Mutation Spectrum analysis of ITGA3 gene associated with Nephrotic syndrome

QuratulAin Nazir
University of Veterinary and Animal Sciences Lahore

Wasim Shehzad
University of Veterinary and Animal Sciences Lahore

Muhammad Asad Ali
University of Veterinary and Animal Sciences Lahore

Muhammad Yasir Zahoor (yasir.zahoor@uvas.edu.pk)
University of Veterinary and Animal Sciences Lahore

Research Article

Keywords: Nephrotic syndrome, ITGA, nephrogenesis, in silico, bioinformatics

Posted Date: October 29th, 2023

DOI: https://doi.org/10.21203/rs.3.rs-3460322/v1

License: © This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Additional Declarations: No competing interests reported.
Background

Nephrotic syndrome is a renal disorder in which the Glomerular Filtration Barrier (GFB) is affected. There is a potential indication that ITGA3 may have a pivotal function in the intricate interaction between cells, morphogens, and the extracellular matrix (ECM) that is essential for the development of the kidneys (nephrogenesis). The present study involves a detailed analysis of the reported missense mutations in the ITGA3 gene through various in silico and bioinformatics tools. The data about reported mutations was collected from HGMD and tools such as PolyPhen-2, SIFT, I-Mutant were used to predict the affects of these mutations. The conservation analysis of this protein was also done by analyzing which mutations fall on the conserved region of the protein and are hence more detrimental.

Results

A total of 7 mutations were identified, out of which one (A349S) was found to have the least detrimental affect on the protein structure. The phylogenetic analysis of the ITGA protein family was also done to determine the relationship of ITGA3 protein with the other proteins in its family.

Conclusion

The data obtained in this study is aimed to facilitate future studies on ITGA3 protein and its role in the development of the Nephrotic syndrome, along with the implication of the mutations on the structure and function of the ITGA3 protein. This study also gives an insight on the detrimental effect of the mutations on the protein.

Abstract

Background

Nephrotic syndrome is a renal disorder in which the Glomerular Filtration Barrier (GFB) is affected. This causes the occurrence of symptoms such as hypo-tension, hypercoagulation and life-threatening infections. The Glomerular Filtration Barrier (GFB) is responsible for the filtration of the blood in kidneys [1]. Damage to this barrier may result in many renal disorders including nephrotic syndrome. This may be caused by immune complex deposition, phospholipase antibody production, or the development of alloantibodies [2]. Mutations in numerous podocyte proteins may target the functioning of the podocyte via various pathological mechanisms by modifying the slit diaphragm structure, perturbing the delicate podocyte cytoskeleton, destroying cell-matrix connections, or obstructing critical signaling pathways [3].

Research has indicated that changes in the manifestation or operation of particular integrins within the glomerulus can play a role in the impairment of podocytes and the disruption of the glomerular barrier. Mutations or dysregulation of ITGA3, ITGA4, ITGA5, ITGA6, and ITGA8 have been linked to different types of glomerular diseases that are associated with nephrotic syndrome. The ITGA3 gene has been linked to a condition known as generalised junctional epidermolysis bullosa with respiratory and renal involvement (JEB-RR), as well as congenital interstitial lung disease, nephrotic syndrome, and epidermolysis bullosa [4].

The genes belonging to the integrin α (ITGA) subfamily are known to have a significant impact on the development and progression of different types of disorders including cancer. Integrins are classified as heterodimeric surface receptors, comprising of non-covalently linked α and β subunits. Currently, our understanding suggests that the integrin family encompasses a total of 18a and 8β members [5]. Numerous studies have demonstrated that integrins possess the ability to act as signalling molecules across the cell membrane in both directions. This includes “inside-out signalling,” which occurs when extracellular stimulation prompts the binding of intracellular linin and kindlin to the cytoskeleton, resulting in the extracellular domain adopting a high affinity state. Additionally, there is “outside-in signalling,” a complex process whereby the heterodimeric adhesion receptors of integrins facilitate cell adhesion to the extracellular matrix (ECM) and subsequently activate integrins to interact with the cytoskeleton. This activation leads to the initiation of various intracellular signalling pathways, which in turn enhance the binding of activated integrin ligands and enable the perception of the intracellular environment [6]. The expression of Integrin α3β1 is prevalent in the epithelial tissues of the skin, lungs, and kidneys [7].

While integrins play a crucial role as receptors for extracellular matrix (ECM) proteins and are widely present throughout kidney development, ITGA3 is predominantly recognized as a passive stabilizer of glioblastoma multiforme (GBM) rather than an active participant in nephrogenesis. In recent studies, it has been demonstrated that genetic alterations occurring within the human ITGA3 gene are responsible for the development of a complex condition known as NEP syndrome (Nephrotic syndrome, Epidermolysis bullosa, and Pulmonary disease).

The deficiency of ITGA3 manifests as an early developmental defect characterized by the dysregulation of multiple crucial pathways involved in nephrogenesis. This dysregulation ultimately gives rise to the renal hypodysplasia/CAKUT phenotype. The transcripts that were found to be down regulated in the kidney, specifically in the differentiated proximal tubules, include genes such as GATM, AGXT2, GSTA1, SLC3A1, SLC13A1, SLC7A9, SLC2A2, and AQP11 [8].

Page 2/9
Several approaches have been employed to explore the mutations in these genes and their impacts on the disease etiology. These include both in vitro and in vivo studies but these methods frequently need numerous technical and financial resources, manpower and are also quite time-consuming. Numerous computational techniques can be used for the prediction of different SNPs and their effects on the functions and structure of proteins [9].

The present study involves a detailed analysis of the reported missense mutations in the ITGA3 gene through various in silico and bioinformatics tools. The severity of the mutations was analyzed along with the conservation score of each mutation to determine which mutations fall on the conserved regions and are consequently more detrimental to the protein structure and function. The 3D structure of the mutant forms of the protein was also investigated to determine the changes in the structure due to these mutations. Finally, a phylogenetic analysis was performed to determine the relationship between different ITGA proteins. The present study demonstrates the potential of using computational methods in predicting the effect of deleterious SNPs on protein structure.

**Methodology**

**Data collection**

The data about the reported missense mutations (nsSNPs) in ITGA3 gene was obtained from Human Gene Mutation Database (HGMD) on July 22, 2023. The protein sequence of ITGA3 was also obtained through UniProt Database.

**Analysis of mutations**

The single nucleotide variations of the ITGA3 gene, obtained using the dbSNP (SNP database) analysis, were submitted for the computational analysis using various online tools. SIFT (Sorting Intolerant From Tolerant), PolyPhen-2 (Polymorphism Phenotyping v2), I-Mutant, and PhD-SNP (Predictor of human Deleterious Single Nucleotide Polymorphisms) were among the methods employed to predict the effects of SNPs in silico.

The SIFT method was used to assess the influence of amino acid replacement on the protein function. To improve the accuracy of computational methods, the PolyPhen-2, I-Mutant, and PhD-SNP tools were used to validate SNPs predicted in SIFT. PolyPhen-2 calculates the PSIC (position-specific independent score) for each input variable. This tool uses Nave Bayes approach for the determination of the implication of changes in allele [10]. The impact of mutations on the function of the proteins was determined through the PhD-SNP software [11] whereas the I-mutant software employs the SVM (Support Vector Machine) to predict the effects on the stability on the protein structure due to these nsSNPs by calculating the DDG (Delta Delta Gibbs free energy) (kcal/mol) and RI (Reliability Index) values [12].

**Determination of conservation sites**

The evolutionary conservation profile of the ITGA3 gene was predicted using the online tool ConSurf. This tool uses the Bayesian technique for detecting the structural and functional residues and to assess the evolutionarily conserved amino acid residues in the protein. This helps in the identification of the areas of the protein that are crucial for the maintenance of its structure and function [13]. This data was used to examine the possibility of high-risk nsSNP in the protein ITGA3 to cause damage.

**3D Protein Structural Analysis**

Phyre-2 (Protein Homology/AnalogY Recognition Engine) was used to generate the 3D structure of the proteins. PyMOL molecular graphics programme was used to examine the models produced by Phyre-2 and to determine the changes in the protein structure due to these mutations.

**Phylogeny analysis**

The phylogenetic analysis is done to thoroughly understand the evolution of different proteins through genetic modifications. So, the NCBI tool (blastp) was used to carry out the phylogenetic tree analysis of the ITGA family. This technique presents the results in the form of tree, that helps in the comparison of the degree of relationship between all the sequences present in the set. The results of the phylogenetic analysis are helpful in classifying whether there are subset(s) of the sequences in BLAST output that can be grouped as a family.

**Results**

**Data obtained from HGMD and UniProt**

SNPs for the protein ITGA3 were identified from HGMD; it identified seven missense mutations in the protein. The FASTA sequence of the ITGA3 protein was obtained using UniProt database.

**Analysis of mutations**
Different algorithms were used to analyse the effects of these 7 missense mutations in the ITGA3 protein. According to the SIFT results, 5 SNPs were deleterious, and two were found to be tolerated in ITGA3. Out of 7 SNPs submitted to PolyPhen-2 analysis, five were predicted to be probably harmful, and two were identified to be possibly damaging. PhD-SNP predicted six SNPs as diseased, and one as neutral in the target gene. The analysis of I-Mutant demonstrates that all of the seven potential nsSNPs decreased activity by lowering its stability to DDG > 0.5 Kcal/mol. The results of these algorithms are presented in table 1.

**Table I** Results obtained from SIFT, Polyphen 2, I mutant and PhD-SNPs for ITGA3 variants

<table>
<thead>
<tr>
<th>Variant ID</th>
<th>Alleles</th>
<th>Consequence</th>
<th>AA</th>
<th>SIFT Prediction</th>
<th>Score</th>
<th>Polyphen 2 Prediction</th>
<th>Score</th>
<th>SVM2 effect</th>
<th>DDG</th>
<th>RI</th>
<th>Effects</th>
<th>RI</th>
</tr>
</thead>
<tbody>
<tr>
<td>CM1612277</td>
<td>G/A</td>
<td>Missense</td>
<td>G125R</td>
<td>deleterious</td>
<td>0.00</td>
<td>Probably damaging</td>
<td>1.00</td>
<td>decrease</td>
<td>-0.74</td>
<td>6</td>
<td>Disease</td>
<td>9</td>
</tr>
<tr>
<td>Rs763799707</td>
<td>G/A</td>
<td>Missense</td>
<td>R143H</td>
<td>deleterious</td>
<td>0.00</td>
<td>Probably damaging</td>
<td>1.00</td>
<td>Decrease</td>
<td>-1.20</td>
<td>9</td>
<td>Disease</td>
<td>5</td>
</tr>
<tr>
<td>Rs745505565</td>
<td>G/A</td>
<td>Missense</td>
<td>R274Q</td>
<td>deleterious</td>
<td>0.00</td>
<td>Probably damaging</td>
<td>1.00</td>
<td>Decrease</td>
<td>-1.24</td>
<td>8</td>
<td>Disease</td>
<td>8</td>
</tr>
<tr>
<td>CM1211323</td>
<td>G/T</td>
<td>Missense</td>
<td>A349S</td>
<td>Tolerated</td>
<td>0.64</td>
<td>Possibly damaging</td>
<td>0.536</td>
<td>Decrease</td>
<td>-0.44</td>
<td>9</td>
<td>Neutral</td>
<td>7</td>
</tr>
<tr>
<td>Rs79704989</td>
<td>C/T</td>
<td>Missense</td>
<td>R463W</td>
<td>deleterious</td>
<td>0.00</td>
<td>Probably damaging</td>
<td>1.00</td>
<td>Decrease</td>
<td>-0.84</td>
<td>7</td>
<td>Disease</td>
<td>6</td>
</tr>
<tr>
<td>Rs140781106</td>
<td>G/C</td>
<td>Missense</td>
<td>R628P</td>
<td>Tolerated</td>
<td>0.07</td>
<td>Possibly damaging</td>
<td>0.820</td>
<td>Decrease</td>
<td>-1.07</td>
<td>6</td>
<td>Disease</td>
<td>5</td>
</tr>
<tr>
<td>Rs1263593458</td>
<td>C/G</td>
<td>Missense</td>
<td>P680A</td>
<td>deleterious</td>
<td>0.00</td>
<td>Probably damaging</td>
<td>0.989</td>
<td>Decrease</td>
<td>-1.96</td>
<td>6</td>
<td>Disease</td>
<td>6</td>
</tr>
</tbody>
</table>

**Determination of conservation sites:**

ConSurf analysis showed four conserved amino acid residues in the domain of ITGA3. In regard to this, the nsSNP situated at the conserved areas are extremely detrimental to protein function as compared to those at the non-conserved regions. According to the results of ConSurf, the nsSNP G125R, R274Q, R463W, and P680A in ITGA falls into the conserved region with the conservation score ranging from seven to nine (Figure I).

**Fig. I** The color grade indicates the degree of the conservation status of amino acid residues. The color grade rises (1 is highly variable and 9 is a highly conserved site) – predictions of nsSNP in ITGA3 showing conservation scores.

**3D Protein Structural Analysis:**

Interpro projected that ITGA3 proteins include a large domain conserved across species. Integrin_alpha-2 is the domain that contains amino acids 462–916.

**Fig. II** ITGA3 mutations classified into different domains. The orange color indicates conserved region across species.

Later, the schematic representation of the wild type of ITGA3 and its' mutants were produced using the online tool Phyre2. The predicted models were examined using the software PYMOL. Each nsSNP in the ITGA3 protein was assigned the potential 3D structure (Figure IV).

**Phylogeny Analysis:**

Figure III represents the phylogenetic analysis of ITGA family. Different proteins of ITGA family were presented in the form of a phylogenetic tree.

**Fig. III** Phylogenetic tree of the ITGA gene family

The tree indicates that ITGA3 gene is closely related to ITGA6 and ITGA7 and is more distantly related with ITGA2.

**Fig. IV** 3D Structural representation of mutant ITGA protein (a) G125R (b) R143H (c) R274Q (d) A349S (e) R463W (f) R628P (g) P680A

**Discussion**

The ITGA gene family encompasses a variety of genes, with each gene encoding a distinct integrin alpha subunit. The ITGA family comprises several prominent members, namely ITGA1, ITGA2, ITGA3, ITGA4, ITGA5, ITGA6, ITGA7, ITGA8, ITGA9, ITGA10, ITGA11, and ITGA11B. The formation of a functional integrin receptor involves the combination of an alpha subunit with a beta subunit [14]. Prior studies have demonstrated...
that the integrin family facilitates signal transduction through its interaction with the extracellular matrix, facilitated by adhesion receptors located on its surface [15]. The integrin α3 protein is a subunit of a transmembrane adhesion receptor that is known to form heterodimers with integrin β1. The expression of Integrin α3β1 is prevalent in the epithelial tissues of the skin, lungs, and kidneys [16].

Researchers conducted an examination of the histological and molecular characteristics of the kidneys obtained from a solitary patient belonging to the original group displaying an ITGA3 mutation [17]. The objective was to shed light on the involvement of ITGA3 in the process of human renal development. The role of ITGA3 in renal extracellular matrix (ECM) protein interactions extends beyond its conventional perception as a passive anchoring molecule. The findings potentially indicate that ITGA3 may have a pivotal function in the intricate interaction between cells, morphogens, and the extracellular matrix (ECM) that is essential for the development of the kidneys (nephrogenesis). Consequently, they propose the inclusion of ITGA3 in the catalog of genes associated with congenital anomalies of the kidney and urinary tract (CAKUT), as documented by Shukrun et al. (2014) [18].

Through SIFT analysis, two mutations, A349S and R628P were predicted to be tolerant whereas the other 5 mutations were predicted to be deleterious, with a value of 0.0, whereas tolerant mutations had a score of more than 0. The same two mutations were predicted to be possibly damaging with a score of less than 1.0, whereas other mutations were predicted to be probably damaging with a score of 1.0 or closer to 1.0. The I-mutant software gave negative values of DDG that indicated that all the mutations had a destabilizing effect on the structure of the protein. When using Phd-SNPs only one of the mutations (A349S) was predicted to be neutral. This indicates that the mutation A349S creates the minimum change in the protein structure and is hence more tolerant and less damaging than other mutations found in this protein.

The 3D structural analysis of the mutant protein and the amino acid changes associated with these mutations also show that there is very little difference in the structure of the wild type amino acid and the mutant amino acid in A349S. Other mutations, on the other hand, show very significant changes in the overall structures of the amino acids which can be considered as the reason for drastic changes in the protein structure.

In a biological system, proteins containing conserved amino acids are involved in various cellular processes, including genome stability. Amino acids that occupy enzyme sites or required for protein-protein interaction are more conserved in proteins than other amino acids in the same molecule. As a result, nsSNP located in conserved parts of the protein are more harmful than nsSNP located in variable sections of the protein. The evolutionary conservation profile of the ITGA3 gene was predicted using the online tool ConSurf. This tool uses the Bayesian technique for detecting the structural and functional residues and to assess the evolutionarily conserved amino acid residues in the protein. This data was used to examine the possibility of high-risk nsSNP in the protein ITGA3 to cause damage. ConSurf analysis showed four conserved amino acid residues in the domain of ITGA3. The evolutionary information is used to detect if a change in an amino acid would affect the activity of the protein. The predicted function or structure of the conserved residues is determined by the position of the residues on the protein surface vs within the core of the protein. Whereas, in this study we only observed the residues whose location was corresponded to those of seven nsSNP that we had found. In regard to this, the nsSNP situated at the conserved areas are extremely detrimental to protein function as compared to those at the non-conserved regions. According to the results of ConSurf, the nsSNP G125R, R274Q, R463W, and P680A in ITGA falls into the conserved region with the conservation score ranging from seven to nine. According to the results of ConSurf, the nsSNP G125R, R274Q, R463W, and P680A in ITGA falls into the conserved region with the conservation score ranging from seven to nine. It was also determined that 3 out of these 7 mutations fall in the conserved domain of the protein, i.e. integrin_alpha_2.

The phylogenetic analysis of the ITGA family revealed the evolutionary relationship between the proteins of this family. The ITGA3 was seen to be most closely related to the ITGA6 and ITGA7 with the least divergence, and the greatest divergence was found to be with ITGA2. Hermitz junctional epidermolysis bullosa, an autosomal recessive disorder, manifests with the recurring development of blisters due to inherent skin fragility resulting from structural abnormalities. Reported are mutations in various genes implicated in the biosynthesis of collagens, integrins, and laminins. Notably, certain mutations have been observed in conjunction with nephrotic syndrome and focal segmental glomerulosclerosis (FSGS). These genes include ITGB4, which encodes the integrin b4 protein; ITGA4, which encodes the integrin a3 protein; ITGA3, which encodes the integrin alpha3 subunit; LAMB3, which encodes laminin-5; and CD151 [19]. Therefore, it can be concluded that most of the types of ITGA proteins are responsible for the manifestation of this disease.

The application of next generation sequencing techniques facilitated the identification of biallelic missense mutations in the ITGA3 gene. Specifically, these mutations were found at positions c.485G > A (resulting in the amino acid change p.C162Y) in exon 4 and c.1382G > A (resulting in the amino acid change p.R461Q) in exon 9. Computational analysis using in silico tools such as PolyPhen-2 and PROVEAN predicted that these mutations are likely to have deleterious effects [20]. However, these mutations have yet to be confirmed and added in the HGMD.

Several other genes are involved in the occurrence of Nephrotic syndrome as well. Genes such as NPHS1, NPHS2, WT1, LAMB2, & PLCE1 were most often altered in congenital or infantile neuropathy cases. Mutations in their autosomal DNA occur. INF2 and TRPC6 mutations are more prevalent in young people, although not everyone is affected the same way, and their consequences are variable [10]. More than 20 genes have so far been identified as having mutations that induce steroidresistant nephrotic syndrome on their own (SRNS).

The ITGA3 gene has recently been identified as an important gene in the onset of the nephrotic syndrome. However, the data related to this gene is still insufficient and further studies need to be carried out to explore its role in the manifestation of different genetic disorders, particularly
nephrotic syndrome.

**Conclusion**

The *ITGA3* gene plays an important role in the formation of integrin and anchoring proteins that play a vital role in the development of Extracellular matrix of the kidneys. Mutations in this gene can lead to the manifestation of nephrotic syndrome and other disorders. Further studies need to be carried out on this gene to further explore its role as an important factor in the development of kidneys.

**Declarations**

**Ethics approval and consent to participate**

Not applicable

**Consent for publication**

Not applicable

**Availability of data and materials**

The datasets analysed during the current study are available in the HGMD repository. The datasets used during this study are available from the corresponding author on reasonable request. All data generated or analysed during this study are included in this published article.

The datasets analysed during the current study are not publicly available due to the restriction of the access to the HGMD webportal but are available from the corresponding author on reasonable request.

**Competing interests**

The authors declare that they have no competing interests.

**Funding**

The present study was funded by the Higher Education Commission Pakistan.

**Authors’ contribution**

Q.N. and M.Y.Z. have worked on the main manuscript. W.S. and M.A.A have worked on the compilation of data

**Acknowledgements**

Not applicable

**References**


Figures

Figure 1

The color grade indicates the degree of the conservation status of amino acid residues. The color grade rises (1 is highly variable and 9 is a highly conserved site) – predictions of nsSNP in ITGA3 showing conservation scores.
Figure 2

*ITGA3* mutations classified into different domains. The orange color indicates conserved region across species.

Figure 3

Phylogenetic tree of the ITGA gene family
Figure 4

3D Structural representation of mutant ITGA protein (a) G125R (b) R143H (c) R274Q (d) A349S (e) R463W (f) R628P (g) P680A