NPHS2 Functional Variant (R229Q) in Primary Nephrotic Syndrome: A Study from South India

*Parveen Jahan, 2Asra Tabassum Jaffer, 3Wali Unnisa, 4Un Mahwish, 5G Sumanlatha, 6S Someswar Rao, 7D Sreebhusan Raju

*Department of Zoology, School of Life Sciences, Maulana Azad National Urdu University, Hyderabad, 2, 3, 4, 5, 6 Department of Genetics, Osmania University, Hyderabad, 6 Department of Nephrology, Nizams Institute of Medical Sciences, Hyderabad

*Email: dr.pjahan@gmail.com

Nephrotic Syndrome (NS) comprises of Heavy Proteinuria, Hypo Albuminemia, Edema and Hyperlipidemia. The worldwide prevalence of NS was reported to be 12-16 per 100,000 individuals; epidemiological evidence of a higher incidence of NS in children from South Asia. The present study focuses on the contribution of a common functional polymorphism R229Q (rs61747728) of NPHS2 gene towards NS susceptibility, the underlying histopathology and steroid response in south Indian population. Blood samples from a total of 484 individuals (234 cases and 250 controls) were subjected to molecular analysis using PCR-RFLP method.

Data analysis revealed mutant A allele frequency to be 2.8% in normal healthy individuals and 11% in patients with NS, conferring a fourfold increased risk towards disease susceptibility. Furthermore, the occurrence of A allele was 2.55% and 23.65% in mild (MCD) and severe (FSGS) histopathology patient groups respectively. However, with respect to treatment response no significant variation was noted in the frequency of mutant allele in steroid sensitive and steroid resistant group of patients. Computational predictions of the potential impact of this variant were found to be possibly damaging. Thus R229Q variant appears to be a probable contributing factor towards NS pathogenesis among south Indians.

Key words: Nephrotic Syndrome, Gene, Polymorphism, Steroid resistance, minimal change disease, focal segmental glomerulosclerosis.

Introduction

Nephrotic syndrome (NS) is a malfunction of the kidney glomerular filter and is one of the commonest causes of renal problems in children encountered by nephrologists (1). The characteristic features of NS are heavy proteinuria, hypoalbuminemia, edema and hyperlipidemia (2). The cause of proteinuria is due to injury to the function or structure of glomerular filtration barrier. The estimated annual incidence of NS ranges from 2-7 per 100,000 children and prevalence from 12-16 per 100,000 worldwide. There is epidemiological evidence of a higher incidence of NS in children from South Asia (3). It appears to be a clinically heterogeneous condition characterized by histological variants (4) and different genetic backgrounds (5, 6) with recessive and dominant inheritances. For these patients the therapeutic option is a standard steroid therapy and on the basis of the patient's response to steroid therapy, it is divided into Steroid-sensitive Nephrotic Syndrome (SSNS) and Steroid-resistant Nephrotic Syndrome (SRNS) which may progress to End-stage Renal Disease (ESRD) (7), which is the final result of various etiologies. Numerous prognostic indicators leading to ESRD have been extensively studied, of which genetic factors remain a subject of great concern (8).

The primary defect in NS is an injury to the glomerular basement membrane (GBM) which is responsible for the size and charge-selective properties of renal filter. One of the key components of GBM are podocytes. Podocytes are highly specialized, terminally differentiated cells with cytoplasmic extensions, foot processes. These cells have an important role in size and charge selective permeability, but also in synthesizing and maintaining the GBM (9). Characterization of mutations in the genes coding for podocyte proteins in patients with inherited nephrotic syndrome has been a breakthrough in both molecular and clinical research (10).

Podocin protein, a member of the stomatin protein family is encoded by NPHS2 gene which is located at chromosome 1q25-q31 and was first mapped by linkage analysis in families with autosomal recessive SRNS. NPHS2 coding region is 1,149 bp in length, has 8 exons and encodes the 42kD integral membrane protein podocin which is expressed in both fetal and mature kidney. Podocin recruits other important podocyte proteins like nephrin and CD2AP in lipid rafts ensuring a stable and proper functioning filtration barrier (11). In human, NPHS2 mutations, are mainly associated with autosomal recessive SRNS, but were also found in sporadic cases of SRNS patients.
Mutated podocin may cause disturbance in recruiting nephrin to the plasma membrane, which is comparable with the lack of nephrin in the slit membrane found in podocin deficient mice (12).

Tsukaguchi et al. studied the R229Q variant that is a common polymorphism occurring at an increased frequency in association with late onset FSGS. They used pull-down experiments with anti-podocin antibodies and found a reduced binding to nephrin in the presence of this polymorphism compared with the wild protein. It was put forward that this common variant may contribute to glomerulosclerosis in association with other mutant NPHS2 alleles or acting in synergism with other factors (13).

The present study focuses on the contribution of common polymorphism R229Q (rs61747728) towards susceptibility to NS, underlying histopathology and steroid response in south Indian population. This kind of study holds importance in defining the molecular basis for the disease.

Materials and methods
For the present study, a total of 484 subjects, of which 234 were clinically proven cases of NS fulfilling criteria of the International Study of Kidney Disease in Children (ISKDC) for the diagnosis of NS visiting the Department of nephrology at Nizam’s Institute of Medical Sciences, Hyderabad, India were recruited. The cases were confirmed as nephrotics by the consultant nephrologist after a series of clinical and laboratory investigations. This study was approved by the Institutional Ethics Committee. Detailed information relating to the age, gender, family history, consanguinity was collected from both controls and patients; in addition, information on histopathology and responsiveness to steroid therapy was recorded from the patients. The cases included 147 males and 87 females belonging to all socioeconomic backgrounds with a south Indian origin. 250 were healthy volunteers without any family history of renal disorders, no proteinuria on dipstick analysis. Controls were preferred from a higher age group in order to rule out the possibility of pediatric onset of nephrotic syndrome in them but were sex matched with the case group. Patients with secondary reasons for kidney disease, lupus nephritis, diabetic nephropathy, renal failure and any other associated conditions were excluded from the study population. Similarly, controls with diabetes, hypertension and autoimmune diseases were not included either. Based on the underlying histopathology the cases were classified as mild if the histopathology was minimal change disease (MCD) and severe if the histopathology was mesangio proliferative glomerulo nephrotis (MPGN) and focal segmental glomerulo sclerosis (FSGS). Steroid Sensitivity (SS) was defined as cessation of proteinuria for at least three consecutive days after standard steroid treatment. No achievement of remission even after four weeks of steroid treatment was classified as Non-steroid sensitive (Non-SS). Based on the age of disease onset the cases were categorized as early onset if the age was less than 18 years and late onset if the age of onset was greater than or equal to 18 years. Detailed information and clinical data from all the cases and controls was collected using well designed Proforma on age, sex, age at onset, histopathology, steroid responsiveness, consanguinity, family history, etc. Written informed consent was obtained from all the participating subjects prior to blood sample collection. Biochemical test values were obtained from the records. Biopsy-proven documentation on histopathology was available for 230 cases. The follow up clinical data to evaluate steroid responsiveness was obtained from 233 cases. Three generation pedigree was generated using the details of an occurrence of consanguinity, family history of NS.

5 milliliters of venous blood were collected from all the participants in sterile EDTA vacutainers, that were immediately stored at either -20°C or -70°C till further molecular analysis. Genomic DNA was extracted from blood collected from all the 484 subjects in our laboratory using rapid, non-enzymatic, salting out method (14). All the DNA samples were subjected to PCR amplification using the forward primer 5'-GGA TTT ACC ACA GGA TTA AGT TGT GCA- 3' and reverse primer 5'AGC TAT GAG CTC CCA AAG GGA TGG - 3' (13). The annealing temperature was 55°C. From these amplified products, 3 microliters of unpurified PCR product was used for restriction digestion. The amplicons were digested overnight with Clai which recognizes the sequence AT^CG_AT, at 37°C for genotyping. The G755A missense substitution results in a loss of Clai site, mutant R229Q alleles give a product of 545 bp as they miss the Clai site, whereas the normal allele gives a digested product of 364 and 181 bps bands (Figure 1). Each batch of restriction digestion contained a positive control (confirmed G allele) to avoid mistyping. 10% of the samples were randomly picked and re-genotyped to give consistent results. The products were resolved on 2% agarose gel and viewed in a gel documentation unit (Bio-Rad).

Results
The present study comprised a total of 484 subjects of which 250 were healthy controls recruited from personal contacts and various hospitals (Table 1). Mean age of the control group was 32.16 ± 8.5 years.
a higher age group was considered in order to rule out the possibility of childhood onset of NS. The percentage of males to females was 51.12% and 48.8%, which was almost equal. During the study period, 234 proven cases of NS were included in the study with a mean age at disease onset of 17.6±11.4 years. The male to female ratio was 2:1, which is in accordance with previous reports (15).

**Molecular analysis**

In NS patients the frequencies of R229Q GG, GA and AA genotypes were 79.5%, 18.8% and 1.7% respectively. The percentage of G allele was 89.9% and A allele was 11%, whereas among the normal subjects, the distribution of GG, GA and AA genotypes was 94.8%, 4.8% and 0.4%. The allele frequency of G and A in controls was 97% and 2.8% correspondingly. The distributions of R229Q genotype and allele frequency of patients and controls are given in Table 2. The frequencies of G and A alleles were tested for goodness of fit and were found to be in Hardy Weinberg equilibrium. The result of the chi-square test for association showed a value of 6.50 (p<0.01). Significant difference was found in the allele and the genotype frequency distribution between patients and controls [G vs. A; Odds ratio =3.9, p = 0.048; and GA+AA vs. GG; Odds ratio = 4.3; p<0.01 at 95% CI].

![Figure 1. Gel image showing the R229Q genotypes](image)

<table>
<thead>
<tr>
<th>Table 1. Demographic features of the study participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>Controls</td>
</tr>
<tr>
<td>Number (n)</td>
</tr>
<tr>
<td>Males</td>
</tr>
<tr>
<td>Females</td>
</tr>
<tr>
<td>Consanguinity</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 2. Distribution of R229Q genotypes among various groups</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Category</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Susceptibility</strong></td>
</tr>
<tr>
<td>Controls 250</td>
</tr>
<tr>
<td>Patients 234</td>
</tr>
<tr>
<td><strong>Age at onset</strong></td>
</tr>
<tr>
<td>Early 116</td>
</tr>
<tr>
<td>Late 118</td>
</tr>
<tr>
<td><strong>Histopathology</strong></td>
</tr>
<tr>
<td>Mild 137</td>
</tr>
<tr>
<td>Severe 93</td>
</tr>
<tr>
<td><strong>Response to steroid therapy</strong></td>
</tr>
<tr>
<td>Non-SS 178</td>
</tr>
<tr>
<td>SS 85</td>
</tr>
</tbody>
</table>

*GA+AA vs. GG
Analysis of R229Q genotypes in different histopathological conditions of NS revealed 57, 38.7, 4.3 and 95, 5, 0% of GG, GA and AA genotypes in severe and mild forms, respectively. The allele frequencies of G and A were 97.4 and 2.6% in mild form, 76.34 and 23.65% in the most severe histopathology group ($\chi^2 = 46.6, p < 0.0000$). Distribution of R229Q genotypes GG, GA and AA in SS and Non-SS groups was observed to be 80, 18, 2 and 79, 19, 2% correspondingly and the difference was not statistically significant.

**Insilico analysis**

A homology model of human Podocin (Figure 2a) was built using the I-TASSER server (16); its structural as well as functional annotations were interpreted by means of PredictProtein (17). The secondary structure composition was found to be 12.7% strand, 49.35% loop and 38.38% helix. Solvent accessibility was found to be 59.79% exposed, 34.2% buried and 6.01% intermediate. Transmembrane helices ranged from amino acid residues 103-123 with a length of 21. No disulphide bridges were noted between pairs of cysteine residues.

Prediction of the impact of Arg to Gln substitution (Figure 2b) gave a notably strong signal for effect on function, with a score of 54, indicating the likelihood of this specific mutation to alter the native protein function. Furthermore, it is predicted that this variation is located in a disordered region with residues ranging from 229-242, making it difficult to predict the exact impact of this SNP as it lies in the intrinsic disordered region. The assignment of disordered or unfolded regions is problematic as they consist of an ensemble of rapidly interconverting conformers that cannot be easily visualized. One reason for the certainty of unstructured regions is their amino-acid compositional bias.

SNP prediction algorithms such as SIFT (18) and PolyPhen (19) were used to predict the effect of the studied variant on protein function. Computational predictions of the potential impact of this variation using SIFT and Polyphen gave a score of 0.05=Tolerated and 0.782=possibly damaging respectively. In addition, change in protein stability upon mutation was analyzed using DUET, a web server that predicts the effects of mutation on protein stability (20). This mutation was found to have a destabilizing effect on the protein structure. The variation in the Gibbs free energy ($\Delta\Delta G$) was observed to be -1.2 kcal/mol, where negative value denotes destabilizing mutations.

It is suggested that podocin accumulates in dimeric or oligomeric forms in lipid raft microdomains at the podocyte slit diaphragm, which is the key component of the glomerular filtration barrier (9). We used the podocin monomer to generate its dimers (Figure 2c), the final models are produced by co-threading of both chains followed by superposition of individual chain templates generated by the COTH web server (21). Scrutiny of the homo- as well as hetero-dimers of podocin with and without the R229Q polymorphism revealed that this alteration did not seem to impair the mode of podocin dimerization.

**Figure 2a. Homology model built for podocin using I-TASSER web server.** **2b Location of R229 residue in the predicted 3D model of podocin.** **2c. Podocin homo-dimer generated by COTH server.**

**Discussion**

The polymorphism R229Q is one of the most commonly reported podocin sequence variations. The arginine residue at protein position 229 is highly conserved across species, and arginine-to-glutamine substitution R229Q was found to possess altered functional properties of podocin in vitro and probably in vivo (22). Analysis of available sequence from species orthologs of NPHS2 genes shows that the arginine residue at position 229 is evolutionarily conserved across invertebrates, vertebrates and invertebrates to vertebrates, suggesting an important role of Arg229 for its structure and function. Based on the high frequency of the R229Q allele in FSGS families and the evolutionary conservation of this residue, it is believed that R229Q allele is a relatively common disease-causing recessive allele, rather than a benign polymorphism. In vitro studies have demonstrated decreased binding of the p.R229Q mutant protein to
nephrin, providing a likely explanation for its pathogenic role (13). This polymorphism was observed frequently with increased incidence in NS patients compared to healthy controls (24), similar to the results reported in this study. Further, investigations have been carried out to see the influence of the R229Q functional variant on microalbuminuria; one of the reports showed an association in the general population (24), however, another study demonstrated apparently no altered risk of proteinuria in both general and diabetic population (25). There is an uneven R229Q allele distribution throughout different populations and more frequent among South Americans, Europeans and European Americans (4-7%) than among Africans, African Americans, and Asians (0-1.5%), suggesting that this variant emerged in Europe, although it is not possible to discern its specific geographic origin (26, 27). The maximum mutant allele frequency 7.3% was described in Chileans and Argentinians (10) and in European inhabitants, the highest occurrence of p. R229Q polymorphism was described in France (4.5%), 3.2% in an Italian population as well as 3.1% in a Spanish population (11, 28, 29).

The pathogenicity of the NPHS2 homozygous p.R229Q variant in SRNS is doubtful. While it has been reported in unaffected controls, it is enriched in patients with SRNS, indicating its certain role in the disease state. A rare association of FSGS to a PAX2 mutation observed in a family of FSGS may reflect the modifier effect of p.R229Q in FSGS. Kerti et al., recommend such an effect can also explain its enrichment in SRNS patients and that patients with homozygous p.R229Q should be screened for the causative mutation in a second gene (23).

We noticed a strong alliance of A allele with a four fold increased risk towards the disease (odds ratio 3.99, 95% CI 1.08-14.8), an observation comparable with the recent report from central India (30). Significantly elevated frequency of mutant A allele noted in severe histopathology signifies its role in the FSGS phenotype. R229Q mutant of podocin is known to cause decreased binding to nephrin and our dimer analysis suggests that this variation does not affect the mode of podocin dimerization and hence the pathogenetic effect could be contributed due to its tainted interaction with other protein partners.

Previously we reported an elevated frequency of the mutant allele in diabetic nephropathy patients with high glycosylated hemoglobin (HbA1c) compared to the diabetic individuals with lower levels of HbA1c suggestive of modifying (gene-environmental interaction) effect of this polymorphism in causing kidney related disorders (31).

**Conclusion**

This is one of the largest studies of Asia and our data with respect to allele frequency and the *insilico* analysis supports the role of the R229Q substitution in susceptibility and severity of NS. It appears to be a good contributing factor in disease pathogenesis in our population. We showed that it may not affect the manner of podocin dimerization however, it was noted to have possibly destablizing effect on the podocin protein, thereby reflect the probable influence of R229Q in the pathogenesis of NS. As it is a common polymorphism, extensive studies in different ethnic populations are warranted in order to understand the gene-gene as well as gene-environment interactions for this particular polymorphism in primary and secondary kidney disease.

**Acknowledgement**

The authors wish to thank the Indian Council of Medical Research, New Delhi for the financial support.

**Conflict of interest**

The authors declare that there is no conflict of interest.

**References**


17. http://ppopen.informatik.tu-muenchen.de/

18. www.sift.bii.a-star.edu.sg/


