

Identification of *OPN1LW* Exon 3 Variants Impairing Red-Cone Function in Color Vision Deficiency

Shaaban Z. Omar^{1,2†} and Karim J. Karim³

¹Department of Biology, Faculty of Science and Health, Koya University,
Koya 44023, Kurdistan Region-F.R. Iraq

²Department of Medical Microbiology, Faculty of Science and Health, Koya University,
Koya 44023, Kurdistan Region-F.R. Iraq

³Department of Medical Laboratory Science, Faculty of Science and Health, Koya University,
Koya 44023, Kurdistan Region-F.R. Iraq

Abstract—The most common form of inherited color blindness is red-green color vision deficiency (CVD), which is frequently caused by mutations in the X-linked *OPN1LW* gene. Red cone malfunction is linked to mutations in exon 3 of this gene. In this study, the Ishihara test was used to evaluate the color vision of 1500 Kurdish students, ages 13–18. Polymerase chain reaction amplification and Sanger sequencing of *OPN1LW* exon 3 were performed on 50 students who had been diagnosed with protanopia or protanomaly. Variants (nucleotide changes) were analyzed using Geneious Prime® software. Functional impact of variants was predicted using PolyPhen-2 and SIFT. The study found 30 different nucleotide variations, comprising 63.3% missense mutations, 23.3% silent mutations, and 13.3% frameshift mutations. The most common variants were found c.30G>A (p.Arg10Arg), c.106T>C (p.His35Pro), and c.161_162insG (p.Asp54Gly). SIFT found (57.8%) of variations as deleterious (scoring ≤0.05), but PolyPhen-2 assessed (63.1%) as potentially damaging (score >0.9). ABO blood type was unrelated to CVD risk, although consanguinity and family history were strongly linked to CVD risk. Our study revealed that people with red-green CVD have frequent and possibly harmful mutations in exon 3 of *OPN1LW*. These results may aid in the molecular characterization of CVD in the Kurdish population and could help develop future diagnostic and treatment approaches.

Index Terms—Color vision deficiency, Exon 3, Nucleotide variants, *OPN1LW* gene, PolyPhen-2, SIFT.

I. INTRODUCTION

Color vision deficiency (CVD) is estimated to affect approximately 300 million people worldwide. Although complete color blindness is uncommon, perception and discrimination of specific colors are substantially impaired in

CVD (Lilja, 2024). An understanding of CVD is important for affected individuals and for society, as accurate color perception underpins activities ranging from transportation safety to professional performance. Hence, reduced or absent color perception imposes barriers to daily tasks and participation in occupations such as driving, teaching, fashion design, and firefighting (Alam, et al., 2022). Red-green CVD – among the most prevalent inherited disorders globally – is associated with educational barriers, restricted access to careers in aviation, medicine, and electrical engineering, and risks in everyday tasks such as interpreting warning lights (Asadi, 2023). Given its worldwide burden, clarification of the genetic causes of this condition is necessary for biological and public-health relevance (Kuo, Tsao and Pei-Chang, 2023).

The human retina comprises three kinds of cone photoreceptors, each expressing a distinct visual opsin with characteristic spectral sensitivity (Segre, 2019). *OPN1LW* encodes the long-wavelength-sensitive L-cone opsin, which has a peak absorption at about 560 nm (red perception); *OPN1MW* encodes the medium-wavelength-sensitive M-cone opsin, which has a peak at about 532 nm (green perception); and *OPN1SW* encodes the short-wavelength-sensitive S-cone opsin, which has a peak at about 420 nm (blue perception) (Atilano, et al., 2020; Cai, 2023). A congenital condition called X-linked color vision deficit (CCVD) affects the ability to distinguish colors because of mutations in the genes that produce cone photoreceptors (Akhtar, et al., 2019). Cone type affects the severity of CCVD; monochromacy affects many cone types, but anomalous trichromacy and dichromacy affect just one kind of cone (Swathi, et al., 2020). Approximately one in 12 males and one in 200 females suffer from red-green blindness (protanopia or deuteranopia) (Haim, Fledelius and Skarsholm, 1988). L and M-opsin genes (*OPN1LW*, *OPN1MW*) reside on Xq28 in a tandem array and share 94% nucleotide identity (Nathans, Thomas and Hogness, 1986). Typically, gene arrays have one L gene, one or more M genes, and a locus control region (LCR). The first two genes are expressed because they are close to the upstream cis-regulatory region (Nathans, 1999). Mutations

ARO-The Scientific Journal of Koya University
Vol. XIII, No.2(2025), Article ID: ARO.12480. 15 pages
DOI: 10.14500/aro.12480

Received: 28 July 2025; Accepted: 05 October 2025
Regular research paper; Published: 10 November 2025

†Corresponding author's e-mail: shaaban.zirar@koyauniversity.org
Copyright © 2025 Shaaban Z. Omar and Karim J. Karim. This is
an open-access article distributed under the Creative Commons
Attribution License (CC BY-NC-SA 4.0).



of the LCR, L, and M opsin genes on Xq28 are connected to a number of X-linked diseases. These conditions include X-linked cone malfunction, blue cone monochromacy, and red-green blindness abnormalities (Aboshiha, et al., 2016) (Gardner, et al., 2014). Genetic variants linked to the L or M opsin genes impair vision in red and green colors. These alterations impair the perception and identification of red and green. As a result of these changes, L and M cones disappear, and aberrant opsin pigments are generated in these cones (Neitz and Neitz, 2011). In the L-M gene cluster, numerous harmful variants have been found (Neitz, Patterson, and Neitz, 2019). They are adjacent to one another and homologous, which causes rearrangements (Neitz, Neitz and Grishok, 1995). Furthermore, exon 3 of the L and M genes contains eight nucleotide variants that encode seven amino acids [c.453G > A; p.(Arg151Arg)], [c.457C > A; p.(Leu153Met)], [c.465G > A; p.(Val155Val)], [c.511G > A and c.513G > T; p.(Val171Ile)], [c.521C > T; p.(Ala174Val)], [c.532 A > G; p.(Ile178Val)], and [c.538G > T; p.(Ser180Ala)], each of which occurs frequently in the population. However, certain rare combinations of these variants are pathogenic because they cause incorrect protein splicing (Greenwald, et al., 2017). LIAVA in the L-gene and MVVVA in the M-gene are examples of combinations denoted by the one-letter amino acid abbreviations of the five amino acids at positions 153, 171, 174, 178, and 180 (Buena-Atienza, et al., 2016). Several studies using the Ishihara test have demonstrated the prevalence of red-green color-vision deficit in the Iraqi Kurdish people. Males have consistently been found to have a higher incidence throughout cohorts of school-aged children and young adults. In the same samples, deutan subtypes were reported more frequently than protan subtypes (deutan: 3.89%, 1.33%, 3.66%; protan: 1.74%, 0.73%, 2.75%). This indicates a recurring pattern in defect type (Hussein and Al-Dabbagh, 2022, Yousif, et al., 2024, Karim and Saleem, 2013). However, this research only used clinical testing, and there are currently no molecular studies in Kurdish communities. Therefore, the purpose of this study was to identify and characterize the nucleotide variants in exon 3 of the *OPN1LW* gene in a Kurdish cohort to broaden the understanding of red-cone dysfunction in CVD, along with their potential functional impact.

II. MATERIALS AND METHODS

A. Study Design and Participants

This study examined the genetic basis and prevalence of CVD among Kurdish students at the high school and secondary school levels. Schools across Erbil Province were randomly selected between February and May 2024, and approximately 100 students at each selected school underwent testing, yielding a total sample of 1500 participants; no formal a priori sample-size calculation was performed, and the sample size was defined operationally as ~100 students per school. Participants were 13–18 years of age (mean $\pm$ SD, 14.9 $\pm$ 1.6); 819 (54.6%) were male and 681 (45.4%) were female. Inclusion criteria were current

school enrollment, provision of parental consent and student assent, and ability to complete color-vision testing; exclusion criteria included any history of ocular disease or surgery and use of systemic medications known to affect vision. The study received approval from Koya University Faculty of Science and Health's Ethics Committee (Approval number 001MMB). The study acquired the agreement of both parents and students before participation. All participants in this study had their color vision evaluated using the Ishihara 38-Plate Tests, which were conducted in a room during daytime under a daylight-equivalent LED lamp ($\approx$ 6500 K) in a fixed setup; direct sunlight was excluded and glare minimized. Each plate was examined monocularly for up to 5 s at a distance of about 75 cm (Birch, 1997). Participants who misread four or more plates were identified with a red-green color vision deficit and underwent a molecular investigation. All assessments were administered by an examiner with eye-clinic testing expertise who completed protocol training and calibration for color-vision assessment before data collection.

B. Blood Sample Collection and Genomic DNA Extraction

The current study investigated 1500 students for protanopia and protanomaly. Based on Ishihara testing, the disorders were identified in only 50 individuals. Venous blood (2–3 mL) was collected into sterile K₂-EDTA vacutainer tubes and gently inverted 8–10 times to ensure anticoagulation from each of these 50 individuals. Immediately, tubes were placed in an insulated icebox and maintained at $\sim$ 4°C during transport; specimens reached the laboratory within 1 h of collection. Genomic DNA extraction was performed the same day ($\leq$ 24 h) using the [Quick-DNA™ Miniprep Plus Kit, Zymo Research, USA], according to the manufacturer's protocol. When immediate processing was not possible, whole blood was stored at -20°C (short term) or -80°C (long term) and thawed once before extraction. DNA quantity and purity were assessed using a Nanodrop spectrophotometer; acceptance thresholds were A260/A280 = 1.8–2.0 and A260/A230 $\geq$ 1.8. Samples outside these ranges were re-purified or re-extracted before downstream analyses. DNA integrity was tested by electrophoresis with a 0.8% agarose gel.

C. Polymerase Chain Reaction (PCR) Amplification and Gel Electrophoresis

To investigate genetic variants associated with color vision impairment, primers targeting *OPN1LW* exon 3 were custom-designed using Vector Bee (Vector Builder, a software tool for primer designing) in each of the 50 samples (Table I). PCR reactions were performed in a total volume of 25 μL , consisting of 50 ng of genomic DNA, forward and reverse primers, and 2 $\times$ Hot-Start Master Mix (ADDBIO INC) (Table II) for reagent details. The thermal cycling conditions are detailed in Table III. The amplified PCR products were electrophoresed on a 1.5% agarose gel that had been stained with 10 μL (0.5 $\mu\text{g/mL}$) of ethidium bromide. 6 $\times$ DNA loading dye (ADDBIO INC) (glycerol-based with bromophenol blue and xylene cyanol) to achieve a final 1.5 $\times$ dye concentration in the well (3 μL PCR product + 1 μL 6 $\times$

TABLE I
OPN1LW EXON 3 PRIMER SETS. PRIMERS WERE CUSTOM-DESIGNED USING VECTOR BEE SOFTWARE (VECTOR BUILDER) FOR PCR AMPLIFICATION

Primer name	Sequence (5'→3')	Length (bp)	Primer binding position	Product size (bp)
OPN1LW-F	GCCCTCATCTGTCTGCTCTC	20	33 bp upstream of exon 3	226
OPN1LW-R	CCACTCCATCTTGCGTCCTC	20	24 bp downstream of exon 3	

TABLE II
PCR REACTION COMPONENTS SPECIFIC TO THE AMPLIFICATION OF
EXON 3 OPN1LW

Component	Volume (μL)	Final concentration
2X Hot-Start Master Mix	12.5	1×
Forward Primer	1	10 pmol/l
Reverse Primer	1	10 pmol/l
DNA template	2	50 ng
DNA nuclease-free water	8.5	The final volume of 25 μL

TABLE III
PCR THERMAL CYCLING CONDITIONS (DAVIDOFF, 2015)

Steps/ Parameter	Initial denaturation	Denaturation	Annealing	Elongation	Final elongation
Temperature	94°C	94°C	59°C	72°C	72°C
Time	3 min	30 s	30 s	30 s	3 min
No. of cycles	1	30	30	30	1

dye). DNA ladder H3 RTU (100 bp) (Gene Dire X, Inc. [size range 100–3000 bp]) was used for fragment-size verification. Electrophoresis was performed in 1× TBE running buffer (89 mM Tris, 89 mM boric acid, 2 mM EDTA; pH ~8.3); gels were run for 50 min at 90 volts. Gels were imaged on a UV transilluminator and documented using an InGenius LHR Syngene Bio Imaging System Laboratory with consistent exposure settings.

D. Genomic Sequencing and Variant Analysis

PCR products from all 50 samples were purified using an enzymatic cleanup protocol (ExoSAP-IT™) to remove residual primers and dNTPs. The purified products were then sent to Macrogen Inc. (Seoul, South Korea). Macrogen performed all subsequent services, including the cleanup of the sequencing reaction and the actual Sanger sequencing. Unidirectional sequencing of exon 3 was performed on an ABI 3730xl sequencer (Applied Biosystems) using the BigDye Terminator v3.1 chemistry and 10 pmol of the reverse primer (5'-CCACTCCATCTTGCGTCCTC-3'). The chromatogram data were processed and analyzed using Geneious Prime® v2025.1.3. Obtained sequences were quality-checked and aligned to RefSeq Gene NG_009105.2 (OPN1LW exon 3) using the built-in NCBI BLASTn/megablast. An E-value cutoff of 1×10⁻²⁰ was applied; hits were accepted only if percent identity ≥97% and aligned length ≥95% of the amplicon. Query coverage was 100% for all amplicons, and no alternative genomic hits of comparable score were observed under these settings. The OPN1LW exon-3 amplicon measured 226 bp (primer pair in Table I). Sanger traces were quality-filtered in Geneious Prime with end-trimming at Q20, requiring ≥90% of bases with Phred ≥30 and a mean read quality ≥30.

All variants were recorded using the HGVS nomenclature format. Variants were then coordinated/normalized to GRCh38 (hg38) and represented as both chromosome HGVS (X: g on NC_000023.11) and cDNA coordinates on NM_020061.6 (MANE Select). Annotation was performed with Ensembl Variant Effect Predictor (VEP) on GRCh38 with RefSeq transcripts enabled and the options “Include existing variants” and “Frequencies” checked. Known variation was assessed via the Existing_variation field (dbSNP rsIDs) and cross-referenced in ClinVar and Gnome AD. Variants without an exact allele match across these resources were designated novel. Full per-variant details (nucleotide changes, HGVS name [NG_009105.2], and coordinates GRCh38 HGVS [NC_000023.11]) are provided in Table IV, with databases (Ensembl Variant Effect Predictor [VEP], ClinVar, and GenomeAD exports for the exon 3 region in supplementary Tables S2–S4 accessed September 2025)]. Bioinformatics predictor tools, SIFT (Sorting Intolerant From Tolerant) and PolyPhen-2 (Polymorphism Phenotyping v2), were used to evaluate whether the identified amino acid substitutions are likely to affect protein function. PolyPhen-2 integrates structural context, physicochemical differences, and evolutionary conservation to classify variants as benign, possibly damaging, or probably damaging. We used the HumVar model (PolyPhen-2 v2.2.2; server v2.1.0 r367) and applied the standard thresholds: benign <0.446, possibly damaging 0.446–0.909, and probably damaging ≥0.909 (Adzhubei, et al., 2010). SIFT considers sequence homology and the degree of evolutionary conservation of amino acids across species. SIFT infers functional impact from sequence homology and cross-species conservation; predictions were interpreted with the conventional cutoff <0.05 = deleterious and ≥0.05 = tolerated (SIFT 4G; GRCh38 database; site updated April 25, 2025) (Ng and Henikoff, 2003).

FASTA sequence files and raw exon-3 Sanger chromatograms, together with participant demographics and clinical features for individuals with CVD, are available from the corresponding author on reasonable request. The OPN1LW exon-3 Sanger sequences (226 bp amplicon; consensus FASTA) will be deposited in GenBank.

E. Statistical Analysis

Data processing and graphics were performed using GraphPad Prism (version 10.3.1) to summarize the frequency of nucleotide and amino acid variants across samples. All categorical comparisons were performed with Fisher’s exact tests. We report effect sizes as odds ratios (OR) with 95% confidence intervals (95% CI) and p-values. Fisher’s exact tests were employed to evaluate gender bias in the distribution of variant types. The relationships between mutation types and the anticipated functional impact (PolyPhen-2 and SIFT)

TABLE IV

COMPREHENSIVE VARIANT ANALYSIS OF *OPN1LW* EXON 3 WITH NORMALIZATION TO GRCh38. VARIANTS DETECTED FROM SANGER CHROMATOGRAMS WERE ALIGNED TO REFSEQ GENE NG_009105.2 (*OPN1LW*, EXON 3) AND ANNOTATED IN HGVS NOMENCLATURE; THE CORRESPONDING GRCh38 GENOMIC COORDINATES ON NC_000023.11 (chrX) ARE PROVIDED. TRANSCRIPT-LEVEL ANNOTATIONS REFER TO NM_020061.6 (MANE SELECT). THESE ENTRIES CONSTITUTE THE PRIMARY VARIANTS ASSESSED FOR NOVELTY

Sample ID	Nucleotide change	HGVS name (NG_009105.2)	GRCh38 HGVS (NC_000023.11)	Amino acid change	Mutation type	Frequency (100%)
L30, L51 L56, L76	c. 161_162insG	NG_009105.2:g. 13668-13669insG	g. 154152918_154152919insG	p.Asp54Gly	Frameshift insertion	13.3
L28, L43L65	c. 106T>C	NG_009105.2:g. 13731T>C	g. 154152981T>C	p.His35Pro	Missense	10
L29, L57	c. 87G>C	NG_009105.2:g. 13739G>C	g. 154152989G>C	p.Arg29His	Missense	6.6
L6, L63	c. 94G>C	NG_009105.2:g. 13743G>C	g. 154152993G>C	p.Gln31His	Missense	6.6
L36, L42L68, L80	c. 30G>A	NG_009105.2:g. 13800G>A	g. 154153050G>A	p.Arg10Arg	Silent	13.3
L3, L21, L50	c. 95T>G	NG_009105.2:g. 13735T>G	g. 154152985T>G	p.Ser32Ser	Silent	10
L71	c. 102T>G	NG_009105.2:g. 13735T>G	g. 154152985T>G	p.Ile34Gln	Missense	3.3
L52, L55	c. 93T>G	NG_009105.2:g. 13736T>G	g. 154152986T>G	p.His31Ser	Missense	6.6
L53, L77	c. 97T>C	NG_009105.2:g. 13732T>C	g. 154152982T>C	p.Ser33His	Missense	6.6
L48	c. 12C>A	NG_009105.2:g. 13817C>A	g. 154153067C>A	p.Ser4Arg	Missense	3.3
L49, L78, L81	c. 48C>T	NG_009105.2:g. 13789C>T	g. 154153039C>T	p.Thr16Asn	Missense	10
L72, L74	c. 85G>C	NG_009105.2:g. 13744G>C	g. 154152994G>C	p.Asp28His	Missense	6.6
L59	c. 39C>T	NG_009105.2:g. 13790C>T	g. 154153040C>T	p.His13Asn	Missense	3.3

were examined using Fisher's exact tests. $p < 0.05$ was regarded as statistically significant.

III. RESULTS AND DISCUSSION

A. Demographic and Clinical Correlates of Red-green CVD

The investigation involved 1500 students, with 50 (3.3%) revealing signs of red-green CVD. The affected individuals were aged between 13 and 18 years, with a mean age of 14.9 ± 1.6 years. Of these, 15 (30%) were female and 35 (70%) were male. It was found that males (70%) had a higher prevalence of CVD than females (30%), with an odds ratio of 2.19 (95% CI: 0.87–5.87; $p = 0.043$) (Table V). These findings are consistent with a previous study conducted among schoolchildren in Duhok, Iraq, which reported CVD in 47 males (9.6%) and eight females (1.64%) out of 978 students (Hussein and Al-Dabbagh, 2022). This is in line with the X-linked recessive inheritance pattern of *OPN1LW* mutations, where hemizygous males expressing a single mutant allele are affected, whereas heterozygous females are typically protected from the null phenotype due to random X-chromosome inactivation (lyonization), which facilitates functional genetic compensation (Neitz and Neitz, 2011). Significantly, in our cohort, affected males exhibited a higher rate of protanopia 23 (76.7%) than protanomaly 12 (60.0%), a finding congruent with global patterns where protanopia constitutes 60–70% of protanopia defects in males (Simunovic, 2010, Birch, 2012). This male prevalence of the more severe phenotype is often attributed to the total absence of functional L-cone photopigment. In contrast, the overall presentation of females in our study was milder, with protanomaly 8 (40.0%) being higher than protanopia 7 (23.3%). This pattern is quite typical in heterozygous females and is a direct outcome of different lyonization effects, where an attenuated phenotype is produced by the retinal mosaic of cells expressing the mutant vs. wild-type allele (Migeon, 2020, Gocuk, et al., 2024). The prevalence

of CVD was also compared between young and older students, but age-related differences were not statistically significant ($p = 0.532$). In addition, 30 cases of protanopia and 20 cases of protanomaly were detected in the CVD-positive group, correlated with family history and parental relationships. Participants with a family history of CVD had 2.45 times greater odds of being diagnosed with protanopia and protanomaly 14 (46.7%) and 8 (40.0%), respectively ($p = 0.039$). In addition, parental relation was associated with a 2.75-fold higher risk of protanopia 22 (73.3%) and protanomaly 10 (50.0%) ($p = 0.018$) (Table V). These results tie with a study from Saudi Arabia (Alnahedh, et al., 2025), where consanguineous families exposed a 1.68-fold increased risk of CVD ($p = 0.03$), although the effect was more pronounced in our Kurdish cohort (73.3% vs. 42.1% protanopia prevalence). The significantly higher occurrence in our population could be due to either a higher level of consanguinity or a higher frequency of founder mutations in the *OPN1LW* gene array, which are prevalent in populations that are relatively isolated. This finding is further supported by the elevated familial risk found in the present cohort, which aligns with an Iranian pedigree analysis that reported a 2.30-fold increased risk of protanopia ($p = 0.01$) and a prevalence of 50.8% in affected families (Jafarzadehpur, et al., 2014). This evidence for a large regional genetic component is reinforced by the constant duplication of these genetic relationships across other groups in the Middle East, including Saudi, Iranian, and now Kurdish. All of these results suggest that genetic inheritance, especially in consanguineous family structures, is a major factor in disease manifestation and that environmental factors are less important in the etiology of inherited red-green color vision deficiencies.

In the present cohort, no significant association was found between ABO blood groups and CVD. Blood group A had the highest prevalence (46%, $p = 0.420$), with B having the second highest (34%, $p = 1.000$), and AB (12%, $p = 0.670$) and

TABLE V
PARTICIPANT DEMOGRAPHICS AND CLINICAL FEATURES OF THOSE WITH PROTANOPIA (n=30) AND PROTANOMALY (n=20)

Parameter	Category	Protanomaly (%)	Protanopia (%)	Total (%)	OR (95% CI)	p-value
Gender	Male	12 (60.0)	23 (76.7)	35 (70)	2.19 (0.87–5.87)	0.043*
	Female	8 (40.0)	7 (23.3)	15 (30)	-	
Age (years)	Mean±SD (Range)	15.2±1.8 (13–17)	14.9±1.6 (13–18)	-	-	0.532
Family History	Yes	8 (40.0)	14 (46.7)	22 (44)	2.45 (1.04–5.78)	0.039*
	No	12 (60.0)	16 (53.3)	28 (56)	-	
Parent Relation	Yes	10 (50.0)	22 (73.3)	32 (64)	2.75 (1.16–6.85)	0.018*
	No	10 (50.0)	8 (26.7)	18 (36)	-	
Blood Group A+	A+	8 (40.0)	15 (50.0)	23 (46)	0.67 (0.25–1.78)	0.420
Blood Group B+	B+	7 (35.0)	10 (33.3)	17 (34)	1.08 (0.38–3.10)	1.000
Blood Group AB+	AB+	3 (15.0)	3 (10.0)	6 (12)	1.59 (0.23–9.87)	0.670
Blood Group O+	O+	2 (10.0)	2 (6.7)	4 (8)	1.56 (0.16–14.1)	0.990

The presentation of continuous variables (age) is mean±standard deviation (range), whereas that of categorical variables is n (%). (*p<0.05) indicates significant differences between groups as determined by Fisher's Exact Test for categorical variables

O (8%, p = 0.990) having the lowest (Table V). According to our findings, ABO blood type does not influence CVD risk in Kurdish people. It contrasts with previous studies (Ebrahim, Shaker and Kadhir, 2016, Tyagi, et al., 2024), which reported higher CVD prevalence in individuals with blood groups A, O, or B in their populations. The inconsistency is likely due to ethnicity, sample size, or genetic background. Therefore, our findings support the view that X-linked opsin gene mutations are the main cause of red–green CVD, rather than ABO antigens (Gudeta and Asrat, 2024). To further investigate CVD in affected individuals, molecular analysis was conducted on exon 3 of the *OPN1LW* gene in order to identify variants that may be associated with protanopia, as well as protanomaly in affected individuals.

B. Genetic Variant Analysis

PCR amplification of the *OPN1LW* gene exon 3 was approved by gel electrophoresis, as shown in Fig. 1. An apparent band of exactly the expected product size, 226 base pairs, was noticed in an amplified sample, displaying valid and targeted amplification. The PCR products had high quality and integrity because there were no non-specific products or smears appearing on the gels, which were used later for Sanger sequencing and downstream variant analyses.

The molecular analysis of this cohort showed that *OPN1LW* exon 3 exhibited a variety of genetic variations, with 30 nucleotide variants found in 50 samples. As illustrated in Fig. 2, the distribution of nucleotide changes, amino acid substitutions, and mutation classes across the cohort highlights a predominance of missense over silent and frameshift changes, an enrichment of substitutions within the central portion of exon 3, and a single frameshift insertion.

Among individuals with variants, 23 (76.6%) were male and 7 (23.3%) were female; the odds of having a variant did not differ significantly by gender (OR [male vs. female] = 1.19; 95% CI, 0.35–4.23; Table VI). The most common alterations were missense mutations (63.3%), frameshift insertions (13.3%), and silent substitutions (23.3%) (Table VI). The high frequency of missense variants is in line with established mechanisms of red-green CVD, which primarily involve amino acid substitutions in the opsin protein to alter spectral sensitivity and impede

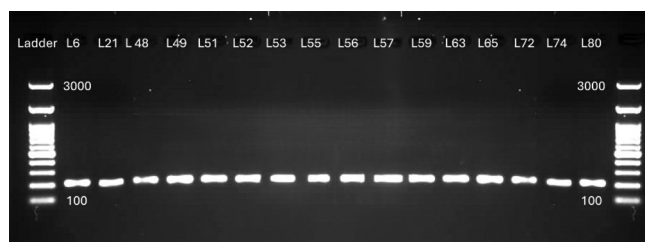


Fig. 1. Agarose gel electrophoresis of polymerase chain reaction (PCR) results displaying clear bands (~220–226 bp) from samples (L6–L80) illustrates the PCR amplification of exon 3 of the *OPN1LW* gene. As a size marker, the 100 bp DNA ladder was used.

phototransduction, whereas frameshift insertions typically result in a truncated, non-functional protein (Neitz and Neitz, 2021). The statistical investigation indicated that missense mutations 19 (63.3%) accounted for the majority of cases, OR: 1.07, 95% CI: [0.29, 4.08]. According to the gender-based analysis, 15 males (75.0%) and 4 females (40.0%) had missense variants (OR: 1.07, 95% CI: [0.29, 4.08], p = 1.000), whereas frameshift mutation frequencies in males were 3 (13.6%) and in females were 1 (10.0%), (OR: 0.80, 95% CI: [0.12, 5.18], p = 1.000) (Table VI). It is consistent with expectations for females carrying heterozygous X-linked genes (Kohl, et al., 2018). The detection of recurrent variants, particularly c.161_162insG (p.Asp54Gly) and c.30G>A (p.Arg10Arg), each found in 13.3% of the cohort, suggests they may represent mutational hotspots within the *OPN1LW* gene (Table VI).

According to Lejeune (2017), a premature termination codon is introduced by an exon-3 frameshift in *OPN1LW*, which predicts nonsense-mediated mRNA decay (NMD) (Lejeune, 2017). The protein would be truncated before many transmembrane helices and the retinal-binding Lys296, even if NMD is partially escaped, but they would destroy the function of L-opsin (Kim, et al., 2004, Robinson, et al., 1992). Hemizygous males are predicted to have a protanopia deficit because *OPN1LW* is X-linked, whereas heterozygous females may have mosaic loss in L-cone function as a result of X-chromosome inactivation (Basta and Pandya, 2020, Haer-Wigman, et al., 2022). The frequency of these variations in exon 3 is very essential since it encodes crucial opsin

TABLE VI
DETAILED ANALYSIS OF VARIANT DISTRIBUTION, GENDER BASIS AND FUNCTIONAL PREDICTIONS IN *OPN1LW* GENE EXON 3

Category	Gender (%)	Total (%)	Odds ratio and 95% CI:	p-value
Variant distribution	Males 23 (76.6) Females 7 (23.3)	30 (100)	OR: 1.19 (Males vs Females having variants) 95% CI: [0.35, 4.23]	0.772
Gender bias in mutation types	Misense:	19 (63.3)	OR: 1.07 95% CI: [0.29, 4.08]	1.000
	Males: 15/20 (75.0) Females: 4/10 (40.0)			
	Silent	7 (23.3)	OR: 0.56 95% CI: [0.09, 3.18]	0.675
	Males: 5/20 (25.0) Females: 2/10 (20.0)			
	Frameshift	4 (13.3)	OR: 0.80 95% CI: [0.12, 5.18]	1.000
Recurrent variants	Variables	Frequency	Samples	
	c. 161_162insG (p. Asp54Gly)	4/30 (13.3)	L30, L51, L56, L76	-
	c. 106T>C (p. His35Pro)	3/30 (10.0)	L28, L43, L65	
	c. 30G>A (p. Arg10Arg)	4/30 (13.3)	L36, L42, L68, L80	
Pathogenicity predictions	Damaging/Deleterious	Frequency	Odds Ratio and 95% CI:	p-value
PolyPhen-2	Damaging: Missense	12/19 (63.1)	OR: 1.71 [95% CI: 0.54, 5.45]	0.240
	Damaging: Silent	0/7 (0)	OR: 0.00 [95% CI: 0.00, 0.59] (infinite protection)	0.003
	Damaging: Frameshift	4/4 (100)	OR: ∞ [95% CI: 1.47, ∞] (infinite risk)	0.041
SIFT	Deleterious: Missense	11/19 (57.8)	OR: 1.38 [95% CI: 0.43, 4.40]	0.407
	Deleterious: Silent	0/7 (0)	OR: 0.00 [95% CI: 0.00, 0.69] (infinite protection)	0.006
	Deleterious: Frameshift	4/4 (100)	OR: ∞ [95% CI: 1.33, ∞] (infinite risk)	0.049

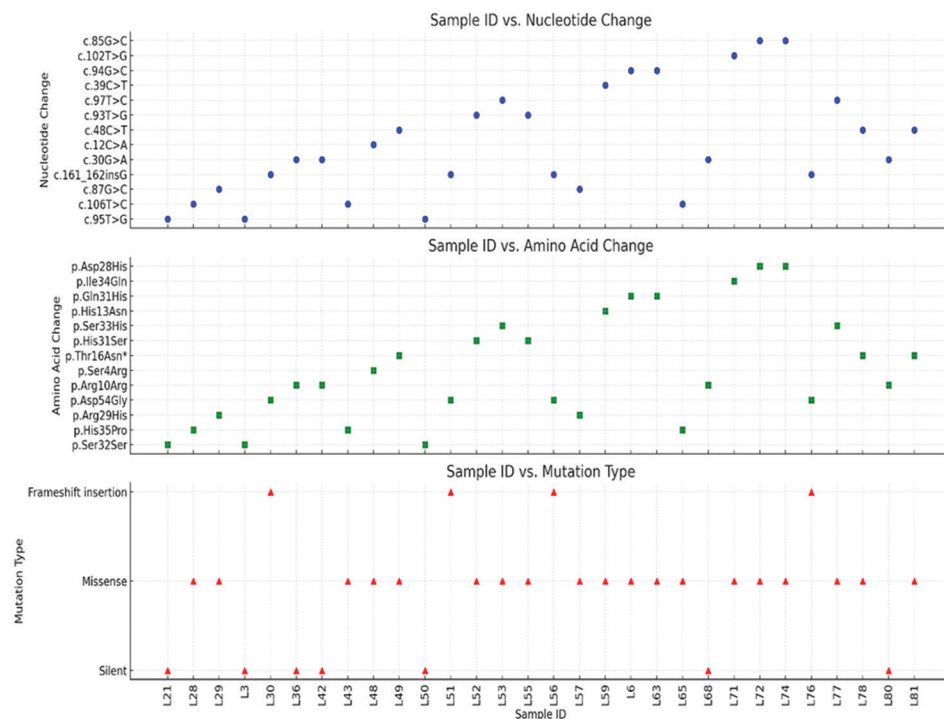


Fig. 2. The spread of nucleotide changes, amino acid substitutions, and mutation types in exon 3 of the *OPN1LW* Gene by samples: The figure is made up of three panels that depict the mutation spectrum for 13 different variants found in exon 3 of the *OPN1LW* gene. Each point represents a specific sample ID associated with: (Top) the relevant nucleotide alteration, (middle) the consequent amino acid substitution, and (bottom) the mutation type (missense, silent, or frameshift).

protein transmembrane domains. This finding aligns with earlier research showing that exon 3 is a mutation-rich locus that plays a crucial role in the pathophysiology of red-green

color vision impairment (Haer-Wigman, et al., 2022). The c.106T>C (p.His35Pro) missense mutation, found in 10% of samples, most likely alters the protein's structure and

functionality. However, the high frequency of the c.30G>A (p.Arg10Arg) variant is fascinating; whereas it does not alter the amino acid sequence, it may have an effect on stability, mRNA splicing, or gene expression, which calls for further functional investigation (Orosz, et al., 2017).

As a result of PolyPhen-2 and SIFT functional impact predictions, missense variants were the most prevalent. Among missense substitutions, SIFT classified 11/19 (57.8%) as deleterious (OR [male vs. female] = 1.38; 95% CI 0.43–4.40; $p = 0.407$), and PolyPhen-2 classified 12/19 (63.1%) as probably damaging (OR = 1.71; 95% CI 0.54–5.45; $p = 0.240$) (Table VI). Frameshift variants (4/4; 100%) were interpreted as loss-of-function based on predicted truncation and are not scored by SIFT or PolyPhen-2, supporting prior evidence linking exon 3 variants to protein dysfunction (Balachandran and Beck, 2020). Silent changes (0/7; 0%) have no predicted protein effect and were treated as benign/tolerated. The frameshift insertion is consistent with loss-of-function, whereas synonymous changes have no predicted protein effect.

To assess prior evidence, we cross-referenced all exon-3 variants against ClinVar, dbSNP, and Gnome AD (Table IV and Supplementary Tables S1–S4). Two missense variants – NG_009105.2:g.13743G>C and NG_009105.2:g.13790C>T – corresponded to existing database entries (rs1557157624 and rs782074812, respectively), whereas the remaining variants were not cataloged as pathogenic or likely pathogenic. We identified no published functional studies of these exact alleles. Therefore, experimental validation (e.g., opsin expression/pigment reconstitution or minigene splicing tests) is necessary to demonstrate biological impact, even if a number of missense alterations are expected to be harmful in silico and the frameshift supports loss-of-function. As a consequence of our findings, there is a high mutation burden in exon 3 of *OPN1LW*, which may lead to silent polymorphisms or potentially truncating lesions. Through this study, the genetics and molecular understanding of red cone dysfunction are enhanced. In addition, a resource for future clinical interpretation and functional validation is provided.

C. Estimated Functional Effects of Exon 3 Variants Using SIFT and PolyPhen-2 Analysis

The functional significance of amino acid substitutions in exon 3 of *OPN1LW* was assessed using in silico predictions based on the PolyPhen-2 and SIFT algorithms (Fig. 3). In the analysis of the 13 amino acid changes made in L-cone opsin, PolyPhen-2 (score ≥ 0.9) classified six variants as “potentially damaging.” These substitutions were p.Asp54Gly, p.His35Pro, p.Arg29His, p.His31Ser, p.Thr16Asn, and p.Tyr11Cys. These mutations are predicted to impair L-cone function, and further experimental validation is required. The remaining five variants, with scores ranging from 0.6 to 0.9, are classified as “possibly damaging.” Variants such as p. Arg10Arg and p. Ser32Ser are predicted to be benign mutations. These observations were corroborated by the SIFT algorithm, which identified 10 out of 13 variants as “deleterious” (score ≤ 0.05), indicating that these substitutions are likely to impact

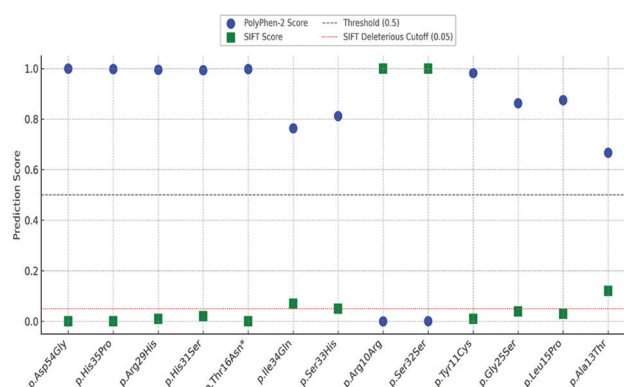


Fig. 3. Evaluation of the Functional Impact of 13 Different Amino Acid Substitutions in *OPN1LW* Gene Exon 3 Using SIFT and PolyPhen-2: PolyPhen-2 scores (shown as blue circles) estimate the probability that a substitution disrupts protein function, with scores ≥ 0.9 labeled as “probably damaging.” Meanwhile, SIFT scores (represented by green squares) predict functional impact, classifying substitutions with scores ≤ 0.05 as “deleterious.” A general functional threshold is marked by the gray dashed line at 0.5, whereas the red dotted line indicates the SIFT cutoff for deleterious effects at 0.05.

protein function. Interestingly, SIFT prediction completely overlapped with Polyphen-2 variants, scoring the highest. In particular, p. Asp54Gly, p. His35Pro, and p. Arg29His all scored highly. In addition to their likely pathogenicity, these substitutions are located in highly conserved transmembrane regions of the opsin protein. The findings of another study of exon 3 support our PolyPhen-2 and SIFT predictions, highlighting the potential pathogenicity of missense variants that disrupt opsin folding, stability, or chromophore interaction (Neitz, et al., 2022). However, missense mutations that affect folding, transport, or chromophore binding include p. His35Pro and p. Arg29His. This is in line with another study, which demonstrated the mechanism of cone-opsin malfunction (Wissinger, et al., 2022). On the other hand, mild substitutions and silent alterations (such as p.Ala13Thr, p.Ser33His, and p.Ile34Gln) demonstrated higher SIFT scores (>0.05), revealing that they were benign or tolerated. As a consequence of these predictions, numerous potential pathogenic mutations are identified and recommended for further experimental investigations, notably in the context of red-green color vision deficit and L-cone malfunction.

The current analysis is based on in-silico predictors SIFT and PolyPhen-2, which might produce inconsistent performance and false positives or negatives; the findings are suggestive rather than conclusive and lack proof of pathogenicity (Grimm, et al., 2015, Vaser, et al., 2016, Dong, et al., 2015). These tools assess missense alternatives only (not frameshift/silent), and splice effects need specialized predictors with imperfect agreement (Jaganathan, et al., 2019). Predictions that frameshift/nonsense alleles trigger nonsense-mediated decay (NMD) are probabilistic because NMD efficiency relies on exon–exon junction context and cellular state (Kurosaki, Popp and Maquat, 2019). Our results should therefore be regarded as preliminary unless

they are confirmed by cellular or biochemical analysis (e.g., heterologous expression of *OPN1LW* with pigment reconstitution) (Asenjo, Rim and Oprian, 1994).

IV. CONCLUSION

The present study identified 13 different nucleotide variants in exon 3 of the *OPN1LW* gene among Kurdish students with red-green color blindness. Two variants corresponded to previously cataloged entries (ClinVar/dbSNP), whereas the remaining eleven were not present in ClinVar, Ensembl, or gnomAD and, to our knowledge, have not been reported in the literature. These data broaden the *OPN1LW* variant spectrum in this population and identify candidates for functional validation in future work. One of the most mutation-prone areas associated with red cone malfunction is exon 3, that contains a significant number of missense and frameshift mutations that are expected to negatively impact protein function. Further evidence that CVD has a strong genetic basis is provided by the association between family history and consanguinity. As a result of these findings, genetic counseling and therapeutic approaches will be more effective and contribute to a broader understanding of red cone dysfunction. Further studies of variations in other exons of the *OPN1LW* gene can improve understanding of the molecular and genetic basis of color blindness in the Kurdish community.

REFERENCES

- Aboshiha, J., Dubis, A.M., Carroll, J., Hardcastle, A.J., and Michaelides, M., 2016. The cone dysfunction syndromes. *British Journal of Ophthalmology*, 100, pp.115-121.
- Adzhubei, I.A., Schmidt, S., Peshkin, L., Ramensky, V.E., Gerasimova, A., Bork, P., Kondrashov, A.S., and Sunyaev, S.R., 2010. A method and server for predicting damaging missense mutations. *Nature Methods*, 7, pp.248-249.
- Akhtar, M.S., Aslamkhan, M., Zar, M.S., Hanif, A., and Haris, A.R., 2019. Dichromacy: Color vision impairment and consanguinity in heterogenous population of Pakistan. *The International Journal of Frontier Sciences*, 3, pp.41-56.
- Alam, F., Salih, A.E., Elsherif, M., Yetisen, A.K., and Butt, H., 2022. 3D printed contact lenses for the management of color blindness. *Additive Manufacturing*, 49, pp.102464.
- Alnahedh, T., Alnahedh, A., Alhawas, A., Marghlani, L., Al Harbi, W., and Alhadlaq, O.S., 2025. The prevalence of color vision deficiency in medical students at King Saud bin Abdulaziz university for health sciences. *The Open Ophthalmology Journal*, 19, p. e18743641353155.
- Asadi, S., 2023. The role of mutations on *Gene NF1* in neurofibromatosis type 1 syndrome. *International Journal of clinical and Medical Case Reports*, 2, pp. 1-6.
- Asenjo, A.B., Rim, J., and Oprian, D.D., 1994. Molecular determinants of human red/green color discrimination. *Neuron*, 12, pp.1131-1138.
- Atilano, S.R., Kenney, M.C., Briscoe, A.D., and Jameson, K.A., 2020. A two-step method for identifying photopigment opsin and rhodopsin gene sequences underlying human color vision phenotypes. *Molecular Vision*, 26, p.158.
- Balachandran, P., and Beck, C.R., 2020. Structural variant identification and characterization. *Chromosome Research*, 28, pp.31-47.
- Basta, M., and Pandya, A.M., 2020. *Genetics, X-Linked Inheritance*. StatPearls Publishing, Treasure Island, FL.
- Birch, J., 1997. Efficiency of the Ishihara test for identifying red-green colour deficiency. *Ophthalmic and Physiological Optics*, 17, pp.403-408.
- Birch, J., 2012. Worldwide prevalence of red-green color deficiency. *Journal of the Optical Society of America A*, 29, pp.313-320.
- Buena-Atienza, E., R  ther, K., Baumann, B., Bergholz, R., Birch, D., De Baere, E., Dollfus, H., Greal, M.T., Gustavsson, P., and Hamel, C.P., 2016. *De novo* intrachromosomal gene conversion from OPN1MW to OPN1LW in the male germline results in blue cone monochromacy. *Scientific Reports*, 6, p.28253.
- Cai, Z.X., 2023. Link color blind: A systematic review of congenital color vision deficiency cognitively and neurologically. *SHS Web of Conferences*, 180, p.03015.
- Davidoff, C., 2015. *Cone Opsin Gene Variants in Color Blindness and other Vision Disorders*. University of Washington, Washington, DC.
- Dong, C., Wei, P., Jian, X., Gibbs, R., Boerwinkle, E., Wang, K., and Liu, X., 2015. Comparison and integration of deleteriousness prediction methods for nonsynonymous SNVs in whole exome sequencing studies. *Human Molecular Genetics*, 24, pp.2125-2137.
- Ebrahim, N., Shaker, I.A., and Kadir, A., 2016. Prevalence of color vision deficiency (CVD) and ABO blood groups in Kannur district of Kerala, India. *International Journal of Bioassays*, 5, pp.4760-4763.
- Gardner, J.C., Liew, G., Quan, Y.H., Ermetal, B., Ueyama, H., Davidson, A.E., Schwarz, N., Kanuga, N., Chana, R., and Maher, E.R., 2014. Three different cone opsin gene array mutational mechanisms with genotype-phenotype correlation and functional investigation of cone opsin variants. *Human Mutation*, 35, pp.1354-1362.
- Gocuk, S.A., Lancaster, J., Su, S., Jolly, J.K., Edwards, T.L., Hickey, D.G., Ritchie, M.E., Blewitt, M.E., Ayton, I.N., and Gouil, Q., 2024. Measuring X-chromosome inactivation skew for X-linked diseases with adaptive nanopore sequencing. *Genome Research*, 34, pp.1954-1965.
- Greenwald, S.H., Kuchenbecker, J.A., Rowlan, J.S., Neitz, J., and Neitz, M., 2017. Role of a dual splicing and amino acid code in Myopia, cone dysfunction and cone dystrophy associated with L/M opsin interchange mutations. *Translational Vision Science and Technology*, 6, p.2.
- Grimm, D.G., Azencott, C.A., Aicheler, F., Gieraths, U., Macarthur, D.G., Samocha, K.E., Cooper, D.N., Stenson, P.D., Daly, M.J., and Smoller, J.W., 2015. The evaluation of tools used to predict the impact of missense variants is hindered by two types of circularity. *Human Mutation*, 36, pp.513-523.
- Gudeta, T.B., and Asrat, T., 2024. Prevalence and genotypic frequency of color vision defects among primary schoolchildren in Adama Town, Eastern Ethiopia. *BMC Pediatrics*, 24, p.72.
- Haer-Wigman, L., Den Ouden, A., Van Genderen, M.M., Kroes, H.Y., Verheij, J., Smailhodzic, D., Hoekstra, A.S., Vijzelaar, R., Blom, J., and Derks, R., 2022. Diagnostic analysis of the highly complex *OPN1LW/OPN1MW* gene cluster using long-read sequencing and MLPA. *NPJ Genomic Medicine*, 7, p.65.
- Haim, M., Fledelius, H.C., and Skarsholm, D., 1988. X-linked myopia in danish family. *Acta Ophthalmologica*, 66, pp.450-456.
- Hussein, A.J., and Al-Dabbagh, S.A., 2022. Prevalence of color vision deficiency among primary school pupils in Duhok city, Kurdistan Region, Iraq. *AMJ (Advanced Medical Journal)*, 7, pp.11-16.
- Jafarzadehpour, E., Hashemi, H., Emamian, M.H., Khabazkhoob, M., Mehravaran, S., Shariati, M., and Fotouhi, A., 2014. Color vision deficiency in a middle-aged population: The Shahroud Eye Study. *International Ophthalmology*, 34, pp.1067-1074.
- Jaganathan, K., Panagiotopoulou, S.K., Mcrae, J.F., Darbandi, S.F., Knowles, D., Li, Y.I., Kosmicki, J.A., Arbelaez, J., Cui, W., and Schwartz, G.B., 2019. Predicting splicing from primary sequence with deep learning. *Cell*, 176, pp.535-548.e24.
- Karim, K.J., and Saleem, M.A., 2013. Prevalence of congenital red-green color

vision defects among various ethnic groups of students in Erbil City. *Jordan Journal of Biological Sciences*, 6, pp.235-238.

Kim, J.M., Altenbach, C., Kono, M., Oprian, D.D., Hubbell, W.L., and Khorana, H.G., 2004. Structural origins of constitutive activation in rhodopsin: Role of the K296/E113 salt bridge. *Proceedings of the National Academy of Sciences*, 101, pp.12508-12513.

Kohl, S., Jägle, H., Wissinger, B., and Zobor, D., 2018. Achromatopsia. In: Adam, M.P., Ardinger, H.H. and Pagon, R.A., Eds. *GeneReviews*®. University of Washington, Seattle (WA). Available from: <https://www.ncbi.nlm.nih.gov/books/NBK1418> [Last accessed on 2025 Sep 29].

Kuo, H.K., Tsao, S.T., and Pei-Chang, W.U., 2023. Prevalence of congenital color vision deficiency in southern taiwan and detection of female carriers by visual pigment gene analysis. *International Journal of Molecular Sciences*, 24, p.15247.

Kurosaki, T., Popp, M.W., and Maquat, L.E., 2019. Quality and quantity control of gene expression by nonsense-mediated mRNA decay. *Nature Reviews Molecular Cell biology*, 20, pp.406-420.

Lejeune, F., 2017. Nonsense-mediated mRNA decay at the crossroads of many cellular pathways. *BMB Reports*, 50, p.175.

Lilja, O., 2024. *Genetic Variations in Human Opsin Genes*. Tampere University, Finland.

Migeon, B.R., 2020. X-linked diseases: Susceptible females. *Genetics in Medicine*, 22, pp.1156-1174.

Nathans, J., 1999. The evolution and physiology of human color vision: Insights from molecular genetic studies of visual pigments. *Neuron*, 24, pp.299-312.

Nathans, J., Thomas, D., and Hogness, D.S., 1986. Molecular genetics of human color vision: The genes encoding blue, green, and red pigments. *Science*, 232, 193-202.

Neitz, J., and Neitz, M., 2011. The genetics of normal and defective color vision. *Vision Research*, 51, pp.633-651.

Neitz, M., and Neitz, J., 2021. Intermixing the *OPN1LW* and *OPN1MW* genes disrupts the exonic splicing code causing an array of vision disorders. *Genes*, 12, p.1180.

Neitz, M., Neitz, J., and Grishok, A., 1995. Polymorphism in the number of genes encoding long-wavelength-sensitive cone pigments among males with normal color vision. *Vision Research*, 35, pp.2395-2407.

Neitz, M., Patterson, S.S., and Neitz, J., 2019. Photopigment genes, cones,

and color update: Disrupting the splicing code causes a diverse array of vision disorders. *Current Opinion in Behavioral Sciences*, 30, pp.60-66.

Neitz, M., Wagner-Schuman, M., Rowland, J.S., Kuchenbecker, J.A., and Neitz, J., 2022. Insight from *OPN1LW* gene haplotypes into the cause and prevention of myopia. *Genes*, 13, p.942.

Ng, P.C., and Henikoff, S., 2003. SIFT: Predicting amino acid changes that affect protein function. *Nucleic Acids Research*, 31, pp.3812-3814.

Orosz, O., Rajta, I., Vajas, A., Takács, L., Csutak, A., Fodor, M., Koložsvári, B., Resch, M., Sényi, K., and Lesch, B., 2017. Myopia and late-onset progressive cone dystrophy associate to LVAVA/MVAVA exon 3 interchange haplotypes of opsin genes on chromosome X. *Investigative Ophthalmology and Visual Science*, 58, pp.1834-1842.

Robinson, P.R., Cohen, G.B., Zhukovsky, E.A., and Oprian, D.D., 1992. Constitutively active mutants of rhodopsin. *Neuron*, 9, pp.719-725.

Segre, L., 2019. *Eye Anatomy: A Closer Look at the Parts of the Eye*. Verywell Health. Available from: <https://www.allaboutvision.com/eye-care/eye-anatomy/overview-of-anatomy> [Last accessed on 2025 Sep 29].

Simunovic, M., 2010. Colour vision deficiency. *Eye*, 24, pp.747-755.

Swathi, B., Koushalya, R., Roshan, J.V., and Gowtham, M., 2020. Color blindness algorithm comparison for developing an android application. *International Research Journal of Engineering and Technology (IRJET)*, 7, p. 3608.

Tyagi, S., Nema, P., Tyagi, M., and Agrawal, M., 2024. Association between color vision deficiency and ABO blood group in young adults: A prevalence study. *European Journal of Cardiovascular Medicine*, 14, pp.173-176.

Vaser, R., Adusumalli, S., Leng, S.N., Sikic, M., and Ng, P.C., 2016. SIFT missense predictions for genomes. *Nature protocols*, 11, pp.1-9.

Wissinger, B., Baumann, B., Buena-Atienza, E., Ravesh, Z., Cideciyan, A.V., Stingl, K., Audo, I., Meunier, I., Bocquet, B., and Traboulsi, E.I., 2022. The landscape of submicroscopic structural variants at the *OPN1LW/OPN1MW* gene cluster on Xq28 underlying blue cone monochromacy. *Proceedings of the National Academy of Sciences*, 119, p.e2115538119.

Yousif, N., Abdullah, M., Abdularazaq, F., Sabir, A., Mohammed, M., and Saleem, N., 2024. Determination of Colorblindness Among Public Secondary School Children in Erbil City. In: 5th International Conference on Biomedical and Health Sciences. Cihan University-Erbil, pp.104-108.

SUPPLEMENTARY TABLES

TABLE S1: STUDY-IDENTIFIED *OPN1LW* EXON 3 VARIANTS AND NORMALIZATION TO GRCh38. VARIANTS DETECTED FROM SANGER CHROMATOGRAMS ALIGNED TO REFSeqGENE NG_009105.2 (*OPN1LW* EXON 3) AND NAMED IN HGVS FORMAT ARE LISTED ALONGSIDE THEIR NORMALIZED GRCh38 COORDINATES ON NC_000023.11 (CHRX). FOR REFERENCE, TRANSCRIPT-LEVEL NOTATIONS CORRESPOND TO NM_020061.6 (MANE SELECT). THESE ENTRIES CONSTITUTE THE PRIMARY CALLS EVALUATED FOR NOVELTY.

Sample ID	Nucleotide Change	HGVS Name (NG_009105.2)	GRCh38 HGVS (NC_000023.11)	Amino Acid Change	Mutation Type	Frequency (%100)	Novel/Known
L30, L51, L56, L76	c. 161_162insG	NG_009105.2:g. 13668-13669insG	g. 154152918_154152919insG	p.Asp54Gly	Frameshift insertion	%13.3	Novel
L28, L43, L65	c. 106T>C	NG_009105.2:g. 13731T>C	g. 154152981T>C	p.His35Pro	Missense	%10	Novel
L29, L57	c. 87G>C	NG_009105.2:g. 13739G>C	g. 154152989G>C	p.Arg29His	Missense	%6.6	Novel
L6, L63	c. 94G>C	NG_009105.2:g. 13743G>C	g. 154152993G>C	p.Gln31His	Missense	%6.6	Known rs1557157624
L36, L42L68, L80	c. 30G>A	NG_009105.2:g. 13800G>A	g. 154153050G>A	p.Arg10Arg	Silent	%13.3	Novel
L3, L21, L50	c. 95T>G	NG_009105.2:g. 13735T>G	g. 154152985T>G	p.Ser32Ser	Silent	%10	Novel
L71	c. 102T>G	NG_009105.2:g. 13735T>G	g. 154152985T>G	p.Ile34Gln	Missense	%3.3	Novel
L52, L55	c. 93T>G	NG_009105.2:g. 13736T>G	g. 154152986T>G	p.His31Ser	Missense	%6.6	Novel
L53, L77	c. 97T>C	NG_009105.2:g. 13732T>C	g. 154152982T>C	p.Ser33His	Missense	%6.6	Novel
L48	c. 12C>A	NG_009105.2:g. 13817C>A	g. 154153067C>A	p.Ser4Arg	Missense	%3.3	Novel
L49, L78, L81	c. 48C>T	NG_009105.2:g. 13789C>T	g. 154153039C>T	p.Thr16Asn*	Missense	%10	Novel
L72, L74	c. 85G>C	NG_009105.2:g. 13744G>C	g. 154152994G>C	p.Asp28His	Missense	%6.6	Novel
L59	c. 39C>T	NG_009105.2:g. 13790C>T	g. 154153040C>T	p.His13Asn	Missense	%3.3	Known rs782074812

TABLE S2: ENSEMBL (VEP) CATALOGUE OF ALL VARIANTS RECORDED WITHIN *OPN1LW* EXON 3 (GRCh38). COMPLETE EXPORT FROM ENSEMBL/VARIATION FOR THE *OPN1LW* EXON 3 REGION ON GRCh38, INCLUDING GENOMIC POSITIONS, ALLELES, CONSEQUENCES, TRANSCRIPT ANNOTATIONS (WITH NM_020061.6 DISPLAYED WHERE AVAILABLE), AND EXISTING IDENTIFIERS (E.G., dbSNP rsIDs). THIS TABLE WAS USED TO CHECK WHETHER ANY STUDY VARIANTS IN TABLE S1 ARE ALREADY PRESENT IN ENSEMBL/dbSNP.

Variant ID	Chr. Location	Alleles	Class	Source	Conseq. Type	sift_class	SIFT	polyphen_class	PolyPhen
rs782722274	X: 154152935	C/T	SNP	dbSNP	splice region variant~splice polypyrimidine tract variant~intron variant	-	-		
rs781816908	X: 154152936	A/G	SNP	dbSNP	splice region variant~splice polypyrimidine tract variant~intron variant	-	-		
COSV104685014	X: 154152936	COSMIC_ MUTATION	somatic SNV	COSMIC	splice region variant~splice polypyrimidine tract variant~intron variant	-	-		
COSV100886868	X: 154152938	COSMIC_ MUTATION	somatic SNV	COSMIC	splice acceptor variant	-	-		
COSV64064441	X: 154152939	COSMIC_ MUTATION	somatic SNV	COSMIC	splice acceptor variant	-	-		
COSV64065931	X: 154152941	COSMIC_ MUTATION	somatic SNV	COSMIC	splice region variant~coding sequence variant	-	-		
rs782484270	X: 154152942	A/G	SNP	dbSNP	missense variant~splice region variant	deleterious	0.02	benign	0.098
rs2067075081	X: 154152948	G/A	SNP	dbSNP	missense variant	deleterious	0.03	benign	0.031
rs1557157619	X: 154152982	G/T	SNP	dbSNP	missense variant	deleterious	0	probably damaging	0.984
rs713	X: 154152987	A/C	SNP	dbSNP	missense variant	tolerated	0.87	benign	0
rs2067075173	X: 154152990	G/C	SNP	dbSNP	missense variant	deleterious	0	probably damaging	0.957
rs1557157624	X: 154152993	G/C	SNP	dbSNP	missense variant	deleterious	0.02	possibly damaging	0.719
rs782340336	X: 154153006	T/G	SNP	dbSNP	missense variant	deleterious	0	possibly damaging	0.716
rs782249764	X: 154153015	T/A	SNP	dbSNP	missense variant	deleterious	0	benign	0.214
rs184047537	X: 154153018	G/T	SNP	dbSNP	missense variant	deleterious	0.03	benign	0.035
rs782526787	X: 154153021	T/G	SNP	dbSNP	missense variant	deleterious	0	probably damaging	0.997
rs373601495	X: 154153024	A/G	SNP	dbSNP	missense variant	tolerated	0.17	benign	0.001
rs782292865	X: 154153026	G/A/T	SNP	dbSNP	missense variant	tolerated	0.08	benign	0.005
rs782292865	X: 154153026	G/A/T	SNP	dbSNP	missense variant	tolerated	0.38	benign	0.001

(Contd...)

TABLE S2: (CONTINUED)

Variant ID	Chr. Location	Alleles	Class	Source	Conseq. Type	sift_class	SIFT	polyphen_class	PolyPhen
rs1557157636	X: 154153027	C/G	SNP	dbSNP	missense variant	tolerated	0.64	benign	0.003
rs782305369	X: 154153032	C/A	SNP	dbSNP	missense variant	tolerated	0.22	benign	0.009
rs1557157639	X: 154153036	C/T	SNP	dbSNP	missense variant	deleterious	0	probably damaging	0.995
rs781931953	X: 154153038	A/G	SNP	dbSNP	missense variant	tolerated	1	benign	0
rs5986963	X: 154153041	G/A/C	SNP	dbSNP	missense variant	tolerated	0.29	benign	0.003
rs5986963	X: 154153041	G/A/C	SNP	dbSNP	missense variant	tolerated	0.22	benign	0
rs1557157644	X: 154153044	G/A	SNP	dbSNP	missense variant	deleterious	0.02	possibly damaging	0.74
rs1557157645	X: 154153048	T/C	SNP	dbSNP	missense variant	deleterious	0	benign	0.073
rs1557157647	X: 154153051	CC/CCC	indel	dbSNP	frameshift variant	-	-	-	-
rs149897670	X: 154153051	C/T	SNP	dbSNP	missense variant	tolerated	1	benign	0.003
rs150391359	X: 154153059	T/C	SNP	dbSNP	missense variant	deleterious	0.01	probably damaging	0.999
rs145009674	X: 154153062	A/G	SNP	dbSNP	missense variant	tolerated	1	benign	0
rs781836898	X: 154153063	T/A	SNP	dbSNP	missense variant	deleterious	0.01	benign	0.108
rs1557157653	X: 154153067	G/A	SNP	dbSNP	stop gained	-	-	-	-
rs1557157655	X: 154153068	G/T	SNP	dbSNP	missense variant	tolerated	1	benign	0.007
rs375863094	X: 154153074	G/C	SNP	dbSNP	missense variant	tolerated	0.77	benign	0
rs782549215	X: 154153080	A/G	SNP	dbSNP	missense variant	tolerated	0.43	benign	0.006
rs782568898	X: 154153083	G/A	SNP	dbSNP	missense variant	tolerated	0.31	benign	0.005
rs782197748	X: 154153087	C/T	SNP	dbSNP	missense variant	deleterious	0.01	benign	0.274
rs782699550	X: 154153092	A/G	SNP	dbSNP	missense variant	tolerated	0.12	benign	0.001
rs1557157668	X: 154153099	G/C	SNP	dbSNP	missense variant	deleterious	0.02	probably damaging	0.999
rs1557157669	X: 154153104	A/G	SNP	dbSNP	missense variant	deleterious	0	possibly damaging	0.901

TABLE S3: ClinVar records for *OPN1LW* exon 3 (GRCh38). ALL ClinVar submissions mapped to *OPN1LW* exon 3 on GRCh38, with accessions, genomic coordinates, reported molecular consequence, and clinical significance (if provided). THIS TABLE WAS USED TO DETERMINE WHETHER ANY STUDY VARIANTS IN TABLE S1 HAVE PRIOR ClinVar ENTRIES.

Variation ID:	Accession:	Type and length	Location	HGVS	Protein	Molecular consequence
1326975	VCV001326975.3	SNV 1	X: 154153041 (GRCh38)	NG_009105.2:g. 13791G>A	NP_064445.2:p.Val171Met	Missense
1206115	VCV001206115.5	SNV 1		NG_009105.2:g. 13768G>T	NP_064445.2:p.Arg163Ile	Missense
403267	VCV000403267.5	SNV 1	X: 154153062 (GRCh38)	NG_009105.2:g. 13812A > G	NP_064445.2:p.Ile178Val	Missense
10504	VCV000010504.7	SNV -	X: 154153068 (GRCh38)	NG_009105.2:g. 13818=	NP_064445.2:p.Ala180=	No consequence alteration
3410096	VCV003410096.1	SNV 1	X: 154153009 (GRCh38)	NG_009105.2:g. 13759G > A	NP_064445.2:p.Gly160Asp	missense
1326976	VCV001326976.3	SNV 1	X: 154153043 (GRCh38)	NG_009105.2:g. 13793G > T	NP_064445.2:p.Val171=	synonymous
3410103	VCV003410103.1	SNV 1	X: 154153044 (GRCh38)	NG_009105.2:g. 13794G > A	NP_064445.2:p.Gly172Ser	missense
3410098	VCV003410098.1	SNV 1	X: 154153075 (GRCh38)	NG_009105.2:g. 13825T > C	NP_064445.2:p.Val182Ala	missense

TABLE S 4: Gnome AD (Genome Aggregation Database) Variants and Allele Frequencies across *OPN1LW* Exon 3 (GRCh38). List of variants observed in Gnome AD for the *OPN1LW* Exon 3 regions on GRCh38, including genomic coordinates, alleles, and population allele frequencies (where available). This table was used to assess population occurrence and to corroborate whether study variants in Table S1 have existing population evidence.

gnomAD ID	Chromosome	Position	rsIDs	Reference	Alternate	Source	Filters - exomes	Filters - genomes	Transcript	HGVs	Protein	Transcript	Allele
										Consequence	Consequence		Number
X-154151000-G-A	X	154151000	rs188572146	G	A	gnomAD Exomes, gnomAD Genomes	PASS	PASS	ENST00000369951.9	c. 409+48G>A		c. 409+48G>A	1173477
X-154151001-C-T	X	154151001	rs1220647160	C	T	gnomAD Exomes	PASS	NA	ENST00000369951.9	c. 409+49C>T		c. 409+49C>T	1174750
X-154151006-T-A	X	154151006	rs782171425	T	A	gnomAD Exomes	PASS	NA	ENST00000369951.9	c. 409+54T>A		c. 409+54T>A	1168490
X-154151006-T-C	X	154151006	rs782171425	T	C	gnomAD Exomes	PASS	NA	ENST00000369951.9	c. 409+54T>C		c. 409+54T>C	1162692
X-154151008-C-T	X	154151008	rs1603345229	C	T	gnomAD Exomes	PASS	NA	ENST00000369951.9	c. 409+56C>T		c. 409+56C>T	1172427
X-154151008-C-A	X	154151008		C	A	gnomAD Exomes	PASS	NA	ENST00000369951.9	c. 409+56C>A		c. 409+56C>A	1172427
X-154151012-A-G	X	154151012	rs11492096	A	G	gnomAD Exomes, gnomAD Genomes	PASS	PASS	ENST00000369951.9	c. 409+60A>G		c. 409+60A>G	1153673
X-154151014-C-T	X	154151014		C	T	gnomAD Exomes	PASS	NA	ENST00000369951.9	c. 409+62C>T		c. 409+62C>T	1172432
X-154151015-C-G	X	154151015	rs146138552	C	G	gnomAD Exomes, gnomAD Genomes	PASS	PASS	ENST00000369951.9	c. 409+63C>G		c. 409+63C>G	1169792
X-154151023-G-C	X	154151023		G	C	gnomAD Exomes	PASS	NA	ENST00000369951.9	c. 409+71G>C		c. 409+71G>C	1167492
X-154154501-G-T	X	154154501	rs2067079386	G	T	gnomAD Exomes	PASS	NA	ENST00000369951.9	c. 579-73G>T		c. 579-73G>T	804845
X-154154506-G-A	X	154154506	rs1469241450	G	A	gnomAD Exomes	NA	PASS	ENST00000369951.9	c. 579-68G>A		c. 579-68G>A	861138
X-154154508-G-A	X	154154508	rs1237683448	G	A	gnomAD Exomes	PASS	NA	ENST00000369951.9	c. 579-66G>A		c. 579-66G>A	873495
X-154154517-G-T	X	154154517	rs782374234	G	T	gnomAD Exomes	PASS	NA	ENST00000369951.9	c. 579-57G>T		c. 579-57G>T	944059
X-154154521-G-A	X	154154521	rs1209229124	G	A	gnomAD Exomes	PASS	NA	ENST00000369951.9	c. 579-53G>A		c. 579-53G>A	1013496
X-154154528-G-T	X	154154528	rs1557157795	G	T	gnomAD Exomes	PASS	NA	ENST00000369951.9	c. 579-46G>T		c. 579-46G>T	1066385
X-154154559-G-T	X	154154559	rs782637696	G	T	gnomAD Exomes	PASS	NA	ENST00000369951.9	c. 579-15G>T		c. 579-15G>T	1154342
X-154154559-G-A	X	154154559	rs782637696	G	A	gnomAD Exomes	NA	PASS	ENST00000369951.9	c. 579-15G>A		c. 579-15G>A	1148574
X-154154561-C-A	X	154154561		C	A	gnomAD Exomes	PASS	NA	ENST00000369951.9	c. 579-13C>A		c. 579-13C>A	1156856
X-154154563-CT-C	X	154154563	rs2067079725	CT	C	gnomAD Exomes	NA	PASS	ENST00000369951.9	c. 579-9del		c. 579-9del	1155596
X-154154565-T-C	X	154154565	rs782319497	T	C	gnomAD Exomes	PASS	NA	ENST00000369951.9	c. 579-9T>C		c. 579-9T>C	1162532
X-154154572-A-C	X	154154572		A	C	gnomAD Exomes	PASS	NA	ENST00000369951.9	c. 579-2A>C		c. 579-2A>C	1167558
X-154154574-G-A	X	154154574	rs782592456	G	A	gnomAD Exomes	PASS	NA	ENST00000369951.9	p.Arg193Arg	p.Arg193Arg	c. 579G>A	1169464
X-154154581-C-G	X	154154581	rs1380050111	C	G	gnomAD Exomes	NA	PASS	ENST00000369951.9	p.Pro196Ala	p.Pro196Ala	c. 586C>G	1165873
X-154154586-C-T	X	154154586	rs782223333	C	T	gnomAD Exomes, gnomAD Genomes	PASS	PASS	ENST00000369951.9	p.His197His	p.His197His	c. 591C>T	1168369
X-154154587-G-A	X	154154587	rs782361699	G	A	gnomAD Exomes	PASS	NA	ENST00000369951.9	p.Gly198Ser	p.Gly198Ser	c. 592G>A	1169417
X-154154588-G-A	X	154154588	rs781989418	G	A	gnomAD Exomes, gnomAD Genomes	PASS	PASS	ENST00000369951.9	p.Gly198Asp	p.Gly198Asp	c. 593G>A	1170070
X-154154589-C-T	X	154154589		C	T	gnomAD Exomes	PASS	NA	ENST00000369951.9	p.Gly198Gly	p.Gly198Gly	c. 594C>T	1176060
X-154154592-G-A	X	154154592	rs1557157804	G	A	gnomAD Exomes	PASS	NA	ENST00000369951.9	p.Leu199Leu	p.Leu199Leu	c. 597G>A	1178512
X-154154602-T-C	X	154154602	rs121434621	T	C	gnomAD Exomes	PASS	NA	ENST00000369951.9	p.Cys203Arg	p.Cys203Arg	c. 607T>C	1180190
X-154154604-C-T	X	154154604	rs782419890	C	T	gnomAD Exomes, gnomAD Genomes	PASS	PASS	ENST00000369951.9	p.Cys203Cys	p.Cys203Cys	c. 609C>T	1172503
X-154154605-G-A	X	154154605	rs2067079911	G	A	gnomAD Exomes	PASS	NA	ENST00000369951.9	p.Gly204Ser	p.Gly204Ser	c. 610G>A	1174725
X-154154609-C-A	X	154154609		C	A	gnomAD Exomes	PASS	NA	ENST00000369951.9	p.Pro205Gln	p.Pro205Gln	c. 614C>A	1179144
X-154154613-C-T	X	154154613	rs1459231886	C	T	gnomAD Exomes	NA	PASS	ENST00000369951.9	p.Asp206Asp	p.Asp206Asp	c. 618C>T	1174657
X-154154614-G-C	X	154154614		G	C	gnomAD Exomes	PASS	NA	ENST00000369951.9	p.Val207Leu	p.Val207Leu	c. 619G>C	1182450
X-154154614-G-A	X	154154614	rs1557157806	G	A	gnomAD Exomes	PASS	NA	ENST00000369951.9	p.Val207Met	p.Val207Met	c. 619G>A	1177265

(Contd...)

TABLE S 4: (CONTINUED)

gnomAD ID	Chromosome	Position	rsIDs	Reference	Alternate	Source	Filters - exomes	Filters - genomes	Transcript	HGVs	Protein	Transcript	Allele
										Consequence	Consequence	Consequence	Number
X-154154622-C-T	X	154154622	rs782047119	C	T	gnomAD Exomes, gnomAD Genomes	PASS	PASS	ENST00000369951.9	p.Ser209Ser	p.Ser209Ser	c. 627C>T	1182205
X-154154623-G-A	X	154154623	rs782715468	G	A	gnomAD Exomes, gnomAD Genomes	PASS	PASS	ENST00000369951.9	p.Gly210Ser	p.Gly210Ser	c. 628G>A	1182840
X-154154627-G-T	X	154154627		G	T	gnomAD Exomes	PASS	NA	ENST00000369951.9	p.Ser211Ile	p.Ser211Ile	c. 632G>T	1189704
X-154154628-C-G	X	154154628	rs1159151639	C	G	gnomAD Exomes	NA	PASS	ENST00000369951.9	p.Ser211Arg	p.Ser211Arg	c. 633C>G	1184533
X-154154630-C-T	X	154154630	rs1399594374	C	T	gnomAD Exomes, gnomAD Genomes	PASS	PASS	ENST00000369951.9	p.Ser212Leu	p.Ser212Leu	c. 635C>T	1184661
X-154154631-G-A	X	154154631	rs142831615	G	A	gnomAD Exomes, gnomAD Genomes	PASS	PASS	ENST00000369951.9	p.Ser212Ser	p.Ser212Ser	c. 636G>A	1184712
X-154154634-C-T	X	154154634	rs146127280	C	T	gnomAD Exomes, gnomAD Genomes	PASS	PASS	ENST00000369951.9	p.Tyr213Tyr	p.Tyr213Tyr	c. 639C>T	1185586
X-154154637-C-T	X	154154637	rs138979991	C	T	gnomAD Exomes, gnomAD Genomes	PASS	PASS	ENST00000369951.9	p.Pro214Pro	p.Pro214Pro	c. 642C>T	1186062
X-154154638-G-A	X	154154638	rs782523727	G	A	gnomAD Exomes, gnomAD Genomes	PASS	PASS	ENST00000369951.9	p.Gly215Arg	p.Gly215Arg	c. 643G>A	1186117
X-154154641-G-A	X	154154641	rs1188983268	G	A	gnomAD Exomes, gnomAD Genomes	PASS	PASS	ENST00000369951.9	p.Val216Met	p.Val216Met	c. 646G>A	1186442
X-154154643-G-A	X	154154643	rs782583807	G	A	gnomAD Exomes, gnomAD Genomes	PASS	PASS	ENST00000369951.9	p.Val216Val	p.Val216Val	c. 648G>A	1186808
X-154154650-T-C	X	154154650		T	C	gnomAD Exomes	PASS	NA	ENST00000369951.9	p.Tyr219His	p.Tyr219His	c. 655T>C	1191889
X-154154654-T-C	X	154154654		T	C	gnomAD Exomes	PASS	NA	ENST00000369951.9	p.Met220Thr	p.Met220Thr	c. 659T>C	1192506
X-154154661-C-A	X	154154661	rs371840734	C	A	gnomAD Exomes	NA	PASS	ENST00000369951.9	p.Val222Val	p.Val222Val	c. 666C>A	1188499
X-154154662-C-A	X	154154662		C	C	gnomAD Exomes	PASS	NA	ENST00000369951.9	p.Leu223Ile	p.Leu223Ile	c. 667C>A	1193146
X-154154664-C-G	X	154154664		C	G	gnomAD Exomes	PASS	NA	ENST00000369951.9	p.Leu223Leu	p.Leu223Leu	c. 669C>G	1193389
X-154154664-C-T	X	154154664	rs1557157823	C	T	gnomAD Exomes	NA	PASS	ENST00000369951.9	p.Leu223Leu	p.Leu223Leu	c. 669C>T	1188885
X-154154665-A-C	X	154154665	rs1557157826	A	C	gnomAD Exomes	PASS	NA	ENST00000369951.9	p.Met224Leu	p.Met224Leu	c. 670A>C	1192731
X-154154667-G-A	X	154154667		G	A	gnomAD Exomes	PASS	NA	ENST00000369951.9	p.Met224Ile	p.Met224Ile	c. 672G>A	1193300
X-154154680-A-G	X	154154680		A	G	gnomAD Exomes	PASS	NA	ENST00000369951.9	p.Ile229Val	p.Ile229Val	c. 685A>G	1194352
X-154154683-A-G	X	154154683	rs144864828	A	G	gnomAD Exomes, gnomAD Genomes	PASS	PASS	ENST00000369951.9	p.Ile230Val	p.Ile230Val	c. 688A>G	1190055
X-154154684-T-C	X	154154684	rs148583295	T	C	gnomAD Exomes, gnomAD Genomes	PASS	PASS	ENST00000369951.9	p.Ile230Thr	p.Ile230Thr	c. 689T>C	1188737
X-154154686-C-A	X	154154686		C	A	gnomAD Exomes	PASS	NA	ENST00000369951.9	p.Pro231Thr	p.Pro231Thr	c. 691C>A	1194529
X-154154688-A-G	X	154154688		A	G	gnomAD Exomes	PASS	NA	ENST00000369951.9	p.Pro231Pro	p.Pro231Pro	c. 693A>G	1193293
X-154154691-C-T	X	154154691	rs782304727	C	T	gnomAD Exomes, gnomAD Genomes	PASS	PASS	ENST00000369951.9	p.Leu232Leu	p.Leu232Leu	c. 696C>T	1190698
X-154154692-G-A	X	154154692	rs781936473	G	A	gnomAD Exomes, gnomAD Genomes	PASS	PASS	ENST00000369951.9	p.Alala233Thr	p.Alala233Thr	c. 697G>A	1190550
X-154154692-G-T	X	154154692	rs781936473	G	T	gnomAD Exomes	NA	PASS	ENST00000369951.9	p.Alala233Ser	p.Alala233Ser	c. 697G>T	1190550
X-154154693-C-T	X	154154693	rs1292153674	C	T	gnomAD Exomes	NA	PASS	ENST00000369951.9	p.Alala233Val	p.Alala233Val	c. 698C>T	1190711
X-154154694-T-C	X	154154694	rs1226772901	T	C	gnomAD Exomes	PASS	NA	ENST00000369951.9	p.Alala233Ala	p.Alala233Ala	c. 699T>C	1194561
X-154154697-C-T	X	154154697		C	T	gnomAD Exomes	PASS	NA	ENST00000369951.9	p.Ile234Ile	p.Ile234Ile	c. 702C>T	1194798
X-154154697-C-G	X	154154697	rs782079976	C	G	gnomAD Exomes	PASS	NA	ENST00000369951.9	p.Ile234Met	p.Ile234Met	c. 702C>G	1194799
X-154154698-A-G	X	154154698	rs782358874	A	G	gnomAD Exomes	PASS	NA	ENST00000369951.9	p.Ile235Val	p.Ile235Val	c. 703A>G	1194460

(Contd...)

TABLE S 4: (CONTINUED)

gnomAD ID	Chromosome	Position	rsIDs	Reference	Alternate	Source	Filters - exomes	Filters - genomes	Transcript	HGVs	Protein	Transcript	Allele
										Consequence	Consequence		Number
X-154154699-T-A	X	154154699		T	A	gnomAD Exomes	PASS	NA	ENST00000369951.9	p.Ile235Asn	p.Ile235Asn	c. 704T>A	1194550
X-154154699-T-C	X	154154699	rs2067080469	T	C	gnomAD Exomes	PASS	NA	ENST00000369951.9	p.Ile235Thr	p.Ile235Thr	c. 705T>C	1194550
X-154154700-C-A	X	154154700		C	A	gnomAD Exomes	PASS	NA	ENST00000369951.9	p.Ile235Ile	p.Ile235Ile	c. 705C>A	1194567
X-154154700-C-T	X	154154700		C	T	gnomAD Exomes	PASS	NA	ENST00000369951.9	p.Ile235Ile	p.Ile235Ile	c. 705C>T	1194567
X-154154701-A-C	X	154154701	rs1065426	A	C	gnomAD Exomes	PASS	PASS	ENST00000369951.9	p.Met236Leu	p.Met236Leu	c. 706A>C	1189711
						gnomAD Genomes							
X-154154701-A-T	X	154154701	rs1065426	A	T	gnomAD Exomes	PASS	PASS	ENST00000369951.9	p.Met236Leu	p.Met236Leu	c. 706A>T	1189675
						gnomAD Genomes							
X-154154701-A-G	X	154154701	rs1065426	A	G	gnomAD Exomes	PASS	PASS	ENST00000369951.9	p.Met236Val	p.Met236Val	c. 706A>G	1187715
X-154154705-T-C	X	154154705	rs2067080545	T	C	gnomAD Exomes	PASS	PASS	ENST00000369951.9	p.Leu237Pro	p.Leu237Pro	c. 710T>C	1189881
X-154154706-C-G	X	154154706		C	G	gnomAD Exomes	PASS	NA	ENST00000369951.9	p.Leu237Leu	p.Leu237Leu	c. 711C>G	1194123
X-154154708-G-T	X	154154708	rs1288768228	G	T	gnomAD Exomes	PASS	PASS	ENST00000369951.9	p.Cys238Phe	p.Cys238Phe	c. 713G>T	1189562
						gnomAD Genomes							
X-154154712-C-T	X	154154712	rs2067080593	C	T	gnomAD Exomes	PASS	PASS	ENST00000369951.9	p.Tyr239Tyr	p.Tyr239Tyr	c. 717C>T	1189487
						gnomAD Genomes							
X-154154714-T-G	X	154154714	rs1437848284	T	G	gnomAD Exomes	NA	PASS	ENST00000369951.9	p.Leu240Arg	p.Leu240Arg	c. 719T>G	1189217
X-154154716-C-A	X	154154716		C	A	gnomAD Exomes	PASS	NA	ENST00000369951.9	p.Gln241Lys	p.Gln241Lys	c. 721C>A	1193536
X-154154718-A-C	X	154154718		A	C	gnomAD Exomes	PASS	NA	ENST00000369951.9	p.Gln241His	p.Gln241His	c. 723A>C	1192823
X-154154719-G-A	X	154154719		G	A	gnomAD Exomes	PASS	NA	ENST00000369951.9	p.Val242Met	p.Val242Met	c. 724G>A	1193030
X-154154723-G-T	X	154154723	rs1349908448	G	T	gnomAD Exomes	NA	PASS	ENST00000369951.9	p.Trp243Leu	p.Trp243Leu	c. 728G>T	1188300
X-154154723-G-C	X	154154723	rs1349908448	G	C	gnomAD Exomes	PASS	NA	ENST00000369951.9	p.Trp243Ser	p.Trp243Ser	c. 728G>C	1192621
X-154154725-C-G	X	154154725	rs1166314366	C	G	gnomAD Exomes	NA	PASS	ENST00000369951.9	p.Leu244Val	p.Leu244Val	c. 730C>G	1188263
X-154154728-G-A	X	154154728	rs782727254	G	A	gnomAD Exomes	PASS	NA	ENST00000369951.9	p.Ala245Thr	p.Ala245Thr	c. 733G>A	1191451
X-154154728-G-C	X	154154728	rs782727254	G	C	gnomAD Exomes	PASS	NA	ENST00000369951.9	p.Ala245Pro	p.Ala245Pro	c. 733G>C	1191451
X-154154735-G-A	X	154154735	rs781846140	G	A	gnomAD Exomes	PASS	PASS	ENST00000369951.9	p.Arg247Gln	p.Arg247Gln	c. 740G>A	1186558
						gnomAD Genomes							
X-154154738-C-T	X	154154738	rs782506559	C	T	gnomAD Exomes	PASS	PASS	ENST00000369951.9	p.Ala248Val	p.Ala248Val	c. 743C>T	1185955
						gnomAD Genomes							
X-154154738-C-G	X	154154738	rs782506559	C	G	gnomAD Exomes	PASS	NA	ENST00000369951.9	p.Ala248Gly	p.Ala248Gly	c. 743C>G	1190440
X-154154739-G-A	X	154154739	rs139583770	G	A	gnomAD Exomes	PASS	PASS	ENST00000369951.9	p.Ala248Ala	p.Ala248Ala	c. 744G>A	1185625
						gnomAD Genomes							
X-154154743-A-G-A	X	154154743	rs1451070275	AG	A	gnomAD Exomes	NA	PASS	ENST00000369951.9	c. 744+5del	c. 744+5del	c. 744+5del	1184588
X-154154745-C-A	X	154154745		C	A	gnomAD Exomes	PASS	NA	ENST00000369951.9	c. 744+6C>A	c. 744+6C>A	c. 744+6C>A	1186746
X-154154745-C-T	X	154154745		C	T	gnomAD Exomes	PASS	NA	ENST00000369951.9	c. 744+6C>T	c. 744+6C>T	c. 744+6C>T	1186747
X-154154750-C-T	X	154154750	rs781890434	C	T	gnomAD Exomes	PASS	PASS	ENST00000369951.9	c. 744+11C>T	c. 744+11C>T	c. 744+11C>T	1179860
						gnomAD Genomes							
X-154154750-C-G	X	154154750	rs781890434	C	G	gnomAD Exomes	PASS	NA	ENST00000369951.9	c. 744+11C>G	c. 744+11C>G	c. 744+11C>G	1184561
X-154154751-G-A	X	154154751	rs782576171	G	A	gnomAD Exomes	PASS	NA	ENST00000369951.9	c. 744+12G>A	c. 744+12G>A	c. 744+12G>A	1177695
X-154154768-C-T	X	154154768	rs1557157853	C	T	gnomAD Exomes	NA	PASS	ENST00000369951.9	c. 744+29C>T	c. 744+29C>T	c. 744+29C>T	1121918
X-154154770-C-T	X	154154770	rs782188648	C	T	gnomAD Exomes	PASS	PASS	ENST00000369951.9	c. 744+31C>T	c. 744+31C>T	c. 744+31C>T	1110695
						gnomAD Genomes							
X-154154771-G-A	X	154154771	rs782481346	G	A	gnomAD Exomes	PASS	PASS	ENST00000369951.9	c. 744+32G>A	c. 744+32G>A	c. 744+32G>A	1111467
						gnomAD Genomes							

(Contd...)

TABLE S 4: (CONTINUED)

gnomAD ID	Chromosome	Position	rsIDs	Reference	Alternate	Source	Filters - exomes	Filters - genomes	Transcript	HGVS Consequence	Protein Consequence	Transcript Consequence	Allele Number
X-154154776-C-T	X	154154776	rs1557157857	C	T	gnomAD Exomes, gnomAD Genomes	PASS	PASS	ENST00000369951.9	c. 744+37C>T		c. 744+37C>T	1100720
X-154154777-T-C	X	154154777		T	C	gnomAD Exomes	PASS	NA	ENST00000369951.9	c. 744+38T>C		c. 744+38T>C	1099625
X-154154789-G-A	X	154154789	rs2067081001	G	A	gnomAD Exomes	PASS	NA	ENST00000369951.9	c. 744+50G>A		c. 744+50G>A	1050381
X-154154790-C-T	X	154154790	rs1557157860	C	T	gnomAD Exomes	PASS	NA	ENST00000369951.9	c. 744+51C>T		c. 744+51C>T	1046880
X-154154801-A-G	X	154154801		A	G	gnomAD Exomes	PASS	NA	ENST00000369951.9	c. 744+62A>G		c. 744+62A>G	949455
X-154154807-T-G	X	154154807		T	G	gnomAD Exomes	PASS	NA	ENST00000369951.9	c. 744+68T>G		c. 744+68T>G	895618
X-154154811-C-T	X	154154811	rs1557157864	C	T	gnomAD Exomes, gnomAD Genomes	PASS	PASS	ENST00000369951.9	c. 744+72C>T		c. 744+72C>T	838160