Yuan H, Chiou JJ, Tseng WH, Liu CH, Liu CK, Lin YJ, Wang HH, Yao A, Chen YT, Hsu CN, (2006). FASTSNP: an always up-to-date and extendable service for SNP function analysis and prioritization, Nucleic Acids Res. 34:W635–W641.
[32] Kumar MD, Bava KA, Gromiha MM, Prabakaran P, Kitajima K, Uedaira H, Sarai A. (2006) ProTherm and ProNIT: thermodynamic databases for proteins and protein-nucleic acid interactions. Nucleic Acids Res. (34); D204-206.
[33] Bava KA., Gromiha MM., Uedaira H, Kitajima K., Sarai a. (2004).'Protherm, version 4.0: thermodynamic database for proteins and mutants' Nucleic Acids Res.32, D120-D121.
[34] Krawczak M, Ball EV, Fenton I, Stenson PD, Abeysinghe S, Thomas N, Cooper DN (2000) Human gene mutation database-a biomedical information and research resource. Hum Mutat 15(1):45–51.
[35] Hudson TJ, (2003). Wanted: regulatory SNPs. Nat Genet, 33:439-440.
[36] Yan H, Yuan W, Velculescu VE, (2002). Allelic variation in human gene expression. Science, 297:1143.