Exome Sequence Analysis of 14 Families With High Myopia

Bethany A. Kloss,[1] Stuart W. Tompson,[1] Kristina N. Whisenhunt,[1] Krystina L. Quow,[2] Samuel J. Huang,[1] Derek M. Pavelec,[3] Thomas Rosenberg,[4,5] and Terri L. Young[1]

[1]Department of Ophthalmology and Visual Sciences, University of Wisconsin, Madison, Wisconsin, United States
[2]Center for Human Genetics, Duke University Medical Center, Durham, North Carolina, United States
[3]Biotechnology Center, University of Wisconsin, Madison, Wisconsin, United States
[4]The National Eye Clinic, Rigshospitalet, Kennedy Center, Glostrup, Denmark
[5]Institute of Clinical Medicine, University of Copenhagen, Denmark

Correspondence: Terri L. Young, Department of Ophthalmology and Visual Sciences, University of Wisconsin, 2828 Marshall Court, Suite 200, Madison, WI 53705, USA; tyoung6@wisc.edu.

Submitted: October 6, 2016
Accepted: February 21, 2017

Citation: Kloss BA, Tompson SW, Whisenhunt KN, et al. Exome sequence analysis of 14 families with high myopia. *Invest Ophthalmol Vis Sci*. 2017;58:1982–1990. DOI: 10.1167/iovs.16-20883

Purpose. To identify causal gene mutations in 14 families with autosomal dominant (AD) high myopia using exome sequencing.

Methods. Select individuals from 14 large Caucasian families with high myopia were exome sequenced. Gene variants were filtered to identify potential pathogenic changes. Sanger sequencing was used to confirm variants in original DNA, and to test for disease cosegregation in additional family members. Candidate genes and chromosomal loci previously associated with myopic refractive error and its endophenotypes were comprehensively screened.

Results. In 14 high myopia families, we identified 73 rare and 31 novel gene variants as candidates for pathogenicity. In seven of these families, two of the novel and eight of the rare variants were within known myopia loci. A total of 104 heterozygous nonsynonymous rare variants in 104 genes were identified in 10 out of 14 probands. Each variant cosegregated with affection status. No rare variants were identified in genes known to cause myopia or in genes closest to published genome-wide association study association signals for refractive error or its endophenotypes.

Conclusions. Whole exome sequencing was performed to determine gene variants implicated in the pathogenesis of AD high myopia. This study provides new genes for consideration in the pathogenesis of high myopia, and may aid in the development of genetic profiling of those at greatest risk for attendant ocular morbidities of this disorder.

Keywords: high myopia, candidate genes, variants, exome sequencing, GWAS

Myopic refractive error, also known as short-sightedness, is the most common eye disorder worldwide. The extreme form, high myopia, is defined as a refractive error of –6.00 diopters (D) or higher. High myopia affects 2% to 5% of Americans, Western Europeans, and Australians aged older than 40 years.[1] In East Asian countries, the prevalence of high myopia has been reported as low as 1.8% to 5% in Chinese adults aged older than 30 years,[2,3] while higher rates are seen in Singaporean children aged 7 to 9 years (14.6%–17.9%)[4] and university students in Taiwan (up to 38%).[5]

High myopia is characterized by axial elongation and thinning of the sclera,[6,7] with predisposition to additional ocular morbidities such as retinal detachment, glaucoma, and premature cataracts.[1,3,8] Recently, exome sequencing of large families with high myopia has identified mutations in the genes *LEPREL1/P3H2* and *SCO2*.[9,10] The disease caused by mutations in *LEPREL1/P3H2* followed an autosomal recessive inheritance pattern.[9,11] Genome-wide association studies (GWAS) have identified over 60 candidate genes associated with refractive error.[12,13] Familial linkage studies have also mapped 23 myopia (MYP) loci associated with high myopia. Significant linkage was first reported in eight families with autosomal dominant (AD) high myopia on chromosome 18p11.31.[14,15] Since then, additional loci have been mapped on several chromosomes (Table 1). Known myopia genes, myopia-associated genes identified by GWAS, and MYP chromosomal regions have been extensively reviewed.[38–40]

Recently, Holden et al.[41] estimated an increase in the prevalence of high myopia (greater than –5.00 D) from 163 million people (2.7% global population) in 2000 to 938 million people (9.8% global population) by 2050. This estimation utilized published data from 145 studies since 1995, covering 2.1 million participants. Vitale et al.[42] also noted an 8-fold increase in high myopia (greater than –7.9 D) in the United States over a 30-year period from 1971 to 1972 to 1999 to 2004.[42] Increased prevalence of myopia was observed regardless of age, race, sex, severity of the phenotype, or years of formal education. While the prevalence of high myopia may be increasing worldwide, the condition is still relatively uncommon with many unanswered questions. Increased availability of formal education and decreased exposure to rural or outdoor environments, compared to earlier decades, may partially account for an earlier onset of a myopic phenotype when combined with genetic susceptibility. Genetic analysis of high myopia within families such as those presented in this study may provide further insight into the intersecting contributions of biology with environment.

In this study, we sought to identify causal gene mutations in our high myopia cohort by examining known myopia genes, susceptibility loci identified through published GWAS data and chromosomal regions within the reported MYP loci coordi-

Table 1. Chromosomal Regions Associated With Myopia

Locus	Chromosomal Location	Reference
MYP1	Xq28	Ratnamala et al.[16]
MYP2	18p11.31	Young et al.[14]; Young et al.[15]
MYP3	12q21-q23	Young et al.[17]
MYP5	17q21–q22	Paluru et al.[18]
MYP6	22q12	Stambolian et al.[19]; Klein et al.[20]
MYP7	11p13	Hammond et al.[21]
MYP8	3q26	Hammond et al.[21]
MYP9	4q12	Hammond et al.[21]
MYP10	8p23	Hammond et al.[21]
MYP11	4q22-q27	Zhang et al.[22]
MYP12	2q37.1	Paluru et al.[23]
MYP13	Xq23-q27.2	Zhang et al.[24]; Zhang et al.[25]
MYP14	1p36	Wojciechowski et al.[26]
MYP15	10q21.1	Nallasamy et al.[27]
MYP16	5p15.33-p15.2	Lam et al.[28]
MYP17/MYP4	7p15	Paget et al.[29]; Ciner et al.[30]
MYP18	14q22.1-q24.2	Yang et al.[31]
MYP19	5p15.1-p13.3	Ma et al.[32]
MYP20	13q12.12	Shi et al.[33]
MYP21	1p22.2	Shi et al.[34]
MYP22	4q35	Zhao et al.[35]
MYP23	4p16.3	Aldahmesh et al.[36]
MYP24	12q13	Guo et al.[37]

nates. Selected members from 14 nonsyndromic high myopia families underwent exome sequencing. As each family pedigree appeared to demonstrate AD inheritance transmission, it was hypothesized that each family would carry a single heterozygous private/rare mutation that caused the disease. We also employed a filtering strategy to identify additional rare and conserved variants, thereby expanding the list of candidate genes associated with high myopia. Determining causative genetic and therefore biologic factors for high myopia development will inform more effective strategies to reduce or minimize the likelihood of developing blindness due to associated ocular risks, and improve quality of life for those affected.

Materials and Methods

Human Subjects

We recruited 14 families with high myopia when DNA could be obtained from at least three living affected relatives (Supplementary Figs. S1–S14). Informed consent was obtained from all participants, with approval by the Institutional Review Board at the University of Wisconsin (Madison, WI, USA) according to the tenets of the Declaration of Helsinki. Spherical equivalent (SE) refractive error measurements were conducted and DNA was extracted from blood and/or saliva from participating family members. In our study, 8 of the 14 families were from Denmark, 4 families were from the United States, 1 family was of English/Canadian descent, and 1 family was of Hawaiian/Chinese descent. The families ranged in size from three individuals with high myopia up to the largest family (family 56) with 14 affected individuals. Ethnicity, SE refractive error ranges, and affection status information for each family is available in Table 2.

The affected high myopia phenotype was determined to be those with SE refractive error greater than –6.00 D. Unaffected persons were defined by having SE refractive error less than –2.00 D. Those refractive errors reported by hearsay are indicated in gray in the pedigree figures. They represent primarily deceased and older generations who were not genotyped in this study: family 8, II:1, II:3, family 56, II:1, II:3, III:13, family 104, I:1, family 126 I:1, II:2, IV:7, and family 130, I:1.

Table 2. Clinical Data From 14 Families Affected With High Myopia

Family	Ethnicity	Ascertained Affected Individuals, *n*	SE Range, D	SE Median, D
8	Hawaiian/Chinese	7	−3.75*/−12.25	−7.875
19	Caucasian (US)	3	−6.875/−19.50	−17.438
56	Caucasian (US)	14	−7.25/−27.00	−14.625
68	Caucasian (English/Canadian)	7	−6.00/−50.00	−7.75
81	Caucasian (US)	4	−4.75*/−24.75	−18.00
101	Danish	5	−8.75/−17.00	−15.00
104†	Danish	3	−9.75/−17.00	−13.875
110†	Danish	5	−5.25*/−16.00	−11.25
111†	Danish	3	−9.00/−10.50	−10.00
115	Danish	3	−6.00/−11.50	−10.00
122	Danish	3	−5.75*/−9.00	−7.25
126	Danish	9	−6.50/−19.50	−10.625
130	Danish	4	−7.00/−10.75	−8.375
155	Caucasian (U.S.)	4	−13.75/−19.25	−17.50

* SE lower than –6.00 D reported in one eye.
† Some refractions in this family are self-reported by patients.

Refractive errors were self-reported for the following patients: family 104, individual I:1 (high myopia in both eyes, cataract removed without intraocular lens placement in the left eye to mitigate myopia); and family 111, individuals I:2, II:1, II:2 (refractive errors obtained prior to LASIK corneal surgery to mitigate their myopia). These individuals were genotyped in this study.

To the best of our knowledge, none of the subjects used in this study presented with systemic disorders associated with myopia, including Marfan syndrome, Ehlers-Danlos syndrome, Stickler syndrome, or Wagner syndrome. Affected subjects had no other known ocular disorders or abnormalities, apart from family 101 where some affected family members developed glaucoma later in life. Individuals with SE refractive errors between –2.00 D and –6.00 D were excluded from the study, as they could not be considered neither affected nor unaffected. We sought to perform an extreme phenotype genotyping screening approach for this quantitative trait.

Exome Sequencing

Exome sequencing of pedigrees selected for this study were performed at three facilities. Samples sequenced by the Hudson Alpha Institute for Biotechnology (Huntsville, AL, USA) used a commercial capture kit (SeqCap EZ Exome Capture Kit v2 and v3 and 2 × 100 bp, paired-end; Roche NimbleGen, Inc., Madison, WI, USA) reads on a sequencing platform (HiSeq2000; Illumina, Inc., San Diego, CA, USA). Samples sequenced by the Center for Human Genome Variation at Duke University (Durham, NC, USA) used a commercial capture kit (Roche NimbleGen, Inc.) and a sequencing platform (HiSeq2000 or HiSeq2500; Illumina, Inc.). Sequences were aligned to the GRCh37/hg19 human genome assembly using Burrows Wheeler Aligner (BWA) and variants called using the Genome Analysis Toolkit (GATK). Samples sequenced by the University of Wisconsin Biotechnology Center DNA Sequencing Facility (Madison, WI, USA) used a commercial kit (Sure Select Exome v5 and 2 × 100 bp, paired-end; Agilent Technologies, Inc., Santa Clara, CA, USA)

Investigative Ophthalmology & Visual Science

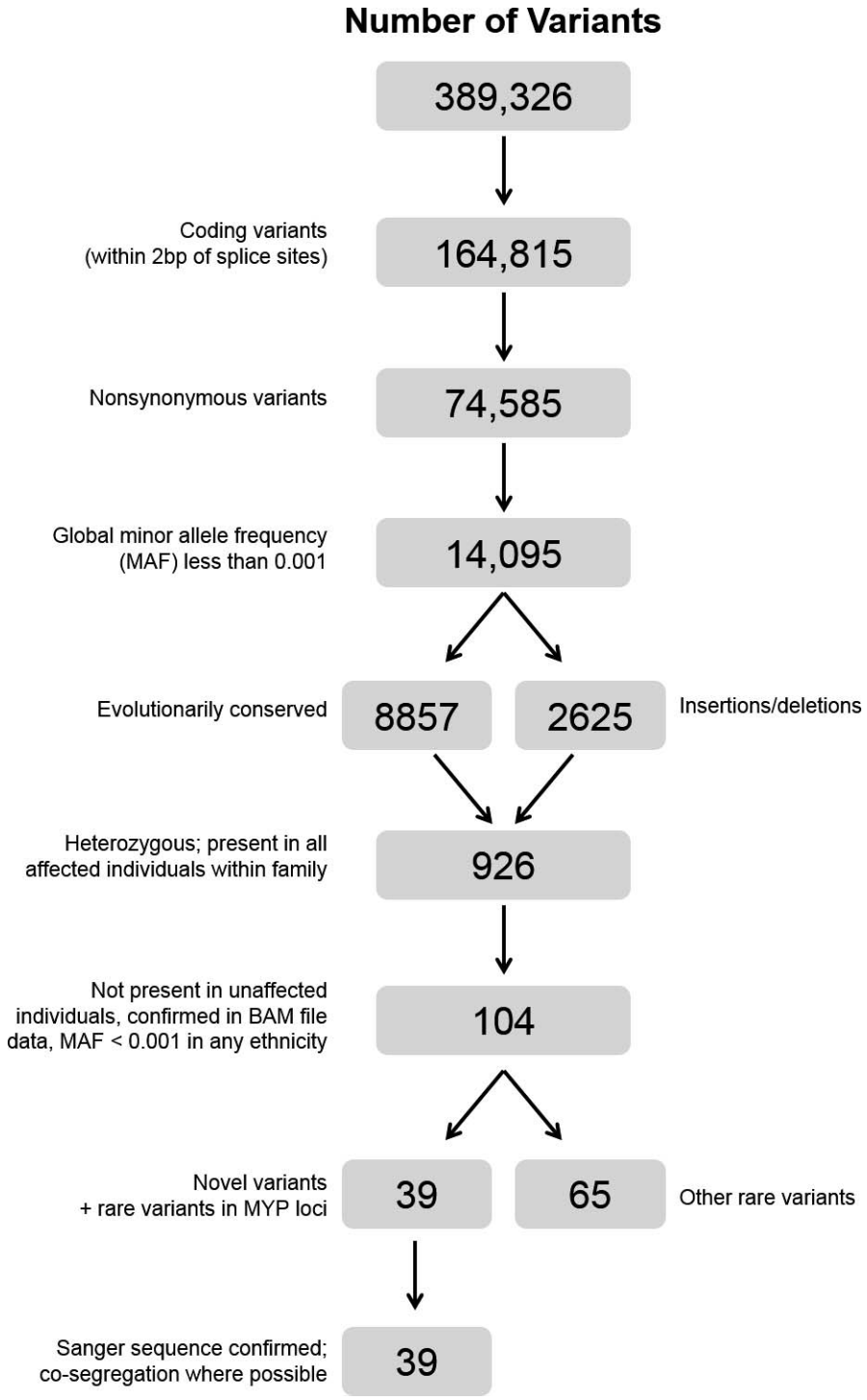

Figure. Overview of the variant filtering strategy of exome sequencing data.

reads on a sequencing platform (Illumina, Inc.). University of Wisconsin Biotechnology Center analysis of DNA sequencing data was performed as follows: Adapter sequences and low-quality sequencing reads were trimmed using Skewer.[43] We used BWA-maximal exact matches (MEM) to align all reads to the reference genome, GRCh37/hg19 assembly.[44] Duplicate alignments were then removed using Picard (http://picard.sourceforge.net, in the public domain). Finally, variants were called using GATK HaplotypeCaller version 3.3 and annotated with SNPeff.[45,46]

Variant Analysis

Single nucleotide polymorphisms (SNPs) and insertion/deletion (indel) changes were filtered and analyzed using SNP and Variation Suite Software v8.3 (Golden Helix, Bozeman, Montana). Variant filtering included the following steps (Fig.):

1. Excluded variants outside of exonic and splice-site regions (within 2 bp; RefSeq Genes 105v2, NCBI);
2. Excluded synonymous variants;

Investigative Ophthalmology & Visual Science

3. Excluded variants with global minor allele frequencies (MAF) greater than 0.001 in the Exome Aggregation Consortium (ExAC) database (ExAC Variant Frequencies 0.3 BROAD)[47];
4. With the exception of indels, excluded variants with negative conservation level across vertebrate species (PhyloP 100-way vertebrate score less than zero)[48];
5. Excluded variants that were not heterozygous in all affected individuals within each family;
6. Excluded variants present in unaffected individuals within the corresponding family, confirmed variant present in raw sequence alignment data, and excluded variants observed in the ExAC database with a MAF greater than 0.001 in any ethnicity.

Additionally, a filtering track was created for use in the Golden Helix software to retrieve all gene variants contained within the 23 MYP chromosomal regions (discussed in OMIM 160700).[49] Confirmation of variants that were novel or located within MYP loci was performed by Sanger sequencing. Cosegregation analysis was also performed when additional family members were available for screening. Primers used for screening *SCO2* were as follows: 5′-GCTTGTTTCCAGGAGCATCA-3′and 5′-TGGCTCAAGACAGGACACT-3′.

Results

After filtering of the exome data, a total of 104 heterozygous variants in 104 genes were identified in 10 families with nonsyndromic high myopia. Variants meeting our filtering criteria were not identified in four families. Each variant identified was located within the coding region or splice site of the gene, reported at a frequency of less than 1 in 1000 alleles in any population in the Exome Aggregation Consortium (ExAC) database, and present in all relevant individuals with high myopia in each family that were available for screening. None of the variants were identified in unaffected family members (SE refractive error of −2.00 D or less severe) or individuals that married into the family. Ten variants were identified within myopia-associated (MYP) chromosomal regions (Table 3). We found 31 of the 104 variants identified were novel changes unreported in genetic population databases (Table 4). Two variants were both novel and located within MYP loci. In our cohort, 73 variants identified were uncommon (less than 1 in 1000 alleles in any population; Supplementary Table S2). Confirmation of the 31 novel variants and 8 rare variants located within MYP loci was achieved by Sanger sequencing.

Discussion

In this study, exome sequencing was utilized to identify gene variants in individuals affected with nonsyndromic high myopia. The association of high myopia and specific genes, chromosomal loci, and SNPs has been previously investigated and reported in several studies. For this analysis, we sought to examine these associations within our own high myopia cohort. We did not observe more than one variant in the same gene in more than one family, and hypothesized that each family harbors its own private pathogenic mutation resulting in high myopia. Since high myopia is a rare disease (2%–5% in the global population),[1] we hypothesized that a causal variant that segregates with high myopia should also be rare. We also examined variants that were novel or observed in less than 1 in 1000 alleles in ~61,000 individuals represented in the ExAC database. We identified potentially pathogenic heterozygous

Table 3. Rare Variants Identified in Families With High Myopia in MYP Loci*

Gene	Position	Transcript	Nucleotide Change	Residue Change	Type of Variant	dbSNP	Conservation	FATHMM	SIFT	Frequency, Global	Frequency, Ethnically Matched	Frequency, Highest in Any Population	Family
AGRN	1:977462	NM_198576.3	c.1304C>T	p.Thr435Met	Nonsynonymous	Not found in dbSNP	4.31	Damaging	Damaging	0.00006173	0.0001135	0.0001135 (EUR)	130
DNAH11	7:21788257	NM_001277115.1	c.8570G>C	p.Gly2857Ala	Nonsynonymous	Not found in dbSNP	6.642	Tolerated	Damaging	0.00004179	6.03E-05	0.00006318 (SAS)	122
EME1	17:48458145	NM_001166131.1	c.1597G>T	p.Asp533Tyr	Nonsynonymous	Not found in dbSNP	5.884	Tolerated	Damaging	0.00004119	7.49E-05	0.00007493 (EUR)	115
ETNPPL/AGXT2L1	4:109667954	NM_031279.3	c.1136C>T	p.Thr379Ile	Nonsynonymous	rs143708546	7.471	Tolerated	N/A	0.0005848	0.00100400	0.001004 (EUR)	104
HOXA2†	7:27142079	NM_006735.3	c.41G>C	p.Ser14Thr	Nonsynonymous	N/A	7.415	Tolerated	Damaging	Novel	Novel	Novel	155
PLOD1	1:12032953	NM_000302.3	c.1927G>A	p.Val643Ile	Nonsynonymous	rs149425237	9.854	Tolerated	Damaging	0.0006343	0.0009892	0.0009892 (EUR)	122
PTPRQ	12:80935492	NM_001145026.1	c.2859A>G	p.Thr1105Ala	Nonsynonymous	rs186746372	3.524	Tolerated	Tolerated	0.0006584	0.001539	0.001539 (EUR)	115
TMPO†	12:98925463	NM_003276.2	c.412A>G	p.Thr138Ala	Nonsynonymous	N/A	7.895	Damaging	Damaging	Novel	Novel	Novel	111
TMTC2	12:83444755	NM_152588.1	c.2225G>A	p.Ser742Asn	Nonsynonymous	rs202237669	2.408	Tolerated	Damaging	0.0001953	0.0003387	0.0003387 (EUR)	115
TOB1	17:48940750	NM_005749.3	c.629G>C	p.Gly210Ala	Nonsynonymous	Not found in dbSNP	7.487	Tolerated	Tolerated	0.00001648	3.00E-05	0.00002997 (EUR)	101

EUR, European; N/A, data not available; SAS, South Asian.
* Chromosome position in accordance with GRCh37/hg19 assembly. Conservation is scored by PhyloP 100-way Vertebrate, in which sites predicted to be conserved are assigned positive scores.
† Variant is both novel and located within MYP locus.

Table 4. Novel Variants Identified in Families With High Myopia

Gene	Position	Transcript	Nucleotide Change	Residue Change	Type of Variant	Conservation	FATHMM	SIFT	Family
AKAP12	6:151671097	NM_005100.3	c.1571G>A	p.Gly524Asp	Nonsynonymous	7.823	Tolerated	Tolerated	111
ANPEP	15:90349571	NM_001150.2	c.244C>G	p.Pro82Ala	Nonsynonymous	9.721	Tolerated	Damaging	110
ARFGEF2	20:47606168	NM_006420.2	c.2761T>A	p.Cys921Ser	Nonsynonymous	7.997	Tolerated	Damaging	104
ATL3	11:63396862	NM_015459.3	c.1555G>A	p.Gly519Ser	Nonsynonymous	0.134	Damaging	Tolerated	122
CAPN8	1:223718199	NM_001143962.1	c.1547A>G	p.Glu516Gly	Nonsynonymous	2.213	Damaging	Damaging	122
CEP128	14:81372334	NM_152446.3	c.326T>G	p.Leu109Trp	Nonsynonymous	6.05	Tolerated	Damaging	115
CPD	17:28750626	NM_001304.4	c.1760C>A	p.Thr587Lys	Nonsynonymous	1.47	Tolerated	Tolerated	101
CSNK2B	6:31637201	NM_001320.5	c.473A>G	p.Tyr158Cys	Nonsynonymous	8.553	N/A	Damaging	111
DNAH12	3:57391521	NM_178504.4	c.6378G>C	p.Glu2126Asp	Nonsynonymous	0.253	Tolerated	Damaging	104
EXOC7	17:74099470	NM_001145297.2	c.98A>G	p.Glu33Gly	Nonsynonymous	5.844	N/A	Tolerated	104
FBXO31	16:87417206	NM_024735.3	c.145G>A	p.Gly49Arg	Nonsynonymous	2.817	N/A	Tolerated	104
*HOXA2**	7:27142079	NM_006735.3	c.41G>C	p.Ser14Thr	Nonsynonymous	7.415	Tolerated	Damaging	155
IBTK	6:82935292	NM_015525.2	c.727G>T	p.Val243Leu	Nonsynonymous	3.594	Damaging	Tolerated	130
IL31RA	5:55206426	NM_139017.5	c.1568T>C	p.Ile523Thr	Nonsynonymous	0.258	Tolerated	Tolerated	155
KIAA1324	1:109743448	NM_001267048.1	c.2638G>A	p.Glu880Lys	Nonsynonymous	9.732	Tolerated	Damaging	122
KLHDC8B	3:49212244	NM_173546.2	c.611G>A	p.Arg204Gln	Nonsynonymous	7.868	Tolerated	Tolerated	155
KMT2A	11:118374678	NM_001197104.1	c.8071G>A	p.Val2691Met	Nonsynonymous	5.921	Damaging	Tolerated	122
MAML2, MIR1260B	11:96074606	NM_032427.1	c.454G>A	p.Gly152Arg	Nonsynonymous	2.428	Tolerated	Tolerated	115
MGRN1	16:4707316	NM_001142289.2	c.513C>G	p.Phe171Leu	Nonsynonymous	3.969	Tolerated	Damaging	104
MICU2	13:22178157	NM_152726.2	c.131T>C	p.Leu44Pro	Nonsynonymous	3.572	Tolerated	Tolerated	122
NR3C1	5:142678245	NM_001024094.1	c.1883T>G	p.Leu628Arg	Nonsynonymous	9.339	Damaging	Damaging	110
ODF2L	1:86841964	NM_001184765.1	c.762_763insA	p.His255fs	Frameshift insertion	N/A	N/A	N/A	122
PROSER2, PROSER2-AS1	10:11911987	NM_153256.3	c.890C>G	p.Pro297Arg	Nonsynonymous	0.212	Tolerated	Damaging	115
PYROXD1	12:21614986	NM_024854.3	c.925G>A	p.Gly309Ser	Nonsynonymous	9.447	N/A	Tolerated	104
RALGPS1, ANGPTL2	9:129870361	NM_012098.2	c.650C>T	p.Pro217Leu	Nonsynonymous	8.674	Tolerated	Tolerated	111
SORBS3	8:22426677	NM_005775.4	c.1322A>G	p.Glu441Gly	Nonsynonymous	5.807	Tolerated	Damaging	111
SPAG17	1:118642247	NM_206996.2	c.811G>C	p.Val271Leu	Nonsynonymous	0.311	Tolerated	Tolerated	155
*TMPO**	12:98925463	NM_003276.2	c.412A>G	p.Thr138Ala	Nonsynonymous	7.895	Damaging	Damaging	111
TRMT44	4:8442843	NM_152544	c.294C>A	p.Cys98Ter	Stopgain	0.038	N/A	Tolerated	104
UPK2	11:118827653	NM_006760.3	c.136_137insT	p.Ala47fs	Frameshift insertion	N/A	N/A	N/A	122
USPL1	13:31233351	NM_005800.4	c.3137C>T	p.Ala1046Val	Nonsynonymous	0.735	Tolerated	Tolerated	122

Chromosome position in accordance with GRCh37/hg19 assembly. Conservation is scored by PhyloP 100-way Vertebrate, in which sites predicted to be conserved are assigned positive scores.

* Variant is both novel and located within MYP locus.

variants in 10 families with high myopia that followed the autosomal dominant inheritance pattern.

A premature stop codon variant (p.Gln53*) in *SCO2* was previously reported in a family with autosomal dominant high myopia.[10] The gene *SCO2* encodes a protein involved in mitochondrial cytochrome c oxidase activity and maps to chromosome 22q13.33, which is located within the MYP6 locus (OMIM 608908).[49] A link between cytochrome c oxidase deficiency and refractive error has not been demonstrated. Tran-Viet et al.[10] also reported three additional *SCO2* mutations in three highly myopic unrelated individuals: p.Arg114His, p.Glu140Lys, and p.Ala259Val. The highest frequencies of each of these genetic variants are reported in the ExAC database at 0.2343% (Latino); 0.01812% (non-Finnish European); and 3.26% (African), respectively. A recent study examined 35 human subjects carrying the p.Glu140Lys mutation, as well as mice carrying heterozygous, compound heterozygous, or homozygous *Sco2* mutations.[50] None of the individuals carrying *SCO2* mutations that were examined demonstrated high myopia (SE refractive error greater than –6.00 D), with no significant differences in refraction found between three p.Glu140Lys carriers and two healthy individuals in the same family. Furthermore, adult *Sco2*-deficient mice revealed no axial elongation indicative of high myopia, suggesting no evidence for *Sco2* mutations causing high myopia. However, as only *SCO2* gene mutations have been reported in families showing dominantly inherited disease, affected individuals in this study were also examined for *SCO2* variants by exome and Sanger sequencing, but no potentially pathogenic mutations were identified.

To date, more than 60 common SNPs implicating loci for myopia/refractive error have been identified through GWAS.[12,13] We examined the gene nearest to each associated SNP for variants in our high myopia cohort (Supplementary Table S1). However, no potentially pathogenic variants were identified within these genes in our exome data. While over 6000 significant SNP-trait associations have been reported through GWAS,[51] these studies were designed to detect associations between common variants in the population and common diseases. Several ophthalmic GWAS have uncovered loci that may help identify molecular pathways associated with ocular diseases such as age-related macular degeneration and glaucoma.[52] However, for the study of less common diseases such as high myopia, it is possible that many loci still remain unidentified.[53,54] Furthermore, approximately 80% of trait-associated SNPs reported are located in intergenic or noncoding regions,[55] raising additional questions about our current understanding of gene regulation in disease pheno-

types. Another consideration is that the age at which GWAS-identified variants may contribute to myopia in the global population is not known. Fan et al.[56] examined the age-of-onset correlation between refractive error and GWAS-identified variants in 5200 children examined longitudinally from ages 7 to 15 years, and found many differences between specific SNPs and their effects during very early life, childhood, or adulthood. Some SNPs showed progressively stronger effects during later childhood while others appeared to have no progressive effect on refractive error. This study also examined the effects of gene-environment interactions involving near work or time spent outdoors, but these associations were rare or absent for most of the GWAS-identified SNPs.

To investigate the variation in our cohort within loci linked to refractive error, the 23 MYP loci were examined. In our exome data, 10 variants were identified within these coordinates, two of which are novel (Table 3). Three variants, in genes *PTPRQ, TMPO*, and *TMTC2*, were identified at the MYP3 locus, which was mapped to chromosome 12q21-q23 in a large German/Italian family segregating autosomal dominant high myopia.[17] Variants in *PTPRQ* and *TMTC2* were identified in family 115, and the variant in *TMPO* was identified in family 111. *PTPRQ* encodes a protein-tyrosine phosphatase implicated in autosomal recessive deafness.[57] A heterozygous missense change in Thymopoietin (*TMPO*) was identified in two brothers with a severe form of cardiomyopathy.[58] The variant identified in *TMPO* in this study was novel. In 2015, a meta-analysis of genome-wide association performed by the International Glaucoma Genetics Consortium identified a SNP near the gene transmembrane and tetratricopeptide repeat containing 2 (*TMTC2*) associated with optic disc morphology and glaucoma,[59] which may indicate a possible role for *TMTC2* in the eye.

In family 104, one variant in the gene *ETNPPL* (alanine-glyoxylate aminotransferase 2-like 1 [*AGXT2L1*]) was identified in MYP11, which was mapped to chromosome 4 in a large Chinese family with autosomal dominant high myopia.[22] The gene *ETNPPL* encodes for the protein AGXT2L1.

Two variants were identified at MYP14 (chromosome 1p26) which was mapped using 49 multigenerational Ashkenazi Jewish families with high myopia.[26] These two variants are nonsynonymous changes in the genes *AGRN* and *PLOD1* identified in families 130 and 122, respectively. Agrin (*AGRN*) plays a role in synapse formation and brain development. A homozygous mutation in *AGRN* was reported in a Swiss brother and sister with congenital myasthenia syndrome (CMS8) which affects skeletal muscle and neuromuscular junctions,[60] while compound heterozygosity for two mutations in *AGRN* was identified in another individual with CMS8.[61] Homozygous mutations in the collagen-modifying gene *PLOD1* have been implicated in Ehlers-Danlos syndrome, a heritable connective tissue disorder that includes several ocular manifestations including myopia (OMIM 153454).[49,62] Common SNPs in *PLOD1* were also studied in a cohort of 600 unrelated Chinese Han individuals (300 with high myopia), but were determined unlikely to result in genetic susceptibility for high myopia.[63]

Originally, MYP4 was mapped to chromosome 7q36 after linkage analysis was performed in 21 French and 2 Algerian families with autosomal dominant high myopia.[64] However, a second study using the same families along with nine newly collected families failed to confirm linkage to 7q36, but rather to 7p15 and the MYP4 locus was renamed MYP17.[29] A study performed by Ciner et al.[30] in the same year also identified linkage to chromosome 7p15 after performing quantitative trait locus linkage analyses in 493 African American individuals and in 96 African American families. In our study, nonsynonymous variants were identified at the MYP17/4 locus in genes *DNAH11* and *HOXA2*. Mutations in the dynein gene *DNAH11* have been reported in Kartagener syndrome, a type of ciliary dyskinesia in which patients suffer from chronic respiratory infections, recurrent bronchitis, and pneumonia.[65,66] Homeobox A2 (*HOXA2*) mutations have been implicated in autosomal dominant and recessive forms of microtia and hearing impairment, sometimes accompanying cleft palate.[67,68] In this study, the variant identified in *HOXA2* was novel.

Two variants in the genes essential meiotic structure-specific endonuclease 1 (*EME1*) and transducer of ERBB2 1 (*TOB1*) were identified at the MYP5 locus, which was mapped to chromosome 17q21-q22 in a multigenerational English/Canadian family with autosomal dominant high myopia.[18] However, no causal mutation was reported in that publication. The same family was utilized for this study (family 68), and notably the proband of family 68 had the most severe SE refractive error (−50.00/−50.00 D) of our cohort. Exome data from three affected and one unaffected individuals were filtered for potentially pathogenic variants. Two variants were identified; one variant identified in the gene FtsJ Homolog 3 (*FTSJ3*) encodes a missense change, while a variant identified in the gene Solute Carrier Family 7 Member 13 (*SLC7A13*) results in a premature stop codon in the third of four exons (Supplementary Table S2). Neither of these variants were located within the MYP5 locus previously linked to this family or within any other MYP loci.

A second large family in our cohort has also undergone prior linkage analysis. Family 56, a Caucasian family from the United States, is a large autosomal dominant high myopia family in which nearly half of the individuals are highly myopic. A novel locus at chromosome 2q37.1 (MYP12) was determined, but causal mutations were not identified.[23] To identify candidate causative variants in this family, we performed exome sequencing on 11 individuals, including 8 that were affected. However, no variants were identified that were present in all 8 affected individuals sequenced. It is interesting to note that within both families 56 and 68, no variants that met our filtering criteria were identified within the chromosomal regions where linkage studies previously mapped myopia-associated loci. While genetic linkage analysis is a powerful tool for mapping disease loci in Mendelian disorders in families, there are limitations. Since proper linkage requires a large family with multiple generations demonstrating the phenotype, most of our families were too small to evaluate using this method. Genetic linkage studies are less helpful for multifactorial or complex traits, where multiple genes are important in disease causation. However, given its severity in young children and infants, and the clearly Mendelian inheritance pattern, there are likely to be strong genetic determinants in high myopia. Nevertheless, myopia is a highly variable disease, with refractive measurements varying greatly between affected individuals in the same family as well as between eyes of the same individual. Numerous studies have examined the correlation between environmental factors and myopia, including levels of education, increased near work, rural versus urban settings, diet, and outdoor activities.[69] A combination of genetic susceptibility and environmental triggers may play a role together in disease manifestation and progression.

Notably for this study, potentially pathogenic variants cosegregating with high myopia were only identified in 10 of 14 families in our cohort. No variants meeting our filtering criteria were identified in families 8, 19, 56, 126. In addition to the added complications given the variability in myopia severity, there are also limitations using exome sequencing as a method to detect causal variants. It was previously estimated that approximately 85% of disease-causing mutations are found in the coding regions of the genome,[70] but our understanding

of the noncoding region's importance has recently rapidly evolved. The protein-coding exome comprises approximately 2% of the human genome, and as functional consequences can occur in response to sequence variation outside of coding regions, whole-genome sequencing may resolve the underlying molecular etiology. Furthermore, exome sequencing is unable to reliably detect large insertions, deletions, or rearrangements, and errors in calling small indel variants are common. Exome capture kits offered by companies such as Agilent or Roche NimbleGen are dependent on our knowledge of the expressed regions of the human genome, and therefore exclude variants in regions of the genome that are not currently known to be transcribed. Exome coverage is also incomplete, with less than 80% of the exome captured at more than ×20 coverage with the Agilent platform (SureSelect) that was utilized for the most recent exome sequencing performed in this study,[69,71] though some groups have reported up to 95% coverage of RefSeq coding exons in their exome data.[72,73] Another potential source of error is the possibility of healthy control populations having unreported high myopia. Many do not consider high myopia a disease and may underestimate their roles as controls in genetic studies. It is therefore difficult to be certain that genomic databases are free of highly myopic individuals when making conclusions in our studies.

While whole genome sequencing remains cost prohibitive for many research groups, it offers advantages over exome sequencing when the pathogenic mutation is suspected to be located in a noncoding region. Variations in upstream elements such as promoters, enhancers, and silencers may be assessed for potential regulatory impact. Noncoding regions that show high conservation among different species as well as predicted transcription factor binding sites can be analyzed for variations. Sequence capture kits are also expected to offer expanded targets in the future which may help to identify variations in noncoding DNA. Advancements in complex prediction algorithms may also aid in the assessment of potentially pathogenic gene variants.

In conclusion, we have identified 104 new candidate variants for high myopia. These variants will require future study in additional patient subjects with high myopia, functional analyses, and/or animal modeling in order to determine their role in this phenotype.

Acknowledgments

The authors thank all patients, subjects, and their families for their participation in this study.

Supported by the National Institutes of Health (NIH), National Eye Institute Grants R01 EY14685 and R01EY018246; NIH National Eye Institute P30 EY016665 CORE Grant for vision research; an unrestricted grant from Research to Prevent Blindness, Inc., to the Department of Ophthalmology and Vision Sciences at the University of Wisconsin; and a Centennial Scholars Grant from the University of Wisconsin School of Medicine and Public Health.

Disclosure: **B.A. Kloss**, None; **S.W. Tompson**, None; **K.N. Whisenhunt**, None; **K.L. Quow**, None; **S.J. Huang**, None; **D.M. Pavelec**, None; **T. Rosenberg**, None; **T.L. Young**, None

References

1. Kempen JH, Mitchell P, Lee KE, et al. The prevalence of refractive errors among adults in the United States, Western Europe, and Australia. *Arch Ophthalmol*. 2004;122:495–505.
2. Liang YB, Wong TY, Sun LP, et al. Refractive errors in a rural Chinese adult population the Handan eye study. *Ophthalmology*. 2009;116:2119–2127.
3. He M, Huang W, Li Y, Zheng Y, Yin Q, Foster PJ. Refractive error and biometry in older Chinese adults: the Liwan eye study. *Invest Ophthalmol Vis Sci*. 2009;50:5130–5136.
4. Saw SM, Tong L, Chua WH, et al. Incidence and progression of myopia in Singaporean school children. *Invest Ophthalmol Vis Sci*. 2005;46:51–57.
5. Wang TJ, Chiang TH, Wang TH, Lin LL, Shih YF. Changes of the ocular refraction among freshmen in National Taiwan University between 1988 and 2005. *Eye (Lond)*. 2009;23:1168–1169.
6. Atchison DA, Jones CE, Schmid KL, et al. Eye shape in emmetropia and myopia. *Invest Ophthalmol Vis Sci*. 2004;45:3380–3386.
7. Rada JA, Shelton S, Norton TT. The sclera and myopia. *Exp Eye Res*. 2006;82:185–200.
8. Burton TC. The influence of refractive error and lattice degeneration on the incidence of retinal detachment. *Trans Am Ophthalmol Soc*. 1989;87:143–155; discussion, 155–157.
9. Mordechai S, Gradstein L, Pasanen A, et al. High myopia caused by a mutation in LEPREL1, encoding prolyl 3-hydroxylase 2. *Am J Hum Genet*. 2011;89:438–445.
10. Tran-Viet KN, Powell C, Barathi VA, et al. Mutations in *SCO2* are associated with autosomal-dominant high myopia. *Am J Hum Genet*. 2013;92:820–826.
11. Guo H, Tong P, Peng Y, et al. Homozygous loss-of-function mutation of the LEPREL1 gene causes severe non-syndromic high myopia with early-onset cataract. *Clin Genet*. 2014;86:575–579.
12. Verhoeven VJ, Hysi PG, Wojciechowski R, et al. Genome-wide meta-analyses of multi-ancestry cohorts identify multiple new susceptibility loci for refractive error and myopia. *Nat Genet*. 2013;45:314–318.
13. Kiefer AK, Tung JY, Do CB, et al. Genome-wide analysis points to roles for extracellular matrix remodeling, the visual cycle, and neuronal development in myopia. *PLoS Genet*. 2013;9:e1003299.
14. Young TL, Ronan SM, Drahozal LA, et al. Evidence that a locus for familial high myopia maps to chromosome 18p. *Am J Hum Genet*. 1998;63:109–119.
15. Young TL, Atwood LD, Ronan S.M, et al. Further refinement of the MYP2 locus for autosomal dominant high myopia by linkage disequilibrium analysis. *Ophthalmic Genet*. 2001;22:69–75.
16. Ratnamala U, Lyle R, Raval R, et al. Refinement of the X-linked nonsyndromic high myopia locus (MYP1) on Xq28 and exclusion of thirteen known positional candidate genes by direct sequencing. *Invest Ophthalmol Vis Sci*. 2011;52:6814–6819.
17. Young TL, Ronan SM, Alvear AB, et al. A second locus for familial high myopia maps to chromosome 12q. *Am J Hum Genet*. 1998;63:1419–1424.
18. Paluru P, Ronan SM, Heon E, et al. New locus for autosomal dominant high myopia maps to the long arm of chromosome 17. *Invest Ophthalmol Vis Sci*. 2003;44:1830–1836.
19. Stambolian D, Ibay G, Reider L, et al. Genomewide linkage scan for myopia susceptibility loci among Ashkenazi Jewish families shows evidence of linkage on chromosome 22q12. *Am J Hum Genet*. 2004;75:448–459.
20. Klein AP, Duggal P, Lee KE, et al. Confirmation of linkage to ocular refraction on chromosome 22q and identification of a novel linkage region on 1q. *Arch Ophthalmol*. 2007;125:80–85.
21. Hammond CJ, Andrew T, Mak YT, Spector TD. 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.

22. Zhang, Q, Guo X, Xiao X, Jia X, Li S, Hejtmancik JF. A new locus for autosomal dominant high myopia maps to 4q22-q27 between D4S1578 and D4S1612. *Mol Vis*. 2005;11:554–560.
23. Paluru PC, Nallasamy S, Devoto M, Rappaport EF, Young TL. Identification of a novel locus on 2q for autosomal dominant high myopia. *Invest Ophthalmol Vis Sci*. 2005;46:2300–2307.
24. Zhang, Q, Guo X, Xiao X, Jia X, Li S, Hejtmancik JF. Novel locus for X linked recessive high myopia maps to Xq23-q25 but outside MYP1. *J Med Genet*. 2006;43:e20.
25. Zhang Q, Li S, Xiao X, Jia X, Guo X. Confirmation of a genetic locus for X-linked recessive high myopia outside MYP1. *J Hum Genet*. 2007;52:469–472.
26. Wojciechowski R, Moy C, Ciner E, et al. Genomewide scan in Ashkenazi Jewish families demonstrates evidence of linkage of ocular refraction to a QTL on chromosome 1p36. *Hum Genet*. 2006;119:389–399.
27. Nallasamy S, Paluru PC, Devoto M, Wasserman NF, Zhou J, Young TL. Genetic linkage study of high myopia in a Hutterite population from South Dakota. *Mol Vis*. 2007;13:229–236.
28. Lam CY, Tam POS, Fan DSP, et al. A genome-wide scan maps a novel high myopia locus to 5p15. *Invest Ophthalmol Vis Sci*. 2008;49:3768–3778.
29. Paget S, Julia S, Vitezica ZG, Soler V, Malecaze F, Calvas P. Linkage analysis of high myopia susceptibility locus in 26 families. *Mol Vis*. 2008;14:2566–2574.
30. Ciner E, Wojciechowski R, Ibay G, Bailey-Wilson JE, Stambolian D. Genomewide scan of ocular refraction in African-American families shows significant linkage to chromosome 7p15. *Genet Epidemiol*. 2008;32:454–463.
31. Yang Z, Xiao X, Li S, Zhang Q. Clinical and linkage study on a consanguineous Chinese family with autosomal recessive high myopia. *Mol Vis*. 2009;15:312–318.
32. Ma JH, Shen SH, Zhang GW, et al. Identification of a locus for autosomal dominant high myopia on chromosome 5p13.3-p15.1 in a Chinese family. *Mol Vis*. 2010;16:2043–2054.
33. Shi Y, Qu J, Zhang D, et al. Genetic variants at 13q12.12 are associated with high myopia in the Han Chinese population. *Am J Hum Genet*. 2011;88:805–813.
34. Shi Y, Li Y, Zhang D, et al. Exome sequencing identifies ZNF644 mutations in high myopia. *PLoS Genet*. 2011;7:e1002084.
35. Zhao F, Wu J, Xue A, et al. Exome sequencing reveals CCDC111 mutation associated with high myopia. *Hum Genet*. 2013;132:913–921.
36. Aldahmesh MA, Khan AO, Alkuraya H, et al. Mutations in LRPAP1 are associated with severe myopia in humans. *Am J Hum Genet*. 2013;93:313–320.
37. Guo H, Jin X, Zhu T, et al. SLC39A5 mutations interfering with the BMP/TGF-beta pathway in non-syndromic high myopia. *J Med Genet*. 2014;51:518–525.
38. Wojciechowski R. Nature and nurture: the complex genetics of myopia and refractive error. *Clin Genet*. 2011;79:301–320.
39. Hawthorne FA, Young TL. Genetic contributions to myopic refractive error: Insights from human studies and supporting evidence from animal models. *Exp Eye Res*. 2013;114:141–149.
40. Stambolian D. Genetic susceptibility and mechanisms for refractive error. *Clin Genet*. 2013;84:102–108.
41. Holden BA, Fricke TR, Wilson DA, et al. Global prevalence of myopia and high myopia and temporal trends from 2000 through 2050. *Ophthalmology*. 2016;123:1036–1042.
42. Vitale S, Sperduto RD, Ferris FL. Increased prevalence of myopia in the United States between 1971-1972 and 1999-2004. *Arch Ophthalmol*. 2009;127:1632–1639.
43. Hongshan J, Rong L, Shou-Wei D, Shuifang Z. Skewer: a fast and accurate adapter trimmer for next-generation sequencing paired-end reads. *BMC Bioinformatics*. 2014;15:182.
44. Li H, Durbin R. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics*. 2010;26:589–595.
45. McKenna A, Hanna M, Banks E, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res*. 2010;20:1297–1303.
46. Cingolani P, Platts A, Wang L, et al. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. *Fly (Austin)*. 2012;6:80–92
47. Lek M, Karczewski K, Minikel EV, et al. Exome Aggregation Consortium. Analysis of protein-coding genetic variation in 60,706 humans. *Nature*. 2016;536:285–291.
48. Siepel A, Haussler D. Phylogenetic hidden Markov models. In: Nielsen R, ed. *Statistical Methods in Molecular Evolution*. New York, NY: Springer; 2005:325–351.
49. Online Mendelian Inheritance in Man, OMIM®. McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University, Baltimore, MD, USA. Available at: http://omim.org/. Accessed April 12, 2016.
50. Piekutowska-Abramczuk D, Kocyła-Karczmarewicz B, Małkowska M, et al. No evidence for association of SCO2 heterozygosity with high myopia or other diseases with possible mitochondrial dysfunction. *JIMD Rep*. 2016;27:63–68.
51. Welter D, MacArthur J, Morales J, et al. The NHGRI GWAS Catalog, a curated resource of SNP-trait associations. *Nucleic Acids Res*. 2014;42(database issue):D1001–D1006.
52. Mackey DA, Hewitt AW. Genome-wide association study success in ophthalmology. *Curr Opin Ophthalmol*. 2014;25:386–393.
53. Manolio TA. Genomewide association studies and assessment of the risk of disease. *N Engl J Med*. 2010;363:166–176.
54. Visscher PM, Brown MA, McCarthy MI, Yang J. Five years of GWAS discovery. *Am J Hum Genet*. 2012;90:7–24.
55. Hardy J, Singleton A. Genomewide association studies and human disease. *N Engl J Med*. 2009;360:1759–1768.
56. Fan Q, Guo X, Tideman JW, et al. Childhood gene-environment interactions and age-dependent effects of genetic variants associated with refractive error and myopia: the CREAM Consortium. *Sci Rep*. 2016;6:25853.
57. Schraders M, Oostrik J, Huygen PLM, et al. Mutations in PTPRQ are a cause of autosomal-recessive nonsyndromic hearing impairment DFNB84 and associated with vestibular dysfunction. *Am J Hum Genet*. 2010;86:604–610.
58. Taylor MRG, Slavov D, Gajewski A, et al. Thymopoietin (lamina-associated polypeptide 2) gene mutation associated with dilated cardiomyopathy. *Hum Mutat*. 2005;26:566–574.
59. Springelkamp H, Mishra A, Hysi PG, et al. Meta-analysis of genome-wide association studies identifies novel loci associated with optic disc morphology. *Genet Epidemiol*. 2015;39:207–216.
60. Huze C, Bauche S, Richard P, et al. Identification of an agrin mutation that causes congenital myasthenia and affects synapse function. *Am J Hum Genet*. 2009;85:155–167.
61. Maselli RA, Fernandez JM, Arredondo J, et al. LG2 agrin mutation causing severe congenital myasthenic syndrome mimics functional characteristics of non-neural (z-) agrin. *Hum Genet*. 2012;131:1123–1135.
62. Hautala T, Heikkinen J, Kivirikko KI, Myllyla R. A large duplication in the gene for lysyl hydroxylase accounts for the type VI variant of Ehlers-Danlos syndrome in two siblings. *Genomics*. 1993;15:399–404.
63. Yip SP, Leung KH, Fung WY, Ng PW, Sham PC, Yap MK. A DNA pooling-based case-control study of myopia candidate genes

Investigative Ophthalmology & Visual Science

COL11A1, COL18A1, FBN1, and PLOD1 in a Chinese population. *Mol Vis*. 2011;17:810–821.

64. Naiglin L, Gazagne C, Dallongeville F, et al. A genome wide scan for familial high myopia suggests a novel locus on chromosome 7q36. *J Med Genet*. 2002;39:118–124.
65. Schwabe GC, Hoffmann K, Loges NT, et al. Primary ciliary dyskinesia associated with normal axoneme ultrastructure is caused by DNAH11 mutations. *Hum Mutat*. 2008;29:289–298.
66. Lucas JS, Adam EC, Goggin PM, et al. Static respiratory cilia associated with mutations in Dnahc11/DNAH11: a mouse model of PCD. *Hum Mutat*. 2012;33:495–503.
67. Alasti F, Sadeghi A, Sanati MH, et al. A mutation in HOXA2 is responsible for autosomal-recessive microtia in an Iranian family. *Am J Hum Genet*. 2008;82:982–991.
68. Brown KK, Viana LM, Helwig CC, et al. HOXA2 haploinsufficiency in dominant bilateral microtia and hearing loss. *Hum Mutat*. 2013;34:1347–1351.
69. Goldschmidt E, Jacobsen N. Genetic and environmental effects on myopia development and progression. *Eye (Lond)*. 2014;28:126–133.
70. Cooper DN, Krawczak M, Antonarakis SE. The nature and mechanisms of human gene mutation. In: Scriver CR, Beaudet AL, Sly WS, Valle D, eds. *The Metabolic and Molecular Bases of Inherited Disease*. 7th ed. New York, NY: McGraw-Hill; 1995:259–291.
71. Agilent Technologies. Agilent SureSelect Performance. http://www.genomics.agilent.com/article.jsp?pageId=2094. Accessed September 13, 2016.
72. Choi M, Scholl UI, Ji W, et al. Genetic diagnosis by whole exome capture and massively parallel DNA sequencing. *Proc Natl Acad Sci U S A*. 2009;106:19096–19101.
73. Meienberg J, Zerjavic K, Keller I, et al. New insights into the performance of human whole-exome capture platforms. *Nucleic Acids Res*. 2015;43:e76.

Investigative Ophthalmology & Visual Science