Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

# Novel Genetic Insights into Early-Onset High Myopia Through Trio-Based Exome Sequencing

**Wenjing Li**
Ningxia Hui Autonomous Region Peoples Hospital
**zhenglai Wang**
Ningxia Hui Autonomous Region Peoples Hospital
**Bingtao Li**
Ningxia Hui Autonomous Region Peoples Hospital
**Keyan Liu**
Ningxia Hui Autonomous Region Peoples Hospital
**Zichun Lin**
Ningxia Hui Autonomous Region Peoples Hospital
**Rui Li**
Ningxia Hui Autonomous Region Peoples Hospital
**Yanan Ma**
Ningxia Hui Autonomous Region Peoples Hospital
**Meiling Su**
Ningxia Hui Autonomous Region Peoples Hospital
**Haixuan Wang**
Ningxia Hui Autonomous Region Peoples Hospital
**Jinjin Zhang**
Ningxia Hui Autonomous Region Peoples Hospital
**Wenjuan Zhuang**
zh_wenj@163.com

Ningxia Hui Autonomous Region Peoples Hospital

**Article**

**Keywords:** Early onset high myopia, Whole exome sequencing, Oligogenic Overlap Model, de novo mutations

**Posted Date:** September 2nd, 2026

**DOI:** https://doi.org/10.21203/rs.3.rs-10253637/v1

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

**Additional Declarations:** No competing interests reported.

# Novel Genetic Insights into Early-Onset High Myopia Through Trio-Based Exome Sequencing

Wenjing Li[†1], Zhenglai Wang[†1], Bingtao Li[1,2], Keyan Liu[1,2], Zichun Lin[1,2], Rui Li[1], Yanan Ma[1,2], Meiling Su[1,2], Haixuan Wang[1,2] , Jinjin Zhang[*1] and Wenjuan Zhuang[*1✉]

[1] People's Hospital of Ningxia Hui Autonomous Region, Ningxia Eye Hospital, Yinchuan, 750002，Ning Xia, China.
[2] The Third Clinical School of Medicine, Ningxia Medical University, Yinchuan, 750004, Ning Xia, China.
[†]Wenjing Li and Zhenglai Wang contributed equally to this work and share the first authors.
[*]Corresponding author: Jinjin Zhang,Wenjuan Zhuang
[✉]email: zh_wenj@163.com

**Abstract**

The genetic etiology of early-onset high myopia (eoHM) remains incompletely understood. This study employed trio-based whole-exome sequencing (WES) to systematically investigate its genetic architecture and elucidate its genetic basis.Trio-based WES was performed on 36 probands with unexplained eoHM and their unaffected parents. Bioinformatic filtering was used to identify potential deleterious variants following multiple inheritance patterns, including de novo, recessive homozygous, and compound heterozygous. Candidate genes were functionally annotated using GO and KEGG pathway enrichment analyses. PPI networks were subsequently constructed to explore potential functional relationships among these genes. Probands were stratified according to genetic burden, defined as the number of candidate genes identified in each individual and the associating between genetic burden and clinical phenotypes—including refractive error, axial length, and fundus lesion severity—were systematically evaluated. Potential pathogenic variants were detected in 61.1% (22/36) of probands, comprising 36 variants across 23 candidate genes, with de novo variants representing 44.4% (16/36). Functional enrichment analysis revealed that the candidate genes were significantly associated with three major biological modules: visual perception, photoreceptor cilium assembly, and extracellular matrix remodeling. Protein interaction networks supported functional connectivity among these genes. Notably, genetic burden showed a significant correlation with the severity of fundus lesions, but not significant correlation with refractive error or axial length. This study establishes de novo mutations as an important contributor to eoHM and reveals a significant genotype–phenotype correlation between genetic burden and the severity of graded fundus lesions. Our findings suggest that eoHM likely follows an oligogenic additive

model rather than a simple Mendelian inheritance pattern.

**Key words** Early onset high myopia, Whole exome sequencing, Oligogenic Overlap Model, de novo mutations

## Introduction

Myopia represents a leading cause of visual impairment worldwide, with a rapidly increasing global prevalence. Currently, high myopia (HM) affects over 4.0% of the population globally, and this figure is projected to rise to 9.8% by 2050[1]. Individuals with HM are at a significantly increased risk of developing severe ocular complications, including cataract, glaucoma, retinal detachment, macular degeneration, and optic neuropathies. These complications often lead to irreversible vision loss or blindness, making HM a major public health concern[2]. HM is a complex disorder resulting from intricate gene-environment interactions. Notably, eoHM predominantly follows a Mendelian inheritance pattern and is minimally influenced by environmental factors[3], thereby offering a valuable model for elucidating its genetic underpinnings.

To date, 17 genes have been implicated in nonsyndromic HM, encompassing diverse inheritance modes: 11 autosomal dominant (ZNF644[4], SCO2[5], SLC39A5[6] CCDC111[7], P4HA2[8], BSG[9], CPSF1[10], NDUFAF7[11], TNFRSF21[12], XYLT1[13] and DZIP[13]), 4 autosomal recessive (LRPAP1[14], CTSH[14], LOXL3[15], and LEPREL1[16]), and 2 X-linked genes (OPN1LW[17], and ARR3[18]) by WES. However, studies from our group have demonstrated that identifiable genetic causes account for fewer than 14% of HM familial cases[19]. Moreover, many reported loci lack consistent replication across diverse populations, suggesting that a substantial portion of the genetic architecture underlying HM remains to be elucidated.

Current genetic investigations have predominantly centered on sporadic cases or multigenerational pedigrees exhibiting clear autosomal dominant inheritance. Consequently, proband-parent trio studies—which are essential for capturing the complete phenotypic spectrum and for identifying de novo, recessive, and compound heterozygous variants—remain scarce for eoHM[20]. This gap significantly hinders our understanding of the genetic etiology in patients, particularly those without a clear familial history. To bridge this knowledge deficit and to methodically explore the genetic foundations of HM, we performed trio-based WES on 36 unrelated probands with unexplained eoHM.

### Material and methods

#### Patient Subjects

A total of 36 unrelated probands diagnosed with eoHM and their unaffected biological parents were enrolled in this study. Probands were consecutively recruited from the Ophthalmology Department of the Ningxia Hui Autonomous Region People's Hospital between January 2022 and October 2024. Inclusion criteria for probands were as follows[21]: (1) myopia with a spherical equivalent (SE) ≤ –6.00 diopters or an

axial length (AL) ≥ 26 mm in at least one eye, and an age of onset before 7 years; (2) no evidence of congenital cataract, congenital ptosis, or retinopathy of prematurity; (3) no clinical signs of inherited retinal diseases or systemic disorders associated with myopia during follow-up evaluations. All probands and their parents underwent a comprehensive ophthalmic assessment[19], including measurement of best-corrected visual acuity (BCVA), cycloplegic refraction, slit-lamp biomicroscopy, ocular axial length, intraocular pressure (IOP), visual field testing, color fundus photography, optical coherence tomography (OCT), color vision testing, and full-field electroretinography (ERG). The study was conducted in accordance with the principles of the Declaration of Helsinki and was approved by the Ethics Committee of the Ningxia Hui Autonomous Region People's Hospital (Approval No. 2024-GZR-72). Written informed consent was obtained from the parents or legal guardians of all minors participating in this study.

**Whole-exome sequencing**

WES was performed on genomic DNA extracted from 36 probands with eoHM and their unaffected parents. DNA quality was assessed using Qubit fluorometry and agarose gel electrophoresis. Exome capture was carried out for all 108 samples using the Agilent SureSelect Human All Exon Enrichment kit V6 (60 Mb; Agilent Technologies, USA) according to the manufacturer’s instructions. Sequencing was conducted on an Illumina HiSeq platform with 2×150 bp paired-end reads, achieving an average coverage depth of >100× across the target regions. Raw sequencing data were processed using the Burrows-Wheeler Aligner (BWA-MEM) for alignment to the hg19 (GRCh37) reference genome[22]. Variant calling, including single-nucleotide polymorphisms (SNPs) and insertions/deletions (indels), was performed using the Genome Analysis Toolkit (GATK) with standard quality control filters applied throughout the pipeline.

**Bioinformatics analysis**

**Candidate genes variant Filtering**

To improve the accessibility and efficiency of genetic testing, we developed a virtual gene panel comprising approximately 917 genes associated with ocular disorders. This panel includes genes associated with eoHM—both in its isolated form and within syndromic presentations—as well as genes crucial for the development, homeostasis, and function of ocular tissues implicated in eoHM, such as the retina, choroid, and sclera. Candidate genes were selected according to the analysis pipeline summarized in Fig.1. Based on the observed pedigree patterns, we hypothesized that the underlying genetic cause could involve de novo mutations or recessive variants segregating within the eoHM trios. Therefore, variants consistent with de novo, recessive homozygous, or compound heterozygous inheritance models were prioritized. Subsequent filtering was applied using the following criteria: (1) minor allele frequency (MAF) < 0.01 in both the 1000 Genomes Project (1000g_all) and gnomAD (gnomAD_ALL and gnomAD_EAS) databases; (2) location within exonic or splice-site regions (±2 bp); (3) predicted impact on splicing or protein sequence; (4)

classification as deleterious by more than half of the in silico prediction tools applied (SIFT, Polyphen-2, MutationTaster, CADD); (5) for splice-site variants, assessment of potential functional consequences using multiple bioinformatic predictors to evaluate clinical relevance.

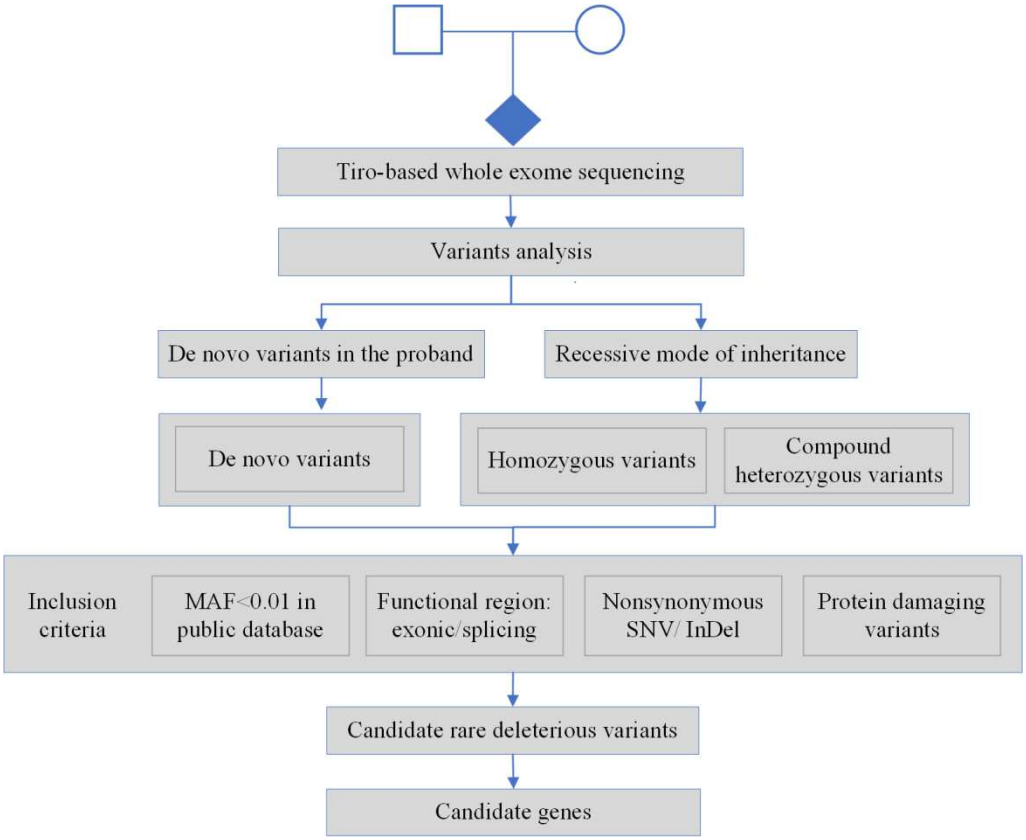



**Fig.1.** Analysis workflow of trio-based exome sequencing data.

**Functional enrichment analysis of candidate genes**

Enrichment analysis of candidate genes associated with eoHM was performed using the Database for Annotation, Visualization and Integrated Discovery (DAVID; https://david.ncifcrf.gov/)[23]. Analyses were conducted based on Gene Ontology (GO) terms, including biological process (BP), cellular component (CC), and molecular function (MF), as well as Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways. Terms with a p-value < 0.05 were considered statistically significant.

**Protein-protein interaction network analysis**

Protein-protein interaction (PPI) networks were constructed using the Search Tool for the Retrieval of Interacting Genes (STRING) database[24]. The candidate gene list was submitted to STRING with "Homo sapiens" set as the organism. Only interactions that were experimentally validated and had a combined confidence score ≥0.4 were

retained for further analysis. Functional enrichment within the PPI network was also assessed using GO annotations through the STRING platform.

**Genotype-phenotype correlation analysis**

Based on the number of candidate disease-associated genes identified, the 36 probands were stratified into three groups: a no-candidate-gene group, a single-candidate-gene group, and a two-candidate-gene group. Phenotypic parameters, including spherical equivalent refraction, axial length, and graded fundus lesions, were subsequently compared across these genetically stratified cohorts. Statistical analyses were performed using SPSS software (version 26.0). Continuous variables are presented as mean ± standard deviation, and categorical variables as frequency (percentage). Group comparisons for categorical variables were conducted using the Chi-square test or Fisher's exact test, as appropriate. Differences among the three groups for continuous variables were assessed by one-way analysis of variance (ANOVA). If the ANOVA result was significant ($p<0.05$), post-hoc pairwise comparisons were performed with Bonferroni correction. A two-sided p-value $< 0.05$ was considered statistically significant.

## Results

**Clinical characteristics of the study cohort**

A total of 36 unrelated probands with eoHM from 36 independent families (19 males, 52.8%; 17 females, 47.2%) were enrolled based on the predefined inclusion criteria. The mean age at diagnosis was 5.08 ± 0.84 years, and the mean age at examination was 8.90 ± 2.35 years (range: 3–12 years). All probands presented with high myopia before the age of 7 years, forming a genetically homogeneous cohort for analysis. Notably, 53% of the subjects were younger than 7 years at the time of evaluation.

Ocular parameters are summarized in Table 1. The mean spherical equivalent (SE) was −9.13 ± 2.96 D in the right eye and −8.64 ± 2.62 D in the left eye. Corresponding axial lengths (AL) were 26.33 ± 1.56 mm and 26.13 ± 1.43 mm, respectively. Two probands exhibited monocular high myopia, with the fellow eye showing moderate myopia (−5.00 D and −3.75 D, respectively). According to the International Photographic Classification and Grading System for Myopic Maculopathy[25], fundus grading of the 36 participants revealed: 7 cases (19%) with no myopic retinal lesions (category C0), 23 (64%) with tessellated fundus (C1), and 6 (17%) with diffuse chorioretinal atrophy (C2) (Fig. 2). For gender-based severity analysis, probands were stratified into more-severe and less-severe eoHM groups, with detailed distribution shown in Table 2.

**Table 1.** Clinical data for 36 probands with eoHM.

| ID | S | AE | ADHM | SE ( D ) | | AL ( mm ) | |
|---|---|---|---|---|---|---|---|
| | | | | OD | OS | OD | OS |
| F1 | M | 7 | 6 | -7.50 | -7.50 | 26.37 | 26.21 |
| F2 | F | 6 | 6 | -6.75 | -6.50 | 26.27 | 26.07 |
| F3 | M | 9 | 5 | -8.00 | -9.50 | 25.61 | 26.13 |
| F4 | F | 12 | 5 | -9.00 | -10.00 | 28.92 | 29.44 |
| F5 | M | 12 | 5 | -10.25 | -10.50 | 27.32 | 27.37 |
| F6 | F | 10 | 5 | -10.25 | -9.75 | 28.06 | 28.01 |
| F7 | F | 7 | 5 | -11.00 | -9.75 | 26.96 | 26.21 |
| F8 | M | 6 | 6 | -7.00 | -6.00 | 25.42 | 25.13 |
| F9 | M | 10 | 5 | -5.00 | -10.75 | 25.41 | 28.98 |
| F10 | M | 10 | 5 | -10.50 | -9.50 | 27.29 | 26.98 |
| F11 | F | 6 | 5 | -7.50 | -7.00 | 24.88 | 24.63 |
| F12 | F | 7 | 4 | -9.00 | -8.50 | 27.45 | 27.27 |
| F13 | F | 6 | 6 | -7.75 | -6.50 | 25.67 | 25.10 |
| F14 | M | 6 | 4 | -4.50 | -9.00 | 23.81 | 25.22 |
| F15 | F | 8 | 6 | -9.00 | -3.75 | 27.03 | 23.33 |
| F16 | M | 6 | 4 | -6.75 | -6.25 | 24.35 | 24.21 |
| F17 | F | 10 | 3 | -13.75 | -13.00 | 28.17 | 27.83 |
| F18 | F | 12 | 6 | -10.50 | -10.00 | 27.60 | 27.42 |
| F19 | M | 7 | 6 | -6.75 | -6.75 | 25.43 | 25.79 |
| F20 | F | 6 | 6 | -6.00 | -6.75 | 24.43 | 24.75 |
| F21 | M | 5 | 4 | -6.00 | -6.50 | 24.53 | 24.87 |
| F22 | M | 7 | 6 | -8.00 | -8.00 | 27.48 | 27.54 |
| F23 | F | 6 | 5 | -8.25 | -9.00 | 24.56 | 24.79 |
| F24 | F | 12 | 4 | -14.75 | -16.00 | 27.57 | 27.90 |
| F25 | F | 12 | 4 | -14.50 | -11.50 | 26.57 | 26.19 |
| F26 | M | 5 | 5 | -6.25 | -6.00 | 25.04 | 24.53 |
| F27 | F | 7 | 6 | -6.75 | -6.25 | 25.44 | 25.23 |
| F28 | F | 10 | 5 | -11.00 | -12.00 | 26.21 | 26.35 |
| F29 | F | 8 | 5 | -9.00 | -7.50 | 24.64 | 24.15 |
| F30 | M | 8 | 4 | -11.00 | -9.00 | 26.15 | 25.57 |
| F31 | M | 10 | 6 | -11.25 | -6.50 | 29.22 | 26.89 |
| F32 | M | 7 | 6 | -16.75 | -14.75 | 27.30 | 27.05 |
| F33 | M | 9 | 5 | -9.75 | -10.25 | 26.85 | 27.06 |
| F34 | M | 11 | 5 | -14.75 | -6.25 | 30.53 | 26.43 |
| F35 | M | 6 | 6 | -7.50 | -6.75 | 24.34 | 24.15 |
| F36 | F | 7 | 4 | -6.50 | -7.50 | 25.30 | 25.97 |

eoHM: early-onset high myopia; S: sex; AE: age at examination; ADHM: age at diagnosis of high myopia; AL: axial length; F: female; M: male; OD: Oculus Dextrus; OS: Oculus Sinister; SE: spherical equivalent.

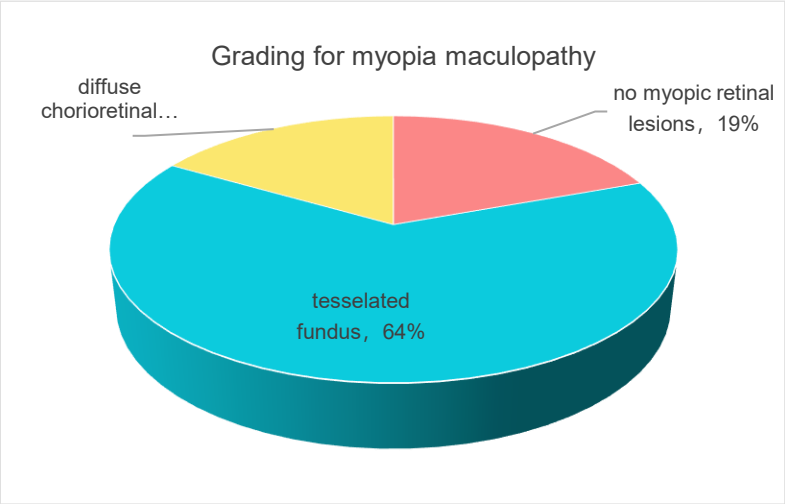



**Fig.2.** Distribution of myopic maculopathy grades in eoHM probands.

**Table 2.** Association Between Gender and Disease Severity in Patients with eoHM.

| | Female | Male | Total Female and Male |
|---|---|---|---|
| More severe≤-10.00 D | 8 | 8 | 16 |
| Less severe＞-10.00 D | 9 | 11 | 20 |
| Tatal More and Less severe | 17 | 19 | 36 |

A chi-square test was applied as all expected frequencies exceeded 5. The analysis revealed no statistically significant association between disease severity and gender ($p$ = 0.765), indicating these variables are independent in the studied cohort.

**Candidate genes identified through trio-based exome sequencing**

Trio-based exome sequencing and subsequent filtering identified 36 rare candidate variants across 23 genes in the eoHM probands (Table 3). According to ACMG guidelines, these variants included 6 classified as pathogenic (P), 1 as likely pathogenic (LP), and 29 as variants of uncertain significance (VUS) (Fig.3). All probands carrying these variants exhibited an isolated myopic phenotype without clinical signs of inherited retinal disease (IRD) upon comprehensive ophthalmic evaluation, including ERG and longitudinal follow-up. Genetic analysis of the trios revealed 16 de novo mutations and 20 compound heterozygous mutations (Fig.4). Among all detected variants, missense changes were predominant (81%), followed by frameshift (8%), stop-gain (8%), and splice-site variants (3%) (Fig.5). Notably, several probands carried candidate variants in two distinct genes (Fig.6).

**Table 3.** Candidate Pathogenic Mutations of Probands with eoHM.

| ID | Gene |
|---|---|
| F2 | HPS6 |
| | HPS6 |
| F3 | ERCC8 |
| F4* | RP1L1 |
| | CNGB1 |
| | CNGB1 |
| F5 | HSPG2 |
| | HSPG2 |
| F7 | CEP290 |
| F9 | USH2A |
| F10 | FLG |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | FLG | NM_002016.2 | c.5258G>A | p.G1753D | missense SNV | VUS | PM2+PP3 | het | Pat |
| F11 | IFIH1 | NM_022168.4 | c.2459G>T | p.R820L | missense SNV | VUS | PM2+PP3 | het | DNV |
| F17 | LTBP2 | NM_000428.3 | c.1399+4A>T | — | splicing | LP | PS2+PM2+PP3 | het | DNV |
| F18 | CLN6 | NM_017882.2 | c.5A>G | p.E2G | missense SNV | VUS | PM2+PP3 | het | DNV |
| F19 | HERC2 | NM_004667.5 | c.2101A>G | p.K701E | missense SNV | VUS | PM2+PP3 | het | DNV |
| F20 | GPR143 | NM_000273.2 | c.36delC | p.P12fs | frameshift | P | PVS1+PM2+PP3 | het | DNV |
| F21* | RAX2 | NM_001319074.1 | c.500G>A | p.R167Q | missense SNV | VUS | PM2+PP3 | het | DNV |
| | CNGB1 | NM_001297.4 | c.1631C>T | p.P544L | missense SNV | VUS | PM2+PP3 | het | DNV |
| F22 | SPTBN1 | NM_003128.2 | c.362A>G | p.Q121R | missense SNV | VUS | PM2+PP3 | het | DNV |
| F23* | GALM | NM_138801.2 | c.77C>A | p.S26X | stopgain | P | PVS1+PM2+PP3 | het | DNV |
| | KIAA0586 | NM_001244189.1 | c.949G>A | p.E317K | missense SNV | VUS | PM2+PP3 | het | DNV |
| F24 | NSD1 | NM_022455.4 | c.2170G>A | p.E724K | missense SNV | VUS | PM2+PP3 | het | Pat |
| | NSD1 | NM_022455.4 | c.3187A>G | p.T1063A | missense SNV | VUS | PM2+PP3 | het | Mat |
| F27 | CSPP1 | NM_001291339.1 | c.109C>T | p.H37Y | missense SNV | VUS | PM2+PP3 | het | Pat |
| | CSPP1 | NM_001291339.1 | c.677C>T | p.A226V | missense SNV | VUS | PM2+PP3 | het | Mat |
| F31* | VCAN | NM_001164097.1 | c.1801C>A | p.P601T | missense SNV | VUS | PM2+PP3 | het | DNV |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | AGRN | NM_198576.4 | c.4055G>A | p.G1352D | missense SNV | VUS | PM2+PP3 | het | Pat |
| | AGRN | NM_198576.4 | c.4132G>A | p.E1378K | missense SNV | VUS | PM2+PP3 | het | Mat |
| F32 | RP1L1 | NM_178857.5 | c.4284G>C | p.Q1428H | missense SNV | VUS | PM2+PP3 | het | Mat |
| | RP1L1 | NM_178857.5 | c.6063delC | p.D2021Efs*3 | frameshift | P | PVS1+PM2+PP3 | het | Pat |
| F33 | FLG2 | NM_001014342.2 | c.733G>A | p.E245K | missense SNV | VUS | PM2+PP3 | het | Mat |
| | FLG2 | NM_001014342.2 | c.359A>T | p.E120V | missense SNV | VUS | PM2+PP3 | het | Pat |
| F34 | TUBA3D | NM_080386.4 | c.107T>G | p.M36R | missense SNV | VUS | PM2+PP3 | het | DNV |
| F36 | HERC2 | NM_004667.6 | c.10474T>C | p.S3492P | missense SNV | VUS | PM2+PP3 | het | Pat |
| | HERC2 | NM_004667.6 | c.8329A>G | p.M2777V | missense SNV | VUS | PM2+PP3 | het | Mat |

AC: ACMG-Guideline Class; VUS: uncertain significance variant; P: pathogenic variant; het: heterozygous; hom: homozygous; hem: hemizygous; DNV: de novo variant; Mat: maternal; Pat: paternal; Bip: biparental. *: probands harbor two candidate genes；RefSeq ID：Reference Sequence Database accession number；

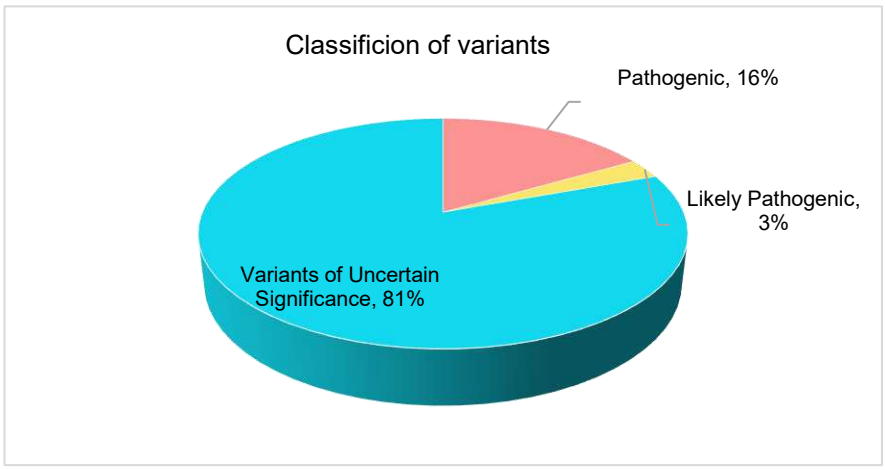



**Fig. 3.** Classification of variant pathogenicity in candidate eoHM-associated genes.

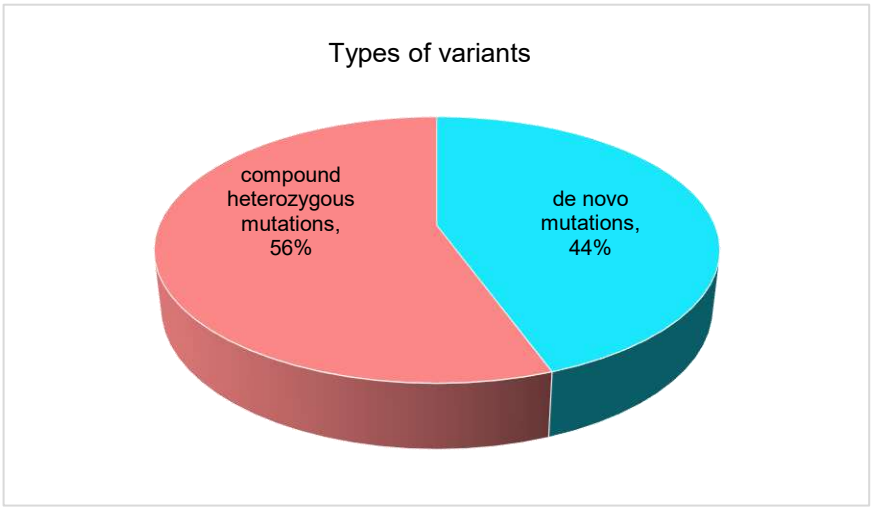



**Fig. 4.** Inheritance patterns of identified candidate variants in eoHM probands.

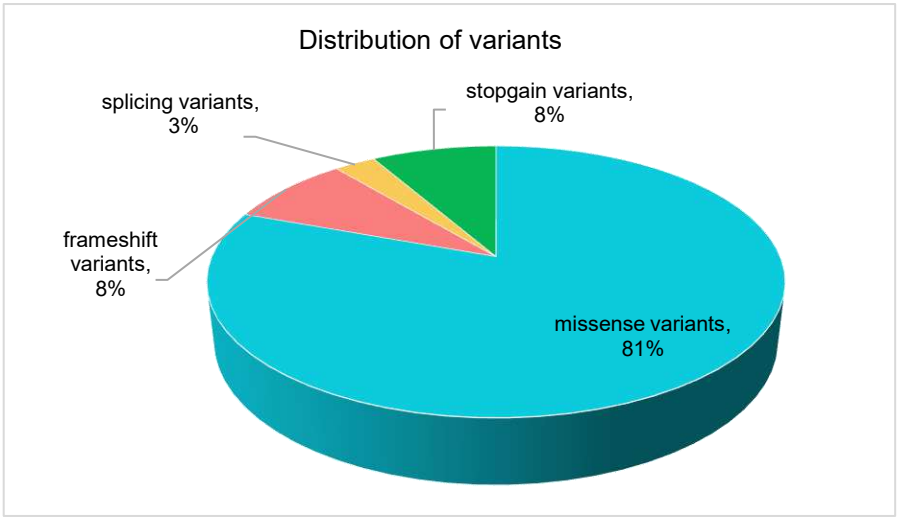



**Fig. 5.** Distribution of variant types among candidate eoHM genes.

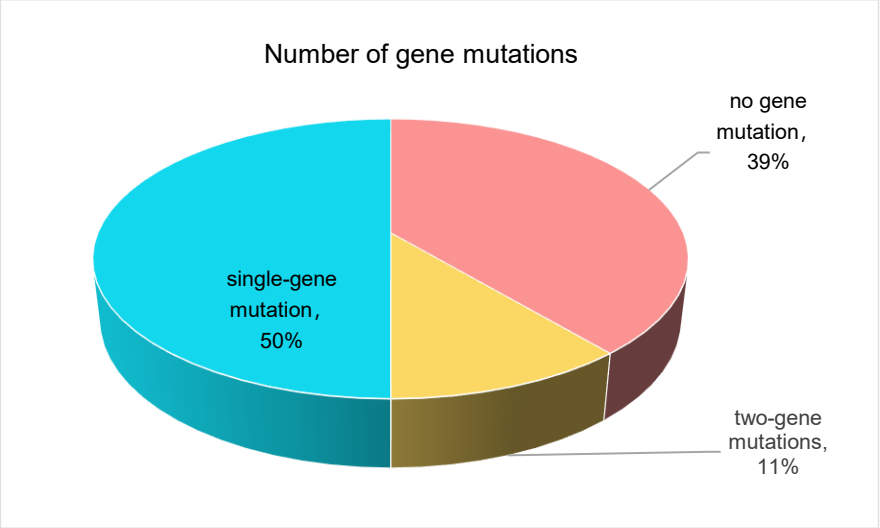



**Fig. 6.** The number of candidate gene mutations associated with eoHM.

### Functional enrichment analysis of candidate genes

Functional enrichment analysis was performed on the candidate gene set using DAVID to identify overrepresented GO terms and KEGG pathways. GO analysis yielded 5 significantly enriched BP terms, 6 CC terms, and 4 MF terms (Fig.7A). Enriched BP terms primarily involved visual perception, photoreceptor cell maintenance, regulation of protein localization, establishment of skin barrier, and microtubule-based processes. Enriched CC terms included the ciliary basal body, collagen-containing extracellular matrix, photoreceptor connecting cilium, cell projection, photoreceptor outer segment, and keratohyalin granule. The MF terms were enriched for structural constituent of cytoskeleton, calcium ion binding, structural constituent of extracellular matrix, and transition metal ion binding. KEGG pathway analysis further indicated significant enrichment for the regulation of actin cytoskeleton pathway (Fig.7B).

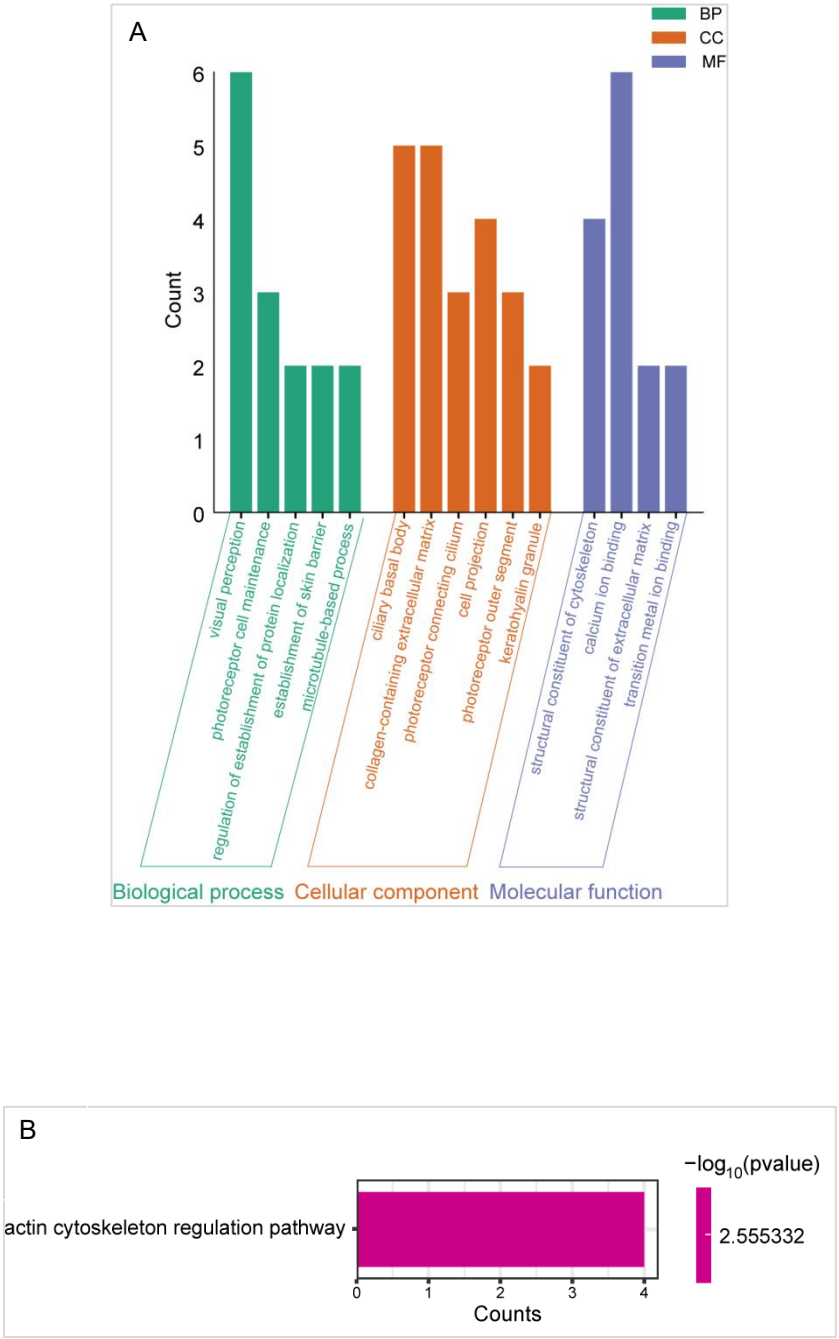



**Fig.7.** (A) Significantly enriched GO terms in candidate genes. A p-value < 0.05 were set as the cut off criteria. (B) Significantly enriched KEGG terms in candidate genes. A p-value < 0.05 were set as the cut off criteria. Image adapted from the KEGG pathway map (map04810) [26].

**Protein-protein interaction network analysis of candidate genes**

A PPI network was constructed for the 23 candidate genes using the STRING database with a confidence threshold > 0.4. The resulting network comprised 23 nodes and 17 edges, with a significant PPI enrichment p-value of $5.1\times10^{-9}$, indicating strong evidence of biological interaction among the candidate proteins. Network topology analysis revealed a local clustering coefficient of 0.522, indicating a

moderately clustered structure. Within this network, 15 genes demonstrated direct interactions (node degree ≥ 1), while 8 remained isolated. GO enrichment analysis of the interacting gene set identified significantly enriched biological processes (FDR < 0.01), primarily associated with visual perception, cilium assembly, and extracellular matrix remodeling. The observed physical interactions and shared functional annotations suggest that these candidate genes form a coherent biological module relevant to eoHM pathogenesis (Fig.8).

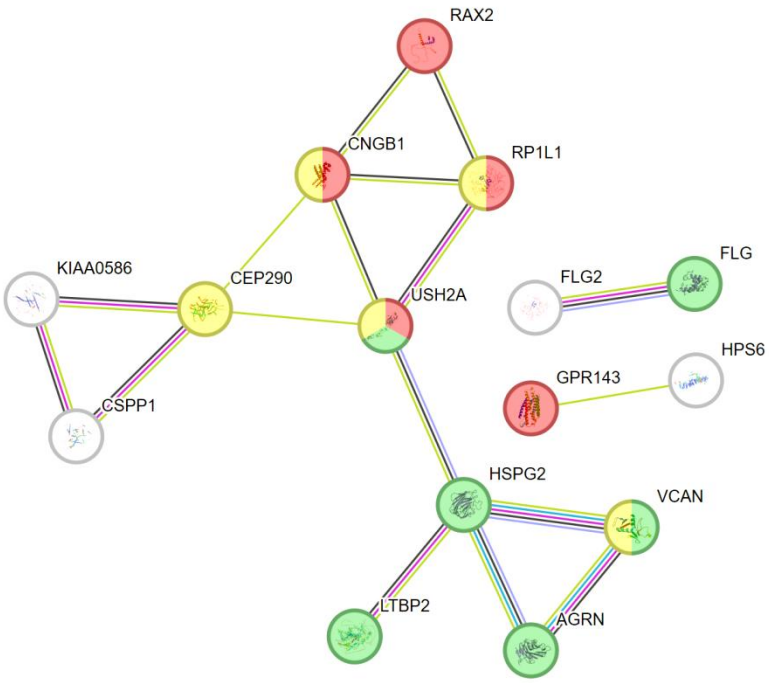



**Fig. 8.** PPI network of candidate eoHM genes. The network was generated using the STRING database, with nodes color-coded by their primary enriched Gene Ontology biological process: red for visual perception, yellow for cilium assembly, green for extracellular matrix remodeling, and white for genes without predominant enrichment in these specific categories.

### Genotype-phenotype correlation analysis

Based on their candidate gene carrier status, the 36 probands were stratified into three groups: a no-candidate-gene group (n=14), a single-candidate-gene group (n=18), and a two-candidate-gene group (n=4). The mean SE was −9.04 ± 2.28 D, −10.18 ± 3.29 D, and −9.69 ± 2.79 D for the three groups, respectively. Corresponding mean AL were 25.90 ± 1.14 mm, 27.00 ± 1.44 mm, and 27.33 ± 2.90 mm. No statistically significant differences in SE or AL were observed across the groups ($p > 0.05$).

Fundus lesions were graded according to the International Photographic Classification and Grading System for Myopic Maculopathy, using the more severely affected eye in cases of interocular asymmetry. The distribution of myopic maculopathy categories (C0: no lesions; C1: tessellated fundus; C2: diffuse

chorioretinal atrophy) across the three genetic groups is summarized as follows. In the no-candidate-gene group, 6 cases (42.9%) were graded C0 and 8 cases (57.1%) C1, with no C2 lesions. In the single-candidate-gene group, lesions were distributed as C0: 1 (5.6%), C1: 13 (72.2%), and C2: 4 (22.2%). The two-candidate-gene group showed no C0 lesions, with 2 cases (50.0%) each in C1 and C2. A significant association was observed between the number of candidate genes and fundus lesion severity by Fisher's exact test ($p = 0.023$).

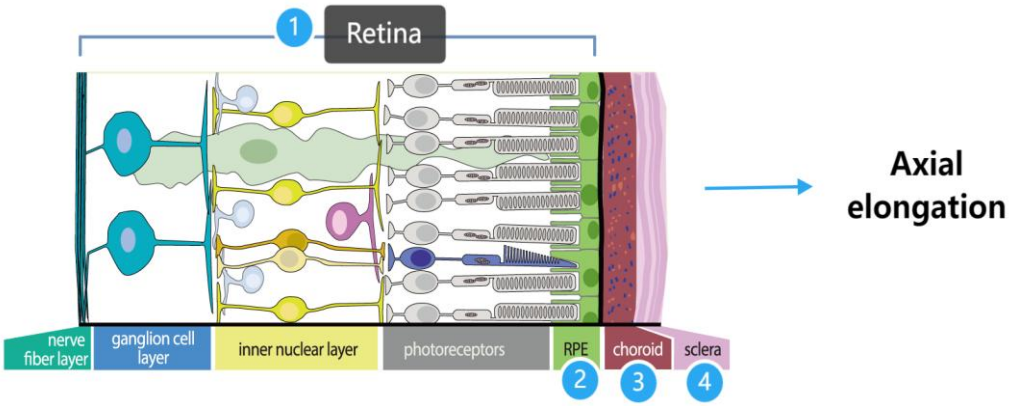



**Fig.9.** Schematic diagram illustrating the multilayer signaling cascade that promotes axial elongation. The order of signal transduction is depicted by the numerals 1 through 4. Adapted from Tedja et al[27].

**Discussion**

In genetic studies of eoHM, prior research has largely focused on large pedigrees following Mendelian inheritance or on known HM-associated genes[28,29]. These investigations have predominantly been conducted under an autosomal dominant inheritance framework, with detection rates often limited to specific families or previously reported gene sets. However, accumulating evidence indicates that eoHM may involve a more complex genetic architecture[30,31], encompassing autosomal recessive inheritance, de novo variants, and polygenic additive effects—patterns that are likely to be underdetected or overlooked by conventional gene-screening strategies. Based on this understanding, the present study employed a trio-based WES approach to systematically investigate the genetic etiology of eoHM. Through analysis of 36 core families with genetically unexplained eoHM, we identified 36 potentially pathogenic variants across 23 candidate genes in 22 probands, yielding an overall detection rate of 61.1%. De novo variants accounted for 44.4% (16/36) of these findings, indicating a potentially substantial contribution of new mutations to the genetic makeup of eoHM. Functional enrichment analysis further showed that these genes are enriched in pathways related to visual perception, photoreceptor cilule function, and extracellular matrix remodeling, suggesting a cooperative multi-pathway mechanism in eoHM pathogenesis.

We observed a relatively high frequency of de novo variants (DNMs), exceeding previously reported rates[19]. While this proportion may be partially attributable to the limited sample size, the results suggest that DNMs constitute a notable etiological factor in eoHM. The trio-based study design employed here provides a distinct advantage for delineating the inheritance patterns of candidate causal variants since Trio-based WES reliably identifies de novo events while also resolving recessive and dominant inheritance[20,32]. The established role of DNMs in complex neurodevelopmental disorders such as schizophrenia and intellectual disability further supports their potential contribution to other complex diseases, including ocular disorders[33-35]. Although DNMs have been implicated in certain inherited eye diseases such as Stargardt disease and retinoschisis[36,37], reported cases remain scarce, and prior pedigree-based studies have typically included fewer than a dozen families. The considerable proportion of DNMs identified in our eoHM cohort underscores the need for larger, systematic investigations to clarify their prevalence and pathogenic role across the spectrum of inherited ocular conditions.

Functional enrichment analysis identified three principal biological modules implicated in eoHM pathogenesis. First, genes including RAX2, CLN6, GRP143, RP1L1, CNGB1, and USH2A are involved in phototransduction—the process that converts light stimuli into neural signals within the retina[38]. While a direct causal relationship with eoHM requires further validation, their established role in visual perception—a function essential for ocular growth feedback—strongly supports their potential contribution to disease pathogenesis, consistent with prior reports[39,40]. Second, enrichment of photoreceptor cilium-related genes (CNGB1, CEP290, RP1L1, USH2A, VCAN) highlights an alternative pathway. The outer segment cilium is a critical structure for phototransduction and photoelectric conversion[41,42]. Mutations in cilium-associated genes may lead to photoreceptor dysfunction, which has been linked to myopia progression[43]. Third, extracellular matrix (ECM) remodeling—closely associated with scleral alterations in myopia—involves genes such as HSPG2, VCAN, LTBP2, AGRN, and USH2A. These genes influence axial elongation by modulating ECM composition. Notably, a subset (HSPG2, SPTBN1, VCAN, AGRN) is also enriched in the regulation of actin cytoskeleton pathway. Scleral hypoxia can activate this pathway via HIF-1α signaling, driving myofibroblast transdifferentiation and ECM remodeling, thereby promoting myopic progression[44,45]. Variants in ECM- and cytoskeleton-related genes may thus directly compromise scleral biomechanical strength, increasing susceptibility to abnormal ocular expansion. Given the central role of scleral remodeling in myopia[46], these genes likely form a structural foundation in eoHM pathology. Several candidate genes exhibit functional pleiotropy, participating in multiple modules simultaneously. For example, some genes contribute to both phototransduction, ciliary function and ECM remodeling, underscoring the interconnected nature of the genetic network underlying high myopia. Ultimately, these processes can be integrated within a unified visual-feedback-regulated emmetropization framework: optical defocus signals are detected by the retina, triggering an intricate intracellular signaling

cascade that depends on proper cilium function and signal transduction. These signals are then relayed through the retinal pigment epithelium (RPE) and choroid to finally act on the sclera, where they modulate actin cytoskeleton dynamics and ECM composition to achieve adaptive axial growth(Fig. 9)[27]. The candidate genes identified in this study are distributed across various levels of this regulatory pathway, and their functional alterations may disrupt emmetropization at distinct nodes, ultimately leading to excessive axial elongation[47]. Therefore, the genetic etiology of eoHM can be summarized as a cooperative dysregulation of multiple interrelated biological processes within the visual-signaling and ocular-growth regulatory network.

This study demonstrates a significant association between genetic burden (the number of candidate genes) and the severity of fundus lesions through genotype–phenotype correlation analysis in an eoHM cohort. Although SE and AL did not differ significantly among groups stratified by genetic burden, fundus-lesion grades showed a distinct distribution. This suggests that excessive axial elongation and structural fundus damage in eoHM may be driven by partially distinct genetic pathways. Specifically, neither SE nor AL differed significantly across groups, irrespective of the number of candidate genes carried. This finding suggests that in the established eoHM phenotype, the primary genetic drivers of abnormal axial elongation may have already reached full expressivity, or that compensatory mechanisms buffer these gross optical parameters against limited variations in genetic burden. In contrast, the number of candidate gene variants showed a positive correlation with the severity of fundus lesions, indicating that increased genetic burden primarily exacerbates tissue vulnerability and impairs reparative capacity in the fundus during disease progression. Collectively, these results suggest that the genetic basis of eoHM aligns more closely with an oligogenic model rather than a purely Mendelian inheritance pattern. To further evaluate this model, we constructed a protein–protein interaction network for the candidate genes. Network analysis of patients carrying two candidate variants revealed oligogenic interaction patterns, predominantly involving combinations of visual signaling with photoreceptor cilium assembly, visual signaling with ECM homeostasis, and ECM homeostasis with photoreceptor cilium assembly. We hypothesize that co-inheritance of variants affecting distinct functional modules may create synergistic deficits in both visual signal transduction and fundus structural maintenance, thereby accelerating the transition from tessellated fundus to irreversible atrophic lesions during myopic progression. By extending the genetic focus beyond axial growth regulation to include susceptibility to fundus complications, this study provides a novel perspective for understanding phenotypic heterogeneity in HM. Limitations of the work include the modest sample size, especially in the subgroup carrying two candidate variants, which may affect statistical power.

The distribution of pathogenic variant types in our eoHM cohort aligns with prior reports. Yang et al[28] reported that missense variants accounted for 78.38% of mutations identified in eoHM patients, with the remaining 21.62% comprising nonsense, frameshift, splice-site, and start-loss variants. In our study, 81% of the identified variants were missense, while 19% were loss-of-function alterations

(including nonsense, frameshift, and splice-site variants), demonstrating a highly consistent mutational spectrum. Although the precise distribution of subtypes may be influenced by our limited sample size, these findings further support the predominant role of missense variants in the genomic architecture of eoHM. Notably, the majority of variants were classified as variants of uncertain significance (VUS). This observation is likely attributable to the high proportion of missense changes, which can exert diverse effects on protein function through single-amino-acid substitutions and are often difficult to interpret definitively without functional evidence. Consequently, they are more frequently assigned as VUS. This underscores the need for subsequent functional validation through in vitro or in vivo studies to clarify the pathogenicity of these variants and to confirm the role of the associated genes in eoHM etiology.

Notably, several candidate genes identified in this study have been associated in the OMIM database with a spectrum of ocular disorders, including retinitis pigmentosa, cone–rod dystrophy, and keratoconus. To date, none of the variant carriers in our cohort have manifested clinical features consistent with these syndromic conditions. Therefore, long-term follow-up of these eoHM patients is warranted to determine whether these genes genuinely contribute to the pathogenesis of non-syndromic HM.

## Conclusions

This study employed WES technology for the entire family to systematically identify candidate pathogenic genes for eoHM, revealing the significant role of new mutations in the genetic etiology of eoHM. It also systematically confirmed for a significant correlation between genetic burden and the severity of graded fundus lesions. The genetic pattern of eoHM may conform to the polygenic superimposition model rather than a single Mendelian inheritance. These findings provide a new perspective for understanding the genetic mechanism of eoHM and lay a theoretical foundation for future precision medicine and individualized intervention strategies.

### References:


1. Holden, B. A. et al. Global Prevalence of Myopia and High Myopia and Temporal Trends From 2000 through 2050. *Ophthalmology*. **123**, 1036-1042 (2016).
2. Haarman, A. E. G. et al. The Complications of Myopia: A Review and Meta-Analysis. *Invest. Ophth. Vis. Sci.* **61**, 49 (2020).
3. Zhang, Q. Genetics of Refraction and Myopia. *Prog Mol Biol Transl Sci*. **134**, 269-279 (2015).
4. Shi, Y. et al. Exome Sequencing Identifies Znf644 Mutations in High Myopia. *Plos Genet.* **7**, e1002084 (2011).
5. Tran-Viet, K. N. et al. Mutations in Sco2 are Associated with Autosomal-Dominant High-Grade Myopia. *Am. J. Hum. Genet.* **92**, 820-826 (2013).
6. Guo, H. et al. Slc39a5 Mutations Interfering with the Bmp/Tgf-Beta Pathway in Non-Syndromic High Myopia. *J. Med. Genet.* **51**, 518-525 (2014).
7. Zhao, F. et al. Exome Sequencing Reveals Ccdc111 Mutation Associated with High Myopia. *Hum.*

*Genet.* **132**, 913-921 (2013).
8. Guo, H. et al. Mutations of P4Ha2 Encoding Prolyl 4-Hydroxylase 2 are Associated with Nonsyndromic High Myopia. *Genet. Med.* **17**, 300-306 (2015).
9. Jin, Z. B. et al. Trio-Based Exome Sequencing Arrests De Novo Mutations in Early-Onset High Myopia. *Proc Natl Acad Sci U S A*. **114**, 4219-4224 (2017).
10. Ouyang, J. et al. Cpsf1 Mutations are Associated with Early-Onset High Myopia and Involved in Retinal Ganglion Cell Axon Projection. *Hum. Mol. Genet.* **28**, 1959-1970 (2019).
11. Wang, B. et al. A Novel Potentially Causative Variant of Ndufaf7 Revealed by Mutation Screening in a Chinese Family with Pathologic Myopia. *Invest Ophthalmol Vis Sci*. **58**, 4182-4192 (2017).
12. Pan, H. et al. Tnfrsf21 Mutations Cause High Myopia. *J. Med. Genet.* **56**, 671-677 (2019).
13. Lee, J. K., Kim, H., Park, Y. M., Kim, D. H. & Lim, H. T. Mutations in Dzip1 and Xylt1 are Associated with Nonsyndromic Early Onset High Myopia in the Korean Population. *Ophthalmic Genet.* **38**, 395-397 (2017).
14. Aldahmesh, M. A. et al. Mutations in Lrpap1 are Associated with Severe Myopia in Humans. *Am. J. Hum. Genet.* **93**, 313-320 (2013).
15. Li, J. et al. Exome Sequencing Identified Null Mutations in Loxl3 Associated with Early-Onset High Myopia. *Mol. Vis.* **22**, 161-167 (2016).
16. Guo, H. et al. Homozygous Loss-of-Function Mutation of the Leprel1 Gene Causes Severe Non-Syndromic High Myopia with Early-Onset Cataract. *Clin. Genet.* **86**, 575-579 (2014).
17. Li, J. et al. Unique Variants in Opn1Lw Cause Both Syndromic and Nonsyndromic X-Linked High Myopia Mapped to Myp1. *Invest Ophthalmol Vis Sci*. **56**, 4150-4155 (2015).
18. Xiao, X., Li, S., Jia, X., Guo, X. & Zhang, Q. X-Linked Heterozygous Mutations in Arr3 Cause Female-Limited Early Onset High Myopia. *Mol. Vis.* **22**, 1257-1266 (2016).
19. Ye, M. et al. Mutational Investigation of 17 Causative Genes in a Cohort of 113 Families with Nonsyndromic Early-Onset High Myopia in Northwestern China. *Mol. Genet. Genomics*. **298**, 669-682 (2023).
20. Brea-Fernandez, A. J. et al. Trio-Based Exome Sequencing Reveals a High Rate of the De Novo Variants in Intellectual Disability. *Eur. J. Hum. Genet.* **30**, 938-945 (2022).
21. Liu, Y. et al. 18 Individual Genes Underwent Variant Screening in a Northwest Chinese Group Comprised 83 Probands Diagnosed with Early-Onset High Myopia. *Plos One*. **20**, e329472 (2025).
22. Li, H. & Durbin, R. Fast and Accurate Long-Read Alignment with Burrows-Wheeler Transform. *Bioinformatics*. **26**, 589-595 (2010).
23. Huang, D. W., Sherman, B. T. & Lempicki, R. A. Systematic and Integrative Analysis of Large Gene Lists Using David Bioinformatics Resources. *Nat. Protoc.* **4**, 44-57 (2009).
24. Szklarczyk, D. et al. String V10: Protein-Protein Interaction Networks, Integrated Over the Tree of Life. *Nucleic Acids Res.* **43**, D447-D452 (2015).
25. Ohno-Matsui, K. et al. International Photographic Classification and Grading System for Myopic Maculopathy. *Am. J. Ophthalmol.* **159**, 877-883 (2015).
26. Kanehisa, M. & Goto, S. KEGG: Kyoto Encyclopedia of Genes and Genomes. Nucleic Acids Res. **28**, 27-30 (2000).
27. Tedja, M. S. et al. Imi - Myopia Genetics Report. *Invest Ophthalmol Vis Sci*. **60**, M89-M105 (2019).
28. Yang, E., Yu, J., Liu, X., Chu, H. & Li, L. Familial Whole Exome Sequencing Study of 30 Families with Early-Onset High Myopia. *Invest. Ophth. Vis. Sci.* **64**, 10 (2023).

29. Ye, M. et al. Mutational Investigation of 17 Causative Genes in a Cohort of 113 Families with Nonsyndromic Early-Onset High Myopia in Northwestern China. *Molecular Genetics and Genomics : Mgg*. **298**, 669-682 (2023).
30. Jin, Z. et al. Trio-Based Exome Sequencing Arrests De Novo Mutations in Early-Onset High Myopia. *P. Natl. Acad. Sci. Usa.* **114**, 4219-4224 (2017).
31. Sánchez-Cazorla, E. et al. Whole-Exome Sequencing of 21 Families: Candidate Genes for Early-Onset High Myopia. *Int. J. Mol. Sci.* **24**, 15676 (2023).
32. Bowling, K. M. et al. Genomic Diagnosis for Children with Intellectual Disability and/Or Developmental Delay. *Genome Med.* **9**, 43 (2017).
33. Howrigan, D. P. et al. Exome Sequencing in Schizophrenia-Affected Parent-Offspring Trios Reveals Risk Conferred by Protein-Coding De Novo Mutations. *Nat. Neurosci.* **23**, 185-193 (2020).
34. Li, C. et al. Trio-Whole Exome Sequencing Reveals the Importance of De Novo Variants in Children with Intellectual Disability and Developmental Delay. *Sci Rep*. **14**, 27590 (2024).
35. Brea-Fernández, A. J. et al. Trio-Based Exome Sequencing Reveals a High Rate of the De Novo Variants in Intellectual Disability. *European Journal of Human Genetics : Ejhg*. **30**, 938-945 (2022).
36. Hu, F. et al. Novel Variants Associated with Stargardt Disease in Chinese Patients. *Gene*. **754**, 144890 (2020).
37. Gao, F. et al. Comprehensive Analysis of Genetic and Clinical Characteristics of 30 Patients with X-Linked Juvenile Retinoschisis in China. *Acta Ophthalmol.* **99**, e470-e479 (2021).
38. Smith, R. L., Sivaprasad, S. & Chong, V. Retinal Biochemistry, Physiology and Cell Biology. *Dev Ophthalmol*. **55**, 18-27 (2016).
39. Wan, L., Deng, B., Wu, Z. & Chen, X. Exome Sequencing Study of 20 Patients with High Myopia. *Peerj*. **6**, e5552 (2018).
40. Zhuang, W. et al. Genotype-Ocular Biometry Correlation Analysis of Eight Primary Angle Closure Glaucoma Susceptibility Loci in a Cohort From Northern China. *Plos One*. **13**, e206935 (2018).
41. Masek, M. & Bachmann-Gagescu, R. Control of Protein and Lipid Composition of Photoreceptor Outer Segments-Implications for Retinal Disease. *Curr. Top. Dev. Biol.* **155**, 165-225 (2023).
42. Corral-Serrano, J. C. et al. Eupatilin Improves Cilia Defects in Human Cep290 Ciliopathy Models. *Cells*. **12**, (2023).
43. Sanchez-Cazorla, E. et al. Whole-Exome Sequencing of 21 Families: Candidate Genes for Early-Onset High Myopia. *Int. J. Mol. Sci.* **24**, (2023).
44. Wu, H. et al. Scleral Hypoxia is a Target for Myopia Control. *Proc Natl Acad Sci U S A*. **115**, E7091-E7100 (2018).
45. Zhao, F. et al. Scleral Hif-1Alpha is a Prominent Regulatory Candidate for Genetic and Environmental Interactions in Human Myopia Pathogenesis. *Ebiomedicine*. **57**, 102878 (2020).
46. McBrien, N. A. & Gentle, A. Role of the Sclera in the Development and Pathological Complications of Myopia. *Prog. Retin. Eye Res.* **22**, 307-338 (2003).
47. Summers, J. A., Schaeffel, F., Marcos, S., Wu, H. & Tkatchenko, A. V. Functional Integration of Eye Tissues and Refractive Eye Development: Mechanisms and Pathways. *Exp. Eye Res.* **209**, 108693 (2021).

**Acknowledgments:** Thanks to all the contributors and participants.

**Author contributions:** These authors contributed equally: Wenjing Li and Zhenglai Wang. Material preparation: Wenjing Li. Bingtao Li. Data Curation: Keyan Liu. Rui Li. Yanan Ma. Meiling Su. Haixuan Wang. Zichun Lin; Formal Analysis: Wenjing Li and Zhenglai Wang; Funding Acquisition: Wenjuan Zhuang; Methodology: Jinjin Zhang. Wenjuan Zhuang; Supervision: Wenjuan Zhuang; Writing-original draft: Wenjing Li.

**Funding:** This work was supported by the National Natural Science Foundation of China: 82460202; Ningxia Key Research and Development Program Project: 2024BEG02016;

**Data availability:** The datasets analysed during the current study have been deposited in the NCBI BioProject database under the accession number PRJNA1506340. The data will be publicly released upon publication. For reviewer verification, a confidential access link is available at: https://dataview.ncbi.nlm.nih.gov/object/PRJNA1506340?reviewer=vu19qgb06hp6nv9585fra3n5vb.

**Conflicts of Interes**t: The authors declare no conflict of interest.

**Ethics approval** This work was covered by ethical approval from the ethical committee of the Ningxia Hui Autonomous Region People's Hospital (2024-GZR-72).Written informed consent was obtained from the parents or legal guardians of all minors participating in this study.