The Systemic Genotype-Phenotype Characterization of *PAX6*-Related Eye Disease in 164 Chinese Families

Yi Jiang, Zhen Yi, Yuxi Zheng, Jiamin Ouyang, Dongwei Guo, Shiqiang Li, Xueshan Xiao, Panfeng Wang, Wenmin Sun, and Qingjiong Zhang

State Key Laboratory of Ophthalmology, Zhongshan Ophthalmic Center, Sun Yat-Sen University, Guangdong Provincial Key Laboratory of Ophthalmology and Visual Science, Guangzhou, China

Correspondence: Qingjiong Zhang, FARVO, Pediatric and Genetic Eye Clinic, Zhongshan Ophthalmic Center, Sun Yat-Sen University, 54 Xianlie Road, Guangzhou 510060, China; zhangqji@mail.sysu.edu.cn.
Wenmin Sun, Ophthalmic Molecular Genetics, Zhongshan Ophthalmic Center, Sun Yat-Sen University, 54 Xianlie Road, Guangzhou 510060, China; sunwenmin@gzzoc.com.

Received: May 21, 2024
Accepted: August 11, 2024
Published: August 30, 2024

Citation: Jiang Y, Yi Z, Zheng Y, et al. The systemic genotype-phenotype characterization of *PAX6*-related eye disease in 164 chinese families. *Invest Ophthalmol Vis Sci.* 2024;65(10):46. https://doi.org/10.1167/iovs.65.10.46

Purpose. This study aims to evaluate the genetic and phenotypic characteristics and elucidate the genotype-phenotype correlations of a large Chinese cohort with *PAX6*-related disorders.

Methods. Variants detected with exome sequencing were filtered through multistep bioinformatic and co-segregation analyses, and validated by Sanger sequencing. The related clinical data were collected, and cluster analysis and statistical analysis of the *PAX6*-related phenotypes across different variant groups were carried out. Parental mosaicism was investigated using cloning analysis and Droplet digital PCR.

Results. A total of 119 pathogenic or likely pathogenic *PAX6* variants, including 74 truncation, 31 missense, and 14 others, were identified in 228 patients from 164 unrelated families. The most common phenotypes were foveal hypoplasia (97.8%), nystagmus (92.6%), aniridia (76.7%), cataract (36.8%), and iris hypoplasia (22.4%). Mosaicism ranging from 13.9% to 18.8% was identified in 3 unrelated patients' parents with relatively mild phenotypes. Missense variants in the linker region of the paired domain were associated with high myopia, whereas truncation variants in the homeodomain and proline-serine-threonine-rich domain were associated with hyperopia. Similarly, the degree of iris defects, visual acuity, and associated ocular comorbidity varied among the different types and locations of *PAX6* variants.

Conclusions. Our data indicate that foveal hypoplasia but not aniridia is the most common sign of *PAX6*-related disorders, contributing to subtle iris changes that might easily be overlooked in clinical practice. Recognition of mosaicism in atypical cases or parents with very mild phenotypes is important in genetic counseling as their offspring are at increased risk of typical aniridia. Recognition of the genotype-phenotype relationship emphasizes involvement of *PAX6* regulation in shaping complex ocular phenotypes.

Keywords: *PAX6*, genotype-phenotype, high myopia, mosaicism, aniridia

PAX6 (OMIM 607108) encodes the human paired box 6, the essential transcription factor, consisting of 2 conserved DNA-binding domains: the homeodomain (HD) and the paired domain (PD) comprising N-terminal (NTS) and C-terminal (CTS) subdomains.[1,2] It plays the vital role in normal eye development but also in embryological development of the olfactory, pancreas, and central nervous system.[3] Heterozygous variants in *PAX6* can lead to broad phenotypic anomalies of the eye, in which aniridia is the most common ocular developmental malformation due to single gene variants. Besides, additional signs are frequently reported to be *PAX6* variant-related phenotypes, including nystagmus, foveal hypoplasia, coloboma, cataract, glaucoma, microphthalmia, high myopia, aniridia-related keratopathy (ARK), etc. Some of these signs may present together or independently. However, recent studies demonstrated that a proportion of patients with *PAX6* variants mainly manifested foveal hypoplasia but without obvious complete or partial defect, so such patients might be misdiagnosed with other conditions without the aid of genetic diagnosis.[4,5] The most striking problem is that a significant proportion of patients with *PAX6* variants are at risk of developing blindness complications, that is, 48% to 80% with ARK and 15% to 66.7% with glaucoma.[6] Therefore, recognition of the most common hallmark of *PAX6*-associated ocular disease would be important in clinical diagnosis, genetic diagnosis, as well as long-term care.

Additionally, *PAX6* has been suggested as the candidate gene for refractive error.[7] Interestingly, the association between *PAX6* polymorphism and refractive error as well as myopia remains controversial. Although *PAX6* single nucleotide polymorphisms (SNPs) have shown significant associations with extreme high myopia in Chinese adults[8] and anisometropia in Chinese children, particularly in the development of anisometropia axial length, no direct association between *PAX6* and myopia in the White population,[9] or high myopia in Chinese children has been observed.[10] So far, no study has systematically investigated the relationship between refractive error and *PAX6* variants.

To date, nearly 500 variants in *PAX6* have been reported, with the most common being the truncation variants (nonsense, frameshift, and splicing variants), accounting for 70% of identified variants.[3] The PAX6 haploinsufficiency, exclusively caused by truncation variants, gene deletion, and cis-regulatory variants, can account for most aniridia cases.[3,11,12] The missense variants tend to lead to atypical-aniridia phenotypes or non-aniridia phenotypes, such as microphthalmia, anophthalmia, and coloboma (MAC) or isolated foveal hypoplasia.[5,13] Besides, the location of missense variants seemed to correlate with some particular phenotypes, such as isolated foveal hypoplasia.[5,14] Compared to the truncation variants, the evaluation of pathogenicity in missense variants has been a long-standing puzzle. Recently, researchers presented the AlphaMissense, a novel missense variant pathogenicity prediction tool that utilizes the artificial intelligence system AlphalFold2 to elucidate the molecular effect of variants on protein function.[27] It may contribute to investigation of pathogenicity evaluation of missense variants in *PAX6*.

Based on previous studies,[15,16] the *PAX6* gene dosage effect and variant type may have influence on the severity of *PAX6*-associated phenotype. Furthermore, the parental mosaicism in *PAX6* has been suggested to explain the intrafamilial variabilities in some sporadic cases.[17,18] Given this multifaceted context, despite the high penetrance of *PAX6* variants, the phenotypic variability is commonly observed, posing a challenge to the establishment of genotype-phenotype correlation, and impeding timely clinical diagnosis and treatment. The study herein aims to investigate the phenotypic and genotypic characteristics, as well as the genotype-phenotype relationship, in a large Chinese cohort of molecularly confirmed patients with *PAX6*-related disorders, comprising 164 unrelated families (117 newly identified and 47 previously reported by our group).

Methods

Patients' Recruitment

This study was approved by the institutional review board of Zhongshan Ophthalmic Center and the study procedure was in accordance with the tenets of the Declaration of Helsinki. A total of 10,530 unrelated individuals with various eye conditions were recruited from the Pediatric and Genetic Clinic, Zhongshan Ophthalmic Center, Guangzhou. The unrelatedness of probands was determined based on their self-reported family histories, including pedigrees through their second-degree relatives and questions regarding any family members with ocular diseases. On the other hand, the issue of related families should be considered for samples with significant founder variants, however, this is unlikely to be a significant issue in our cohort because most of our patients are sporadic cases and only a small portion of families share a limited number of common variants. In addition, some of the shared variants likely occurred independently, as they are de novo in different families. Prior to the clinical data and peripheral venous blood sample collection, all participants provided signed informed consent. Genomic DNA was extracted from peripheral venous blood samples using previously described methods.[19]

All patients recruited in this study received routine ophthalmic examinations. After that, specific ophthalmic examinations were performed on patients, including anterior segment examination, fundus photography, optical coherence tomography (OCT), and scanning laser ophthalmoscopy. The grade of foveal hypoplasia was assessed using the Leicester Grading System for foveal hypoplasia.[20]

Variant Detection and Screening

The variants in *PAX6* were detected by exome sequencing, using whole-exome sequencing (WES) and/or targeted exome sequencing (TES), on the genomic DNA of 10,530 unrelated patients, as previously described.[21] The detected variants were sequentially screened using the strategy described in our previous study. The variants with low sequencing quality with coverage depth lower than 5 and high minor allele frequency (MAF) of more than 0.01 in the gnomAD database (https://gnomad.broadinstitute.org/) were excluded. The splicing effect of synonymous variants and variants located in the non-coding region was evaluated through several bioinformatic tools including the Human Splicing Finder program (HSF, https://hsf.genomnis.com/mutation/analysis), Splice AI (https://spliceailookup.broadinstitute.org/), MaxEntScan, NNSPLICE, and Splicing Sequence Finder from Alamut software. The missense variants were assessed through multiple in silico tools, including SIFT (http://sift.jcvi.org/), Polyphen-2 (http://genetics.bwh.harvard.edu/pph2/), PROVEAN (http://provean.jcvi.org/seq_submit.php), CADD (http://cadd.gs.washington.edu), and REVEL (https://sites.google.com/site/revelgenomics/). Additionally, missense variants underwent the evaluation of the Human Gene Mutation Database (HGMD) and the AlphaMissense Database. The MetaDome analysis (https://stuart.radboudumc.nl/metadome/) was performed to predict the genetic tolerance at the PAX6 protein residues. The potential pathogenic variants were further identified by the Sanger sequencing and co-segregation analysis.

Droplet Digital PCR

Droplet digital PCR (ddPCR) was used for quantification of the mutant allele fraction to investigate the parental mosaicism. The ddPCR was carried out in families meeting two criteria suggesting mosaicism or anticipation: (1) intrafamilial phenotypic variability, with milder phenotypes in parents, particularly in cases with two or more affected siblings; and (2) Sanger sequencing results showing a lower peak height of the affected allele in the paternal blood sample compared to those of heterozygous variant carriers. In this cohort, ddPCR was only performed in two families with nonsense variants. The primers and probes for the genotyping *PAX6* variant c.676G>T (p. Glu226*) and c.702T>A (p. Tyr234*) were designed by Primer Express software (Applied Biosystems) and generated by Thermos Fisher Scientific. Genomic DNA extracted from peripheral blood samples was analyzed using ddPCR with the ddPCR QX200 System (Bio-Rad, Hercules, CA, USA). According to the manufacturer's protocol (Bio-Rad), the 20 µL ddPCR mixture was composed of 2 × ddPCR Supermix for probes and primers (no UTP; Bio–Rad), target-specific PCR primers (final concentration 900 nM), FAM-labeled (final concentration 250 nM, wild-type–allele) probe, VIC-labeled (final concentration 250 nM, mutant-type allele) probe, and 20 ng of DNA sample. The droplet generation was performed using a QX200 Droplet Generator and the PCR cycling was run in a C1000 Touch Thermal Cycler (BioRad) using the following conditions: enzyme activation at 95°C for 10 minutes, the 40 cycles of denaturation at 94°C for 30 seconds, and

annealing/extension at 55°C for 1 minute, and the enzyme deactivation at 98°C for 10 minutes. After PCR amplification, the product was read using QX200 droplet reader (Bio-Rad) and analyzed using QuantaSoft Analysis Pro software.

Cloning Sequencing

The genomic DNA was extracted from the peripheral blood samples and was amplified by PCR reaction using primers (available on request). The PCR amplicons were cloned into pGEMT easy cloning vector (Promega Corporation, USA) following the manufacturer's instructions and transformed into the competent DH5α Escherichia coli (*E. coli*) cells (Takara). The T4 DNA ligase (NEB) was used to ligate the fragment and vector. The transformed cells were grown on LB agar plates supplemented with Ampicillin, IPTG, and X-gal. The positive clones were selected through the blue-white colony screening and sequenced using primers (available on request) on an ABI3100 Genetic Analyzer (Applied Biosystems).

Statistical Analysis and Visualization

The statistical analysis and visualization in the current study were carried out using IBM SPSS Statistics V26.0 software (IBM Corp., Armonk, NY, USA), open-source R software (version 4.2.1), and RStudio. The refractive error was defined according to the previous studies.[22–24] Myopia was defined as the spherical equivalent refraction (SER) ≤ -0.50 diopters (D), hyperopia was defined as the SER $\geq$ 0.50 D (or $\geq$ 2.00 D when the age less than 7 years old), emmetropia was defined as -0.50 D $<$ SER $<$ 0.50. The myopia classification included high myopia (SER ≤ -6.00 D), moderate myopia (-6.00 D $<$ SER ≤ -3.00 D), and mild myopia (-3.00 D $<$ SER ≤ -0.50 D). The different types of refractive error were color coded in patients with *PAX6* variants. Correspondingly, the variant locations were color-coded, including NTS, Link region between CTS and NTS, CTS, Link region between PD and HD domain, HD domain, and proline-serine-threonine-rich domain (PSTD). The hierarchical cluster analysis and heatmap generation, using Euclidean distance, were conducted with R packages and the ggplots library to better demonstrate and investigate the relationship between refractive error and the location of *PAX6* variants (details were described in Supplementary Table S1). When the Manhattan distance (sum of the absolute differences) was used as the distance metric, the result of the cluster analysis showed no significant difference compared to using Euclidean distance (root sum-of-squares of differences) as the distance metric (Supplementary Fig. S2). A Kruskal-Wallis test or Mann-Whitney *U* test was applied to compare best corrected visual acuity (BCVA) across different variant groups. A Bonferroni correction was applied for the pairwise multiple comparisons when a significant difference was detected. A Chi-Square test or Fisher's exact test depending on the number of samples and expected number of elements in each cell were used to analysis the categorical data, including the distribution of different ocular findings related to the *PAX6* variant, refraction types, and iris absence types across various variant types or locations. The generalized estimating equation (GEE) models was applied to account for inter-eye correlation. Adjusted standardized residuals analysis (Z-score) with a Bonferroni correction was performed as a post hoc test following the Chi-Square test. The value of adjusted residuals above 1.96 indicated a statistically significant positive correlation, while those below -1.96 represent a reverse relationship.[6,25,26] The statistical significance was assessed at a 0.05 significance level.

Results

Molecular Characteristics of Patients With Heterozygous *PAX6* Variants

A total of 119 different *PAX6* variants were identified in 228 patients of 164 unrelated pedigrees in our cohort, including 169 individuals of 117 pedigrees newly identified in this study and 59 individuals of 47 pedigrees previously reported by our group (Supplementary Table S2). The 119 variants were classified as missense variant (31/119, 26.1%), truncation variant (74/119, 62.2%, consisting of 20 nonsense, 18 splicing, and 36 frameshift), C-terminal extension variant (CTE, 2/119, 1.7%), in-frame insertion-deletion (indel) variants (2/119, 1.7%), synonymous variant (1/119, 0.8%), start-loss variant (1/119, 0.8%), and gene or exon deletion (8/119, 6.7%). The majority of these variants (114/119) was absent in the gnomAD database, whereas four of the remaining five had been previously reported and the other one was a newly identified variant. The allele frequency of the five variants presented in gnomAD was less than 1×10^{-5}. Furthermore, the missense variants in newly identified families were predicted as deleterious by five online bioinformatic tools, including SIFT, Polyphen-2, PROVEAN, CADD, and REVEL. The 18 splicing variants, two in-frame variants, and one synonymous variant were predicted to destroy wild-type acceptor or donor site (15/21) or introduce a new acceptor site or donor site (6/21) by at least two splicing prediction tools, including HSF, Alamut, and SpliceAI (see Supplementary Table S2). Except for the previously reported variants by our group, the variants in 117 newly identified families were further validated by Sanger sequencing and underwent co-segregation analysis (see Supplementary Fig. S1). Among these families, de novo variants had been identified in 19.7% (23/117) of the cases, and co-segregation had been observed in the remaining families.

In addition, all 31 missense variants in the current study were evaluated as pathogenic by the AlphaMissense.[27] In comparison, the 127 disease-causing missense (DM) variants in HGMD were evaluated as pathogenic (116, 91.3%), ambiguous (4, 3.1%), and benign (7, 5.6%) by the AlphaMissense, respectively. Meanwhile, the 13 likely disease-causing missense (DM?) variants in HGMD were classified as pathogenic variants (6, 46.2%) or benign (7, 53.8%) by the AlphaMissense (Supplementary Table S3, Fig. 1).

All patients were assigned to missense, truncation, in-frame indel, gene deletions, and CTE variant groups. The truncation variant group was subdivided into the splicing, frameshift, and nonsense variant subgroups. The newly identified synonymous variant c.249T>G (p.Val83Val) was predicted to introduce a new donor site by the in silico tools HSF and demonstrated to result in partial skipping of 108 nucleotides of exon 6 and in-frame protein deletion p.Lys86_Ser121del by a previous study.[28] Hence, it was classified into the in-frame variant group.

The General Clinical Characteristics of Patients in This Cohort

The 456 eyes of the 228 patients with *PAX6* variants (127 male patients and 101 female patients) in unrelated 164 fami-

Investigative Ophthalmology & Visual Science

Figure 1. The AlphaMissense evaluation for the *PAX6* missense variants in this cohort and in HGMD database. (**A**) The schematic above shows that the results of Alpha Missense prediction are highly consistence with the molecular confirmed findings in this cohort. The schematic below is the variation pathogenicity prediction at each *PAX6* protein residues using AlphaMissense analysis (NP_000271.1). The y-axis represents the prediction score of AlphaMissense. The color bar below is corresponding to the score of pathogenicity, whereas the *purple* indicates the pathogenicity, the *beige* indicates the benign, and the *white* indicates the ambiguous. The color bar above is also corresponding to the prediction score of AlphaMissense prediction results. The *circle pattern* represents the molecular confirmed missense variants in this cohort and the *filled color* represents the prediction results of AlphaMissense. (**B**) The diagram shows the results of AlphaMissense prediction in HGMD database. The *circle pattern* with the *white border* represents the likely disease-causing variants (DM?) in HGMD database. The *circle pattern* without the *white border* represents the disease-causing variants (DM). The *filled color* in the *circle pattern* indicates the prediction results of AlphaMissense.

lies were clinically evaluated in this study (see the Table, Supplementary Table S2). The mean age of all patients at the first examination was 12.19 ± 13.20 years (median = 6.00 years, range = 0.10–62.00 years). Among these patients, the mean BCVA was 1.14 ± 0.92 logMAR (median = 1.00, range = 0.00–4.00) in the right eye (OD) and 1.13 ± 0.84 logMAR (median = 1.00, range = 0.10–4.00) in the left eye (OS). In addition, based on the available clinical data, complete or partial aniridia was observed in 76.7% of eyes (342/446), whereas iris hypoplasia was observed in 22.4% of eyes (100/446). The iris data was lacked in remaining 4 eyes (4/446). Meanwhile, foveal hypoplasia and nystagmus had been found in 97.8% (348/356) and 92.6% (403/435) of eyes. OCT scanning results were available in 92 eyes of 46 patients. According to the Leicester Grading System, foveal hypoplasia was classified as grade 4 (40/92, 43.5%), grade 3 (32/92, 34.8%), and grade 1 (18/92, 19.6%), respectively. In the one remaining patient, bilateral eyes exhibited a normal foveal structure. Among 42 eyes of 21 individuals with iris hypoplasia phenotype (complete iris with subtle structure changes), 32 eyes of 16 individuals (76.2%) were found to have grade 4 or grade 3 foveal hypoplasia.

The Correlation Between Variant Groups and Iris Defect Degree

Iris abnormalities were identified in 442 eyes of 221 patients with *PAX6* variants. Of them, a detailed description or photographs of the iris were available in 362 eyes of 182 patients for following analysis. The complete aniridia was observed in 58.0% (210/362) of eyes, whereas the partial aniridia, including nasal iris absence, temporal iris absence, superior iris absence, and inferior iris absence, was seen in 8.0% (29/362), 2.8% (10/362), 2.5% (9/362), and 1.1% (4/362) of eyes, respectively. Additionally, the iris hypoplasia was present in 27.6% (100/362) of eyes (Fig. 2, Supplementary Fig. S3). The anterior segment morphology of the inferior iris

Investigative Ophthalmology & Visual Science

Table. Demographic Characteristics and Baseline Data of 228 Patients With *PAX6* Variants

Characteristic	*n* (%)	Mean ± SD [Median (Range)]
Sex		
Male	126 (55.2)	
Female	102 (44.8)	
Age at examination, y	220	12.19 ± 13.20 [6.00 (0.10 to 62.00)]
VA (LogMAR)		
OD	109	1.14 ± 0.92 [1.00 (0.00 to 4.00)]
OS	113	1.13 ± 0.84 [1.00 (0.10 to 4.00)]
Refractive error (diopters)		
OD	71	−2.73 ± 8.27 [−1.25 (−29.00 to 11.50)]
OS	73	−3.56 ± 8.78 [−1.50 (−30.00 to 12.50)]
Axial length, mm		
OD	30	22.22 ± 2.92 [22.70 (16.00 to 26.80)]
OS	31	22.06 ± 3.13 [21.97 (15.51 to 28.68)]
Clinical findings		
Aniridia	342/446 (76.7)	
Iris hypoplasia	100/446 (22.4)	
Foveal hypoplasia	348/356 (97.8)	
Nystagmus	403/435 (92.6)	
Cataract	168/456 (36.8)	
Microcornea	82/456 (18.0)	
Glaucoma	17/456 (3.7)	
MAC	22/456 (4.8)	
Aniridia related keratopathy	52/456 (11.4)	
Optic disc hypoplasia	18/456 (3.9)	
Retinal detachment	5/456 (1.1)	
Foveal hypoplasia grade*		
Normal	2	
Grade 1	18	
Grade 2	0	
Grade 3	32	
Grade 4	40	

All clinical examination results were based on available clinical data.

* The grade of foveal hypoplasia evaluation was according to the Leicester Grading System for foveal hypoplasia (Thomas MG, Kumar A, Mohammad S, et al. Ophthalmology 2011;118:1653–1660).

MAC, microphthalmia, anophthalmia, and coloboma.

absence showed a unique sectorial defect, differing from the keyhole shape observed in iris coloboma due to variants in genes other than *PAX6*.

There was a statistical significance association between different variant subgroups and iris defect degrees (Fisher's Exact Test, $P < 0.001$; GEE model, $P < 0.001$; Supplementary Table S4-1). In the complete aniridia group (210 eyes), the contribution of the truncation variant (172/210, 81.9%) was more than missense (5/210, 2.4%), CTE (14/210, 6.7%), in-frame indel variants (3/210, 1.4%), and gene deletions (16/210, 7.6%) subgroups. The gene deletions (Z-score = 2.7), frameshift (Z-score = 3.7), and nonsense variants (Z-score = 6.9) tend to lead to complete iris absence, as opposed to missense (Z-score = −11.0, P = 3.56E-04, P = 1.26E-20, and P = 4.13E-28, respectively), CTE (Z-score = −2.4, P = 1.00E-03, P = 9.80E-05, and P = 2.86E-09), and in-frame indel variants (Z-score = −1.8, P = 6.00E-03, P = 1.30E-02, and P = 1.11E-05; see Fig. 2, Supplementary Table S4-3, Bonferroni corrected P value (Pc) threshold = 2.38E-03). Conversely, in the iris hypoplasia group (100 eyes), the missense (70/100, 70%) and CTE variant (18/100, 18%) subgroups accounted for a greater proportion compared with other variant subgroups. Compared to the gene deletions, splicing, frameshift, and nonsense variants, the missense variants tended to result in iris hypoplasia (Z-score = 13.1, P = 4.69E-12, P = 3.17E-19, P = 4.95E-22, and P = 3.99E-29, respectively), followed by the CTE variants (Z-score = 3.2, P = 2.39E-04, P = 2.00E-06, P = 8.08E-08, and P = 4.30E-13), and in-frame indel variants (Z-score = 0.9, P = 2.00E-02, P = 1.25E-04, P = 7.00E-03, and P = 5.40E-05; see Supplementary Table S4-4, Pc threshold = 2.38E-03).

Similarly, a significant correlation was also observed between different variant location subgroups and the degree of iris defects (Chi-square test, P = 1.09E-12 and GEE model, $P < 0.001$; Supplementary Table S5-1). Variants in the linker region between the two subdomains of PD tended to cause iris hypoplasia (Z score = 7.23, P = 4.80E-13) compared to other variant location subgroups (Pc threshold = 1.67E-03). The variants in the HD region (Z-score = −2.97, P = 2.98E-03) showed a comparatively lower contribution. Specifically for missense variants, those located in the CTS-NTS linker region (Z score = 2.3) and CTS region (Z-score = 1.5) were more prevalent than in the PD-HD linker (Z-score = −2.3) and PSTD region (Z-score = −2.6; see Supplementary Table S5-3). Conversely, the contribution to all iris absence of variants in HD or PSTD was greater than other subgroups (Z score = 3.09, P = 2.00E-03 and Z-score = 3.31, P = 9.33E-04, respectively), whereas the contribution of variants in Linker region between CTS and NTS was less than other subgroups (Z score = −5.24, P = 1.61E-07). In the partial aniridia group, no clear difference was observed among the contribution of the variant types, but statistical significance was observed

Investigative Ophthalmology & Visual Science

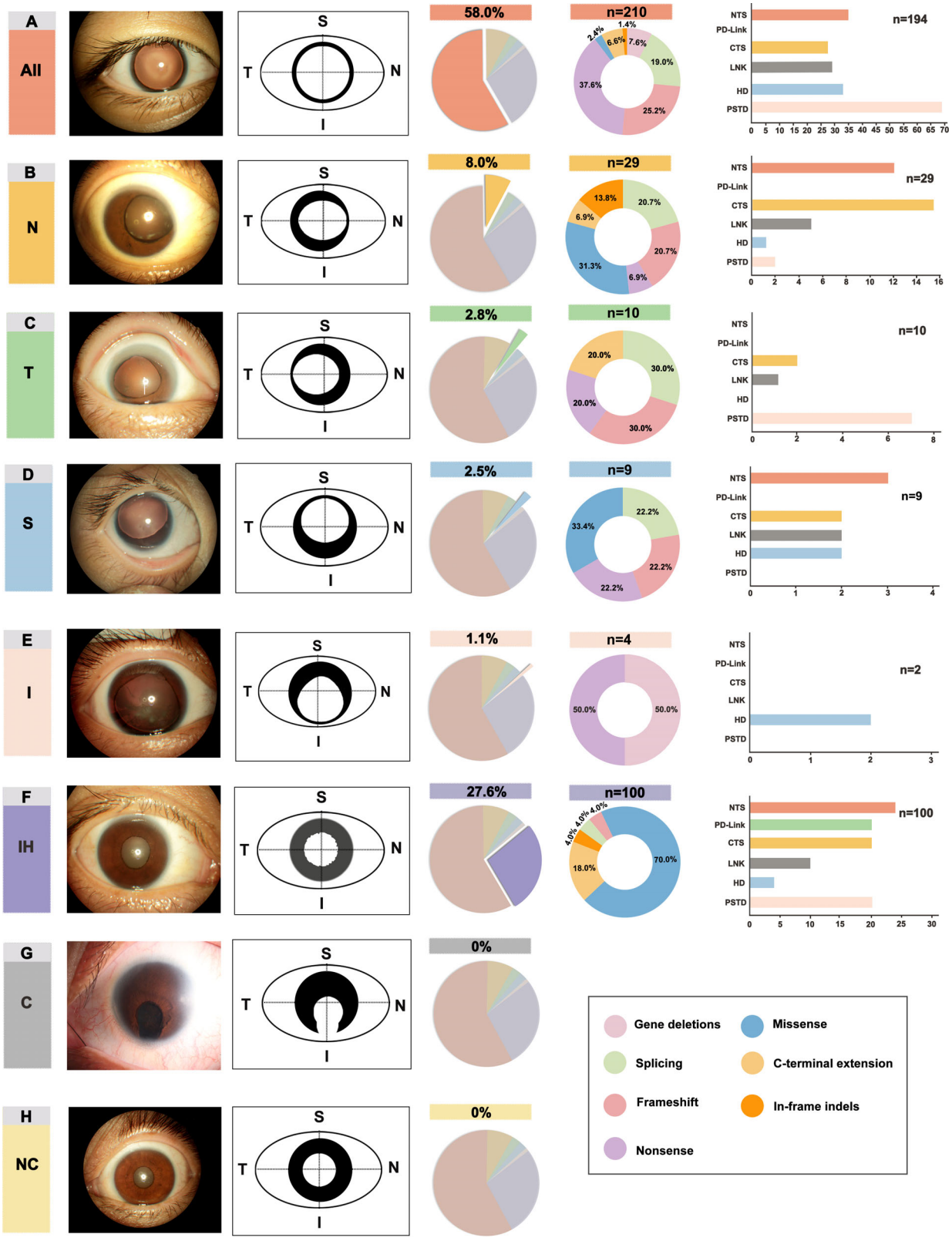

Figure 2. The distribution of the different iris defect degrees and the correlation between the iris defect degrees and different variant subgroups in our in-house data. (**A–F**) The representative anterior segment images of different iris defect subgroups and corresponding schematic diagrams are shown in the *second* and *third columns* on the *left*, respectively. The *fourth column* shows the distribution of different iris defect degrees in this cohort, with colors representing iris defect subgroups: all iris absence (*red*, 58.0%), nasal iris absence (*yellow*, 8.0%), temporal iris absence (*green*, 2.8%), superior iris absence (*blue*, 2.5%), inferior iris absence (*pink*, 1.1%), and iris hypoplasia (*purple*, 27.6%). The distribution of different variant types among iris defect subgroups is shown in the *fifth column*. In the all iris absence subgroup (**A**), the proportion of truncation variants including splicing (19.0%), frameshift (25.2%), and nonsense (37.6%) is higher than

Investigative Ophthalmology & Visual Science

the missense (2.4%), CTE (6.6%), in-frame indel variants (1.4%), and gene deletions (7.6%) subgroups. In the iris hypoplasia subgroup (**F**), missense (70%) and CTE variants (18%) subgroups are accounted for a greater proportion compared with in-frame indel (4%), the splicing (4%), and the gene deletions variants (4%). The *rightmost column* shows the distribution of iris defect among different variant locations. In the all iris absence subgroup (**A**), the proportion of HD ($n = 33$) and PSTD ($n = 69$) domain was higher than other subgroups. In the iris hypoplasia subgroup (**F**), the contribution of paired domain, including NTS ($n = 26$), linker region between CTS and NTS ($n = 20$), and CTS ($n = 20$), was higher than other subgroups. (**G**) The representative anterior segment image of iris coloboma in the in-house database exhibited the keyhole pupil in the inferior nasal quadrant, which is different from the unique sectorial defect or water droplet-like defect caused by the inferior iris absence in the aniridia (**E**). (**H**) The representee anterior segment photography and corresponding schematic diagrams of the normal control. All, all iris defect; C, iris coloboma; CTS, C-terminal subdomain; HD, homeodomain; I, inferior iris absence; IH, iris hypoplasia; LNK, linker region between paired domain and homeodomain; N, nasal iris absence; NC, normal control; NTS, N-terminal subdomain; PD-link, linker region between CTS and NTS of the paired domain; PSTD, proline-serine-threonine-rich domain; S, superior iris absence; T, temporal iris absence.

among the contribution of variant locations (Fisher's exact test, $P = 0.13$, and $P = 0.017$ respectively; see Fig. 2, Supplementary Table S4-2, Supplementary Table S5-2).

The Parental Mosaicism in Patients With *PAX6* Variants

In two families, probands (F1-II:1 and F2-II:1) exhibited low visual acuity, typical aniridia, and nystagmus. The grades 3 to 4 foveal hypoplasia has been observed in these probands. Two nonsense variants in *PAX6*, c.676G>T (p. Glu266*) and c.702T>A (p. Tyr234*), were identified in these two probands (F1-II:1 and F2-II:1) by Sanger sequencing, respectively. The mother of family F1 (F1-I:2) and the father of family F2 (F2-I:1) both showed milder clinical phenotypes compared to their children, including rudimentary iris remnant, better visual acuity (range from 0.32 to 0.5), and lower grade foveal hypoplasia (grade 1; see Supplementary Table S2, Fig. 3A).

In the other family F3, the proband F3-II:2 showed iris hypoplasia, and his affected brother F3-II:1 had partial aniridia. The frameshift variant in *PAX6*, c.112del (p. Arg38Glyfs*16), was identified in these two patients by Sanger sequencing. Both patients had nystagmus, high-grade foveal hypoplasia, and poor visual acuity. Surprisingly, the parents in this family showed no signs of iris anomalies, anterior segment dysgenesis, or other abnormalities typically linked to aniridia, such as cataracts, glaucoma, and keratopathy. Moreover, neither nystagmus nor foveal hypoplasia was observed in the parents.

The parental mosaicism or the intrafamilial phenotypic variability may explain the phenotype difference between parents and children. Notably, the Sanger sequencing results of paternal blood samples (F1-I:2, F2-I:1, and F3-I:1) and sperm sample (F3-I:1) showed that lower peak height of the affected allele compared to the children with the corresponding heterozygous variant (see Figs. 3A1, 3B1, Supplementary Fig. S1), indicating the possibility of parental mosaicism in these cases.

For further investigation of mosaicism, ddPCR was specifically performed in two families with nonsense *PAX6* variants (F1 and F2). DNA from blood samples of the two probands (F1-II:1 and F2-II:1) and their family members (F1-I:2 and F2-I:1) was used to accurately assess the mutant allele fraction. The mutant allele fraction of the blood samples in 2 affected individuals (F1-I:2 and F2-I:1) was 14.9% and 13.9%, respectively (see Fig. 3). Unfortunately, no sperm sample was available from the affected father (F2-I:1). Due to technological constraints, cloning sequencing was used instead of ddPCR to quantify the mutant allele fraction in family F3, using a sperm sample from the proband F3-II:1's father. The gonosomal fraction had been confirmed in the father (F3-I:1) of the family F3 with frameshift variant c.112del (p. Arg38Glyfs*16), and the affected allele fraction of the sperm sample was 18.8%.

The Relationship Between Variant Groups and Refractive Error as Well as BCVA

The refractive error among this cohort was classified into four categories, including hyperopia, emmetropia, low or moderate myopia, and high myopia. The distribution of four categories was as follows: hyperopia (34.9%), emmetropia (17.5%), low or moderate myopia (24.3%), and high myopia (23.3%; Fig. 4). The distribution of different refractive error subgroups among truncation, missense, and CTE variant groups was significantly different (Fisher's exact test, $P =$ 1.46E-04; GEE model, $P = 0.007$; Supplementary Table S6-1). The truncation variants group showed increased hyperopia compared to the missense variants group (Fisher's exact test, $P =$ 1.20E-03), whereas the missense variants group was associated with high myopia (Fisher's exact test, $P =$ 2.13E-04). Additionally, the CTE variants group exhibited a tendency toward emmetropia compared to missense (Fisher's exact test, $P =$ 4.1E-03) and truncation variants (Fisher's exact test, $P =$ 4.5E-03). The cluster analysis has identified a significant difference in the distribution of refractive errors among different variant location. The hyperopia was highly related to variants in the NTS, HD, and PSTD regions. High myopia showed a strong correlation with the CTS-NTS linker region. Emmetropia also had strong relationship with CTS region (Fig. 4A). The remaining one subgroup was distributed more evenly among all regions.

Of the 90 myopic eyes, including 44 high myopic eyes and 46 low/moderate myopic eyes, the missense variant c.622C>T (p. Arg208Trp) could account for 10 myopic eyes (11.1%, 10/90), comprising 5 high myopic eyes and 5 low/moderate myopic eyes (Fig. 4B). Additionally, the missense variant c.233T>G (p. Val78Gly) was associated with 6 extremely highly myopic eyes, whereas 2 other missense variants near the amino acid location c.229A>G (p.Arg77Gly), c.235G>C (p.Ala79Pro) were observed in 2 highly myopic eyes and 2 low/moderate myopic eyes, respectively. Similarly, the variant c.295G>C (p.Ala99Pro) was linked to 4 extremely highly myopic eyes, and another missense variant c.296C>T (p.Ala99Val) affecting the same amino acid location was found in 4 low/moderate myopic eyes.

Based on the available clinical data of 113 patients with *PAX6* variants, the BCVA for the majority of people (84/113, 74.3%) was below 0.52 logMAR (Snellen equivalent of 0.30, low vision). Among three different variant groups, the age

Investigative Ophthalmology & Visual Science

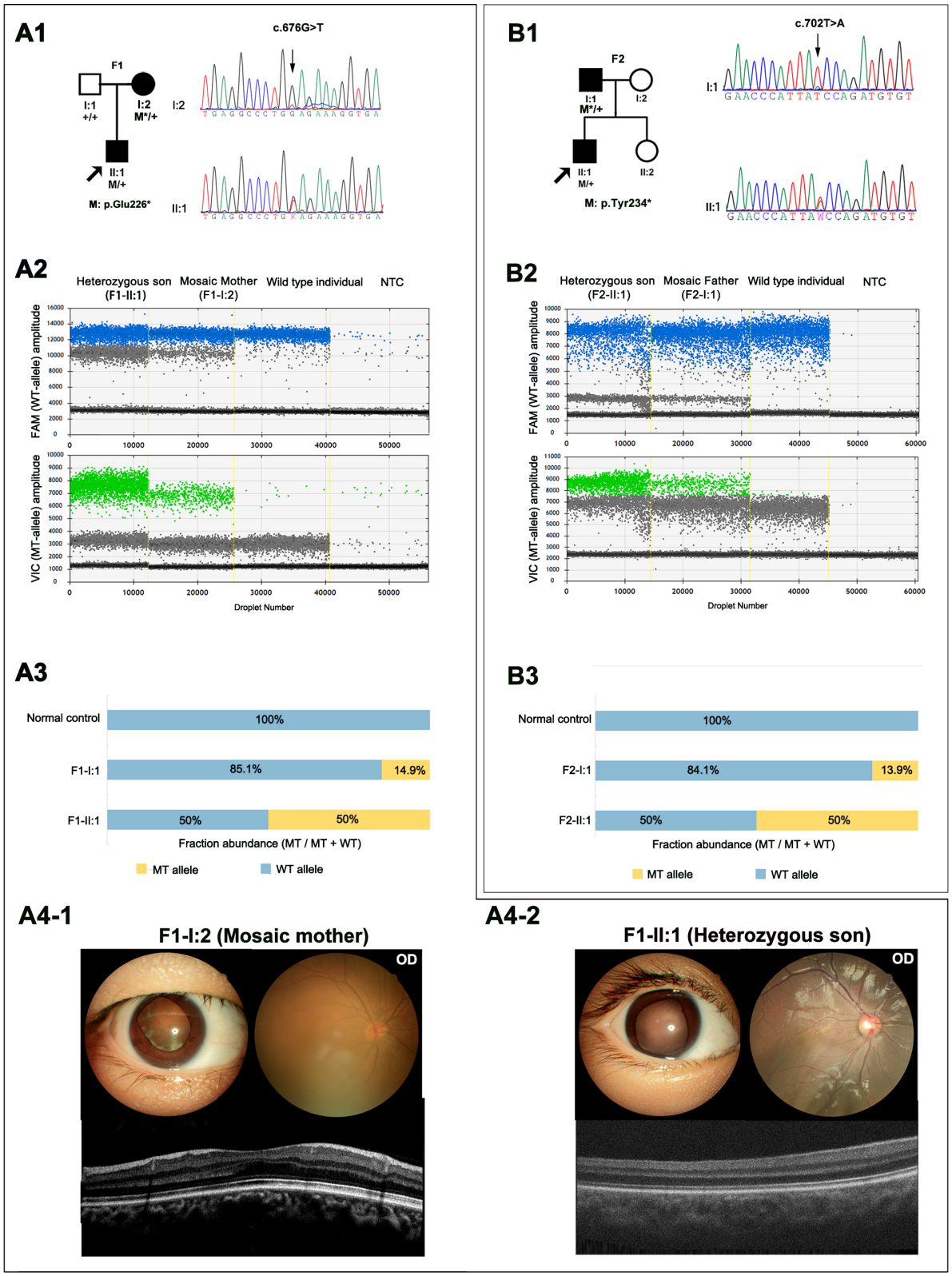

Figure 3. Pedigrees, DNA sequencing, mosaicism, and clinical analysis of two families with *PAX6* variants. The family diagrams and sequencing chromatograms of family F1 with the mosaic and heterozygous *PAX6* variants, c. 676G>T (p.Glu266*) (**A1**) and family F2 with the mosaic and heterozygous variant, c.702T>A (p. Tyr234*) (**B1**) are shown. *PAX6* variants numbering is based on the NCBI reference sequence NM_000280.5. Female patients are represented by *circles*, and male patients by *squares*. The *black pattern* indicates aniridia. M indicates mutant allele, + indicates the normal allele, and * indicates the mosaic alleles. The *black arrows* point to probands. The fluorescence amplitude plots of the family F1 (**A2**) and family F2 (**B2**) are shown. The VIC-labeled droplets in *green* show the presence of the mutant allele in the heterozygous carriers (II:2) and (F2-II:1), the mosaic mother (F1-I:2) or mosaic father (F2-I:1), and absence in the wild-type individual, and the negative control (no template control, NTC). The FAM-labeled droplets in *blue* represented wildtype allele, and were detected in all individuals. The negative droplets are represented in *grey*. The quantification of allele abundance by digital droplet PCR

Investigative Ophthalmology & Visual Science

(ddPCR) for the family F1 (**A3**) and the family F2 (**B3**). The mutant allele percentages of heterozygous carriers (F1-II:1 and F2-II:1), the mosaic variant carriers (F1-I:2 and F2-I:1), and the normal control are shown in horizontal bar plots. (**A4**) The anterior segment photography, the fundus images, and optical coherence tomography scans of the son with heterozygous *PAX6 v*ariant and the mother with mosaic *PAX6* variant in the family F1.

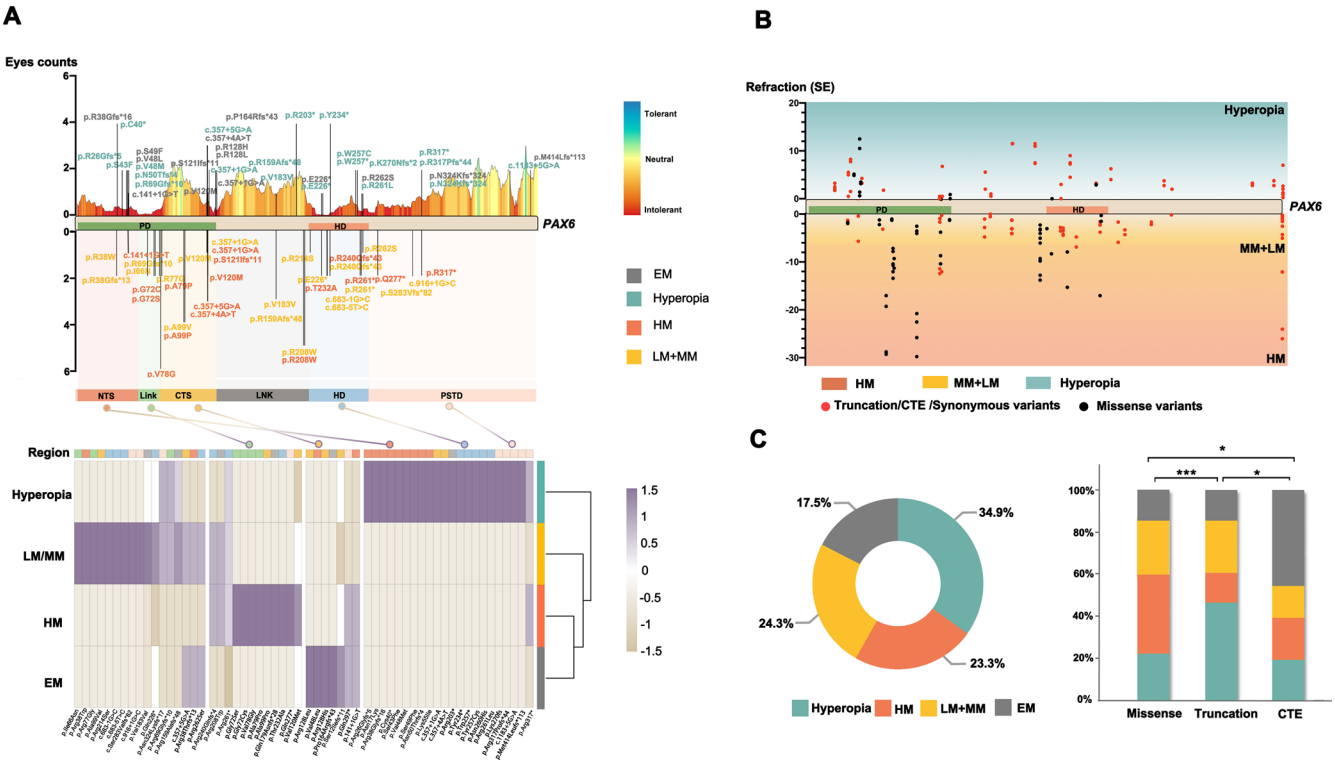

Figure 4. The computational clustering and distribution of refractive error in this cohort with *PAX6* variants. (**A**) The diagram above illustrates the frequency and location of *PAX6* variants associated with refractive error in our cohort. In this diagram, the variants associated with hyperopia (*green*) and emmetropia (*grey*) are shown *above* and variants associated with moderate or low myopia (*yellow*) and high myopia (*red*) are shown *below*. Above this diagram, the predicted tolerance landscape of *PAX6* by MetaDome analysis is shown (red = intolerant; yellow = neutral; and blue = tolerant). Below this diagram, the schematic of structure of functional domains of the *PAX6* isoform (NM_000280.5). The graph below shows the quantitative and clustering analysis of refractive error in this cohort. The different refractive error subgroups are listed in the y-axis and different variant locations are listed in the x-axis. The *PAX6* variant associated refractive error subgroups were clustered according to the variant locations using R packages cluster and gplots, and the function heatmap.2. The functional domains are shown in color bar: the N-terminal subdomain of the paired domain (NTS, *red*), the linker region between CTS and NTS (link, *green*), the C-terminal subdomain of the paired domain (CTS, *yellow*), the linker region between paired domain and homeodomain (LNK, *grey*), the homeodomain (HD, *blue*), and the proline-serine-threonine-rich domain (PSTD, *pink*). The Heatmap colors are corresponding to the trait-Z score, as the frequency histogram showing *purple* represents the high values and *beige* represents the low values for each trait. The clustering analysis demonstrates significant differences in the distribution of various refractive error subgroups across different functional domains. (**B**) The distribution of refractive error severity across different variant groups in this cohort. The y-axis is corresponding to the degrees of refractive error. The *red circle pattern* indicates the truncation, CTE, or synonymous variant. The *black circle pattern* indicates the missense variant. (**C**) The proportions of four refractive error subgroups in this cohort and their distribution across the different variant types. * $P < 0.05$, *** $P < 0.001$.

distribution was not statistically different (Kruskal-Wallis Test, $P = 0.29$). The BCVA of the three variant groups showed no difference in distribution with the increasing age (Fig. 5A). However, compared to the truncation variants (mean BCVA ± SD = 1.10 ± 0.84), the BCVA of the missense variants (mean BCVA ± SD = 0.75 ± 0.37) was better (Mann-Whitney *U* test, $P = 0.033$; Fig. 5B).

Associated Ocular Comorbidities in Patients With *PAX6* Variants

The distribution of associated ocular comorbidities in 456 eyes of 228 patients in this cohort showed cataract as the most common (168/456, 36.8%), followed by microcornea (82/456, 18.0%) and ARK (52/456, 11.4%). Subsequently, other findings, such as MAC, glaucoma, optic disc hypoplasia, and retinal detachment, were observed (see the Table, Figs. 6, 7). Detailed information on the distribution was provided in the Table. Regarding the correlation between the different variant subgroups and associated ocular findings, there was statistical significance in the distribution of variant subgroups among various ocular-associated findings (Fisher's exact test, $P = 1.00E-04$; Supplementary Table S7-1). In the cataract group (Fisher's exact test, $P = 5.90E-02$ and GEE model, $P = 3.48E-01$), the missense (Z-score = −3.08) variants were negatively associated with cataracts when compared to splicing variants (Z-score = 1.37, $P = 1.00E-03$, Pc threshold = 2.38E-03)

Investigative Ophthalmology & Visual Science

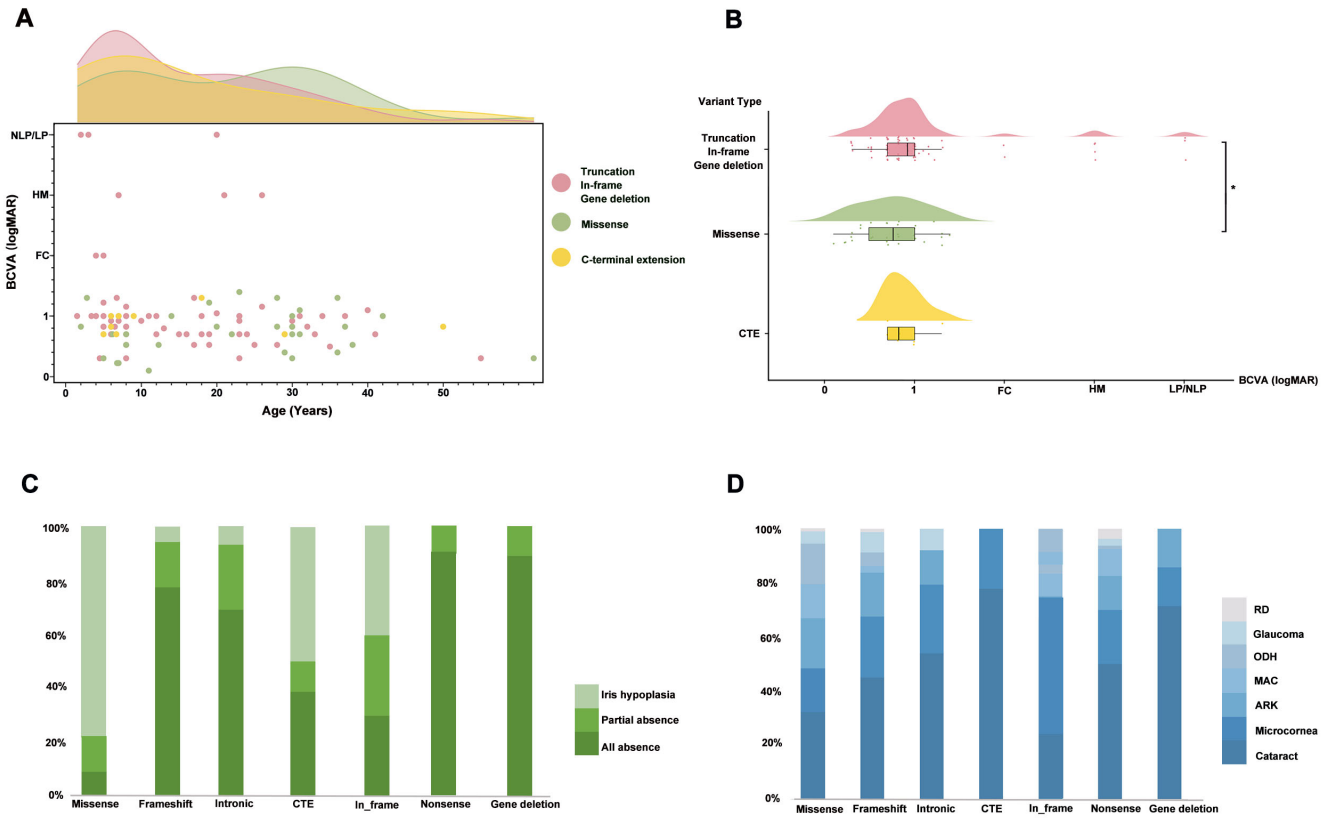

Figure 5. The distribution and comparison of best corrected visual acuity (BCVA), *PAX6* associated ocular comorbidity, and iris defect degree across different variant subgroups. (**A**) The variation of BCVA with age for three variant subgroups. The *red circle pattern* indicates truncation, in-frame indels, or gene deletions variants. The *green circle pattern* indicates the missense variants. The *yellow* circle pattern indicates the C-terminal extension variants. (**B**) The comparison of BCVA across three variant subgroups. * $P < 0.05$. (**C**) The distribution of different iris defect degrees among seven variants subgroups. (**D**) The distribution of seven ocular comorbidities across seven variant subgroups. ARK, aniridia-related keratopathy; FC, finger counting; HM, hand movement; LP, light perception; MAC, microphthalmia, anophthalmia, and coloboma; NLP, no light perception; ODH, optic disc hypoplasia; RD, retinal detachment.

subgroup (see Supplementary Table S7-3). Additionally, in the MAC group (Fisher's exact test, $P = 1.63E-03$), the missense variant group (Z-score = 3.14) showed a tendency toward the MAC, whereas the splicing (Z-score = −2.21) variant group did not exhibit a preposition toward MAC. In the optic disc hypoplasia group (Fisher's exact test, $P = 1.00E-03$ and GEE model, $P = 3.00E-02$), the missense variants (Z-score = 3.97) group tended to be associated to optic disc hypoplasia compared to the nonsense variants group (Z-score = −1.77, $P = 2.36E-03$, Pc threshold = 2.38E-03) (see Supplementary Table S7-6). The in-frame (Z-score = 3.33) tended to be associated with microcornea when compared to other variants groups (see Supplementary Table S7-4). There was no statistical significance in the other three groups.

In terms of anterior segment changes, aside from the different degrees of lens opacification, the ARK predominantly consisted of corneal opacification and fibrovascular pannus in this cohort (see Figs. 6A–H). For the fundus findings, except for the common foveal hypoplasia, the optic disc hypoplasia, including tilted optic disc, morphology abnormalities, and optic disc coloboma, was also observed in 18 eyes of 10 patients (18/456, 3.9%) in this cohort (see Figs. 6, 7). Based on the available clinical data of newly identified families with *PAX6* variants, the tessellated fundus changes below the optic disc have been identified in 11 patients (see Figs. 7I–L).

Discussion

In this large Chinese cohort, we have performed clinical and genetic evaluation of 228 patients from 164 families with *PAX6* variants, 117 of which are newly identified and 47 were previously reported. Paternal mosaicism of *PAX6* variants has been identified in 3 of the 164 families, in which phenotypic differences have been observed between carrier parents and affected children. Additionally, the genotype-phenotype analyses suggest that types and locations of the variants are associated with different types of refractive errors, various degrees of iris defects, and other associated signs and symptoms. It is interesting that foveal hypoplasia and nystagmus, rather than iris defects, are the most common initial signs in this cohort of patients with *PAX6* variants. Severe blindness complications, such as glaucoma and keratopathy, frequently occur in patients with *PAX6* variants, in addition to common symptoms like low vision, nystagmus, and photophobia. Therefore, recognition of atypical phenotypic variations beyond typical aniridia is very important. The insights gained from this study hold significant value for clinical diagnosis, treatment strategies in *PAX6*-related disorders, as well as for genetic counseling for family members. To our knowledge, this is the first study to explore the relationship between refraction and *PAX6* coding variants systematically.

Investigative Ophthalmology & Visual Science

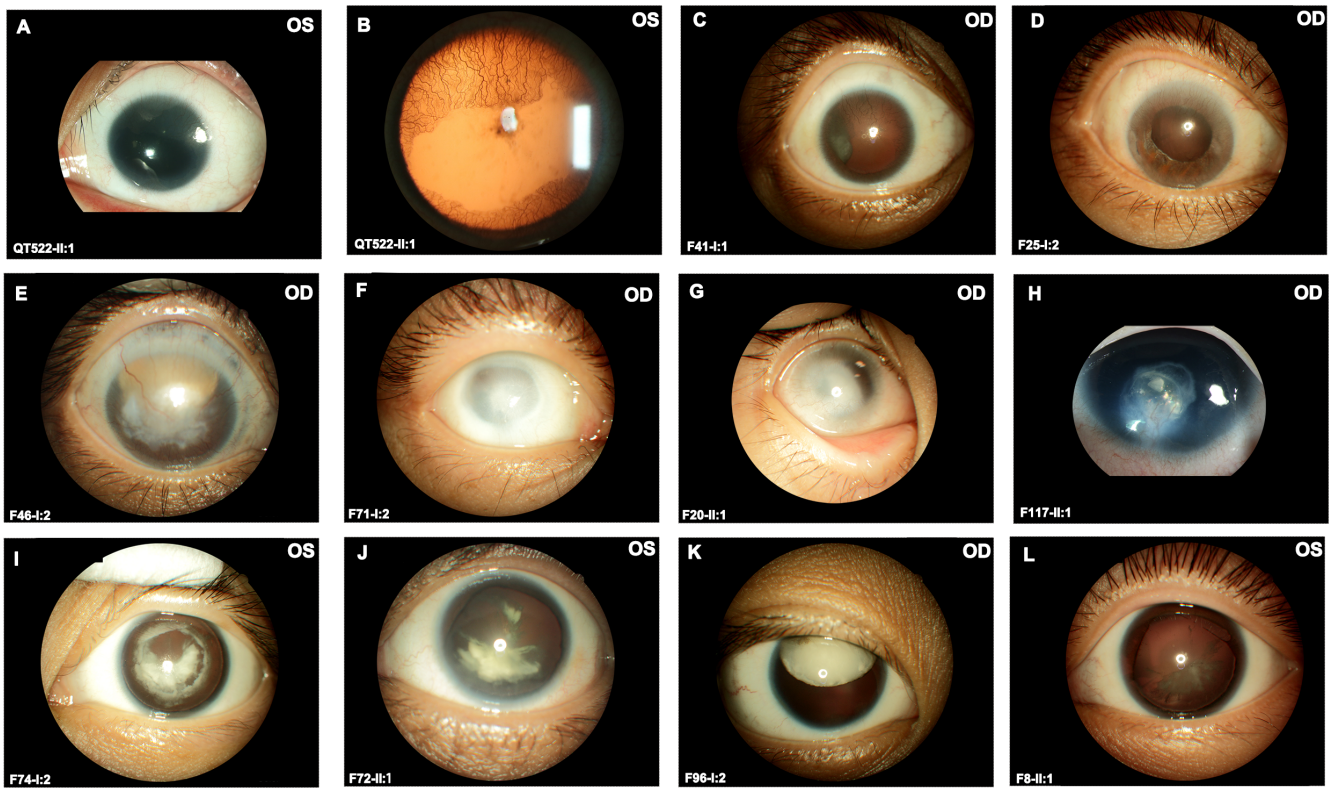

Figure 6. The anterior segment photography of patients with *PAX6* variants in this cohort. (**A–D**) The anterior segment images of three patients with cornea pannus. The conjunctival tissue with vessels invading the peripheral and mid-peripheral cornea was observed in patient QT522-II:1 (**A, B**) and patient F41-I:1 (**C**). The complete translucent and vascularized peripheral cornea was observed in the patient F25-I:2 (**D**). (**E–H**). The anterior segment images of four patients with varying degrees of corneal opacity and corneal vascularization. (**I–L**) The anterior segment images of four patients with different degrees of cataract.

Clarifying the pathogenicity of *PAX6* missense variants is still a challenge in the genetic diagnosis although genotype-phenotype correlation has been discussed before,[3,29] especially for sporadic cases. In this cohort, missense variants have been found to account for about one-fourth of the pathogenic or likely pathogenic variants in *PAX6*. Recently, the AlphaMissense database has been identified in hematological malignancies.[30] Here, all missense *PAX6* variants identified in our cohort have been predicted to be pathogenic by the AlphaMissense, showing a relative stringent classification of disease-associated variants comparing to those *PAX6* variants listed in HGMD,[31] in which about 91.3% disease-causing variants are predicted to be pathogenic, whereas only about 46.2% likely disease-causing variants are predicted to be pathogenic by the AlphaMissense. It suggests that the AlphaMissense can be an alternative complementary pathogenicity evaluation tool for *PAX6* missense variants.

The prevalence of sporadic cases is about 40% in the *PAX6* mutation database,[3] but the sporadic occurrence of newly identified cases in this cohort is 55.6% (65/117 families), with verified de novo mutations in 19.7% (23/117). Parental mosaicism has long been suspected to be one of the potential explanations for sporadic cases with *PAX6* variants,[32,33] and has been recently reported in several sporadic cases with *PAX6*-related disorders including aniridia, microphthalmia,[18] and isolated foveal hypoplasia.[17] Here, parental postzygotic mosaicism has been confirmed in three families with estimated affected allele fractions of 13.9% to 18.8% in sperm or blood tissue. As previously reported, the variable expressivity of phenotype in *PAX6*-related disorders may be partially explained by parental mosaicism.[34] The phenotype of the 3 persons with mosaic variants is milder than that of affected family members with heterozygous variants, expressing primarily in the following 3 aspects: (1) grade 0 to grade 1 foveal hypoplasia, (2) normal to partial iris defect, and (3) no signs of nystagmus. This may be attributed to the *PAX6* dosage effect.[35,36] Notably, despite an affected allele frequency of 18.9%, the mosaic individual (F3-I:1) displayed insignificant ocular abnormalities, whereas those with lower-level mosaicism (14.9% and 13.9%, respectively, in F1-I:2 and F2-I:1) exhibited partial aniridia and milder foveal hypoplasia compared to their heterozygous children. Hence, the variable expressivity of phenotype may not only be influenced by the *PAX6* dosage effects in various tissues. Other factors, such as epigenetic control mechanism,[37] environmental modifier, or perhaps spatiotemporal expression pattern of *PAX6*,[36] which are difficult to be verified in our cohort as most of our patients are sporadic. The above findings suggest that clinicians should contemplate the possibility of paternal mosaicism in sporadic *PAX6* cases, especially in cases with 2 or more affected siblings with "healthy" parents. The ocular manifestations of the parents in these cases warrant careful evaluation of milder atypical changes in the iris and/or fovea. Recognizing these cases holds clinical significance for the genetic counseling of their siblings or offspring.

Investigative Ophthalmology & Visual Science

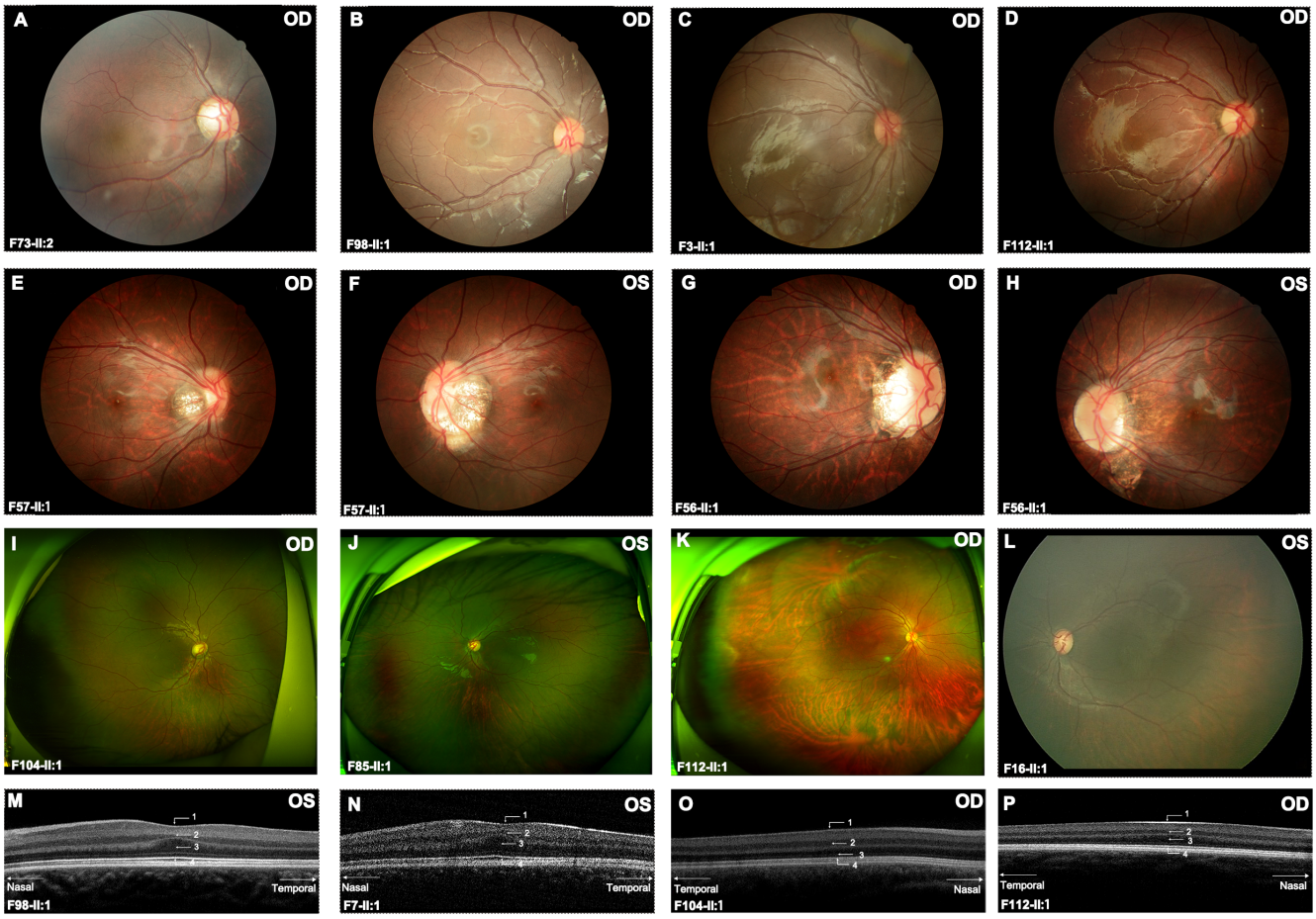

Figure 7. The fundus images and optical coherence tomography scans of patients with *PAX6* variants in this cohort. (**A–D**) The posterior fundus photographs of the right eyes of four patients reveals the different grades of foveal hypoplasia. (**E–H**) The posterior fundus photographs of four patients shows the optic nerve hypoplasia, including tilted disc, optic disc coloboma, and optic disc pit. (**I–L**) The ultra-widefield fundus and RetCam images of four patients show the tessellated fundus changes beneath the optic disc. (**M–P**) The OCT results of four patients show the different grade of foveal hypoplasia, including grade 1 (**M, N**), grade 3 (**O**), and grade 4 (**P**).

Aniridia has long been considered to be the major clinical feature in *PAX6*-related disorders.[38] However, some studies found that foveal hypoplasia without obvious iris abnormalities was common in individuals harboring *PAX6* variants.[4,5] Correspondingly, the current study shows a greater incidence of foveal hypoplasia (97.8%) than aniridia (76.7%) in our large Chinese cohort, suggesting that foveal hypoplasia rather than aniridia may be the most common clinical indication of *PAX6*-related disorders. In addition, partial aniridia was observed in 14.3% (52/362) of patients in this study. Partial aniridia, especially absence of the inferior iris, is easily confused with the iris coloboma.[38] Iris coloboma is known as a developmental abnormality typically located in the inferior nasal quadrant, resulting in a keyhole pupil (see Fig. 2G). We found the shape of the inferior iris absence in aniridia exhibited a unique sectorial defect or water droplet-like defect (see Fig. 2E), differing from the morphology of iris coloboma, suggesting that it can be used as the clinical sign for differential diagnosis.

Furthermore, the findings in this study suggest that the degree of the iris defect could be related to the type and location of the *PAX6* variant. Missense variants located in the NTS-CTS linker and CTS region have a tendency to result in iris hypoplasia, whereas truncation variants are usually associated with complete aniridia, consistent with previous studies.[5,13,39,40] The mechanism behind missense variants resulting in non-aniridia phenotypes remains unclear, but it may relate to perturbation of the DNA-binding activity and alteration of transcription activity,[5,13,39,41,42] and variants in the NTS-CTS linker region have been reported to lead to a non-aniridia phenotype through altering the flexibility in the linker region and affecting DNA binding ability.[39] Additionally, missense variants in the CTS region were predominantly associated with isolated foveal hypoplasia in our previous study.[5] The CTS region and exon 5a (the extra 14 amino acids inserting in the NTS region) have been suggested to be involved in fovea development.[43] Moreover, the CTS and NTS regions have been shown to regulate each other negatively, and the changes in ratio of DNA binding and transactivation activity in these two regions may impact ocular development.[44,45] Conversely, patients with truncation variants tend to exhibit complete aniridia in this cohort, likely due to *PAX6* haploinsufficiency.[3,38] The truncation variants located at the 3′ end of the *PAX6* may evade nonsense-mediated decay and lead to truncated proteins with dominant negative effects, resulting in severe ocular phenotypes.[3,46] The C-terminus of *PAX6* is highly conserved and appears to play a vital role in the DNA binding ability of the HD domain.[47] Previous stud-

ies have noted mild iris abnormalities but severe abnormalities in other ocular structures in patients with CTE variants.[48] Similarly, in our cohort, patients with CTE variants also tend to present iris hypoplasia, distinct from truncation variant phenotypes, suggesting potentially differing mechanism.

Andreas Schedl et al. found that the abnormal expression of *PAX6* can lead to developmental defects, including microphthalmia and anophthalmia, in mutant mice models.[15] Furthermore, *PAX6* missense variants, predominantly located in the PD region, are highly associated with microphthalmia.[13] Besides, expression of PAX6 was inhibited in the form-deprivation chicken model[49] and was enhanced in the hyperopic defocused infant rhesus monkey retinal.[50] Hammond et al. proposed that *PAX6* may play a vital role in eye growth related to myopia based on the strong linkage between the *PAX6* gene region and refractive errors in dizygotic twins, although there was no evidence that those common variants in *PAX6* had direct relationship with myopia.[7] Subsequently, researchers found that several SNPs in *PAX6* were closely associated with high myopia[51] or extreme myopia[8,52] in the Chinese population. However, as described previously, in a British population, no association between *PAX6* and myopia have been observed.[9] Thus, the relationship between specific *PAX6* variants and refractive errors is complex and incompletely understood. In the current study, our data show that different types of refractive errors are associated with different types and locations of *PAX6* variants. Interestingly, the association between *PAX6* variants and refractive errors mirrors the association between *PAX6* variants and the degree of iris defects, suggesting that similar mechanisms, such as changes in DNA affinity activity, may underlie the development of myopia or hyperopia associated with *PAX6* variants. The most common variant related to high myopia in this cohort, c.622C>T (p. Arg208Trp), has also been identified in other studies. In one large African pedigree, 11 patients with the same variant showed myopia, including 6 patients with high myopia.[53] Similarly, varied degrees of myopia have been described in all patients of a large Chinese family[54] as well as in a sporadic case.[55] It might be of interest to explore the underlying mechanism of myopia due to the c.622C>T (p. Arg208Trp) variant in *PAX6*.

In this cohort, the cataract is the most common ocular comorbidity, followed by microcornea, ARK, and glaucoma, in line with the previous studies.[55] The detection rate of cataract (36.8%), glaucoma (3.1%), and ARK (11.4%) in this cohort is lower than the reported frequency, which ranges from 50% to 90% for cataract,[6,56,57] 15% to 66.7% for glaucoma,[6,56,58,59] and 48% to 80% for ARK,[6,60,61] respectively. This may be attributed to the mean age of probands recruited in this cohort being 12.14 ± 13.20 years, because the majority of the patients who visited our Pediatric and Genetic Eye Clinic are children, whereas the reported age of onset of ARK 19 to 33 years of age,[6,57,62] of glaucoma is highly variable, with 70% of patients diagnosed with glaucoma being between 20 and 59 years old.[6] Similarly, congenital cataracts may be present from birth but cannot be easily observed due to the mild lens opacity until they impair vision in adults.

In the current study, we also find a significant correlation between the ocular comorbidities associated with *PAX6* and different variant types. The frequency of cataract is comparatively higher in the splicing variants group than the missense variants group. *PAX6* is identified as the core gene with key role in the lens induction.[63,64] It has been shown that the loss of function of Pax6 during lens induction, after the placode formation has occurred, arrests lens development.[65] Additionally, Pax6 and Bmi1 have been shown to be necessary for lens epithelial stem/progenitor cells (LECs) renewal and loss of LECs homeostasis may result in cataract.[64] Hence, a haploinsufficiency mechanism has been suggested as a possible explanation for *PAX6* variants resulting in the cataract.[66] Conversely, optic disc hypoplasia in this study is more frequently related to the missense variant group compared to the nonsense variant, suggesting that the potential mechanism behind the relationship may involve altering the protein stability and subsequently changing the DNA binding affinity and specificity of *PAX6* rather than a true haploinsufficiency mechanism.[13,67] A similar pattern also has been observed in other studies.[66] In addition, the occurrence of inferior tessellated fundus is not rare, indicating that it may be a characteristic clinical manifestation of fundus changes related to *PAX6* variants.

One limitation of this study is the small family size. The effects of modifying genes and environmental factors on the phenotype caused by the same genetic change is more easily observed in larger families than in small pedigrees. Several previous studies have reported multiple members of the same large family carrying the same variant, showing a wide range of ocular comorbidities, different visual outcomes, and varying severity of iris defects.[68,69] Regarding refractive error, some patients with *PAX6* variants had high myopia, whereas others with the same or a similar variant did not.[70] This intra-pedigree phenotypic variability suggests that the genotype-phenotype relationship of *PAX6* variants is not a simple correspondence. Other modifiers may be involved, such as transcription factor regulation, epigenetic regulation,[71] and subtle changes in the expression of *PAX6* downstream targets and coregulation factors.[72–74] Although the current study lacks large pedigrees, these factors do not seem to have a major impact on the conclusion of genotype-phenotype correlation in the patients described in this study, perhaps because we still have many small pedigrees carrying the same variants in this cohort. To some degree, the number of families can compensate for the limitation of small family sizes and contribute to the generalizability of findings. For example, we describe 10 eyes of 5 individuals across 4 families with same missense variant (c.622C>T), showing different degrees of myopia: 5 eyes with high myopia and 5 eyes with low or moderate myopia. Admittedly, the degree of refractive error may be influenced by multiple factors, but there is no doubt that this missense variant is correlated with myopia. Similarly, the severity of foveal hypoplasia may be affected by a variety modifiers, resulting in the variable degree of foveal hypoplasia among the different families. However, the presence of foveal hypoplasia seems not to be easily changed by factors other than the genotype. The location and degree of iris defects may be modified by environmental and other factors or epigenetic regulation, but the morphology of this iris defect is fundamentally different from the usual iris coloboma. Although these factors may affect this study's conclusions on the degree of genotype-phenotype correlations to some extent, recognizing the relationship can guide judgments on variant pathogenicity and provide valuable insights for future studies on genotype and phenotype characterization.

Another limitation of this study is the absence of the investigation of the variants in the deep intronic region. With the advancement of next-generation sequencing technology, a higher proportion of noncoding and structural variants have been identified as causative variants.[75–77]

Investigative Ophthalmology & Visual Science

Therefore, the possibility of potential variants outside the exome contributing to the observed phenotypes cannot be excluded. However, although whole genome sequencing (WGS) could detect variants throughout the entire genome and potentially identify causal noncoding variants, distinguishing functional causal variants from millions of noncoding variants remains a challenge. Furthermore, compared to non-coding variants, identifying pathogenic variants in the coding region is more feasible because there are well-characterized and validated methods to predict the consequences of each variant and assess their pathogenicity. Thus, for the monogenic disorders like *PAX6*-related disorders, causative pathogenic variants are typically screened from coding region variants first. In fact, based on the HGMD database, the detection rate of potential pathogenic deep-intronic variants is about 3.5% (21/658). We have also performed WGS on 14.8% (9/61) of probands with *PAX6*-related phenotypes from unsolved families in this cohort. Unfortunately, no deep intronic potential pathogenic variants were identified among these cases. We expect more targeted studies will explore this puzzle in the future. Although this does not exclude the possibility of intronic or other variants modifying the phenotype and penetrance, it does suggest that the frequency of such variants causing disease in our cohort is low.

In conclusion, we present a comprehensive clinical and genetic analysis of 228 patients with potential pathogenic *PAX6* variants, including 169 newly identified and 59 previously reported cases. The detection of parental mosaicism and related clinical manifestations in mosaic individuals enhances understanding of the role of parental postzygotic mosaicism in *PAX6*-related disorders and provides valuable insights for genetic counseling. Our study extends beyond the examination of the most common *PAX6*-related disease, aniridia, to encompass a broader spectrum of *PAX6*-related conditions, recognizing that foveal hypoplasia may emerge as the predominant clinical hallmark associated with *PAX6* variants, rather than aniridia. Furthermore, this study not only is the first to analyze the relationship between pathogenic *PAX6* variant type/location and refractive error in a systematic fashion, but also investigates genotype-phenotype relationship in terms of the degree of iris defects, BCVA, and associated ocular comorbidities. These findings support *PAX6* as a susceptibility gene for high myopia, deepen our insight into molecular mechanisms underlying refractive error, and offer novel insights into the clinical diagnosis and treatment of ocular diseases related to *PAX6* variants. The current study illuminates the complexity of *PAX6* regulation in ocular development, paving the way for the future exploration in this area.

Acknowledgments

The authors are appreciative of and thank all the patients and their family members for participating in this study.

Supported by grants from the National Natural Science Foundation of China (82171056), the Science and Technology Planning Projects of Guangzhou (SL2024A03J00525), and the Fundamental Research Funds of the State Key Laboratory of Ophthalmology.

Disclosure: **Y. Jiang**, None; **Z. Yi**, None; **Y. Zheng**, None; **J. Ouyang**, None; **D. Guo**, None; **S. Li**, None; **X. Xiao**, None; **P. Wang**, None; **W. Sun**, None; **Q. Zhang**, None

References

1. Xu HE, Rould MA, Xu W, et al. Crystal structure of the human Pax6 paired domain-DNA complex reveals specific roles for the linker region and carboxy-terminal subdomain in DNA binding. *Genes Dev*. 1999;13:1263–1275.
2. Wilson DS, Guenther B, Desplan C, et al. High resolution crystal structure of a paired (Pax) class cooperative homeodomain dimer on DNA. *Cell*. 1995;82:709–719.
3. Lima Cunha D, Arno G, Corton M, et al. The spectrum of PAX6 mutations and genotype-phenotype correlations in the eye. *Genes (Basel)*. 2019;10:1050.
4. Daruich A, Robert MP, Leroy C, et al. Foveal hypoplasia grading in 95 cases of congenital aniridia: correlation to phenotype and PAX6 genotype. *Am J Ophthalmol*. 2022;237:122–129.
5. Jiang Y, Li S, Xiao X, et al. Genotype-phenotype of isolated foveal hypoplasia in a large cohort: minor iris changes as an indicator of PAX6 involvement. *Invest Ophthalmol Vis Sci*. 2021;62:23.
6. Kit V, Cunha DL, Hagag AM, et al. Longitudinal genotype-phenotype analysis in 86 patients with PAX6-related aniridia. *JCI Insight*. 2021;6:e148406.
7. Hammond CJ, Andrew T, Mak YT, et al. A susceptibility locus for myopia in the normal population is linked to the PAX6 gene region on chromosome 11: a genomewide scan of dizygotic twins. *Am J Hum Genet*. 2004;75:294–304.
8. Tang SM, Ma L, Lu SY, et al. Association of the PAX6 gene with extreme myopia rather than lower grade myopias. *Br J Ophthalmol*. 2018;102:570–574.
9. Simpson CL, Hysi P, Bhattacharya SS, et al. The roles of PAX6 and SOX2 in myopia: lessons from the 1958 British Birth Cohort. *Invest Ophthalmol Vis Sci*. 2007;48:4421–4425.
10. Li FF, Lu SY, Tang SM, et al. Genetic associations of myopia severities and endophenotypes in children. *Br J Ophthalmol*. 2021;105:1178–1183.
11. Vincent MC, Pujo AL, Olivier D, et al. Screening for PAX6 gene mutations is consistent with haploinsufficiency as the main mechanism leading to various ocular defects. *Eur J Hum Genet*. 2003;11:163–169.
12. Plaisancie J, Tarilonte M, Ramos P, et al. Implication of non-coding PAX6 mutations in aniridia. *Hum Genet*. 2018;137:831–846.
13. Williamson KA, Hall HN, Owen LJ, et al. Recurrent heterozygous PAX6 missense variants cause severe bilateral microphthalmia via predictable effects on DNA-protein interaction. *Genet Med*. 2020;22:598–609.
14. Azuma N, Nishina S, Yanagisawa H, et al. PAX6 missense mutation in isolated foveal hypoplasia. *Nat Genet*. 1996;13:141–142.
15. Schedl A, Ross A, Lee M, et al. Influence of PAX6 gene dosage on development: overexpression causes severe eye abnormalities. *Cell*. 1996;86:71–82.
16. Glaser T, Jepeal L, Edwards JG, et al. PAX6 gene dosage effect in a family with congenital cataracts, aniridia, anophthalmia and central nervous system defects. *Nat Genet*. 1994;7:463–471.
17. Lima Cunha D, Owen N, Tailor V, et al. PAX6 missense variants in two families with isolated foveal hypoplasia and nystagmus: evidence of paternal postzygotic mosaicism. *Eur J Hum Genet*. 2021;29:349–355.
18. Tarilonte M, Morin M, Ramos P, et al. Parental mosaicism in PAX6 causes intra-familial variability: implications for genetic counseling of congenital aniridia and microphthalmia. *Front Genet*. 2018;9:479.
19. Zhou L, Xiao X, Li S, et al. Frequent mutations of RetNet genes in eoHM: further confirmation in 325 probands and

comparison with late-onset high myopia based on exome sequencing. *Exp Eye Res*. 2018;171:76–91.
20. Thomas MG, Kumar A, Mohammad S, et al. Structural grading of foveal hypoplasia using spectral-domain optical coherence tomography a predictor of visual acuity? *Ophthalmology*. 2011;118:1653–1660.
21. Jiang Y, Zhou L, Wang Y, et al. The genetic confirmation and clinical characterization of LOXL3-associated MYP28: a common type of recessive extreme high myopia. *Invest Ophthalmol Vis Sci*. 2023;64:14.
22. French AN, Morgan IG, Burlutsky G, et al. Prevalence and 5- to 6-year incidence and progression of myopia and hyperopia in Australian schoolchildren. *Ophthalmology*. 2013;120:1482–1491.
23. Flitcroft DI, He M, Jonas JB, et al. IMI - defining and classifying myopia: a proposed set of standards for clinical and epidemiologic studies. *Invest Ophthalmol Vis Sci*. 2019;60:M20–M30.
24. Tang Y, Chen A, Zou M, et al. Prevalence and time trends of refractive error in Chinese children: a systematic review and meta-analysis. *J Glob Health*. 2021;11:08006.
25. Garcia-Perez MA, Nunez-Anton V. Cellwise residual analysis in two-way contingency tables. *Educ Psychol Meas*. 2003;63:825–839.
26. Beasley TM, Schumacker RE. Multiple regression approach to analyzing contingency tables: post hoc and planned comparison procedures. *J Exp Educ*. 1995;64:79–93.
27. Cheng J, Novati G, Pan J, et al. Accurate proteome-wide missense variant effect prediction with AlphaMissense. *Science*. 2023;381:eadg7492.
28. Tarilonte M, Ramos P, Moya J, et al. Activation of cryptic donor splice sites by non-coding and coding PAX6 variants contributes to congenital aniridia. *J Med Genet*. 2022;59:428–437.
29. Hingorani M, Hanson I, van Heyningen V. Aniridia. *Eur J Hum Genet*. 2012;20:1011–1017.
30. Chabane K, Charlot C, Gugenheim D, et al. Real life evaluation of AlphaMissense predictions in hematological malignancies. *Leukemia*. 2024;38:420–423.
31. Stenson PD, Mort M, Ball EV, et al. The Human Gene Mutation Database (HGMD((R))): optimizing its use in a clinical diagnostic or research setting. *Hum Genet*. 2020;139:1197–1207.
32. Riera M, Wert A, Nieto I, et al. Panel-based whole exome sequencing identifies novel mutations in microphthalmia and anophthalmia patients showing complex Mendelian inheritance patterns. *Mol Genet Genomic Med*. 2017;5:709–719.
33. Reed TE, Falls HF. A pedigree of aniridia with a discussion of germinal mosaicism in man. *Am J Hum Genet*. 1955;7:28–38.
34. Deml B, Reis LM, Lemyre E, et al. Novel mutations in PAX6, OTX2 and NDP in anophthalmia, microphthalmia and coloboma. *Eur J Hum Genet*. 2016;24:535–541.
35. Davis N, Yoffe C, Raviv S, et al. Pax6 dosage requirements in iris and ciliary body differentiation. *Dev Biol*. 2009;333:132–142.
36. Shaham O, Menuchin Y, Farhy C, et al. Pax6: a multi-level regulator of ocular development. *Prog Retin Eye Res*. 2012;31:351–376.
37. Wang L, Liu Z, Lin H, et al. Epigenetic regulation of left-right asymmetry by DNA methylation. *EMBO J*. 2017;36:2987–2997.
38. Moosajee M, Hingorani M, Moore AT. *PAX6-related aniridia*. In: Adam MP, Feldman J, Mirzaa GM, et al. (eds). Seattle, WA; GeneReviews: 1993.
39. Lee S, Lee SH, Heo H, et al. Impaired DNA-binding affinity of novel PAX6 mutations. *Sci Rep*. 2020;10:3062.
40. Tzoulaki I, White IM, Hanson IM. PAX6 mutations: genotype-phenotype correlations. *BMC Genet*. 2005;6:27.
41. Azuma N, Yamaguchi Y, Handa H, et al. Mutations of the PAX6 gene detected in patients with a variety of optic-nerve malformations. *Am J Hum Genet*. 2003;72:1565–1570.
42. Matsushita I, Izumi H, Ueno S, et al. Functional characteristics of diverse PAX6 mutations associated with isolated foveal hypoplasia. *Genes (Basel)*. 2023;14:1483.
43. Azuma N, Tadokoro K, Asaka A, et al. The Pax6 isoform bearing an alternative spliced exon promotes the development of the neural retinal structure. *Hum Mol Genet*. 2005;14:735–745.
44. Yamaguchi Y, Sawada J, Yamada M, et al. Autoregulation of Pax6 transcriptional activation by two distinct DNA-binding subdomains of the paired domain. *Genes Cells*. 1997;2:255–261.
45. Lee H, Khan R, O'Keefe M. Aniridia: current pathology and management. *Acta Ophthalmol*. 2008;86:708–715.
46. Singh S, Tang HK, Lee JY, et al. Truncation mutations in the transactivation region of PAX6 result in dominant-negative mutants. *J Biol Chem*. 1998;273:21531–21541.
47. Singh S, Chao LY, Mishra R, et al. Missense mutation at the C-terminus of PAX6 negatively modulates homeodomain function. *Hum Mol Genet*. 2001;10:911–918.
48. Hingorani M, Williamson KA, Moore AT, et al. Detailed ophthalmologic evaluation of 43 individuals with PAX6 mutations. *Invest Ophthalmol Vis Sci*. 2009;50:2581–2590.
49. Bhat SP, Rayner SA, Chau SC, et al. Pax-6 expression in posthatch chick retina during and recovery from form-deprivation myopia. *Dev Neurosci*. 2004;26:328–335.
50. Zhong XW, Ge J, Deng WG, et al. Expression of PAX-6 in rhesus monkey of optical defocus induced myopia and form deprivation myopia. *Chin Med J (Engl)*. 2004;117:722–726.
51. Han W, Leung KH, Fung WY, et al. Association of PAX6 polymorphisms with high myopia in Han Chinese nuclear families. *Invest Ophthalmol Vis Sci*. 2009;50:47–56.
52. Tsai YY, Chiang CC, Lin HJ, et al. A PAX6 gene polymorphism is associated with genetic predisposition to extreme myopia. *Eye (Lond)*. 2008;22:576–581.
53. Goolam S, Carstens N, Ross M, et al. Familial congenital cataract, coloboma, and nystagmus phenotype with variable expression caused by mutation in PAX6 in a South African family. *Mol Vis*. 2018;24:407–413.
54. Shen T, Qiu X, Lin X, et al. Missense mutation in the PAX6 gene can cause a complex mild variable phenotype predominated by concomitant strabismus. *Ophthalmic Genet*. 2022;43:88–96.
55. Huang L, Peng J, Xie Y, et al. Diversity of clinical phenotypes in a cohort of Han Chinese patients with PAX6 variants. *Front Genet*. 2023;14:1011060.
56. Park SH, Park YG, Lee MY, et al. Clinical features of Korean patients with congenital aniridia. *Korean J Ophthalmol*. 2010;24:291–296.
57. Souzeau E, Rudkin AK, Dubowsky A, et al. PAX6 molecular analysis and genotype-phenotype correlations in families with aniridia from Australasia and Southeast Asia. *Mol Vis*. 2018;24:261–273.
58. You B, Zhang X, Xu K, et al. Mutation spectrum of PAX6 and clinical findings in 95 Chinese patients with aniridia. *Mol Vis*. 2020;26:226–234.
59. Vasilyeva TA, Marakhonov AV, Voskresenskaya AA, et al. Analysis of genotype-phenotype correlations in PAX6-associated aniridia. *J Med Genet*. 2021;58:270–274.
60. Jin C, Wang Q, Li J, et al. A recurrent PAX6 mutation is associated with aniridia and congenital progressive cataract in a Chinese family. *Mol Vis*. 2012;18:465–470.
61. Mayer KL, Nordlund ML, Schwartz GS, et al. Keratopathy in congenital aniridia. *Ocul Surf*. 2003;1:74–79.

Investigative Ophthalmology & Visual Science

62. Yazdanpanah G, Bohm KJ, Hassan OM, et al. Management of congenital aniridia-associated keratopathy: long-term outcomes from a tertiary referral center. *Am J Ophthalmol*. 2020;210:8–18.
63. Altmann CR, Chow RL, Lang RA, et al. Lens induction by Pax-6 in Xenopus laevis. *Dev Biol*. 1997;185:119–123.
64. Lin H, Ouyang H, Zhu J, et al. Lens regeneration using endogenous stem cells with gain of visual function. *Nature*. 2016;531:323–328.
65. Ashery-Padan R, Marquardt T, Zhou X, et al. Pax6 activity in the lens primordium is required for lens formation and for correct placement of a single retina in the eye. *Genes Dev*. 2000;14:2701–2711.
66. Vasilyeva TA, Sukhanova NV, Khalanskaya OV, et al. An unusual presentation of novel missense variant in PAX6 Gene: NM_000280.4:c.341A>G, p.(Asn114Ser). *Curr Issues Mol Biol*. 2023;46:96–105.
67. Alibes A, Nadra AD, De Masi F, et al. Using protein design algorithms to understand the molecular basis of disease caused by protein-DNA interactions: the Pax6 example. *Nucleic Acids Res*. 2010;38:7422–7431.
68. Wang GM, Prasov L, Al-Hasani H, et al. Phenotypic variation in a four-generation family with aniridia carrying a novel PAX6 mutation. *J Ophthalmol*. 2018;2018:5978293.
69. Sale MM, Craig JE, Charlesworth JC, et al. Broad phenotypic variability in a single pedigree with a novel 1410delC mutation in the PST domain of the PAX6 gene. *Hum Mutat*. 2002;20:322.
70. Hewitt AW, Kearns LS, Jamieson RV, et al. PAX6 mutations may be associated with high myopia. *Ophthalmic Genet*. 2007;28:179–182.
71. Xu PX, Zhang X, Heaney S, et al. Regulation of Pax6 expression is conserved between mice and flies. *Development*. 1999;126:383–395.
72. Baumer N, Marquardt T, Stoykova A, et al. Retinal pigmented epithelium determination requires the redundant activities of Pax2 and Pax6. *Development*. 2003;130:2903–2915.
73. Kamachi Y, Uchikawa M, Tanouchi A, et al. Pax6 and SOX2 form a co-DNA-binding partner complex that regulates initiation of lens development. *Genes Dev*. 2001;15:1272–1286.
74. Planque N, Leconte L, Coquelle FM, et al. Specific Pax-6/microphthalmia transcription factor interactions involve their DNA-binding domains and inhibit transcriptional properties of both proteins. *J Biol Chem*. 2001;276:29330–29337.
75. Valente EM, Bhatia KP. Solving mendelian mysteries: the non-coding genome may hold the key. *Cell*. 2018;172:889–891.
76. Filatova A, Reveguk I, Piatkova M, et al. Annotation of uORFs in the OMIM genes allows to reveal pathogenic variants in 5'UTRs. *Nucleic Acids Res*. 2023;51:1229–1244.
77. Wakeling MN, Owens NDL, Hopkinson JR, et al. Non-coding variants disrupting a tissue-specific regulatory element in HK1 cause congenital hyperinsulinism. *Nat Genet*. 2022;54:1615–1620.

Investigative Ophthalmology & Visual Science