HHS Public Access

Author manuscript

Am J Med Genet A. Author manuscript; available in PMC 2023 August 01.

Published in final edited form as:

Am J Med Genet A. 2022 August ; 188(8): 2376–2388. doi:10.1002/ajmg.a.62874.

Exome Sequencing Identifies Genetic Variants in Anophthalmia and Microphthalmia

Jingjing Li[1,*,#], **Wei Yang**[2,*], **Yuejun Jessie Wang**[1,*], **Chen Ma**[2], **Cynthia J. Curry**[3], **Daniel McGoldrick**[4], **Deborah A. Nickerson**[4,16], **Jessica X. Chong**[15,16], **Elizabeth E. Blue**[16,17], **James C. Mullikin**[5], **Jennita Reefhuis**[6], **Wendy N. Nembhard**[7], **Paul A. Romitti**[8], **Martha M. Werler**[9], **Marilyn L. Browne**[10,11], **Andrew F. Olshan**[12], **Richard H. Finnell**[13], **Marcia L. Feldkamp**[14], **Faith Pangilinan**[5], **Lynn M. Almli**[6], **Mike J. Bamshad**[4,15,16], **Lawrence C. Brody**[5], **Mary M. Jenkins**[6], **Gary M. Shaw**[2,#], **NISC Comparative Sequencing Program**[18], **University of Washington Center for Mendelian Genomics**[4], **National Birth Defects Prevention Study**

[1]The Eli and Edythe Broad Center of Regeneration Medicine and Stem Cell Research, the Bakar Computational Health Sciences Institute, the Parker Institute for Cancer Immunotherapy, and the Department of Neurology, School of Medicine, University of California, San Francisco, CA, USA

[2]Department of Pediatrics, Stanford University School of Medicine, Stanford, CA, USA

[3]Genetic Medicine, Department of Pediatrics, University of California, San Francisco/Fresno, CA, USA

[4]Department of Genome Sciences, University of Washington, Seattle, WA, USA

[5]National Human Genome Research Institute, National Institutes of Health, Bethesda, MD, USA

[6]National Center on Birth Defects and Developmental Disabilities, Centers for Disease Control and Prevention, Atlanta, GA, USA

[7]Department of Epidemiology, Fay W. Boozman College of Public Health, University of Arkansas for Medical Sciences, Little Rock, AR, USA

[8]Department of Epidemiology, University of Iowa College of Public Health, Iowa City, IA, USA

[9]Department of Epidemiology, Boston University School of Public Health, Boston, MA, USA

[10]Birth Defects Registry, New York State Department of Health, Albany, NY, USA

[#] To whom correspondence should be addressed: gmshaw@stanford.edu or Jingjing.Li@ucsf.edu.
[*] Contributed equally to this work.

Author Contributions

Gary M. Shaw supervised the study. Lynn M. Almli, Marilyn L. Browne, Cynthia J. Curry, Marcia L. Feldkamp, Richard H. Finnell, Mary M. Jenkins, Wendy N. Nembhard, Andrew F. Olshan, Jennita Reefhuis, Paul A. Romitti, Gary M. Shaw, Martha M. Werler, NBDPS, contributed to the NBDPS study design and data and specimen acquisition. Mike Bamshad, Elizabeth E. Blue, Lawrence C. Brody, Jessica X. Chong, Daniel McGoldrick, James Mullikin, Deborah A. Nickerson, Faith Pangilinan, NISC Comparative Sequencing Program, and UWCMG conducted experimental studies. Jingjing Li, Chen Ma, Yuejun Jessie Wang, and Wei Yang analyzed the data, and Mike Bamshad, Elizabeth E. Blue, Lawrence C. Brody, Jessica X. Chong, Jingjing Li, Chen Ma, Daniel McGoldrick, Deborah A. Nickerson, Gary M. Shaw, Yuejun Jessie Wang, Wei Yang, and UWCMG interpreted the data. Jingjing Li, Yuejun Jessie Wang, and Wei Yang created the initial draft of the article and reviewed the literature and Lynn M. Almli, Cynthia J. Curry, Mary M. Jenkins, Faith Pangilinan, Paul A. Romitti and Gary M. Shaw critically revised the article. All authors read and approved the final article.

[11]Department of Epidemiology and Biostatistics, School of Public Health, University at Albany, Rensselaer, NY, USA

[12]Department of Epidemiology, Gillings School of Global Public Health, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

[13]Center for Precision Environmental Health, Departments of Molecular and Cellular Biology, Molecular and Human Genetics and Medicine, Baylor College of Medicine, Houston, TX, USA

[14]Division of Medical Genetics, Department of Pediatrics, 295 Chipeta Way, Suite 2S010, University of Utah School of Medicine, Salt Lake City, UT, USA

[15]Department of Pediatrics, Division of Genetic Medicine, University of Washington, Seattle, WA

[16]Brotman Baty Institute for Precision Medicine, Seattle, WA, USA

[17]Division of Medical Genetics, Department of Medicine, University of Washington, Seattle, WA, USA

[18]NIH Intramural Sequencing Center, National Human Genome Research Institute, Bethesda, MD

Abstract

Anophthalmia and microphthalmia (A/M) are rare birth defects affecting up to 2 per 10,000 live births. These conditions are manifested by the absence of an eye or reduced eye volumes within the orbit leading to vision loss. Although clinical case series suggest a strong genetic component in A/M, few systematic investigations have been conducted on potential genetic contributions owing to low population prevalence. To overcome this challenge, we utilized DNA samples and data collected as part of the National Birth Defects Prevention Study (NBDPS). The NBDPS employed multi-center ascertainment of infants affected by A/M. We performed exome sequencing on 67 family trios and identified numerous genes affected by rare deleterious nonsense and missense variants in this cohort, including *de novo* variants. We identified 9 nonsense changes and 86 missense variants that are absent from the reference human population (Genome Aggregation Database), and we suggest that these are high priority candidate genes for A/M. We also performed literature curation, single cell transcriptome comparisons, and molecular pathway analysis on the candidate genes and performed protein structure modeling to determine the potential pathogenic variant consequences on PAX6 in this disease.

Keywords

Congenital abnormalities; newborn eye abnormalities; genetic epidemiology

Introduction

Anophthalmia and microphthalmia (A/M) are complex eye malformatons (Dong et al., 2015; Skalicky et al., 2013; Verma & Fitzpatrick, 2007). Human eye development is initiated at three weeks gestation from the anterior neural plate (Sinn & Wittbrodt, 2013; Zagozewski, Zhang, & Eisenstat, 2014). Anophthalmia is defined as a total absence of the eye tissue or any structures associatied with the eye (Plaisancie et al., 2019). True anophthalmia is an abnormality of eye development that occurs at the time of developing

optic vesicle at 3–4 weeks of gestation leading to absence of the eye, optic nerve and chiasm (Harding & Moosajee, 2019). Clinical anophthalmia is considered in the absence of ocular structures with histologically detected remnants (Plaisancie et al., 2019). Microphthalmia is believed to arise early in pregnancy as well, but at least one report has suggested identification of microphthalmia development in midpregnancy (Blazer, Zimmer, Mezer, & Bronshtein, 2006). Microphthalmia is defined as reduction in the volume of the eye, usually characterized by corneal diameters less than 10 mm or anteroposterior globe diameter less than 20 mm, which is part of the spectrum of clinical anophthalmia and includes coloboma, anterior segment abnormalities and cataract (Skalicky et al., 2013; Verma & Fitzpatrick, 2007). A/M are rare congenital malformations with a prevalence of up to 2 per 10,000 live births (Mai et al., 2019; Shaw et al., 2005).

Current understanding of suspected A/M risk factors has been largely derived from clinical case series and experimental models (Graw, 2019; Kumar, Sandell, Trainor, Koentgen, & Duester, 2012; Richardson, Tracey-White, Webster, & Moosajee, 2017). A significant genetic component of A/M has emerged from such clinical series. Several chromosome abnormalities have been implicated, such as trisomies 9, 13, and 18, 14q22.1-q23.2 deletion and 3q, 4p, and 10q duplications (Verma & Fitzpatrick, 2007).

The genetic basis of these conditions is further supported by identification of several candidate genes through clinical genomic sequencing. Approximately 20 genes are associated with A/M, but collectively these only explain a very small proportion of A/M cases (Bardakjian & Schneider, 2011; Williamson & FitzPatrick, 2014). Among these genes, one of the most extensively studied is *PAX6* given its primary role in regulating eye development (Harding, Brooks, FitzPatrick, & Moosajee, 2020; Patel & Sowden, 2019; Verma & Fitzpatrick, 2007). Compound heterozygotic nonsense variants in *PAX6* were previously associated with anophthalmia (Glaser et al., 1994). However, the prevalence of *PAX6* nonsense variants in A/M cases is extremely rare, and *PAX6* variants are more commonly seen in many other human ocular phenotypes, such as aniridia (Patel & Sowden, 2019; Verma & Fitzpatrick, 2007). *SOX2* has also been implicated in A/M in many studies; but may only be associated in approximately 10% of A/M cases (Dash et al., 2020; Fantes et al., 2003; Matias-Perez et al., 2018). Thus, known genetic risk factors explain a very small fraction of A/M cases. An important observation made from previous genetic analyses is the genetic heterogeneity in A/M, where different genes contribute to the molecular etiologies of the A/M phenotype and showed that different individuals with A/M have different sets of variants.

Such genetic heterogeneity poses a significant challenge for classical family segregation analysis for Mendelian traits and candidate gene studies, and suggests the need for an unbiased genomic screen if we are to identify genetic underpinnings in A/M. What has made the investigation more challenging is the rarity of A/M cases, with existing knowledge mostly derived from nonsystematic small collections of affected probands. Exome sequencing analyses have been previously attempted in A/M cases but were limited to small patient cohorts (<30) or to a small set of preselected genes with previously suggestive evidence of microphthalmia, anophthalmia, and coloboma (Deml et al., 2016; Haug et al., 2021). Here we are able to capitalize on the systematic data collection of

the National Birth Defects Prevention Study (NBDPS), the largest birth defects study conducted in the United States. We performed exome sequencing on 67 family trios who participated in the NBDPS and were affected by A/M. We integrated exome analysis, single-cell transcriptome comparison, molecular pathway analysis, protein structure modeling and literature curation, resulting in the identification of numerous genes involved in A/M, and specific retinal cell types potentially implicated in the pathogenesis of A/M.

Methods and Materials

Study populations

The NBDPS is a multi-center case-control study of over 30 major birth defects. This study included data on pregnancies with estimated due dates between October 1997 and December 2011. Complete study methods have been described previously (Reefhuis et al., 2015). Cases and controls were ascertained from 10 birth defects surveillance centers: Arkansas, California, Georgia, Iowa, Massachusetts, New Jersey, New York, North Carolina, Texas, and Utah (Reefhuis et al., 2015). The institutional review boards of each study center provided approval for the NBDPS.

Infants and fetuses diagnosed with A/M had their medical records reviewed for eligibility by clinical geneticists. All included cases required diagnosis by an ophthalmologist or confirmation by surgical pathology or autopsy. Infants were ineligible if their only diagnosis was “small eyes” or “small palpebral fissures.” Infants were also ineligible if their A/M was associated with anterior encephalocele, anencephaly, holoprosencephaly, or amniotic band syndrome. Cases strongly suspected to have a chromosomal abnormality or single-gene condition were also excluded (Rasmussen et al., 2003). Cases were classified as “isolated” or “with accompanying malformations” based on the presence and type of accompanying malformations. Cases with no other major malformations, i.e., only minor anomalies or related anomalies, were considered isolated. Cases with at least one additional, believed to be unrelated major malformation, were considered A/M with accompanying malformations. The term “unrelated” refers to defects in different body parts or systems and not a part of sequence.

Following a computer-assisted telephone interview, each woman was mailed cytobrushes to collect buccal cell specimens for herself, her child (if living), and her child’s biological father (if available). Further details about cytobrush collection can be found in Jenkins et al. 2019 (Jenkins et al., 2019). Out of 235 A/M cases with completed maternal interviews, there were only 76 trios (proband, mother, father) with DNA available from cytobrushes, and eight trios were removed from the 76 because at least one family member had DNA amounts less than 200 ng. Out of the remaining 68 trios, another proband was removed due to poor sample quality, leading to exclusion of the entire family. Therefore, we had 67 probands for sequence analysis as part of a trio, and they served as the analytic base for the current work.

Exome sequencing

Details can be found in Jenkins et al. 2019 (Jenkins et al., 2019). Briefly, exome sequencing was performed at the National Institutes of Health Intramural Sequencing Center (NISC,

Rockville, MD; https://www.nisc.nih.gov/). We used SeqCap® EZ Human Exome + UTR kit v3.0 (Roche NimbleGen, Madison, WI) for exome capture. Sequencing was performed using Illumina's HiSeq 2500 system (Illumina, San Diego, CA), which generated 126 bp paired-end reads. Illumina Genome Analyzer Pipeline software (version 1.18.64.0) was used for downstream analysis. Illumina's Efficient Large-scale Alignment of Nucleotide Databases (ELAND) was used to map the paired-end reads onto the reference human genome (hg19). We used the Burrows-Wheeler Aligner (BWA, v0.7.10) for read alignment (H. Li & Durbin, 2009). We followed the Genome Analysis Toolkit (GATK) Best Practice protocol for variant call involving "duplicate removal" (Picard MarkDuplicates; v1.111), indel realignment (GATK IndelRealigner; v3.2–2), and base quality recalibration (GATK BaseRecalibrator; v3.2–2). Raw Variant Call Format files were generated by GATK HaplotypeCaller (v3.2). Called raw variants were further refined by the variant quality score recalibration (VQSR) procedure. We only considered 117,608 high-confidence single nucleotide variants (SNVs) with "PASS" tags assigned by GATK VQSR.

Data quality control (QC) included an assessment of (1) total reads (minimum of 50 million PE50 reads); (2) library complexity; (3) capture efficiency; (4) coverage distribution: 90% at 8X required for completion; (5) capture uniformity; (6) raw error rates; (7) transition/transversion ratio (Ti/Tv); (8) distribution of known and novel variants relative to dbSNP (typically < 7%); (9) fingerprint concordance > 99%; (10) sample homozygosity and heterozygosity; and (11) sample contamination validation. Exome completion was defined as having > 90% of the exome target at > 8X coverage and >80% of the exome target at > 20X coverage. Typically, this required mean coverage of the target at 50–60X. Variant annotation was based on Annovar v2017Jul17 (Wang, Li, & Hakonarson, 2010), which returns annotations that include dbSNP rsIDs, gene names, and predicted functional effects (e.g., nonsynonymous, stopgain, or stoploss).

Functional genomic analysis

To identify genes affected by nonsense variants, we required variants to be absent in the Genome Aggregation Database (gnomAD) v2.1.1 and the probability of being loss-of-function intolerant (pLI) score (Lek et al., 2016) to be > 0.9, as a predictor of extreme haplo-insufficiency. In analyses of missense variants, our criteria called for variants to be absent in the gnomAD database v2.1.1, a Phred-scaled Combined Annotation Dependent Depletion (CADD) score (Kircher et al., 2014) > 20, Variant Effect Scoring Tool 3.0 (VEST3) score (Carter, Douville, Stenson, Cooper, & Karchin, 2013) ≥ 0.7 and Meta Likelihood ratio (LR) score (Dong et al., 2015) ≥ 0.7, as predictors of extreme deleteriousness. We used Enrichr (Kuleshov et al., 2016) for functional enrichment analysis. We analyzed single-cell transcriptomic data across retina developmental stages from gestational week 9 to 27 (Lu et al., 2020). We first downloaded preprocessed data from GEO with the accession number GSE138002. Then we followed cell type definitions (Lu et al., 2020) and standardized gene expression in each cell type by total number of sequenced reads. We then performed quantile normalization across all cell types. For each gene, we averaged its expression across all cells in a given cell type, and for each cell type, we compared expression differences between the candidate genes and control genes (background) using a Wilcoxon rank-sum test, followed by Benjamini-Hochberg correction. For protein structure analysis, we implemented Site

Directed Mutator (Pandurangan, Ochoa-Montano, Ascher, & Blundell, 2017) and DynaMut (Rodrigues, Pires, & Ascher, 2018) to compute physicochemical properties associated with any residues in a given protein structure.

Sanger sequencing

Sanger sequencing was performed by ACGT, Inc. (http://www.acgtinc.com/) to validate a select group of the rare nonsense (loss-of-function) variants (n=9) detected during exome sequencing. The primer set was synthesized based on the reference sequence. Each locus of interest underwent PCR amplification of genomic DNA (>100 ng/reaction) and bidirectional sequencing. Sequence data from each locus were aligned, and the variants of interest were independently called by at least two individuals. Among the nine variants, Sanger sequencing was attempted in five probands and two trios, i.e., total of seven probands and parents of two of them with sufficient DNA quantities remaining.

Results

Among the 67 probands with A/M, over 90% had microphthalmia, 82% were considered isolated in their phenotype, and 72% occurred unilaterally (Table 1). Demographic characteristics of the probands and their parents are also shown in Table 1. Although no parent of any proband self-reported to have A/M, there were three probands with first degree relatives who self-reported eye conditions. These conditions could be part of the A/M spectrum. The first proband's father was reported as having coloboma, the second proband's mother and sibling were reported to be blind at birth, and the third proband's mother reported Duane syndrome (data not shown).

We performed exome sequencing on 67 trios, generating high quality sequence data for each of the 201 individuals. Subsequent QC analysis confirmed high quality of the called variants (Figure S1). We identified 102,793 SNVs in coding sequences among the 67 probands. We had less confidence on indel and structural variant calls, therefore, we focused our analyses on SNVs.

Given the low population prevalence of A/M (2/10,000 live births), we followed the rare disease, rare variant model and studied the role of rare variants in A/M. These rare variants also included *de novo* variants in the analyses. Studying all rare variants, instead of only focusing on *de novo* variants, enabled us to study private variants in affected families. For exonic variants identified in the 67 probands with A/M, we excluded those observed in the global human population represented by exome and genome sequences from 125,748 and 15,708 individuals, respectively (the gnomAD database v2.1.1), resulting in an entire collection of 2,686 rare exonic variants identified in the probands with A/M. Thus, these variants represent those rare in the human population.

Mapping these rare variants onto their respective protein coding sequences, we identified 72 nonsense (loss-of-function) and 1,776 missense variants. For loss-of-function variants, we considered them to have deleterious effects if the affected proteins were intolerant to copy loss, thereby receiving extreme probability of loss of function intolerance (pLI) scores (Lek et al., 2016) (pLI>0.9). For missense variants, we implemented a set of machine learning

algorithms to jointly predict their mutational consequences and only considered those variants receiving consensus predicted mutational deleteriousness. Following the procedures outlined in Figure 1, we identified 9 and 86 rare deleterious nonsense and missense variants respectively.

We first analyzed the nine rare nonsense variants affecting eight genes (Table 2). Gene *CHD7* was implicated in two unrelated probands. We performed literature curation for each of the genes in Table 2 and found evidence that each is involved in eye development (Bergman et al., 2012; Du et al., 2015; Ergun, Akay, Ergun, & Percin, 2017; G. Li, Jin, & Zhong, 2019; Meijer, 1993; Papakostas et al., 2018). For example, *LRP5*, *NUMA1* and *TUBGCP3* have been identified to be involved in retinal development and *KRT31* to be involved with eye morphological alterations (Du et al., 2015; Ergun et al., 2017; G. Li et al., 2019). Thus, this agnostic analysis of rare nonsense variants successfully identified candidate genes in A/M.

Previous work has identified the retina as an important tissue of origin in A/M (Chao et al., 2010; Garcia-Llorca, Aspelund, Ogmundsdottir, Steingrimsson, & Eysteinsson, 2019). To derive mechanistic insights, we analyzed single-cell transcriptome data across retina developmental stages (Lu et al., 2020), where the global transcriptomes were profiled in each of the 10 cell types in the developing retina from gestational week 9 to 27. As a background control set, we identified a list of 71 genes following the same procedure in the probands (Figure 1), which were affected by rare nonsense variants, without considering their tolerance to gene copy loss. In this way, predicted mutational pathogenicity could be directly associated with functional specificity. We observed that across seven out of the ten retinal cell types, the identified eight genes in Table 2 displayed significantly increased expression in aggregate relative to the background control genes (false discovery rates, FDRs<0.05, Wilcoxon rank-sum test, followed by Benjamini-Hochberg correction), and the pattern was consistent in both the first and second trimesters (Figure 2a and 2b). Therefore, the single-cell analysis confirmed the increased aggregate expression of the identified genes in retinal development, and the comparisons revealed potential cell types that may be involved in development of A/M. We used the detected genes, *CHD7* and *XKR4,* for example, to reveal potential cell types implicated in A/M. *CHD7* is more ubiquitously expressed across all retinal cell types, despite a stronger pattern in retinal progenitor cells, cones and bipolar cell /photoreceptor precursors (Figure 2c). However, *XKR4* displayed strong specificities in the developing retina in only a few cell types, including retinal ganglion cells, amacrines, horizontal cells and cones (Figure 2d). Future functional experiments seem warranted to examine the role of *XKR4* in retinal development through its cell-type-specific expression.

We investigated the parent of origin for these identified nonsense variants (Table 2). We identified two putative *de novo* variants *VPS13D* (c.C12634T, p.Q4212X) and *CHD7* (c.C2959T, p.R987X). While additional DNA samples were unavailable for verifying the *de novo* status of the *VPS13D* (c.C12634T, p.Q4212X) variant (Family #46), we were able to use Sanger sequencing and successfully confirmed the *de novo* variant c.C2959T, p.R987X in *CHD7* (Family #55, Figure S2). In addition to these *de novo* variants, other variants were observed to be either paternally or maternally transmitted. We

additionally performed Sanger sequencing and confirmed the presence of four identified loss-of-function variants in the corresponding probands, including variants in *CHD7* (c.C4480T, p.R1494X), *NUMA1* (c.C1654T, p.Q552X), *TUBGCP3* (c.C1376A, p.S459X), and *FAM192A* (c.C175T, p.Q59X). The *LRP5* (c.C1489T, p.R497X) variant in the proband was not confirmed due to failed PCR. We were not able to perform Sanger validation on *XKR4* (c.C14G, p.S5X) and *KRT31* (c.G331T, p.E111X) due to insufficient DNA remaining post exome sequencing studies.

We implemented three machine learning algorithms to predict potential missense variant effects (Carter et al., 2013; Dong et al., 2015; Kircher et al., 2014). We considered a subset of 86 missense variants receiving consensus deleteriousness predictions from all three algorithms, which represented rare deleterious missense variants of high confidence in the A/M proband cohort (Figure 1). These 86 deleterious variants affected 85 genes. We performed pathway analysis and observed functional enrichment for eye physiologies among these affected genes (Table 3a). Mouse mutants of these genes displayed abnormal horizontal cell morphology and blindness (e.g., UPS38 and WNK4, Table 3b). In humans, these 85 identified genes exhibited phenotypic enrichment in hypoplasia of the fovea, the macula and the retina, as well as in opacification of the corneal stroma and optic atrophy (Table 3a). We performed the same analyses on genes affected by rare potential missense variant without considering mutational consequences, and the same functional enrichment could not be observed, thereby correlating predicted mutational pathogenicity with functional specificities to eye physiologies. Investigating rare missense variants captured additional genes that could be contributing to A/M. Among these rare deleterious missense variants, seven were likely *de novo* variants, 4 were unknown in origin, and 75 were parentally transmitted (Table S1b) further indicating either the variants were not pathologic or the potential heterogeneity in the mode of inheritance underlying A/M.

We also manually examined each of the identified 85 unique genes (Table S1b) and our literature curation uncovered their overall functions in eye development. Some exemplar genes are shown in Table 3b, which were not included as member genes in the above enriched functional ontology categories (Table 3a). We particularly note *UPS38* (c.C2281A, p.L761I) and *WNK4* (c.A151G, p.T51A), whose mutants in model organisms displayed a loss of eyes (Table 3b)(Shimizu, Goto, Sato, & Shibuya, 2013; Tsai et al., 2019), replicating the expected A/M traits in humans.

Among all identified genes, *PAX6* was of prime interest given its role as a key transcription factor driving eye development (Glaser et al., 1994). In our analysis, a potential deleterious missense variant in the N-terminus of the PAX6 protein sequence was detected in one A/M proband (potential pathogenic variant was absent from the reference human population). As shown in Figure 3a, this potential pathogenic variant (c.G56A, p.G19E) [refseq ID: NM_001368911.2] altered the wildtype amino acid from a glycine (G) residue into a glutamic acid (E), localized in the paired DNA binding domain of the transcription factor PAX6. To gain mechanistic insights into the mutational impact on this main eye development driver protein, we modeled the 3D structure of PAX6 (PDB ID: 6PAX) and mapped the p.G19E potential pathogenic variant onto the PAX6 crystal structure. Thermodynamic analysis of the p.G19E potential pathogenic variant revealed that the

mutant allele is predicted to substantially reduce the solvent accessibility associated with the wildtype allele from 113.9% to 66.7%, leading to increased packing density in this local position. As such, the residue-occluded packing density at this position displayed about a 40% increase from 0.16 to 0.22. By burying this mutant amino acid into the protein interior, as opposed to the wildtype residue, which is exposed to the solvent, this potential pathogenic variant in fact destabilized the PAX6 structure with $\Delta\Delta G=-3.34$ kcal/mol. We further computed the alteration of vibrational entropy in the PAX6 crystal structure that resulted from this amino acid change. As shown in Figure 3b, the mutated residue had a significant effect on inducing a rigidification (loss of molecular flexibility) of the PAX6 structure in the N-terminus.

For the three probands with first degree relatives reported as having eye malformations, there were identified missense variants inherited from their affected parents (Table S1b).

Finally, we investigated whether various factors reported during the maternal interview appeared more or less frequently amongst case infants with versus without the identified gene variants. That is, we compared each of the three groups, 9 with *de novo*, 9 with nonsense, and 48 with nonsense or missense variants to the group of 19 case infants without identified variants (some of the infants overlapped between the three groups). More than 200 factors (such as maternal/paternal demographic information, maternal pregnancy history, lifestyle factors, multivitamin/folic acid intake, and nutrient intake from food) were investigated. The sample sizes were expectedly sparse, with none of the assessed factors observed to occur with higher frequency in one group vs. another (data not shown).

Discussion

We employed an agnostic search strategy based on exome sequences in 201 individuals from 67 trio families where the proband had A/M. This search strategy confirmed a few known genes in A/M and also identified novel candidate genes of interest. Contributions of the novel genes identified were orthogonally supported by single-cell transcriptomic data (Lu et al., 2020), molecular pathway analysis, protein structure modeling, and curated literature reports.

We identified 9 nonsense and 86 deleterious missense variants that are absent from gnomAD. Our functional analysis validated their physiological roles in regulating eye development, thereby providing potential mechanistic insights for the findings. We observed recurrent disruptions of *CHD7* in two independent probands, the gene for CHARGE syndrome, in which colobomatous microphthalmia is a component of the syndrome (Haug et al., 2021). For these two probands with a *CHD7* variant, the phenotype of one also included microtia and unilateral cleft lip and palate, and the phenotype of the other did not include any additional malformations. For other candidate genes, two have been implicated in model organisms, such as the loss of eye phenotypes associated with *UPS38* and *WNK4* mutants in mouse or other model organisms. To gain further possible mechanistic insights, we analyzed single-cell transcriptomic data across retinal developmental stages and identified several cell types displaying aggregate expression enrichment of the identified candidate genes, thereby suggesting vulnerable cell types specific to A/M etiologies. However, it is important

to note that in addition to the retina, many other eye components are also implicated in A/M, and future inquiry is required to determine A/M-associated cell types in different eye components.

We especially focused on the analysis of a potential pathogenic missense variant in *PAX6*, a key transcription factor driving eye development (Glaser et al., 1994). The association of *PAX6* with A/M was previously established by a case study identifying loss-of-function variants in *PAX6*, and our work leveraging protein structure modeling now provides evidence for the involvement of missense variants in *PAX6* in A/M. Intriguingly, at the same position, mutating the wildtype glycine residue into two different amino acids resulted in two different eye development disorders: the glycine (G) to tryptophan (W) alteration resulted in aniridia (Wolf et al., 1998), whereas the glycine (G) to glutamic acid (E) was associated with microphthalmia (Figure 3). We postulated that this phenotypic disparity likely resulted from the different side chain structures of the tryptophan and the glutamic acid residues, leading to varying degrees of perturbing the PAX6 protein structure.

The mutated residue, p.G19E, destabilized the local structure in the paired DNA binding domain of *PAX6*, thereby likely perturbing the *PAX6*-mediated gene regulatory network and, thus, leading to eye developmental disorders. Intriguingly, a previous report also identified several missense variants in *PAX6* associated with bilateral microphthalmia (Williamson et al., 2020), which, despite different locations from our observation, all resided in the paired DNA binding domain of PAX6 (Figure 3a). Our observation, in conjunction with previous work, indicates that the PAX6 paired DNA binding domain is a potential pathogenic variant hotspot in bilateral microphthalmia. We also observed that the mutated residue in the proband in our study was maternally transmitted. Further review of maternal interview data indicated blindness in both the proband's mother and a sibling.

In addition, we conducted exploratory analyses of indels that could potentially affect splice sites. However, none of the identified splicing events would be considered reportable based on ACMG practice guidelines(Miller et al., 2021; Richards et al., 2015). While this study is the largest sequencing effort for A/M and the data derived from population-based case ascertainment, it is still of relatively small size. Such limited sample size obviated a robust exploration of possible pregnancy exposures that may , in combination with gene variants, modify risks of A/M. Further, the analysis was based on exome sequence data and not whole genome sequence data rendering us unable to explore potential contributions from non-coding regions of the genome. Therefore, future work with larger studies will help confirm the candidate genes identified here and pursue some of the hypothetical directions offered by our observations.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

This work was supported through cooperative agreements under PA 96043, PA 02081 and FOA DD09–001 from the Centers for Disease Control and Prevention to the Centers for Birth Defects Research and Prevention

participating in the NBDPS. Sequencing data was reprocessed and analyzed by the University of Washington Center for Mendelian Genomics and was funded by the National Human Genome Research Institute (NHGRI) and National Heart, Lung, and Blood Institute grants UM1 HG006493 and U24 HG008956. This work was supported by the Division of Intramural Research of NHGRI , National Institutes of Health. The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the Centers for Disease Control and Prevention, the National Institutes of Health, or the California Department of Public Health.

Disclaimer

Dr. Richard Finnell formerly held a leadership position in TeratOmic Consulting LLC, a now defunct organization. Further, he serves on the editorial board of the journal Reproductive and Developmental Medicine and receives travel funding to attend quarterly editorial board meetings.

Data Availability Statement

The data are from a CDC-funded multi-site study and not publicly available at this time due to privacy or ethical restrictions.

References

Bardakjian TM, & Schneider A (2011). The genetics of anophthalmia and microphthalmia. Current Opinion in Ophthalmology, 22(5), 309–313. doi:10.1097/ICU.0b013e328349b004 [PubMed: 21825993]

Bergman JE, Janssen N, van der Sloot AM, de Walle HE, Schoots J, Rendtorff ND, … Hofstra RM (2012). A novel classification system to predict the pathogenic effects of CHD7 missense variants in CHARGE syndrome. Human Mutation, 33(8), 1251–1260. doi:10.1002/humu.22106 [PubMed: 22539353]

Blazer S, Zimmer EZ, Mezer E, & Bronshtein M (2006). Early and late onset fetal microphthalmia. American Journal of Obstetrics and Gynecology, 194(5), 1354–1359. doi:10.1016/j.ajog.2005.11.010 [PubMed: 16647921]

Carter H, Douville C, Stenson PD, Cooper DN, & Karchin R (2013). Identifying Mendelian disease genes with the variant effect scoring tool. BMC Genomics, 14 Suppl 3, S3. doi:10.1186/1471-2164-14-S3-S3

Chao R, Nevin L, Agarwal P, Riemer J, Bai X, Delaney A, … Slavotinek A (2010). A male with unilateral microphthalmia reveals a role for TMX3 in eye development. PloS One, 5(5), e10565. doi:10.1371/journal.pone.0010565 [PubMed: 20485507]

Dash S, Brastrom LK, Patel SD, Scott CA, Slusarski DC, & Lachke SA (2020). The master transcription factor SOX2, mutated in anophthalmia/microphthalmia, is post-transcriptionally regulated by the conserved RNA-binding protein RBM24 in vertebrate eye development. Human Molecular Genetics, 29(4), 591–604. doi:10.1093/hmg/ddz278 [PubMed: 31814023]

Deml B, Reis LM, Lemyre E, Clark RD, Kariminejad A, & Semina EV (2016). Novel mutations in PAX6, OTX2 and NDP in anophthalmia, microphthalmia and coloboma. European Journal of Human Genetics, 24(4), 535–541. doi:10.1038/ejhg.2015.155 [PubMed: 26130484]

Dong C, Wei P, Jian X, Gibbs R, Boerwinkle E, Wang K, & Liu X (2015). Comparison and integration of deleteriousness prediction methods for nonsynonymous SNVs in whole exome sequencing studies. Human Molecular Genetics, 24(8), 2125–2137. doi:10.1093/hmg/ddu733 [PubMed: 25552646]

Du J, Aleff RA, Soragni E, Kalari K, Nie J, Tang X, … Wieben ED (2015). RNA toxicity and missplicing in the common eye disease fuchs endothelial corneal dystrophy. Journal of Biological Chemistry, 290(10), 5979–5990. doi:10.1074/jbc.M114.621607 [PubMed: 25593321]

Ergun SG, Akay GG, Ergun MA, & Percin EF (2017). LRP5-linked osteoporosis-pseudoglioma syndrome mimicking isolated microphthalmia. European Journal of Medical Genetics, 60(3), 200–204. doi:10.1016/j.ejmg.2017.01.007 [PubMed: 28111184]

Fantes J, Ragge NK, Lynch SA, McGill NI, Collin JR, Howard-Peebles PN, … FitzPatrick DR (2003). Mutations in SOX2 cause anophthalmia. Nature Genetics, 33(4), 461–463. doi:10.1038/ng1120 [PubMed: 12612584]

Garcia-Llorca A, Aspelund SG, Ogmundsdottir MH, Steingrimsson E, & Eysteinsson T (2019). The microphthalmia-associated transcription factor (Mitf) gene and its role in regulating eye function. Scientific Reports, 9(1), 15386. doi:10.1038/s41598-019-51819-0 [PubMed: 31659211]

Glaser T, Jepeal L, Edwards JG, Young SR, Favor J, & Maas RL (1994). PAX6 gene dosage effect in a family with congenital cataracts, aniridia, anophthalmia and central nervous system defects. Nature Genetics, 7(4), 463–471. doi:10.1038/ng0894-463 [PubMed: 7951315]

Graw J (2019). Mouse models for microphthalmia, anophthalmia and cataracts. Human Genetics, 138(8–9), 1007–1018. doi:10.1007/s00439-019-01995-w [PubMed: 30919050]

Harding P, Brooks BP, FitzPatrick D, & Moosajee M (2020). Anophthalmia including next-generation sequencing-based approaches. European Journal of Human Genetics, 28(3), 388–398. doi:10.1038/s41431-019-0479-1 [PubMed: 31358957]

Harding P, & Moosajee M (2019). The molecular basis of human anophthalmia and microphthalmia. Journal of Developmental Biology, 7(3). doi:10.3390/jdb7030016

Haug P, Koller S, Maggi J, Lang E, Feil S, Wlodarczyk A, … Berger W (2021). Whole exome sequencing in coloboma/microphthalmia: Identification of novel and recurrent variants in seven genes. Genes (Basel), 12(1). doi:10.3390/genes12010065

Jenkins MM, Almli LM, Pangilinan F, Chong JX, Blue EE, Shapira SK, … National Birth Defects Prevention Study (2019). Exome sequencing of family trios from the National Birth Defects Prevention Study: Tapping into a rich resource of genetic and environmental data. Birth Defects Research, 111(20), 1618–1632. doi:10.1002/bdr2.1554 [PubMed: 31328417]

Kircher M, Witten DM, Jain P, O'Roak BJ, Cooper GM, & Shendure J (2014). A general framework for estimating the relative pathogenicity of human genetic variants. Nature Genetics, 46(3), 310–315. doi:10.1038/ng.2892 [PubMed: 24487276]

Kuleshov MV, Jones MR, Rouillard AD, Fernandez NF, Duan Q, Wang Z, … Ma'ayan A (2016). Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. Nucleic Acids Research, 44(W1), W90–97. doi:10.1093/nar/gkw377 [PubMed: 27141961]

Kumar S, Sandell LL, Trainor PA, Koentgen F, & Duester G (2012). Alcohol and aldehyde dehydrogenases: retinoid metabolic effects in mouse knockout models. Biochimica et Biophysica Acta, 1821(1), 198–205. doi:10.1016/j.bbalip.2011.04.004 [PubMed: 21515404]

Lek M, Karczewski KJ, Minikel EV, Samocha KE, Banks E, Fennell T, … Exome Aggregation, C. (2016). Analysis of protein-coding genetic variation in 60,706 humans. Nature, 536(7616), 285–291. doi:10.1038/nature19057 [PubMed: 27535533]

Li G, Jin D, & Zhong TP (2019). Tubgcp3 is required for retinal progenitor cell poliferation during zebrafish development. Frontiers in Molecular Neuroscience, 12, 126. doi:10.3389/fnmol.2019.00126 [PubMed: 31178691]

Li H, & Durbin R (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics, 25(14), 1754–1760. doi:10.1093/bioinformatics/btp324 [PubMed: 19451168]

Lu Y, Shiau F, Yi W, Lu S, Wu Q, Pearson JD, … Clark BS (2020). Single-cell analysis of human retina identifies evolutionarily conserved and species-specific mechanisms controlling development. Developmental Cell, 53(4), 473–491 e479. doi:10.1016/j.devcel.2020.04.009 [PubMed: 32386599]

Mai CT, Isenburg JL, Canfield MA, Meyer RE, Correa A, Alverson CJ, … National Birth Defects Prevention Network (2019). National population-based estimates for major birth defects, 2010–2014. Birth Defects Research, 111(18), 1420–1435. doi:10.1002/bdr2.1589 [PubMed: 31580536]

Matias-Perez D, Garcia-Montano LA, Cruz-Aguilar M, Garcia-Montalvo IA, Nava-Valdez J, Barragan-Arevalo T, … Zenteno JC (2018). Identification of novel pathogenic variants and novel gene-phenotype correlations in Mexican subjects with microphthalmia and/or anophthalmia by next-generation sequencing. Journal of Human Genetics, 63(11), 1169–1180. doi:10.1038/s10038-018-0504-1 [PubMed: 30181649]

Meijer IA (1993). VPS13D movement disorder. In Adam MP, Ardinger HH, Pagon RA, Wallace SE, Bean LJH, Mirzaa G, & Amemiya A (Eds.), GeneReviews((R)). Seattle (WA).

Miller DT, Lee K, Chung WK, Gordon AS, Herman GE, Klein TE, … Group, A. S. F. W. (2021). ACMG SF v3.0 list for reporting of secondary findings in clinical exome and genome sequencing:

a policy statement of the American College of Medical Genetics and Genomics (ACMG). Genetics in Medicine, 23(8), 1381–1390. doi:10.1038/s41436-021-01172-3 [PubMed: 34012068]

Pandurangan AP, Ochoa-Montano B, Ascher DB, & Blundell TL (2017). SDM: a server for predicting effects of mutations on protein stability. Nucleic Acids Research, 45(W1), W229–W235. doi:10.1093/nar/gkx439 [PubMed: 28525590]

Papakostas TD, Morrison MA, Lane AM, Awh C, DeAngelis MM, Gragoudas ES, & Kim IK (2018). Genetic risk factors for radiation vasculopathy. Investigative Ophthalmology and Visual Science, 59(3), 1547–1553. doi:10.1167/iovs.17-22791 [PubMed: 29625478]

Patel A, & Sowden JC (2019). Genes and pathways in optic fissure closure. Seminars in Cell and Developmental Biology, 91, 55–65. doi:10.1016/j.semcdb.2017.10.010 [PubMed: 29198497]

Plaisancie J, Ceroni F, Holt R, Zazo Seco C, Calvas P, Chassaing N, & Ragge NK (2019). Genetics of anophthalmia and microphthalmia. Part 1: Non-syndromic anophthalmia/microphthalmia. Human Genetics, 138(8–9), 799–830. doi:10.1007/s00439-019-01977-y [PubMed: 30762128]

Rasmussen SA, Olney RS, Holmes LB, Lin AE, Keppler-Noreuil KM, Moore CA, & National Birth Defects Prevention Study (2003). Guidelines for case classification for the National Birth Defects Prevention Study. Birth Defects Research. Part A: Clinical and Molecular Teratology, 67(3), 193–201. doi:10.1002/bdra.10012 [PubMed: 12797461]

Reefhuis J, Gilboa SM, Anderka M, Browne ML, Feldkamp ML, Hobbs CA, … National Birth Defects Prevention Study (2015). The National Birth Defects Prevention Study: A review of the methods. Birth Defects Research. Part A: Clinical and Molecular Teratology, 103(8), 656–669. doi:10.1002/bdra.23384 [PubMed: 26033852]

Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, … Committee, A. L. Q. A. (2015). Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genetics in Medicine, 17(5), 405–424. doi:10.1038/gim.2015.30 [PubMed: 25741868]

Richardson R, Tracey-White D, Webster A, & Moosajee M (2017). The zebrafish eye-a paradigm for investigating human ocular genetics. Eye (London, England), 31(1), 68–86. doi:10.1038/eye.2016.198

Rodrigues CH, Pires DE, & Ascher DB (2018). DynaMut: predicting the impact of mutations on protein conformation, flexibility and stability. Nucleic Acids Research, 46(W1), W350–W355. doi:10.1093/nar/gky300 [PubMed: 29718330]

Shaw GM, Carmichael SL, Yang W, Harris JA, Finnell RH, & Lammer EJ (2005). Epidemiologic characteristics of anophthalmia and bilateral microphthalmia among 2.5 million births in California, 1989–1997. American Journal of Medical Genetics. Part A, 137(1), 36–40. doi:10.1002/ajmg.a.30840 [PubMed: 16007635]

Shimizu M, Goto T, Sato A, & Shibuya H (2013). WNK4 is an essential effector of anterior formation in FGF signaling. Genes to Cells, 18(6), 442–449. doi:10.1111/gtc.12048 [PubMed: 23517227]

Sinn R, & Wittbrodt J (2013). An eye on eye development. Mechanisms of Development, 130(6–8), 347–358. doi:10.1016/j.mod.2013.05.001 [PubMed: 23684892]

Skalicky SE, White AJ, Grigg JR, Martin F, Smith J, Jones M, … Jamieson RV (2013). Microphthalmia, anophthalmia, and coloboma and associated ocular and systemic features: understanding the spectrum. JAMA Ophthalmol, 131(12), 1517–1524. doi:10.1001/jamaophthalmol.2013.5305 [PubMed: 24177921]

Tsai IC, Adams KA, Tzeng JA, Shennib O, Tan PL, & Katsanis N (2019). Genome-wide suppressor screen identifies USP35/USP38 as therapeutic candidates for ciliopathies. JCI Insight, 4(22). doi:10.1172/jci.insight.130516

Verma AS, & Fitzpatrick DR (2007). Anophthalmia and microphthalmia. Orphanet Journal of Rare Diseases, 2, 47. doi:10.1186/1750-1172-2-47 [PubMed: 18039390]

Wang K, Li M, & Hakonarson H (2010). ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Research, 38(16), e164. doi:10.1093/nar/gkq603 [PubMed: 20601685]

Williamson KA, & FitzPatrick DR (2014). The genetic architecture of microphthalmia, anophthalmia and coloboma. European Journal of Medical Genetics, 57(8), 369–380. doi:10.1016/j.ejmg.2014.05.002 [PubMed: 24859618]

Williamson KA, Hall HN, Owen LJ, Livesey BJ, Hanson IM, Adams GGW, … FitzPatrick DR (2020). Recurrent heterozygous PAX6 missense variants cause severe bilateral microphthalmia via predictable effects on DNA-protein interaction. Genetics in Medicine, 22(3), 598–609. doi:10.1038/s41436-019-0685-9 [PubMed: 31700164]

Wolf MT, Lorenz B, Winterpacht A, Drechsler M, Schumacher V, Royer-Pokora B, … Wildhardt G (1998). Ten novel mutations found in aniridia. Human Mutation, 12(5), 304–313. doi:10.1002/(SICI)1098-1004(1998)12:5<304::AID-HUMU3>3.0.CO;2-D [PubMed: 9792406]

Zagozewski JL, Zhang Q, & Eisenstat DD (2014). Genetic regulation of vertebrate eye development. Clinical Genetics, 86(5), 453–460. doi:10.1111/cge.12493 [PubMed: 25174583]

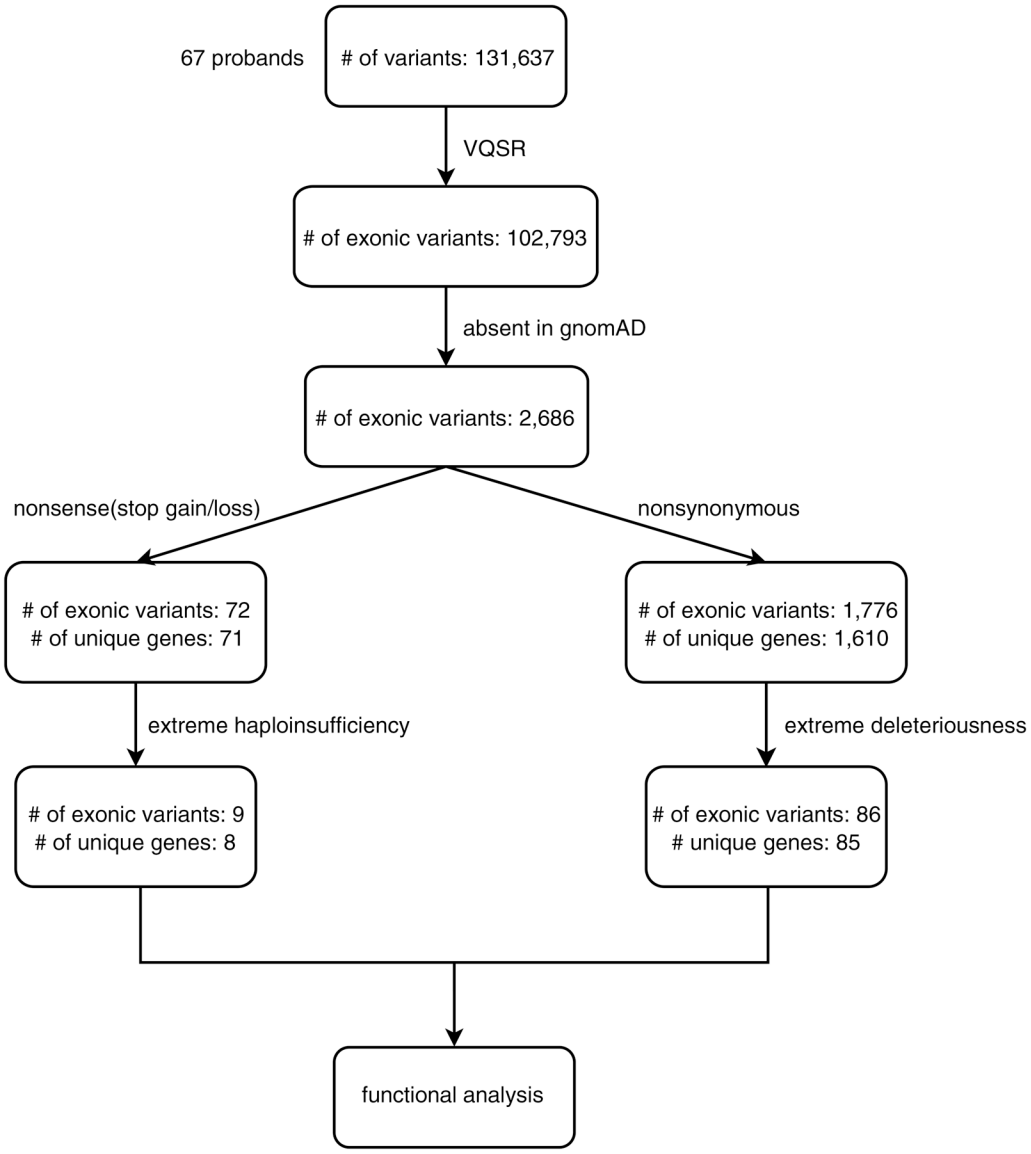


Figure 1. Analysis schematic.

Overview of the approach to identifying candidate genes from 67 probands. Following GATK Best Practice, we applied the variant quality score recalibration (VQSR) procedure to refine the raw variants from the initial genotype calls. To define rare variants, we considered those not seen in the Genome Aggregation Database (gnomAD).

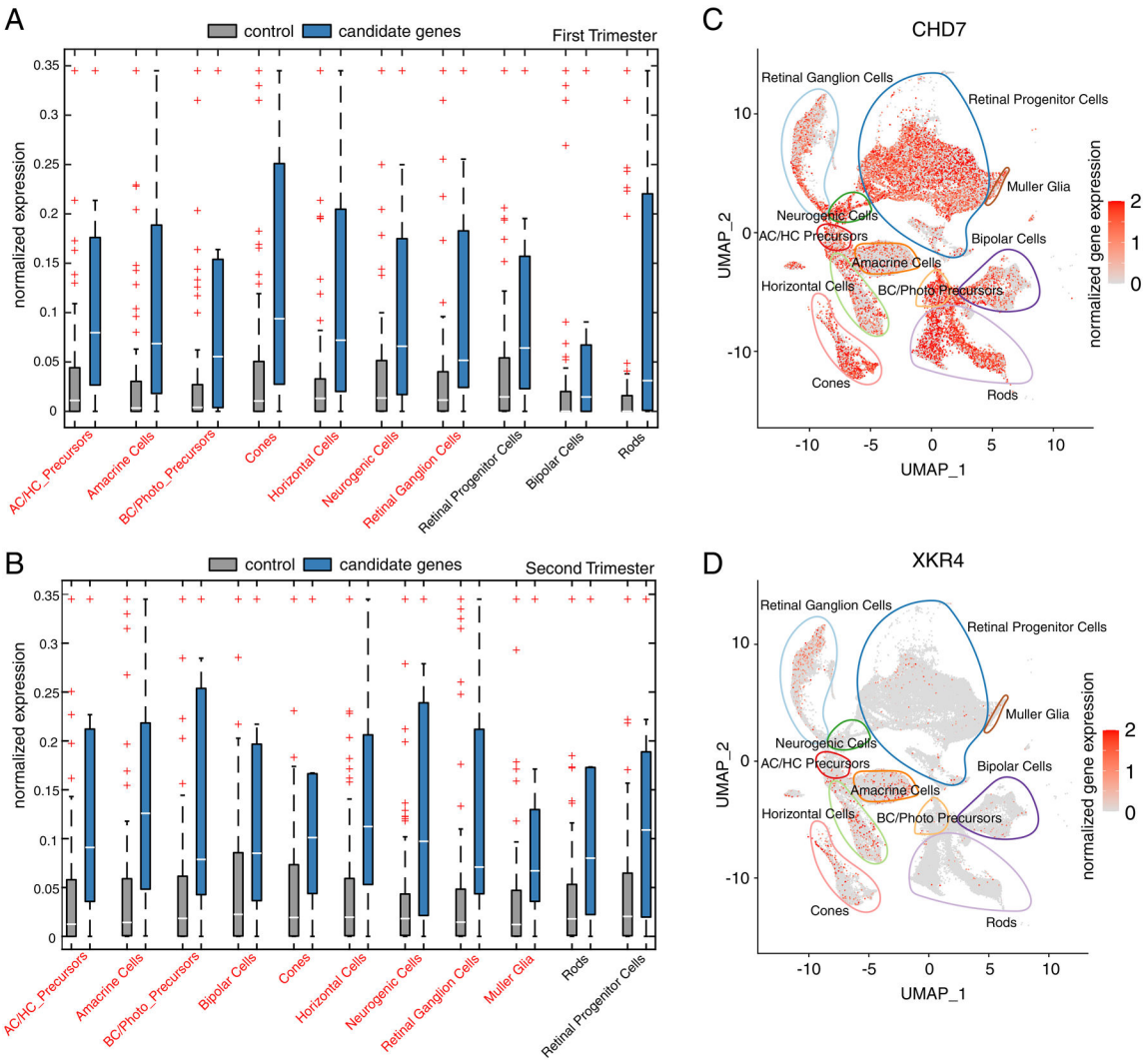


Figure 2. Single-cell transcriptome data analysis of candidate genes across retina developmental stages.

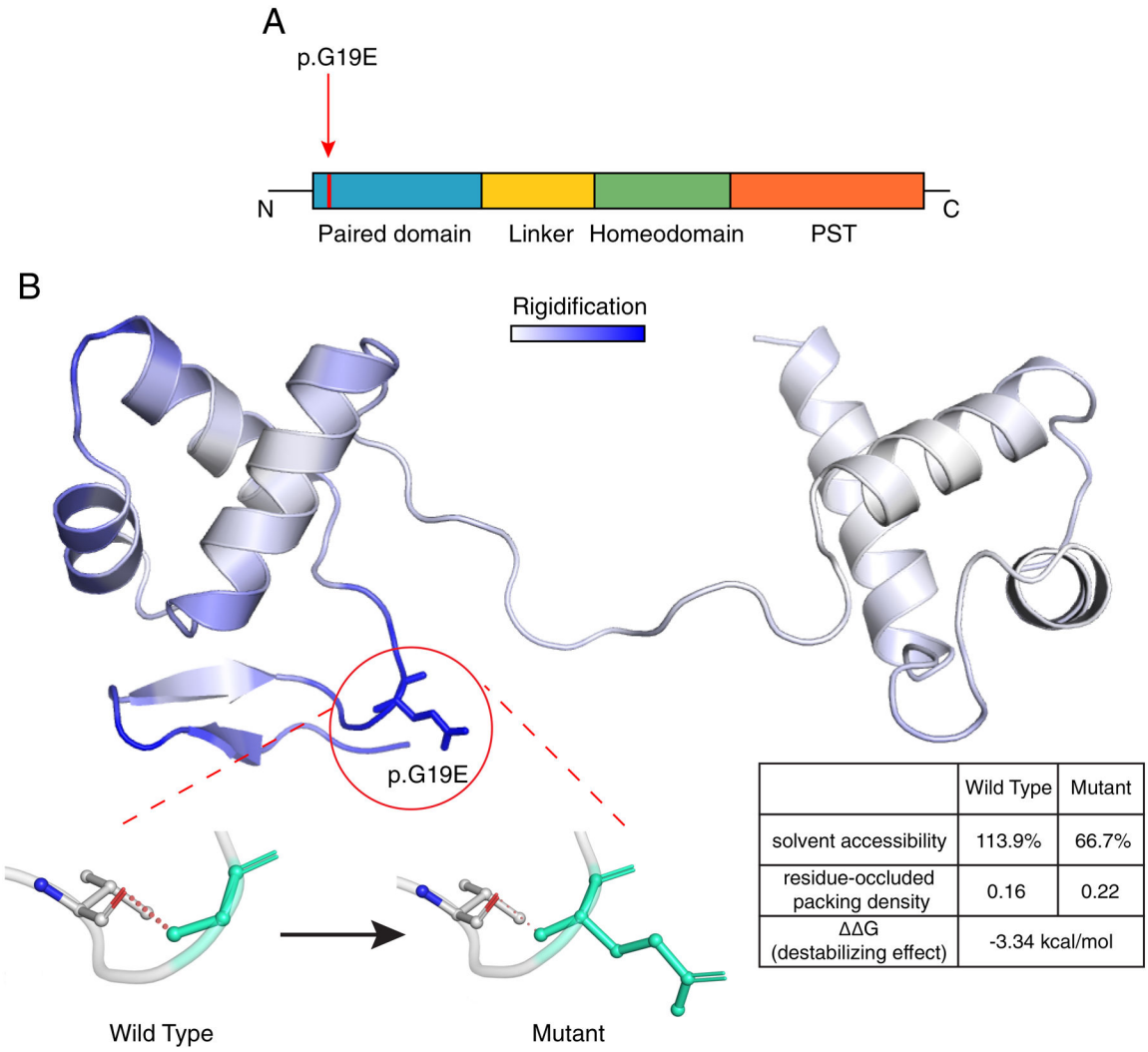
a. In the first trimester, the identified eight genes displayed significantly increased expression relative to the background control genes in AC/HC_Precursors (N= 695), Amacrine Cells (N= 253), BC/Photo_Precursors, Cones (N= 689), Horizontal Cells (N= 1,883), Neurogenic Cells (N= 1,123) and Retinal Ganglion Cells (N= 6,297). The p values were calculated by Wilcoxon rank-sum test, followed by Benjamini-Hochberg correction. Cell types in red font along the x-axis had a false discovery rate (FDR) <0.05. Abbreviations: AC/HC, amacrine cell/horizontal cell; BC/Photo, bipolar cell/photoreceptor.

b. In the second trimester, the identified eight genes displayed significantly increased expression relative to the background control genes in AC/HC_Precursors (N= 1,039), Amacrine Cells (N= 10,706), BC/Photo_Precursors (N= 1,825), Bipolar Cells (N= 5,949), Cones (N= 3,985), Horizontal Cells (N= 4,618), Neurogenic Cells (N= 2,197), Retinal Ganglion Cells (N= 2,662) and Muller Glia (N= 343). The p values were calculated by Wilcoxon rank-sum test, followed by Benjamini-Hochberg correction. Cell types in red font along the x-axis had a false discovery rate (FDR) <0.05.

a. & b. The bottom and top of the boxes denote the first and third quartiles, respectively. The whiskers indicate the minimal value within 1.5 interquartile range (IQR) of the lower quartile and the maximum value within 1.5 IQR of the upper quartile. The plus symbols represent outliers. The black dashed lines indicate the lower and upper limits of the regions with regular scale. Outliers outside of the black dashed lines are visualized with compressed scale in the regions surrounded by gray lines for better visualization.

c. *CHD7* is ubiquitously expressed across all retinal cell types, despite a stronger pattern in retinal progenitor cells, cones and BC/photo precursors. Red and grey colors represent high and low read number in *CHD7* expression, respectively. Uniform manifold approximation and projection (UMAP) is an efficient dimension reduction algorithm commonly used in single-cell RNA sequencing analysis.

d. XKR4 displayed strong specificity in the developing retina only in very few cell types, including retinal ganglion cells, amacrine cells, horizontal cells, and cones. Red and grey colors represent high and low XKR4 expression, respectively.



	Wild Type	Mutant
solvent accessibility	113.9%	66.7%
residue-occluded packing density	0.16	0.22
ΔΔG (destabilizing effect)	-3.34 kcal/mol	

Figure 3. Protein structure analysis of a potential deleterious missense variant in *PAX6*.

a. The gene structure of transcription factor *PAX6*. The potential pathogenic variant (p.G19E) altered the wildtype amino acid from a glycine (G) residue into a glutamic acid (E) in the paired DNA binding domain of *PAX6*. PST represents a carboxyl-terminal transactivation domain rich of proline (P), serine (S), and threonine (T).

b. The p.G19E potential pathogenic variant of PAX6 had a significant effect in loss of molecular flexibility in the N-terminus. 3D structure of mutated PAX6, blue and white colors stand for high and low rigidification. The zoomed in region indicates the differences between wild type and mutant in mutated residue location, red represents hydrogen bonds and light-green represents residues. The table shows the prediction outcomes of interatomic interactions.

Table1.

Descriptive characteristics (percentages) [a] of sequenced anophthalmia/microphthalmia probands, National Birth Defects Prevention Study, 1997–2011

	Total (N=67)
Infant sex	
Male	36 (53.7)
Female	31 (46.3)
Gestational age (weeks)	
Preterm (< 37)	15 (22.4)
Term (≥ 37)	52 (77.6)
Plurality	
Singleton	63 (94.0)
Multiple	4 (6.0)
Phenotypes	
Both	1 (1.5)
Anophthalmia	5 (7.5)
Microphthalmia	61 (91.0)
Isolated/with accompanying malformations	
Isolated	55 (82.1)
With accompanying malformations	12 (17.9)
Laterality	
Unilateral	48 (71.6)
Bilateral	19 (28.4)
Maternal age at delivery (years)	
< 25	22 (32.8)
25–29	22 (32.8)
≥ 30	23 (34.3)
Maternal race/ethnicity	
Non-Hispanic White	41 (61.2)
Non-Hispanic Black	3 (4.5)
Hispanic	19 (28.4)
Other	4 (6.0)
Paternal age at delivery (years)	
< 25	14 (20.9)
25–29	20 (29.9)
≥ 30	32 (47.8)
Paternal race/ethnicity	
Non-Hispanic White	40 (59.7)
Non-Hispanic Black	4 (6.0)
Hispanic	18 (26.9)
Other	3 (4.5)

[a]Sums of percentages may not equal 100 owing to rounding.

Table2.

Nonsense variants (stopgain) in probands with anophthalmia/microphthalmia identified by exome sequencing

Gene	[Chromosome: Position]	[REF/ALT]	Predicted protein change	PMID	Eye development function	Origin	Family ID	Phenotype
VPS13D [a]	1:12557600	C/T	p.Q4212X	30789691	Eye movement abnormalities	*De novo*	#46	Unilateral microphthalmia with accompanying malformations
XKR4	8:56015062	C/G	p.S5X	29625478	Mutation showed increase risk of poor vision	Unknown	#54	Unilateral microphthalmia, isolated
CHD7 [b]	8:61735063	C/T	p.R987X	22539353	Known in A/M	*De novo*	#55	Unilateral microphthalmia, isolated
CHD7	8:61750761	C/T	p.R1494X	22539353	Known in A/M	Unknown	#20	Unilateral microphthalmia with accompanying malformations
LRP5	11:68191161	C/T	p.R497X	28111184	Retinal development	Paternal	#17	Bilateral microphthalmia, isolated
NUMA1	11:71726895	G/A	p.Q552X	25593321	Eye tissue development	Paternal	#45	Bilateral microphthalmia, isolated
TUBGCP3	13:113181728	G/T	p.S459X	31178691	Retinal progenitor cell proliferation; Knockout showed small eye phenotype	Paternal	#16	Unilateral microphthalmia, isolated
FAM192A	16:57206739	G/A	p.Q59X	The Human Protein Atlas	Expression in the retina	Maternal	#64	Unilateral microphthalmia, isolated
KRT31	17:39553461	C/A	p.E111X	International Mouse Phenotyping Consortium	Eye morphology	Maternal	#61	Unilateral microphthalmia, isolated

Abbreviations: ALT, alternate; A/M, anophthalmia/microphtalmia; PMID, PubMed unique identifier; REF, reference.

[a] *de novo* not confirmed by Sanger sequencing (proband result was confirmed, but parent samples were not available)

[b] *de novo* confirmed by Sanger sequencing

Table3a.

Functional enrichment summary for missense variants

Function	FDR (candidate genes)	FDR (control)
Detection of visible light (GO:0009584)	8.82e-3	0.25
Visual perception (GO:0007601)	8.66e-3	0.28
Abnormal horizontal cell morphology (MP:0006068)	0.05	0.72
Blindness (MP:0002001)	0.048	0.76
Hypoplasia of the fovea (HP:0007750)	7.2e-4	0.29
Aplasia/hypoplasia of the macula (HP:0008059)	1.51e-3	0.41
Aplasia/hypoplasia affecting the retina (HP:0008061)	1.53e-3	0.44
Opacification of the corneal stroma (HP:0007759)	1.53e-3	0.68
Optic atrophy (HP:0000648)	1.59e-3	0.55

Abbreviations: FDR, false discovery rate; GO, gene ontology; HP, human phenotype; MP, mammalian phenotype

Table3b.

Selected missense variants in probands with anophthalmia/microphthalmia identified by exome sequencing

Gene	[Chromosome:Position]	[REF/ALT]	Predicted protein change	PMID	Eye development function	Family ID	Origin	Phenotype
AGRN	1:985283	G/C	p.G1582A	17196957;28221305	Overexpression showed eye development defects; Mutant showed restrictive eye.	#42	Maternal	Unilateral microphthalmia, isolated
USP38	4:144135410	C/A	p.L761I	31723061	Mutant showed loss of eyes	#60	Paternal	Unilateral microphthalmia with accompanying malformations
WASHC5	8:126068971	C/A	p.R507L	31971710	Eye anomalies in Ritscher-Schinzel syndrome	#62	Paternal	Unilateral anophthalmia with accompanying malformations
PAX6	11:31824340	C/T	p.G19E	7951315	A key regulator in eye development	#59	Maternal	Bilateral microphthalmia, isolated
COL4A1	13:110829240	C/T	p.G954E	24628545	Known in A/M	#36	*De novo*	Bilateral microphthalmia with accompanying malformations
WNK4	17:40936497	A/G	p.T51A	23517227	Mutant showed loss of eyes in Xenopus embryos	#11	Maternal	Unilateral microphthalmia, isolated

Abbreviations: ALT, alternate; A/M, anophthalmia/microphthalmia; PMID, PubMed unique identifier; REF, reference.