Natarajan *et al.*
Egyptian Journal of Medical Human Genetics (2025) 26:29
https://doi.org/10.1186/s43042-025-00659-x

Egyptian Journal of Medical Human Genetics

RESEARCH Open Access

Whole-exome sequencing and molecular dynamics confirm pathogenicity of a novel SLC6A6 mutation in Leber congenital amaurosis

Srikrupa N. Natarajan[1†], Samdani Ansar[2], Sarangapani Sripriya[1], Sen Parveen[3], Ravi Gupta[4], Umashankar Vetrivel[5] and Mathavan Sinnakarupan[1*]

Abstract

Introduction Inherited retinal diseases (IRDs) are a clinically and genetically heterogenous group where the robust advancement of next-generation sequencing technologies has facilitated genotype-assisted diagnosis. Leber congenital amaurosis (LCA) is a severe form of inherited retinal dystrophy that causes congenital blindness or near-blindness with a global prevalence of 3 per 100,000 live births.It is characterized by a loss of vision at birth or within the first few years of life with overlapping phenotypes to many syndromic and non-syndromic IRDs. With India's rich genetic heterogeneity, WES is a valuable tool for uncovering novel gene mutations linked to LCA. This genetic diversity expands our understanding of the disease's spectrum in the Indian population.

Methods In our previous study, 92 Indian LCA families were screened through targeted resequencing, and 80% of probands exhibited mutations in known genes. Hence, the remaining 20% probands with additional family members (n = 40) were subjected to whole-exome sequencing. An in-house standard bioinformatics pipeline was used for variant calling and annotation. Homology modeling (Modeller-9.23) and molecular simulation were performed on an identified *SLC6A6* gene variant that has not yet been associated with LCA to investigate its potential pathogenicity.

Results Disease-causing pathogenic variants were identified in 15/20 families (75%) across 11 genes with 33% variants being novel. Among the identified 17 variants in 15 families, 35% were missense, 29% nonsense, 29% frameshift and 6% splice variants. Segregation analysis, control screening and *in silico* predictions confirmed the variant's pathogenicity. All variants were classified as pathogenic according to ACMG guidelines. Homology modeling and molecular simulation in the membrane system for the p.Pro82Leu mutant in SLC6A6 protein showed significant modification in helical characteristics around the TM2 helix in the mutant, which could potentially hinder the regular function and cause disruption in taurine transport across the membrane leading to the disease.

Conclusion Taurine being an essential amino acid for photoreceptor development and maintenance, our study suggests that mutation identified in *SLC6A6* gene may cause LCA. This is the first report of *SLC6A6* gene association with LCA and also the first case report in the Indian population.

Keywords Next-generation sequencing, SLC6A6, Leber congenital amaurosis, Whole-exome sequencing, Homology modeling

[†]First author: Srikrupa N. Natarajan.

*Correspondence:
Mathavan Sinnakarupan
mathavans@gmail.com; mathavans@snmail.org
Full list of author information is available at the end of the article

Introduction

Inherited retinal dystrophies (IRDs) are a large, complex group of retinal disorders that lead to congenital or severe progressive retinal degeneration in children [1, 2]. About 280 genes and 316 loci are reported to be associated with IRDs till date [3]. Leber congenital amaurosis (LCA) is a heterogeneous, most severe non-syndromic form of IRDs. The clinical characteristics include congenital or near-blindness, nystagmus, photophobia, reduced or absent light perception, Franceschetti's oculo-digital signs and an extinguished electroretinogram. While it is primarily a non-syndromic condition, LCA has also been associated with a few syndromes, indicating its involvement in broader genetic pathways [4–6]. The global prevalence of LCA is estimated to be between 1 in 30,000 and 1 in 81,000 individuals [7]. In recent years, whole-exome sequencing (WES) has revolutionized the diagnosis of IRDs, including LCA. By analyzing the entire protein-coding region of the genome, WES has facilitated the identification of numerous novel genes associated with LCA. This has significantly expanded our understanding of the genetic landscape of this disease [8–12] with 29 candidate genes reported LCA so far.

WES has been particularly valuable in countries with diverse genetic populations, such as India, where a wide range of LCA-causing mutations may exist. In India, the mutation prevalence rate for known LCA candidate genes is relatively high, reaching 61%. This suggests that genetic factors play a significant role in the development of LCA within the Indian population [13]. In this study, we conducted whole-exome sequencing (WES) on a group of 20 individuals along with their family members who had been diagnosed with LCA. Understanding the specific genetic mutation causing LCA in an individual can inform personalized treatment approaches and potentially guide the development of targeted therapies.

Materials and methods

Ethics statement

The research was conducted in accordance with the principles outlined in the Declaration of Helsinki and was approved by the Institutional Review Board (Ethics Committee- 708–2018-P). Written informed consent was obtained from all patients (or their parents if they were minors) and participating family members.

Patients selection and WES

This study utilized archived DNA samples from 20 Indian families with leber congenital amaurosis (LCA). Forty-five percent of these families reported consanguineous marriages. While previous targeted resequencing of 20 candidate genes had yielded negative results for these patients, whole-exome sequencing was conducted in this study. All patients underwent a thorough ophthalmic examination, including electroretinography, fundus photography, fundus autofluorescence, and optical coherence tomography (OCT) when possible. The DNA quantity was checked using Nanodrop 1000 spectrophotometer (ThermoFisher Scientific) and quality by agarose gel electrophoresis. The samples were outsourced for WES where the exonic regions including 25 bases of the intron covering the splice-site were enriched using the SureSelectXT Human All Exon V5 + UTRs enrichment kit. The resulting enriched DNA libraries were index tagged by multiplexing, amplification, followed by purification. The indexed captured library DNA was finally sequenced using Illumina HiSeq 2500 (Illumina, San Diego, USA) with 80–100 × depth.

Analysis and validation of WES data

Approximately 8 Gb data were obtained for each sample. The sequenced data were processed to generate FASTQ files. The data were pre-processed for removing adapters, primers and low-quality sequence reads using open source softwares. BWA aligner was used for read alignment to GRCh37/hg19 genome built. Picard toolkit was used for PCR duplicate removal. Re-alignment to InDels and re-calibration of the base quality was performed using GATK-lite program. GATK-lite Unified Genotype caller was used to predict the variants. The variants were filtered for variant quality and depth. Only variants with Q30 score were taken for further annotation using VariMAT (internal data analysis pipeline of MedGenome labs, Bangalore).

The variants were compared with various disease-relevant mutation resources like OMIM, SNPedia, and population databases such as dbSNP (db135), 1000 genome project, GenomAD, LOVD, Ensemble variation table, ClinVar, VarSome and SAGE to filter variants with MAF ≤ 0.01. The variants in the coding and splice region of the protein-coding transcripts of the genes were considered for further analysis. The variants in the candidate gene(s) for retinal degenerative disease and that which is expressed in the eye were considered as possible candidates. While considering proband and parent data, homozygous variants in proband that are heterozygous in parent were shortlisted. Apart from this, compound heterozygous variants and heterozygous variants in known dominant and recessive genes were also scrutinized for validation in the proband.

Sanger sequencing

All the scrutinized variants in the analysis were validated by Sanger sequencing in the proband(s) and co-segregated in the parents' and sibling samples. Primers were designed for protein-coding transcripts from Ensembl

GRCh37 browser using Primer 3 (v.0.4.0). PCR was standardized using 10 pmol/µl primers (forward and reverse), 50 ng DNA template, 1X taq buffer, 0.5U Taq polymerase, and 0.5 mM dNTP. The PCR products were enzymatically purified using Exo-SAP (Thermo Fischer-scientific, USA) and cycle sequenced using a Big dye terminator v3.1 kit (Applied Biosystems, USA). Sequencing was performed using ABI 3500 Avant genetic analyzer.

Homology modeling and structure refinement

We performed homology modeling and molecular simulation to validate the pathogenicity of *SLC6A6* mutation in the WES-RD3. The structure of human SLC6A6 is yet to be elucidated, and hence, the 3D structure was modeled by homology modeling using Modeller-9.23 [14]. To start with, the protein sequence of SLC6A6 was searched against the PDB database using Blastp, wherein PDB ID: 4M48 (DAT-Sodium-dependent dopamine transporter) was the chosen template, as it showed a sequence identity of 47.3% and 87% of query coverage. Moreover, earlier studies also have utilized 4M48 as a template for performing structural studies [15]. Subsequently, homology modeling was performed for SLC6A6 using the modeler-9.23 version by generating 100 models. Further, the model with the lowest DOPE score proceeded for structural refinement. Following this, the mutant structure was also predicted by using the modeled wild structure as a template. Both the protein structures were finally refined using the GalaxyRefine web server [16, 17] in order to fix the atomic clashes. The structure quality of the refined proteins was checked using the tools such as Ramachandran plot, ProQ [18], and ProSA [19]. The mutational effect of Pro82Leu on the wild structure stability was predicted using the SDM (site-directed mutator) [20] and mCSM-membrane stability prediction [21]. Mutpred2 was also used to predict the pathogenicity of the mutation [22].

Molecular dynamics simulation

The molecular dynamics simulation for both Wild and Mutant forms was performed in the membrane system. The proteins were prepared using the Charm-GUI server [23] in order to embed it in the lipid bilayer. The molecular composition for building the SLC6A6 membrane system was adopted from earlier studies on SLC6A6 mutation [15]. During the simulation process, CHARMM36 forcefield was used and the system was solvated in the TIP3P water model. Further, the charge of the system was neutralized by the addition of Na^+ and Cl^- counter ions. Following this, six consecutive steps of minimization and equilibration were performed by slowly removing the positional restraints which were applied on the heavy atoms of protein and the membrane. During the system equilibration steps, the temperature was controlled at 300 K, while pressure was maintained at 1 bar with Berendsen thermostat and barostat [24], respectively. In the final production run, the Nose–Hoover [25, 26] and the Parrinello-Rahman barostat [27] were used for maintaining temperature and pressure, respectively. The bonds involving hydrogen atoms were constrained using the LINCS algorithm [28], while the electrostatic interactions were calculated using the Particle Mesh Ewald (PME) method [29]. The production run was performed for 200 ns, and the coordinates were saved for each 10 ps to perform further analysis. All the molecular dynamics simulations were executed using the Gromacs-5.1.4 [30]. The protein backbone root mean square deviation (RMSD), root mean square fluctuation, radius of gyration, and secondary structure variations were calculated using the Gromacs utility tools. The intra-H-bond was calculated using the H-bond analysis plug-in available through virtual molecular dynamics (VMD) software. The intra-hydrophobic contacts formed by the side chain of the wild (P82) and mutant (L82) within its protein residues were calculated using the PLIP (protein–ligand interaction profiler)-2.4 [31]. The protein coordinates at each time frame was sampled, for which the number of hydrophobic contacts formed within the protein residues during the simulation were counted.

Results

Validation of genetic variants

The DNA samples from 20 patients exhibiting typical characteristics of LCA underwent WES, followed by data filtration using our in-house pipeline. We obtained an average of 100X coverage per sample. After comparing with the unaffected family member samples, from around 100,000 variants, scrutinized through minor allele frequency (MAF) filtering, we were able to shortlist about 10–20 variants per sample for further validation. Totally Seventy-eight variants were validated by Sanger sequencing, segregation analysis in extended family members and control screening in 100 unrelated, well-characterized healthy controls (wherever necessary) to identify the disease-causing variant in each family.

We were able to report the disease-causing pathogenic variants in 15 of 20 patients resulting in a mutation detection rate of 75%. The mutations were identified across 11 genes- *AIPL1* (c.844G > T), *AHI1* (c.2087A > T), *CNGA3* (c.847C > T), *CRB1* (c.2132_2141delATGTGGCAGG,c.2143_2144insAAATT,c.2278_2293delTATCAATATATCCGTG), *CNNM4* (c.1475G > A), *LCA5* (c.1062_1068delGTTTTCG), *MYO7a* (c.3638G > A), *RPGRIP1* (c.922_923insCTCCAGAA,c.1480delA,c.2668C > T, c.930 + 3A > G) RPGRIP1L (c.2030C > T), *SPATA7* (c.253C > T, c.1171C > T), and

SLC6A6 (c.245C > T) (Table 1). Segregation within the family members confirmed the inheritance of all 15 mutation-positive cases (Fig. 1). Of the total variants, 29% were novel (5/17) and were not reported in any public databases like dbSNP, 1000 genome project, and GenomAD. These novel variants include c.2087A > T in *AHI1,* c.2132_2141delATGTGGCAGG and c.2143_2144insAAATT in *CRB1*, c.922_923insCTC CAGAA and c.1480delA in *RPGRIP1*. Among the identified 17 variants in 15 families, 35% were missense, 29% nonsense, 29% frameshift and 6% splice variants.

Apart from co-segregation analysis and control screening, the variants were validated for pathogenicity by in silico analysis (Table 1) and classified according to ACMG guidelines [32]. All variants were categorized as pathogenic by ACMG classification according to PP4 (patient's phenotype highly specific for gene), PS4 (prevalence in affected statistically increased over controls), PP1 (increased segregation data) and PM3(for recessive disorders detected in trans with a pathogenic variant). Additionally, the novel variants (p.His696Leu, p.Tyr711Serfs*40, p.Arg715Lysfs*41, p.Leu308Profs*11 and p.Asn495Thrfs*11) were also categorized under PM2 (absent in population database) and the nonsense and frameshift variants identified are categorized under PM4 (protein length changing variant).

SLC6A6- novel gene in LCA

The proband WES-RD20, is a confirmed LCA case with phenotypes of reduced vision from birth, nystagmus, photophobia, and an extinguished ERG under both scotopic and photopic conditions. The pedigree shows consanguinity but no family history of retinal diseases. By WES, we identified a novel missense variant g.14487240 C > T; c.245C > T; p.Pro82Leu, in *SLC6A6,* which segregated in the family (Fig. 2) and were found to be pathogenic by *in silico* tools and ACMG guidelines.

To further confirm the pathogenicity, we performed homology modeling and molecular dynamics simulation of the SLC6A6 protein. The initial homology model of SLC6A6 protein featured the lowest DOPE score energy among 100 models, but showed a few residues in the disallowed regions of the Ramachandran plot. Therefore, the structure was further refined using the Galaxy Refine webserver. The refined structure showed 94.7% of residues to be in the favored region (Suppe S1A).

The refined structure also obtained a PROSA Z-score of -5 and ProQ LGScore of 5.719 and MaxSub score of 0.433, which is indicative of the higher plausibility of the modeled structure. The mutant structure was modeled using the refined Wild structure as a template, wherein it showed 96.2% residues in the favored region (Fig. S1B). The mutant structure obtained a PROSA Z-score of -5.07 and ProQ LGScore of 5.463 & MaxSub score of 0.413, thereby reassuring the higher plausibility of the predicted model. The mutation Pro82Leu introduced the additional hydrophobic side chain in the transmembrane region and was found to have increased the protein stability, as predicted using the SDM2 server, wherein it scored a pseudo ΔΔG value of 3.4 kcal/mol and the mCSM-membrane also showed stabilizing ΔΔG value of 0.559 kcal/mol. The mutation Pro82Leu also showed the Mutpred2 score of 0.954 and predicted the gain of the helix at the region of mutation. The modeled protein structure was composed of 12 transmembrane helices spanning the membrane with extensive intra- and extracellular loops connecting the helices as shown in Fig. 3.

From the RMSD graph, it can be observed that the wild type maintains the lesser backbone deviation of 0.24 ± 0.02 nm until the time period of ~ 150 ns, and a slight increase in RMSD deviation was observed beyond it. However, the wild protein was observed to have an average RMSD of 0.25 ± 0.03 nm deviation for the total simulation time period. The mutant Pro82Leu showed the RMSD deviation to be similar to that of the wild until 80 ns, with an average RMSD of 0.22 ± 0.02 nm which was slightly higher than the wild form. The mutant structure also showed an average RMSD of 0.28 ± 0.06 nm for the total simulation period which is comparatively higher than the wild (Fig. 4A).

The radius of gyration throughout the simulation period for both wild and mutant type was calculated, wherein it was observed that the wild had an average deviation of 2.48 ± 0.01 nm gyration deviation, while the mutant showed an average deviation of 2.46 ± 0.01 nm (Fig. 4B). This indicates that both the wild and mutant types tend to maintain the structural compactness for the entire production run. The intra-H-bond within the proteins during the simulation were calculated, wherein no significant change in the number of H-bonds between the wild (average of 366.47 ± 12.22 H-bonds) and the mutant (Pro82Leu)(average of 365.7 ± 12.74 H-bonds) (Fig. 6A). This finding also confirms the structural compactness with concurrence to the radius of gyration plots.

In the RMSF graph, it can be observed that residues spanning the intracellular and extracellular loop region in the mutant structure were found to fluctuate more than the wild structure. The regions (EL3 and TM2) around the mutation site L82 showed more fluctuations in the wild when compared to the mutant form during the simulation (Fig. 4C). The secondary structure changes during the simulation period were plotted as shown in Fig. 5A & B. In case of mutant, it can be observed that the secondary structure changes to occur around the mutation site region spanning the TM2(78–104) helix. In case of wild, the transition of secondary

Table 1 List of variants identified by WES in the current study

Family NO	Gene; genome ID; cDNA; protein	Exonic/intronic; zygosity; mutation type	dbSNP/ClinVar ID	In silico prediction tools				
				SIFT	PolyPhen	Mutation taster	CADD	Provean
WES-RD5	*AHI1* g.135754344 T > A; c.2087A > T; p.His696Leu	Exonic; Homozygous; Missense	Novel	D	D	Disease causing	27	Deleterious
WES-RD11	*SPATA7* g.88883069C > T; c.253C > T; p.Arg85Ter	Exonic; homozygous nonsense	rs140287375 VCV000030806	NA	NA	Disease causing	37	NA
WES-RD14	*SPATA7* g.88903897C > T; c.1171C > T; p.Arg391Ter	Exonic; homozygous; Nonsense	rs374268850 VCV000489379	NA	NA	Disease causing	41	NA
WES-RD16	*CRB1* g.197396587_197396596delATG TGGCAGG; c.2132_2141delATGTGG CAGG; p.Tyr711Serfs*40	Exonic; homozygous; frameshift	Novel	NA	NA	Disease causing	NA	NA
	CRB1 g.197396599_197396603insAAATT; c.2143_2144 insAAATT; p.Arg715Lysfs*41	Exonic; homozygous; frameshift	Novel	NA	NA	Disease causing	NA	NA
WES-RD17	*CRB1* g.197396733_197396748delTATCAA TATATCCGTG; c.2278_2293delTAT CAATATATCCGTG; p.Tyr760Serfs*3	Exonic; homozygous; Frameshift	HGMD CM130791	NA	NA	Disease causing	NA	NA
WES-RD20	*MYO7A* g.76901072G > A; c.3638G > A; p.Arg1213Gln	Exonic; homozygous; Missense	rs372642675 RCV001111013.2	N	PD	Disease causing	26.9	Tolerated
WES-RD3	*SLC6A6* g.14487240 C > T; c.245C > T; p.Pro82Leu	Exonic; homozygous; Missense	rs1221379240	D	PD	Disease causing	26.8	Deleterious
WES-RD6	*RPGRIP1* g.21778758_21778759insCTC CAGAA; c.922_923insCTCCAGAA; p.Leu308Profs*11	Exonic; homozygous; frameshift	Novel	NA	NA	Disease causing	NA	NA
WES-RD18	*RPGRIP1* g.21789430delA; c.1480delA; p.Asn495Thrfs*11	Exonic; compound heterozygous; Frameshift	Novel	NA	NA	Disease causing	NA	NA
	g.21794290C > T; c.2668C > T; p.Arg890*	Nonsense	rs780587095 RCV002518810.2	NA	NA	Disease causing	37	NA
WES-RD2	*AIPL1* g.6329091C > A; c.844G > T; p.Glu282*	Exonic; homozygous; nonsense	rs779454542	NA	NA	Disease causing	42	NA
WES-RD9	*RPGRIP1* g.21778769A > G; c.930 + 3A > G	Intronic; homozygous; splice	rs150107283 RCV000585329.29	NA	NA	Disease causing	21.6	NA
WES-RD10	*LCA5* g.80201335_80201341delGTT TTCG; c.1062_1068delGTTTTCG; p.Cys353_Tyr354insTer	Exonic; homozygous; nonsense	rs1769845495 RCV001030781.6	NA	NA	Disease causing	NA	NA
WES-RD1	*CNGA3* g.99012480C > T; c.847C > T; p.Arg283Trp	Exonic; homozygous; missense	rs104893613 VCV000009474.31	D	PD	Disease causing	26.5	Deleterious
WES-RD4	*RPGRIP1L* g.53686569G > A; c.2030C > T; p.Thr677Ile	Exonic; homozygous; missense	rs532768944	D	PD	Disease causing	23.7	Deleterious
WES-RD15	*CNNM4* g.97462821G > A; c.1475G > A; p.Gly492Asp	Exonic; homozygous; missense	rs2153349278 VCV0013427021	D	PD	Disease causing	29	Deleterious

This table shows the variants identified by WES along with the bioinformatics analyses of these variants using in silico prediction tools. D—damaging, PD—probably damaging, N—neutral, and NA—not applicable

Fig. 1 Showing electrophoretogram of the identified WES variants in each family validated by Sanger sequencing

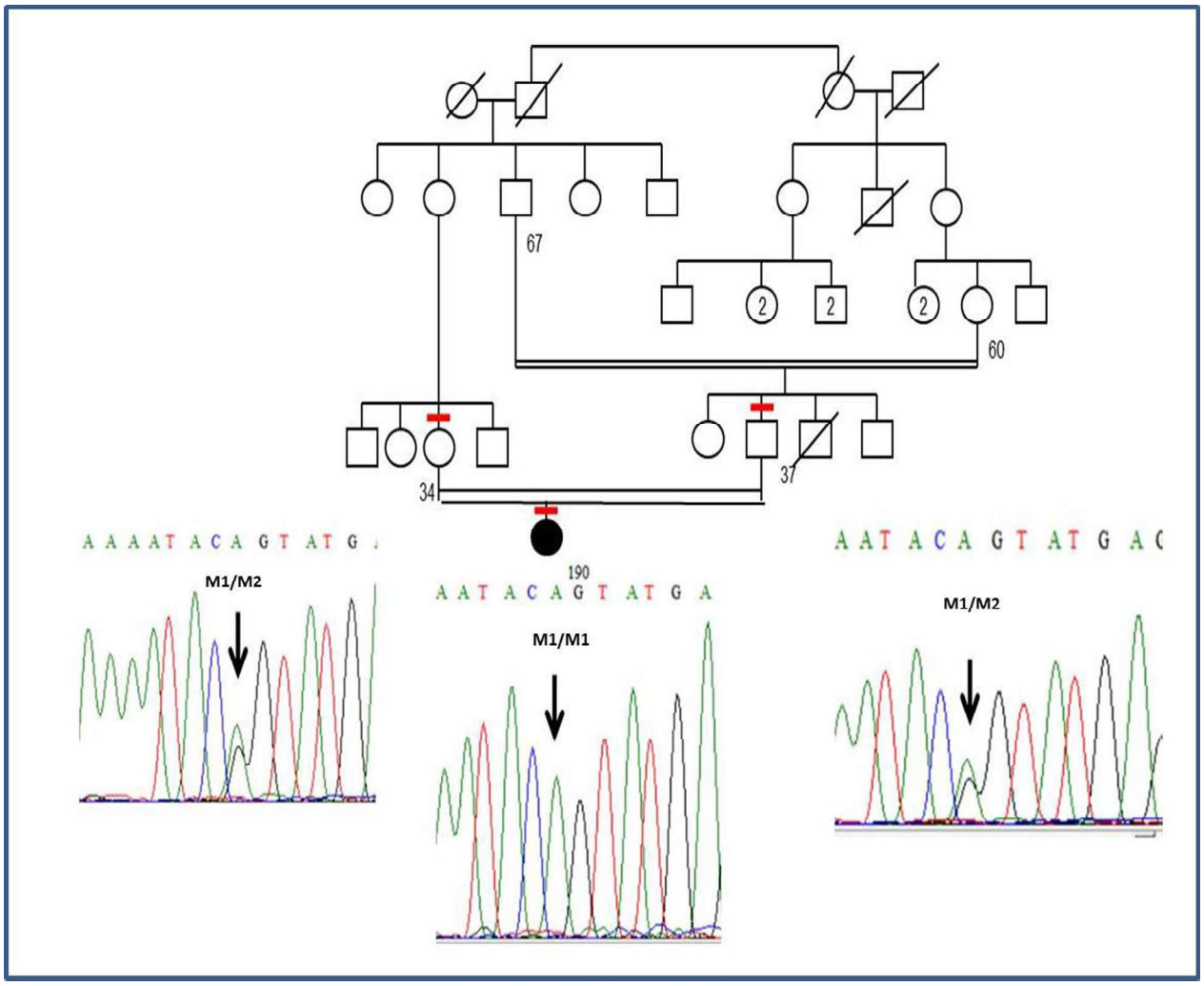

Fig. 2 Pedigree of WES-RD3 family—mutation in *SLC6A6* gene- g.14487240 C > T; c.245C > T; p.Pro82Leu is indicated in arrow (reverse sequence). The proband is homozygous for the variant while the parents are heterozygous. M1-c.245C > T, M2-wild type

structure between turn to α-helix was observed during the simulation, whereas in the case of mutant, a higher number of α-helical characteristics were maintained throughout the time period in the TM2 region.

As the mutation Pro82Leu introduced an additional aliphatic hydrophobic side chain, we calculated the number of additional hydrophobic contacts formed around the region of the mutational site during the simulation period. The intra-hydrophobic contacts formed by the side chain of wild P82 and Mutant L82 during the simulation period were also calculated (Fig. 6B). It was observed that the mutant L82 formed an increased number of hydrophobic contacts than the wild with residues 285 and 290, while other hydrophobic interactions were also observed with other residues such as 291and 294 (Fig. 6C).

Discussion

In the current study, we performed whole-exome sequencing in 20 LCA cases that were negative for candidate gene mutations by targeted resequencing. Mutations were identified in 75% of the cases with 33% variants being novel.The candidate genes identified in our current study were encompassed within our previous targeted panel. However, our previous pipeline failed to detect the variants that were subsequently identified through WES. This emphasizes the necessity of addressing the limitations of NGS technology in our data interpretation, particularly in terms of differences in capture arrays, chemistry, and/or bioinformatics pipelines [33, 34]. Upon reanalysis, we observed a notable improvement in read quality, depth, and coverage in these regions within the corresponding patient's

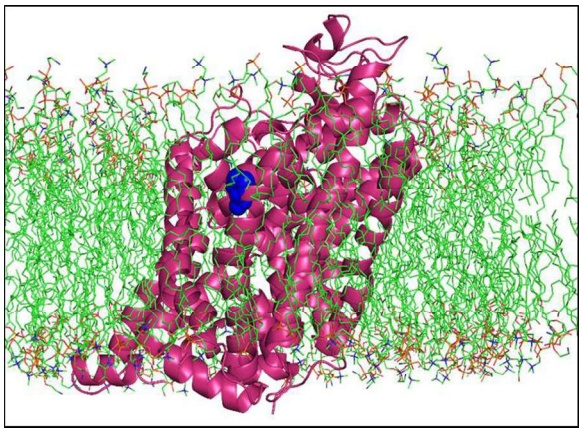

Fig. 3 The protein structure SLC6A6 embedded in the lipid membrane for molecular dynamics simulation. The SLC6A6 shown in cartoon representation colored as hot pink color and the mutation site Pro82Leu as sphere representation with blue color. The membrane are shown in line representation with green as element color

WES data compared to our targeted NGS chemistry. In the current study, for 33% cases (5 out of 15), the novel variants were identified in known LCA candidate genes (*CRB1, RPGRIP1*). We found that the gene coverage for *LCA5* (WES-RD10) was 99.04% with an exon depth of 32.9 in the mutation-identified region. Similarly, *RPGRIP1* (in WES-RD6, WES-RD18, WES-RD9) were ≈98.9% (3 samples) with an average exon depth of 50, *SPATA7* (WES-RD11, WES-RD14) with 95.8% (exon depth 52) and 79.12% (exon depth 43), respectively. This clearly explains the inability of our targeted resequencing data to identify the variants.

Phenotypically, all 20 cases displayed the characteristic features of LCA, including extinguished or non-recordable electroretinogram (ERG) responses, nystagmus, and onset of symptoms from birth. Of these, 55% (11/20) had mutation in known candidate genes of LCA which were non-syndromic and 20% (4/20) cases had a mutation in other inherited retinal degeneration genes like *CNGA3, RPGRIP1L, AHI1* and *CNNM4*. The remarkable fundus features noticed were attenuated vessels and RPE mottling in midperiphery in almost all, peripheral white spots in WES-RD10 with *LCA5* gene mutation, altered reflex in WES-RD11 and WES-RD14 with *SPATA7* gene mutations, nummular pigmentation and macular atrophy in WES-RD16 with *CRB1* gene mutations, keratoconnus, atrophy with scalloped margin, yellow pigmentation in the center and coloboma like phenotype in WES-RD2 with *AIPL1* gene mutation.

WES-RD1 presented to the clinic at 6 months after birth with jerk nystagmus and photophobia since birth.

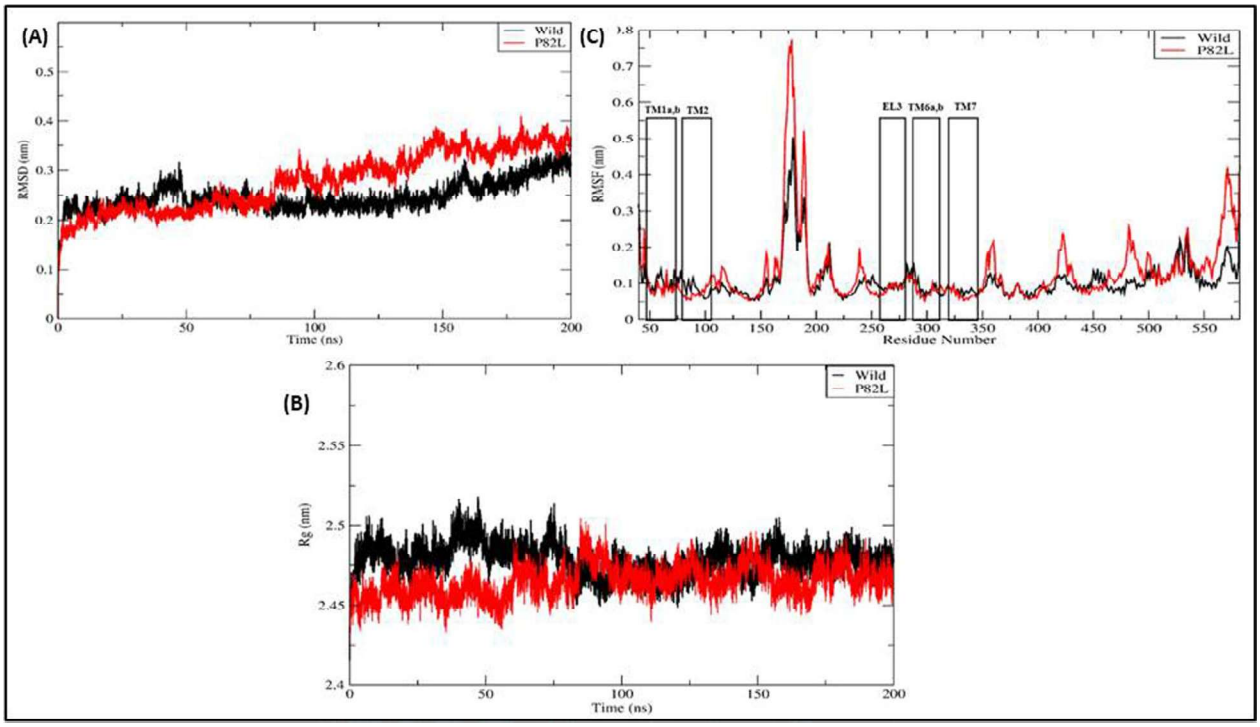

Fig. 4 **A** Root Mean Square Deviation (RMSD), **B** Radius of Gyration of SLC6A6 protein structure wild (black) and mutant Pro82Leu (red) during the simulation period. **C** Root Mean Square Fluctuation (RMSF)

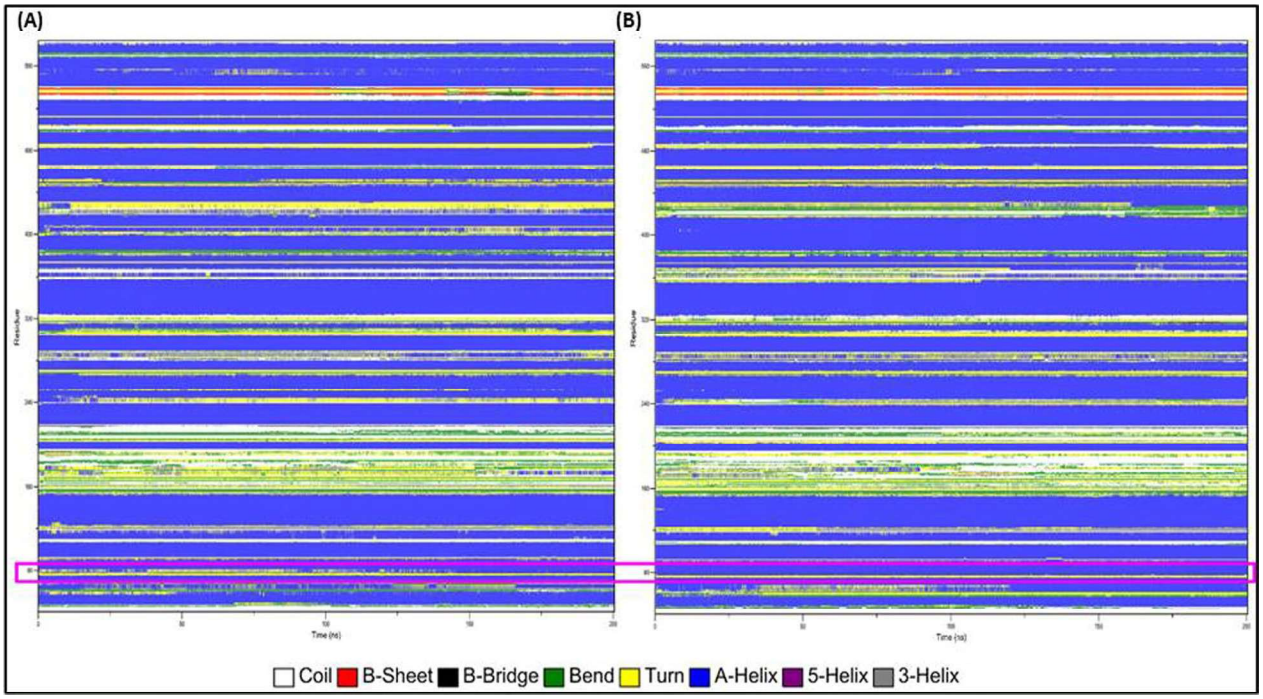

Fig. 5 Secondary structure plot of SLC6A6 (**A**) wild (**B**) mutant Pro82Leu during the simulation period. The mutant site Pro82Leu around the transmembrane helix region (TM2) is highlighted in the pink box

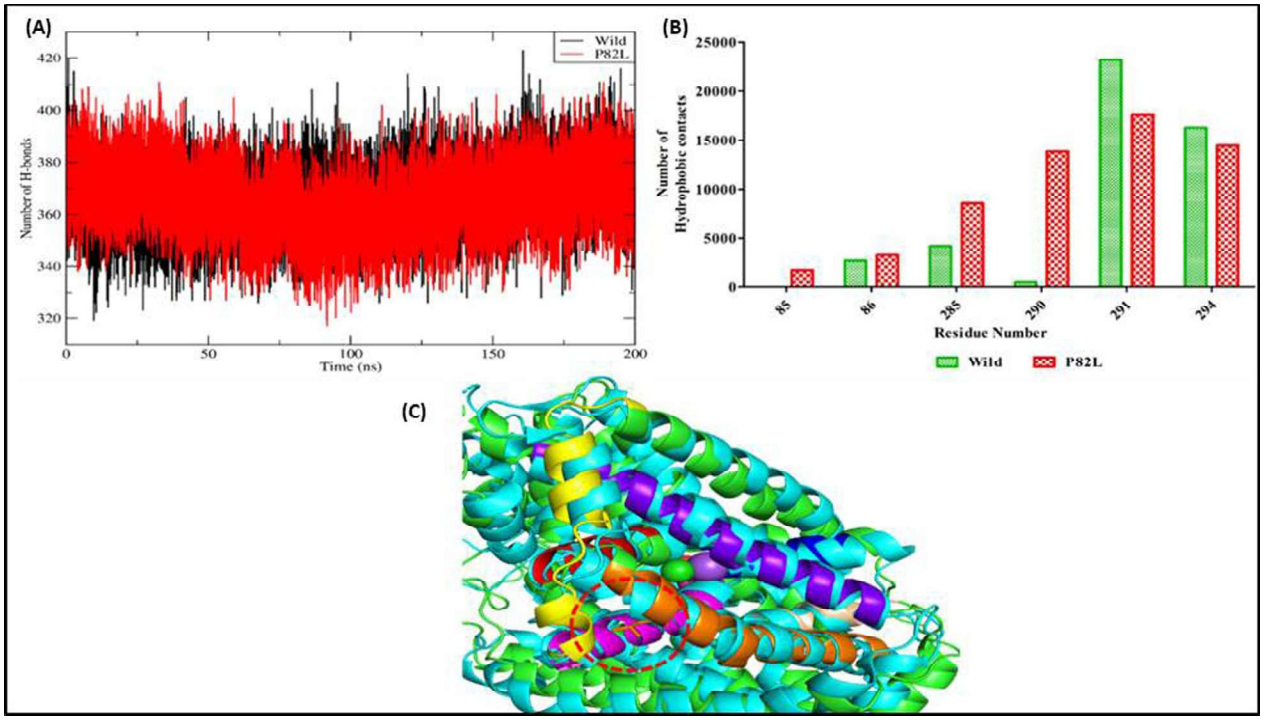

Fig. 6 **A** Intra-H-bond of wild (black) and mutant Pro82Leu (red) color, **B** Number of hydrophobic contacts formed by the side of P82 (wild) and L82 (mutant) during the simulation (**C**) The cartoon representation of the mutant Pro82Leu structure (Green Color) overlay with LeuTAa structure PDB ID:4HOD (Cyan color) structure. The binding site regions in SLC6A6 are colored as follows TM1a (blue); TM1b (red); TM2 (orange); EL3 (yellow); TM6a (magenta); TM6b (wheat); TM7 (purple-blue). The 82L shown in stick representation and Na^+, Cl^- ions are shown as a sphere

There was a history of second-degree consanguinity in the family. Whole-exome sequencing identified a homozygous mutation p.Arg283Trp in *CNGA3* gene segregating in the family. Although *CNGA3* gene mutations have been shown to cause LCA [35], we found that the same variant p.Arg283Trp has been shown to be associated with achromatopsia [36] and cone-rod dystrophy cases [37]. Re-evaluation of the ERG of the patient showed a non-recordable photopic response and reduced/delayed single flash rod and combined response thus rediagnosing the case to be cone-rod dystrophy.

WES-RD4 with *RPGRIP1L* p.Thr677Ile mutation presented with typical features of LCA with a non-recordable scotopic and photopic response, the fundus showed RPE mottling in the midperiphery and tapetal reflex in the macula. This particular mutation was also previously detected in LCA cases [38] and Joubert syndrome type B patient [39]. Co-immunoprecipitation studies have shown a reduced interaction of the p.Thr677Ile mutant *RPGRIP1L* with nephrocystin-4 leading to the disease phenotype of Jouberts [39]. But follow-up of our patient did not show any other symptoms apart from the typical LCA phenotype.

At the age of 4 months, WES-RD5 carrying the *AHI1* p.His696Leu mutation exhibited a lack of responsiveness to bright light and did not display any fixation on light stimuli. The ERG revealed a complete absence of scotopic and photopic responses with no changes in the macula. By the age of 9, the child exhibited delayed developmental milestones and significant hyperopia, which collectively pointed toward a likely diagnosis LCA. This is the first report of *AHI1* mutation in LCA. In cases of RP, a distinct nucleotide alteration within the same region (c.2087A > G) resulted in a different codon and, consequently, an amino acid change (p.His696Arg), leading to decreased protein enrichment in ciliated cells. Like the cases described, the WES-RD5 patient also displayed non-syndromic RD [40].

The current study identified only one syndromic case among the subjects examined. WES-RD15, presented at the age of 2 years, with extinguished ERG photophobia and high hyperopia.The fundus presented with attenuated vessels, granular midperiphery and atrophic macula. Following genetic testing and the identification of the c.1475G > A variant in *CNNM4* gene, a re-evaluation of the patient uncovered a diagnosis of amelogenesis imperfecta, leading to revised diagnosis, now indicating Jalli syndrome. The same mutation has been reported in a patient with hypoplastic-hypocalcified amelogenesis imperfecta and cone-rod dystrophy [41].

In WES-RD3, we have reported for the first time a pathogenic mutation in the *SLC6A6* gene. The patient reported with typical features of LCA and has been on follow-up ever since for about a decade now. The fundus showed bull's eye lesion and RPE atrophy. WES identified p.Pro82Leu mutation in *SLC6A6* gene which is segregated in the consanguineous family. *SLC6A6* is a sodium and chlorine dependent taurine transporter multi-pass membrane protein and has been shown to be present in neutrophils, heart, skeletal muscle, brain, liver, and most abundantly in the retina. Taurine plays a role in many biological activities, including antioxidation, neurotransmission, osmoregulation, and membrane stabilization [42, 43]. In the retina, taurine is shown to be present in abundant and is important for the protection/survival of photoreceptors from oxidative stress caused by light damage [44]. SLC6A6/ TAUT supplies taurine to retinal pigment epithelium and thereby to photoreceptors thus helping in its maintenance. Thus mutation in *SLC6A6* has been shown to cause pan-retinal degeneration in two brothers where supplementation of taurine had shown to reverse cardiomyopathy and retard retinal degeneration in the younger brother [15]. Similar reports are available for two more families thereby collectively reporting only 2 variants in three families so far [45, 46].

In the current study, we reported a *SLC6A6* mutation p.Pro82Leu in a consanguineous south Indian family WES-RD3. Upon characterizing the variant by in silico analysis, we observed that the regions such as TM6a/b, TM1a/b, and TM7 which contain Na^+ and Cl^- binding sites involving the taurine transport cycle were found to show lesser deviation in mutant (Pro82Leu) form than the wild-type structure. These residues forming additional hydrophobic interactions in the mutant are found to be in the TM6a region which involves in the Na^+Cl^- binding sites (Fig. 6C). The additional hydrophobic interactions formed were also found to restrict fluctuation of the TM6a region, which is apparent as per the RMSF plot. An earlier study on the SLC6A6 A78E mutation has also reported having a similar residual restriction of fluctuation around the Na^+ & Cl^- binding sites which was due to additional salt bridge formation upon mutation [15]. The secondary structure around the mutation TM2 helix was observed to have more helical characteristics in the mutant form around the TM2 helix which could potentially hinder the regular function of the taurine transport cycle. The transmembrane helices will be undergoing structural conformational changes during the taurine transport cycle. The stabilization of the TM2 helix to more helical characteristics and increased hydrophobic interaction gained around the mutation region will restrict the conformation changes for the helices needed for the ions and taurine transport. The decrease in conformational change might lead to reducing intake of taurine into the cell and could lead to its deficiency.

In conclusion, our study highlights the power of whole-exome sequencing (WES) in unraveling the genetic causes of leber congenital amaurosis (LCA), especially in cases where targeted resequencing fails and identified mutations in 75% of cases, including 33% novel variants. By identifying a novel mutation in the SLC6A6 gene, we have expanded our understanding of the genetic landscape of this complex disease. The SLC6A6 gene encodes a sodium and chlorine-dependent taurine transporter, crucial for retinal function. Our in silico analysis revealed that the identified mutation, p.Pro82Leu, likely disrupts the taurine transport cycle by altering the structure and dynamics of the protein. This suggests a potential mechanism for the development of LCA in the affected family. The identification of a novel mutation in SLC6A6 has significant implications for our understanding of LCA and the potential for future therapeutic interventions. Taurine, the substrate of the SLC6A6 transporter, plays a vital role in retinal health. It is possible that therapeutic supplementation with taurine could mitigate the effects of the SLC6A6 mutation and improve visual function in patients with LCA.

While WES has proven to be a powerful tool, it is not without limitations. In our study, five cases remained unsolved, suggesting the presence of mutations in non-coding regions or regulatory elements that are not detectable by WES. Whole-genome sequencing may be necessary to identify these elusive genetic variants. Future studies should focus on functional validation of the identified SLC6A6 mutation to confirm its pathogenicity and explore the potential therapeutic benefits of taurine supplementation. Additionally, larger-scale studies are needed to determine the prevalence of SLC6A6 mutations in LCA patients and to evaluate the effectiveness of targeted therapies.

Abbreviations

IRD	Inherited retinal dystrophies
LCA	Leber congenital amaurosis
WES	Whole-exome sequencing
MAF	Minimum allele frequency
ACMG	American college of medical genetics and genomics
ERG	Electroretinogram
RMSD	Root mean square deviation
RP	Retinitis pigmentosa
RD	Retinal degeneration

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s43042-025-00659-x.

Supplementary file 1. Figure S1(A-B): Ramachandran plot for the modelled SLC6A6 protein structure of (A) wild and (B) mutant Pro82Leu

Acknowledgements
We thank the patients and their families for their participation in the study. We acknowledge the support provided by MedGenom Labs Pvt. Ltd, Bangalore, India, for whole-exome sequencing and initial analysis. We are grateful to SciGenom for funding (SGRF grant 2018).

Author contributions
SM, SP, SS, and VUS were responsible for the conception and design of the study. Data acquisition was carried out by SNN, SA, SP, RG, and VUS. The analysis and interpretation of data were performed by SNN, SA, RG, SS, and VUS. SNN and SA drafted the manuscript, while SM, SNN, and VUS critically revised it for important intellectual content. All authors read and approved the final manuscript.

Funding
The whole-exome sequencing was supported by the students' grant received under the 2018 SGRF Genomics Project Grant Award. The analysis, validation, and interpretation of data, as well as the manuscript writing, were conducted using the reagents and manpower available from the PI's completed project. No additional funding was available for these activities.

Availability of data and materials
The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate
The project was approved by the Institutional Review Board—Ethics Committee of the Vision Research Foundation (708–2018-P).

Consent for publication
We have recruited the patients' families for the project with their informed consent. All consent forms are in our possession and can be provided upon request.

Competing interests
The authors declare that they have no competing interests.

Author details
[1]SNONGC Department of Genetics and Molecular Biology, Vision Research Foundation, Sankara Nethralaya Campus, No. 18, College Road, Nungambakkam, Chennai 600006, Tamil Nadu, India. [2]Formerly With Centre for Bioinformatics, Kamalnayan Bajaj Institute for Research in Vision and Ophthalmology, Vision Research Foundation, Sankara Nethralaya, Chennai, India. [3]Formerly With Medical Research Foundation, Shri Bhagwan Mahavir Vitreoretinal Services, Sankara Nethralaya, Chennai, India. [4]MedGenome Labs Pvt. Ltd, Bangalore, India. [5]Department of Health Research, (Govt. of India), National Institute of Traditional Medicine, Indian Council of Medical Research, Belagavi, Karnataka, India.

Received: 20 May 2024 Accepted: 31 January 2025
Published online: 17 February 2025

References

1. Berger W, Kloeckener-Gruissem B, Neidhardt J (2010) The molecular basis of human retinal and vitreoretinal diseases. Prog Retin Eye Res 29(5):335–375. https://doi.org/10.1016/j.preteyeres.2010.03.004
2. Suppiej A, Marino S, Reffo ME et al (2019) Early onset retinal dystrophies: Clinical clues to diagnosis for pediatricians. Ital J Pediatr. https://doi.org/10.1186/s13052-019-0760-5
3. Daiger S, Rossiter B, Greenberg J, et al (1998) Data services and software for identifying genes and mutations causing retinal degeneration. Investig Ophthalmol Vis Sci. 39.
4. Estrada-Cuzcano A, Koenekoop RK, Coppieters F et al (2011) IQCB1 mutations in patients with leber congenital amaurosis. Investig Ophthalmol Vis Sci. https://doi.org/10.1167/iovs.10-5221
5. King MD, Stephenson JBP (1984) Association of Joubert's syndrome with Leber's congenital Amaurosis. Arch Neurol 41(12):1235. https://doi.org/10.1001/archneur.1984.04050230013004

6. Srikrupa NN, Meenakshi S, Arokiasamy T, Murali K, Soumittra N (2014) Leber's congenital amaurosis as the retinal degenerative phenotype in thiamine responsive megaloblastic anemia: a case report. Ophthalmic Genet 35(2):119–124. https://doi.org/10.3109/13816810.2013.793363
7. Shukla R, Chitra Kannabiran JS (2012) Genetics of Leber congenital amaurosis: an update. Expert Rev Ophthalmol. 7:141–151
8. Falk MJ, Zhang Q, Nakamaru-Ogiso E, Kannabiran C, Fonseca-Kelly Z, Chakarova C, Audo I et al (2012) NMNAT1 mutations cause Leber congenital amaurosis. Nat Genet 44(9):1040–1045
9. Keen TJ, Mohamed MD, McKibbin M, Rashid Y, Jafri H, Maumenee IHIC (2003) dentification of a locus (LCA9) for Leber's congenital amaurosis on chromosome 1p36. Eur J Hum Genet 11(5):420–423
10. Wang X, Wang H, Cao M, Li Z, Chen X, Patenia C, Gore A, Abboud EB, Al-Rajhi AA et al (2011) Whole-exome sequencing identifies ALMS1, IQCB1, CNGA3, and MYO7A mutations in patients with Leber congenital amaurosis. Hum Mutat. 32(12):1450–1459
11. Minegishi Y, Sheng X, Yoshitake K, Sergeev Y, Iejima D, Shibagaki Y, Monma N et al (2016) CCT2 Mutations evoke leber congenital amaurosis due to chaperone complex instability. Sci Rep 6:33742
12. Soens ZT, Li Y, Zhao L, Eblimit A, Dharmat R, Li Y, Chen Y et al (2016) Hypomorphic mutations identified in the candidate Leber congenital amaurosis gene CLUAP1. Genet Med 18(10):1044–1051
13. Srikrupa NN, Srilekha S, Sen P, Arokiasamy T, Meenakshi S, Bhende M et al (2018) Genetic profile and mutation spectrum of Leber congenital amaurosis in a larger Indian cohort using high throughput targeted re-sequencing. Clin Genet 93(2):329–339
14. Webb B, Sali A (2016) Comparative protein structure modeling using MODELLER. Curr Protoc Bioinforma 54(1):5–6
15. Preising MN, Görg B, Friedburg C, Qvartskhava N, Budde BS, Bonus M et al (2019) Biallelic mutation of human SLC6A6 encoding the taurine transporter TAUT is linked to early retinal degeneration. FASEB J 33(10):11507–11527
16. Heo L, Park HSC (2013) GalaxyRefine: Protein structure refinement driven by side-chain repacking. Nucleic Acids Res. 41:384–8
17. Lee GR, Heo LSC (2016) Effective protein model structure refinement by loop modeling and overall relaxation. Proteins 84(Suppl 1):293–301
18. Wallner BEA (2003) Can correct protein models be identified? Protein Sci 12(5):1073–1086
19. Wiederstein MSM (2007) ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. Nucleic Acids Res. 35:W407-10
20. Pandurangan AP, Ochoa-Montaño B, Ascher DBBT (2017) SDM: a server for predicting effects of mutations on protein stability. Nucleic Acids Res 45(W1):W229–W235
21. Pires DE, Rodrigues CH, Ascher DB (2020) mCSM-membrane: predicting the effects of mutations on transmembrane proteins. Nucleic Acids Res 48(W1):W147–W153
22. Pejaver V, Urresti J, Lugo-Martinez J, Pagel KA, Lin GN, Nam H et al (2020) Inferring the molecular and phenotypic impact of amino acid variants with MutPred2. Nat Commun 11(1):5918
23. Jo S, Kim T, Iyer VGIW (2008) CHARMM-GUI: a web-based graphical user interface for CHARMM. J Comput Chem 29(11):1859–1865
24. Berendsen HJC, Postma JPM, van Gunsteren WF, DiNola A, Haak JR (1984) Molecular dynamics with coupling to an external bath. J Chem Phys 81(8):3684–3690
25. Nosé S (1984) A molecular dynamics method for simulations in the canonical ensemble. Mol Phys 52(2):255–268
26. Hoover WG (1985) Canonical dynamics: equilibrium phase-space distributions. Phys Rev A 31(3):1695
27. Parrinello M, Rahman A (1980) Crystal structure and pair potentials: a molecular-dynamics study. Phys Rev Lett 45(14):1196
28. Hess B, Bekker H, Berendsen HJ, Fraaije JG (1997) LINCS: a linear constraint solver for molecular simulations. J Comput Chem 18(12):1463–1472
29. Darden T, York D, Pedersen L (1993) Particle mesh Ewald: an N· log (N) method for Ewald sums in large systems. J Chem Phys 98(12):10089–10092
30. M.J. Abraham, T. Murtola, R. Schulz, S. Páll, J.C. Smith, B. Hess EL. GROMACS: high performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX*. 2015;1–2:19–25.
31. Salentin S, Schreiber S, Haupt VJ, Adasme MF, Schroeder M (2015) PLIP: fully automated protein–ligand interaction profiler. Nucleic Acids Res 43(W1):W443–W447
32. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, et al. (2015) ACMG Laboratory Quality Assurance Committee. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med*. 17(5):405–24.
33. DE Abel HJ (2013) Detection of structural DNA variation from next generation sequencing data: a review of informatic approaches. Cancer Genet 206(12):432–440
34. LaDuca H, Farwell KD, Vuong H, Lu HM, Mu W et al (2017) Exome sequencing covers >98% of mutations identified on targeted next generation sequencing panels. PLoS ONE 12(2):e0170843
35. Wang X, Wang H, Cao M, Li Z, Chen X et al (2011) Whole-exome sequencing identifies ALMS1, IQCB1, CNGA3, and MYO7A mutations in patients with Leber congenital amaurosis. Hum Mutat 32(12):1450–1459
36. Zobor D, Werner A, Stanzial F, Benedicenti F, Rudolph G et al (2017) The clinical phenotype of CNGA3-related achromatopsia: pretreatment characterization in preparation of a gene replacement therapy trial. Invest Ophthalmol Vis Sci 58(2):821–832
37. Wissinger B, Gamer D, Jägle H, Giorda R, Marx T et al (2001) CNGA3 mutations in hereditary cone photoreceptor disorders. Am J Hum Genet 69(4):722–737
38. Khanna H, Davis EE, Murga-Zamalloa CA, Estrada-Cuzcano A, Lopez I et al (2009) A common allele in RPGRIP1L is a modifier of retinal degeneration in ciliopathies. Nat Genet 41(6):739–745
39. Delous M, Baala L, Salomon R, Laclef C, Vierkotten J et al (2007) The ciliary gene RPGRIP1L is mutated in cerebello-oculo-renal syndrome (Joubert syndrome type B) and Meckel syndrome. Nat Genet 39:875–881
40. Nguyen TT, Hull S, Roepman R et al (2017) Missense mutations in the WD40 domain of AHI1 cause non-syndromic retinitis pigmentosa. J Med Genet. 54(9):624–632. https://doi.org/10.1136/jmedgenet-2016-104200)
41. Rattanapornsompong K, Gavila P, Tungsanga S et al (2023) Novel CNNM4 variant and clinical features of Jalili syndrome. Clin Genet 103(2):256–257. https://doi.org/10.1111/cge.14258
42. Heller-Stilb B, van Roeyen C, Rascher K, Hartwig HG, Huth A et al (2002) Disruption of the taurine transporter gene (taut) leads to retinal degeneration in mice. FASEB J 16(2):231–233
43. Schaffer SKH (2018) Effects and mechanisms of taurine as a therapeutic agent. Biomol Ther (Seoul) 26(3):225–241
44. Froger N, Moutsimilli L, Cadetti L, Jammoul F, Wang QP et al (2014) Taurine: the comeback of a neutraceutical in the prevention of retinal degenerations. Prog Retin Eye Res 41:44–63
45. Antonarakis SE (2020) Taurine newborn screening to prevent one form of retinal degeneration and cardiomyopathy. Eur J Hum Genet 28(11):1479–1480
46. Ansar M, Ranza E, Shetty M, Paracha SA, Azam M et al (2020) Taurine treatment of retinal degeneration and cardiomyopathy in a consanguineous family with SLC6A6 taurine transporter deficiency. Hum Mol Genet 29(4):618–623

Publisher's Note