

# Myopia in Chinese families shows linkage to 10q26.13

Anthony M. Musolf,<sup>1</sup> Claire L. Simpson,<sup>2,1</sup> Kyle A. Long,<sup>1</sup> Bilal A. Moiz,<sup>1</sup> Deyana D. Lewis,<sup>1</sup> Candace D. Middlebrooks,<sup>1</sup> Laura Portas,<sup>3</sup> Federico Murgia,<sup>3</sup> Elise B. Ciner,<sup>4</sup> Joan E. Bailey-Wilson,<sup>1</sup> Dwight Stambolian<sup>5</sup>

(The first two authors contributed equally to this study)

(Last two authors considered senior authors)

<sup>1</sup>Computational and Statistical Genomics Branch, National Human Genome Research Institute, National Institutes of Health, Baltimore, MD; <sup>2</sup>Department of Genetics, Genomics and Informatics, University of Tennessee Health Science Center, Memphis, TN; <sup>3</sup>Institute of Population Genetics, CNR, Li Punti, Sassari, Italy; <sup>4</sup>The Pennsylvania College of Optometry at Salus University, Elkins Park, PA; <sup>5</sup>Department of Ophthalmology, University of Pennsylvania, Philadelphia, PA

**Purpose:** To determine genetic linkage between myopia and Han Chinese patients with a family history of the disease. **Methods:** One hundred seventy-six Han Chinese patients from 34 extended families were given eye examinations, and mean spherical equivalent (MSE) in diopters (D) was calculated by adding the spherical component of the refraction to one-half the cylindrical component and taking the average of both eyes. The MSE was converted to a binary phenotype, where all patients with an MSE of -1.00 D or less were coded as affected. Unaffected individuals had an MSE greater than 0.00 D (ages 21 years and up), +1.50 (ages 11–20), or +2.00 D (ages 6–10 years). Individuals between the given upper threshold and -1.00 were coded as unknown. Patients were genotyped on an exome chip. Three types of linkage analyses were performed: single-variant two-point, multipoint, and collapsed haplotype pattern (CHP) variant two-point. **Results:** The CHP variant two-point results identified a significant peak (heterogeneity logarithm of the odds [HLOD] = 3.73) at 10q26.13 in *TACC2*. The single-variant two-point and multipoint analyses showed highly suggestive linkage to the same region. The single-variant two-point results identified 25 suggestive variants at *HTRAI*, also at 10q26.13. **Conclusions:** We report a significant genetic linkage between myopia and Han Chinese patients at 10q26.13. 10q26.13 contains several good candidate genes, such as *TACC2* and the known age-related macular degeneration gene *HTRAI*. Targeted sequencing of the region is planned to identify the causal variant(s).

Myopia, or nearsightedness, is one of the most common ocular abnormalities in the world. Myopia is caused by light focusing in front of the retina instead of directly on it, resulting in blurred images. Myopia affects approximately 25% of Americans and has reached epidemic proportions in Southeast Asia, with nearly 80% of the entire population affected with the disease.

Myopia is known to be a complex trait affected by a variety of genetic and environmental factors [1,2]. Factors such as hours of near work and education level have been shown to positively correlate with the disease, while factors such as outdoor activity have been shown to have protective effects [1]. Genetic studies have used population-based genome-wide association studies (GWASs) and family-based linkage studies. GWASs became popular in the mid- to late 2000s with the advent of commercially available genotyping arrays and have been effective in identifying multiple risk variants for negative refractive error [3-8]. Most risk variants identified in GWASs have a common minor allele frequency

(MAF; >0.05) and have a small to moderate effect on risk of myopia.

Family-based linkage studies are complementary to GWASs and have several unique characteristics. These studies have a natural advantage for finding rare, highly penetrant causal variants. Family-based studies do not need as many samples as population-based studies to obtain sufficient power to detect rare variants because rare variants may be common within a particular family even if they are rare in the general population. Family-based studies are not affected by population stratification as each family is analyzed as a single unit and not in a group. Each family receives its own logarithm of the odds (LOD) score, which can be added across families. Finally, family-based studies can exploit large linked haplotypes within a family unit. Individuals in a population-based study are distantly related, and the massive number of meioses through countless generations has created exceedingly small haplotypes, with only the variants that are closest together exhibiting linkage disequilibrium (LD). In family studies, the founding haplotypes for each family member are determined solely by the haplotypes of the family founders, which are only a subset of all the

Correspondence to: Joan E. Bailey-Wilson, 333 Cassell Dr, Suite 1200, Baltimore, MD 21224, Phone: (443) 740-2921, FAX: (443) 740-2165; email: jebw@mail.nih.gov,

haplotypes that exist in the general population. Because the majority of family-based studies are based on small families (two to four generations), only a limited number of recombinations can occur between variants that are moderately far apart on a chromosome. This results in a longer linked haplotype that can provide more power to detect variants along the haplotype (even if the causal variant is rare or not genotyped). Association studies have low power to identify rare variants so if a single gene has multiple rare variants in the population (each of which causes relatively large increases in risk of a disease), then linkage studies observe different linked haplotypes in this region across multiple families, but the association may not be detectable due to low power; see Figure 1 for a visual example. One drawback of the linkage approach is that the linked haplotype along the significant linkage region may contain a large number of candidate genes and potential causal variants because the limited number of meioses observed in each linked family may not be adequate to narrow the linked region around each causal variant to a small region. For an excellent introduction to these concepts, see Nussbaum et al. [9].

The development of cheaper genotyping and sequencing methods within the past few years has made family-based studies more useful as these studies have more power (compared to population-based association studies) to detect individual rare variants that have a large effect on the phenotype of interest. Generally, linkage studies have better power to detect causal variants which are highly penetrant while association studies have better power to detect common causal genetic variants of a small or modest effect. More detailed discussion on family-based linkage studies versus association studies can be found in several reviews [10,11].

Complex traits such as myopia have high amounts of locus heterogeneity, meaning they have multiple variants across the genome contributing to risk. Most of these variants are expected to be common variants of low effect (best identified by association), but one can also expect some rarer variants that confer a greater effect (best identified by linkage). This is what we have seen in many complex traits the most prominent of which is breast cancer. Linkage analyses of highly aggregated families (ascertained because of many closely related women with breast cancer in the families) have identified several genes that each harbor many individually rare high effect risk variants for breast cancer, such as the well-known *BRCA1* (Gene ID 672, OMIM 113705) gene [12,13], while population-based case-control association analyses (which include affected women with and without a family history of the disease) have identified a large number of lower penetrance, common risk variants [14,15]. Other complex traits where risk variants have been identified by linkage and association studies include hypertension [16,17] and cataracts [18,19], to name just a few.

Myopia, too, has had association [3-8] and linkage [20-29] studies that have identified risk variants and loci. Linkage studies are useful because segregation analysis [30] has shown evidence for the high probability of a rare, high penetrance, autosomal dominant risk variant(s) in some highly aggregated families. Multiple family-based studies have used this model to identify high risk loci in pathogenic myopia (mean spherical equivalent (MSE) < -6 diopters [D]) and non-pathogenic myopia (MSE < -1 D) [20-29] in multiple ethnic groups, including Ashkenazi Jews, African Americans, and Caucasians.

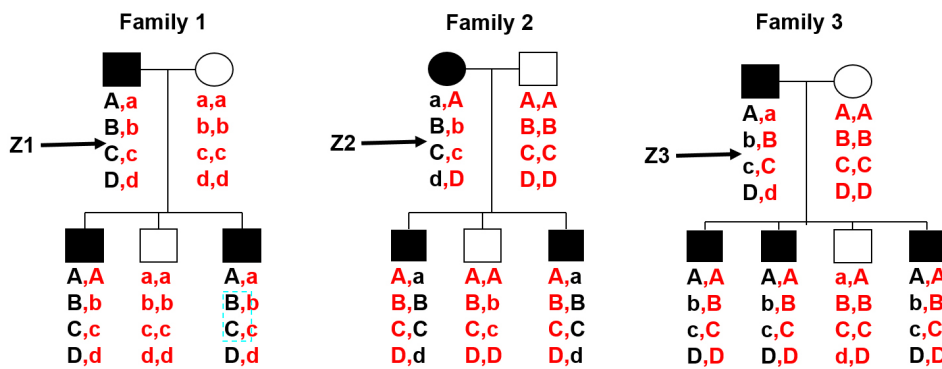


Figure 1. Example of linked disease markers with no association. The diagram illustrates the scenario where common variants can be linked along a haplotype within families but will not be identified by an association study. Here, we place theoretical disease variants Z1, Z2, and Z3 in families 1, 2, and 3, respectively. We assume that the disease-causing variants Z are rare and are not genotyped. Markers A, B, C, and D are assumed to have a common minor allele frequency (MAF) and have been genotyped. All four markers are also located on the same haplotype close to each other, so that it is unlikely any crossing over will occur within families. Looking within each family, it is clear that all markers are linked to the disease marker because in each family the black haplotype cosegregates with the disease phenotype. However, looking across each family, A, B, C, and D are not associated with the disease phenotype, because both alleles of each marker are present in affected and unaffected individuals.

Looking within each family, it is clear that all markers are linked to the disease marker because in each family the black haplotype cosegregates with the disease phenotype. However, looking across each family, A, B, C, and D are not associated with the disease phenotype, because both alleles of each marker are present in affected and unaffected individuals.

This study uses genotype data from Han Chinese families affected with myopia to identify genes that may harbor risk variants of a larger effect on myopia risk than are generally detectable in association studies. In an effort to increase the chance that this study might identify causal variants, we used a dense exome single nucleotide polymorphism (SNP) genotype array that is enriched for variants within exons. Asian populations have a much higher prevalence of myopia than other ethnicities [31]. Association studies have identified several risk variants in Chinese populations, including variants in the 15q14 [32], 11q24.1 [33], and 22q12 [34] regions. Family-based studies are less common but have been used to successfully map evidence for linkage of high myopia to Xq28 [35] and 5p13–15 [36].

## METHODS

*Study design:* This study was a retrospective, observational family study of a binary trait (affected or unaffected with myopia) that involved the recruitment, eye examination, exome chip genotyping, and parametric linkage analysis of 176 Han Chinese individuals from 34 extended families who had recently emigrated from China to Pennsylvania. The study was approved by the institutional review boards of the University of Pennsylvania and the National Human Genome Research Institute. Protocols were performed in accordance with the Declaration of Helsinki of 1976 and adhered to the ARVO statement on human subjects. All study participants provided informed written consent. Families were identified through mailings, eye clinic interviews, and referrals from private eye doctors. To be eligible for this study, all families were required to have at least three participants, at least one myopic parent, and at least two myopic siblings. Children had to be at least 5 years old to be eligible. These families were carefully ascertained to be consistent with an autosomal dominant mode of inheritance given the high prevalence of myopia in this population. In most of the families, only one parent was affected; in only seven families were both parents affected. However, these seven families were all small nuclear families with two parents and two children and did not contribute much information to the analyses (e.g., no family had a LOD score greater than 0.2). Most of the linkage information was obtained from the families that best fit a high-penetrance autosomal dominant model, as would be expected in a parametric linkage analysis that assumed such a model. Careful ascertainment of rare, highly aggregated families that appear to be segregating a high-penetrance genetic risk factor is part of a good study design for a linkage study. Such ascertainment is crucial to achieve high power in a linkage study, particularly for complex traits.

All participants received a comprehensive eye examination that included medical and ocular health history, visual acuity, slit-lamp biomicroscopy, dilated fundus exams, and manifest refraction. For subjects younger than 41 years of age, cycloplegic refraction was measured using 0.5% cyclopentolate or 1% tropicamide. For subjects older than 41 years, refraction was measured via manifest refraction. The refraction measurement used in this study was MSE, measured in diopters. This score is obtained by adding the spherical component of the refraction to one-half the cylindrical component and taking the average of both eyes.

*Genotyping and quality control:* The 176 patient samples were genotyped by the Center for Inherited Disease Research (CIDR) at Johns Hopkins University (Baltimore, MD) using an Illumina ExomePlus array. This exome-based array has a GWAS backbone of common variants that runs throughout the genome. This backbone is critical for linkage analysis, as the analysis can use these more common GWAS variants to tag long linked haplotypes. The array also contains a good number of exonic rare variants, which provides the potential not only to find rare causal variants but also to perform gene-based analyses incorporating these rare variants.

CIDR coded any genotypes with a genotype quality score of 0.15 or less as missing. We further filtered to a call rate of 95% for a mean call rate of 99% per marker. Additional quality control was performed using PLINK [37]. Outliers in heterozygosity rates across samples were excluded, and autosomal SNPs with a sex difference in allelic frequency >0.2 or a sex difference in heterozygosity >0.3 were dropped. All monomorphic variants were removed.

We added an additional 39 individuals who did not have genotype data to connect disjointed pedigrees and ensure proper familial relationships. These individuals were known to exist based on family history but were unable or unwilling to participate in the study. They were coded as having an unknown phenotype and unknown genotypes at all markers. We used PLINK and PREST-PLUS [38] to confirm familial relationships by checking identity-by-descent (IBD) values. Sib-pair was used to check for Mendelian inconsistencies. Any marker with a single Mendelian error was dropped in the offending family; any marker with more than one Mendelian inconsistency was dropped for the entire data set. The remaining SNPs were mapped on the Rutgers Genetic Map version 3 [39] using GRCh37 physical positions, and data set-level allelic frequencies were calculated for each marker using Sib-pair.

*Myopia classification:* Individuals with an MSE of -1.00 D or less were coded as affected, while individuals with an MSE of 0.00 D or greater were coded as unaffected. Individuals

with an MSE that fell within the range of 0.00 D and  $-1.00$  D and individuals with no measurements were coded as unknown. We treated children more stringently than adults because normal developmental changes in refractive error during childhood cause potential misclassification. All children with an MSE of  $-1.00$  D or less were coded as affected, identical to adults. Children were coded as unaffected if they had an MSE of  $+2.00$  D or greater (ages 6–10 years) or  $+1.50$  D or greater (ages 11–20 years). This conservative approach balances the power loss that results from the lack of a good segregation-analysis model of age-dependent penetrance and the concomitant confusion about appropriate genotype probabilities for young unaffected subjects, with the power loss resulting from the classification of normal children as unknown.

The final data set consisted of 215 individuals (176 genotyped individuals) from 34 extended families with 52,253 SNPs across 22 autosomes. The data set was 53.5% female and contained 145 affected, 10 unaffected, and 60 unknown individuals. The average MSE was  $-4.09$  D with a standard deviation of 2.92.

**Parametric linkage analysis:** All parametric linkage analyses were performed assuming an autosomal dominant model with a disease allele frequency of 1%. Penetrance was assumed to be 90% for carriers and 10% for non-carriers. This is a well-tested inheritance model that has been verified with segregation analysis [30] and has been used successfully to identify linked risk loci in families in a similar manner as was used to ascertain the Han Chinese families here, including Ashkenazi Jewish [22], Amish [29], and Caucasian highly aggregated families [27]. Standard single-variant two-point linkage analysis using the Elston-Stewart algorithm was performed with *TwoPointLods*. Multipoint analysis was performed using *SimWalk2* [40-42]. As inter-marker LD between markers in a dense map can cause inflation in type I error rates, we pruned the genotype data before we ran the multipoint analysis. The SNPs were collapsed into 1 cM bins, and the SNP with the highest MAF was selected to represent the bins in the multipoint analysis. *Haploview* [43] was used to remove any additional SNPs with an  $r^2$  greater than 0.2 after pruning leaving 3,422 SNPs for multipoint analysis.

Linkage analysis was also performed using the collapsed haplotype pattern (CHP) method, implemented via the *SEQLinkage* [44] software, using all variants with a minor allele frequency of less than 15% in the founders of these families. This method builds short haplotypes within families using genotypes at individual uncommon variants (which usually do not occur more than once in any mating). These family-specific haplotypes are used as multiallelic

pseudomarkers that correspond to specific genetic regions, such as genes or segments of genes, as determined with *RefSeq*. This approach does not require pruning of the rare variants (which may have larger effect sizes than common variants) and allows for different rare variants in the same gene to be easily captured and combined into a joint test of linkage for each gene. This can result in more power for linkage analysis than can be obtained from a multipoint analysis using only a sparse pruned map of common SNPs that are not in LD with each other. Two-point linkage analysis on the pseudomarkers was then performed via *MERLIN* [45]. All variants were annotated using *ANNOVAR* [46,47], and further functional annotation was provided by *SIFT* [48-52], *ClinVar* [53], and *PolyPhen2* [54].

In addition to the linkage analyses, we performed two types of association analyses. We used the family-based association test *FBAT 2.0.4* [55,56] to examine all variants and *rv-TDT* [57] to examine rare (MAF = 0.05) variants after we selected trios from each pedigree.

## RESULTS

The CHP variant two-point linkage results identified one significant linkage peak at 10q26.13, centered on the *TACC2* (Gene ID 10579, OMIM 605302) gene (HLOD = 3.73; Figure 2). This study used the Lander and Kruglyak values of  $HLOD \geq 3.3$  and  $HLOD \geq 1.9$  as the respective significant and suggestive thresholds [58]. Seventeen other genes also showed suggestive evidence of linkage to myopia, six of which were also located on 10q, ranging from 10q24.2 to 10q26.2. The second highest overall HLOD (2.77) was located at 10q26.2 and centered on *DOCK1* (Gene ID 1793, OMIM 601403). Two suggestive linkage signals were also observed on 1p, 9q, and 15q while three suggestive signals were observed on 12q. A full list of all significant and suggestive linkage results can be found in Table 1.

Multipoint analysis resulted in genome-wide suggestive linkage evidence at 21 SNPs (Table 2). All were located between either 9q33.1 and 9q34.11 (12 SNPs) or 10q26.11 and 10q26.13 (nine SNPs; Figure 3). The most strongly linked SNP in the 9q region (HLOD = 3.13) was localized to an intergenic region between *TLR4* (Gene ID 7099, OMIM 603030) and *BRINP1* (Gene ID 1620, OMIM 602865) at 9q33.1. The most strongly linked SNP in the 10q region (HLOD = 2.35) was located in an intron of the *CACUL1* (Gene ID 143384) gene at 10q26.11.

The single-variant two-point analysis did not yield any significant results, although it identified 62 SNPs with suggestive evidence of linkage to myopia (Figure 4). Thirty-four of the suggestive SNPs localized to the 10q24.2 to



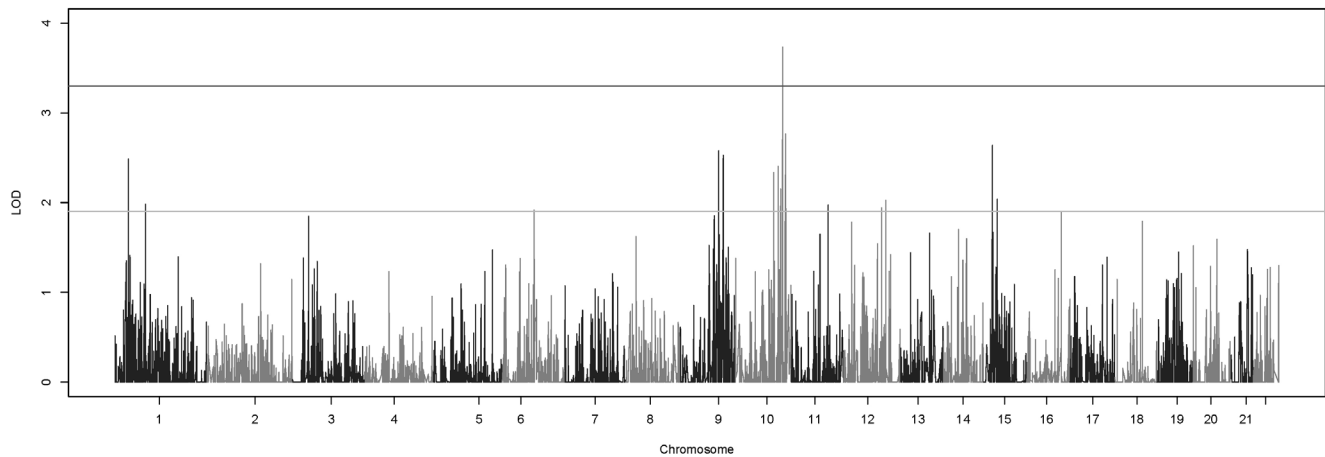


Figure 2. Genome-wide collapsed haplotype pattern variant two-point HLOD scores. Plot showing the genome-wide collapsed haplotype pattern (CHP) variant two-point heterogeneity logarithm of the odds (HLOD) scores produced by SEQLinkage and MERLIN. The lines at 1.9 and 3.3 represent the suggestive and significant thresholds, respectively, recommended by Lander and Kruglyak.

**TABLE 1. VARIANTS THAT EXHIBIT GENOME-WIDE SIGNIFICANT AND SUGGESTIVE EVIDENCE OF LINKAGE TO MYOPIA FROM THE COLLAPSED HAPLOTYPE PATTERN LINKAGE ANALYSIS.**

CHR	HLOD	ALPHA	LOD	POS	GENE
10q26.13	3.73	1.00	3.73	146.7658	<i>TACC2</i>
10q26.2	2.77	1.00	2.77	156.932	<i>DOCK1</i>
15q13.3	2.64	0.87	2.42	27.82131	<i>RYR3</i>
9q31.3	2.58	1.00	2.58	117.7108	<i>SVEP1</i>
9q33.2	2.53	1.00	2.53	131.6962	<i>CDK5RAP2</i>
1p13.1	2.49	1.00	2.49	40.94913	<i>IGSF21</i>
10q25.3	2.41	1.00	2.41	133.0324	<i>VWA2</i>
10q24.31	2.34	1.00	2.34	119.6265	<i>CWF19L1</i>
10q26.11	2.15	1.00	2.15	141.0359	<i>EIF3A</i>
10q24.2	2.08	1.00	2.08	119.2729	<i>ABCC2</i>
15q15.2	2.04	1.00	2.04	42.58454	<i>STARD9</i>
12q24.13	2.03	1.00	2.03	129.8192	<i>CCDC42B</i>
1p31.3	1.98	1.00	1.98	94.08912	<i>INADL</i>
11q22.3	1.97	1.00	1.97	113.0644	<i>EXPH5</i>
12q23.2	1.94	1.00	1.94	117.4295	<i>IGF1</i>
12q24.13	1.94	1.00	1.94	129.8897	<i>IQCD</i>
6q15	1.92	1.00	1.92	99.52311	<i>GABRR1</i>

Table displaying the genome-wide significant and suggestive linkage signals from the collapsed haplotype pattern (CHP) variant two-point linkage analysis sorted by heterogeneity LOD (HLOD). CHP variants are multi-allelic pseudo-markers corresponding to a gene and created from SNPs with a MAF <0.15. The genome-wide significance threshold is 3.3 and the genome-wide suggestive threshold is 1.9, as recommended by Lander and Kruglyak. CHR=chromosomal region, HLOD=heterogeneity LOD score of the marker, ALPHA=estimated proportion of informative families showing evidence of linkage for each HLOD, LOD=cumulative LOD across all families, POS=position in cM of the gene, GENE=Gene location of the CHP marker.

10q26.2 region (Figure 5). Of those 34 SNPs, 25 were located in an intron of *HTRA1* (Gene ID 5654, OMIM 602194; 10q26.13), a well-known susceptibility gene for age-related macular degeneration (AMD) [59-65]. The HLOD scores of these SNPs ranged from 2.52 to 1.9. All of the *HTRA1* SNPs were common (the MAF was about 0.37), and they appeared to be on the same linked haplotype in the intron. However, there are rarer SNPs interspersed along the *HTRA1* intron. Although the majority of the families individually showed evidence of linkage to the common SNPs, three families who showed linkage to the common SNPs also had stronger linkage to the rarer SNPs in this region. The top LOD score of these families was in *HTRA1* in SNPs with an MAF of 0.03, 0.05, and 0.12.

The 9q31.1 to 9q33.2 region contained the second highest number of suggestive SNPs with nine (Figure 6). The highest overall HLOD score (2.56) was located in an exon of the

*CDK5RAP2* (Gene ID 55755, OMIM 608201) gene. It is a fairly common SNP with an MAF of 0.18. It is nonsynonymous but not predicted to be damaging. Most interestingly, the linkage peak was driven primarily by a single family who had an LOD score of 1.33 (the highest overall LOD score in the family). A full list of all suggestive SNPs identified by the single-variant two-point analysis can be found in the Appendix 1.

We did not find any significant associations using the FBAT when we examined individual variants or collapsing rare variants. We found a nominally statistically significant ( $p=0.000533$ ) signal at rs1980493 within the *BTNL2* (Gene ID 56244, OMIM 606000) gene at 6p21.32 using rv-TDT. This gene has been found to be associated with sarcoidosis in Japanese patients [66]. Sarcoidosis causes granulomas to form in different tissues, including eye tissue where the condition can cause blurred vision.

TABLE 2. SNPs WITH GENOME-WIDE SUGGESTIVE EVIDENCE OF LINKAGE FROM MULTIPPOINT ANALYSIS.

CHR	HLOD	rsID	POS	MAF	FUNCTION	GENE
9q33.1	3.14	rs1158810	121,809,519	0.4631	intergenic	<i>TLR4, BRINP1</i>
9q33.1	3.07	rs1009470	122,289,187	0.4688	intergenic	<i>BRINP1, MIR147A</i>
9q33.1	2.97	rs16909449	122,861,297	0.4799	intergenic	<i>BRINP1, MIR147A</i>
9q33.2	2.78	rs17611	123,769,200	0.4927	exonic	<i>C5</i>
9q33.1	2.71	rs1927321	121,066,417	0.4765	intergenic	<i>TLR4, BRINP1</i>
9q33.2	2.67	rs7871736	124,650,611	0.4052	intronic	<i>TLL1</i>
9q33.3	2.48	rs4838334	128,843,105	0.4248	intergenic	<i>PBX3, LOC101929116</i>
10q26.11	2.35	rs7099523	120,493,700	0.4998	intronic	<i>CACUL1</i>
10q26.12	2.33	rs746832	122,348,065	0.4615	ncRNA_intronic	<i>MIR5694</i>
10q26.11	2.31	rs185020036	120,095,773	0.0061	exonic	<i>FAM204A</i>
9q33.2	2.30	rs1536929	125,391,369	0.4291	exonic	<i>OR1B1</i>
9q33.3	2.24	rs10987504	129,613,454	0.466	intergenic	<i>ZBTB43), ZBTB34</i>
10q26.11	2.24	rs1419138	119,731,281	0.4295	intergenic	<i>EMX2, RAB11FIP2</i>
10q26.12	2.22	rs2289337	122,649,482	0.1519	exonic	<i>WDR11</i>
10q26.11	2.14	rs1925283	119,533,345	0.4789	intergenic	<i>EMX2), RAB11FIP2</i>
10q26.12	2.10	rs4752384	121,871,353	0.3317	intergenic	<i>MIR4682, PPAPDC1A</i>
10q26.13	2.07	rs11598592	123,043,389	0.4611	intergenic	<i>MIR5694), FGFR2</i>
9q33.3	2.05	rs4130590	130,107,964	0.4424	intronic	<i>GARNL3</i>
9q33.1	2.02	rs10817896	119,232,655	0.4729	intronic	<i>ASTN2</i>
10q26.13	2.00	rs2672592	124,230,750	0.4857	intronic	<i>HTRA1</i>
9q34.11	1.93	rs2275260	131,285,955	0.4639	exonic	<i>GLE1</i>
10q26.13	1.89	rs2981579	123,337,335	0.4623	intronic	<i>FGFR2</i>

Table displaying the genome-wide significant and suggestive signals from the multipoint linkage analysis sorted by HLOD. The genome-wide significance threshold is 3.3 and the genome-wide suggestive threshold is 1.9, as recommended by Lander and Kruglyak. CHR=chromosomal region, HLOD=heterogeneity LOD score, POS=position in basepairs of each SNP, MAF=minor allele frequency as calculated from the data set, FUNCTION=functional annotation of the SNP, GENE=genic location of the SNP or closest genes in the case of intergenic SNPs. Annotations performed by ANNOVAR.

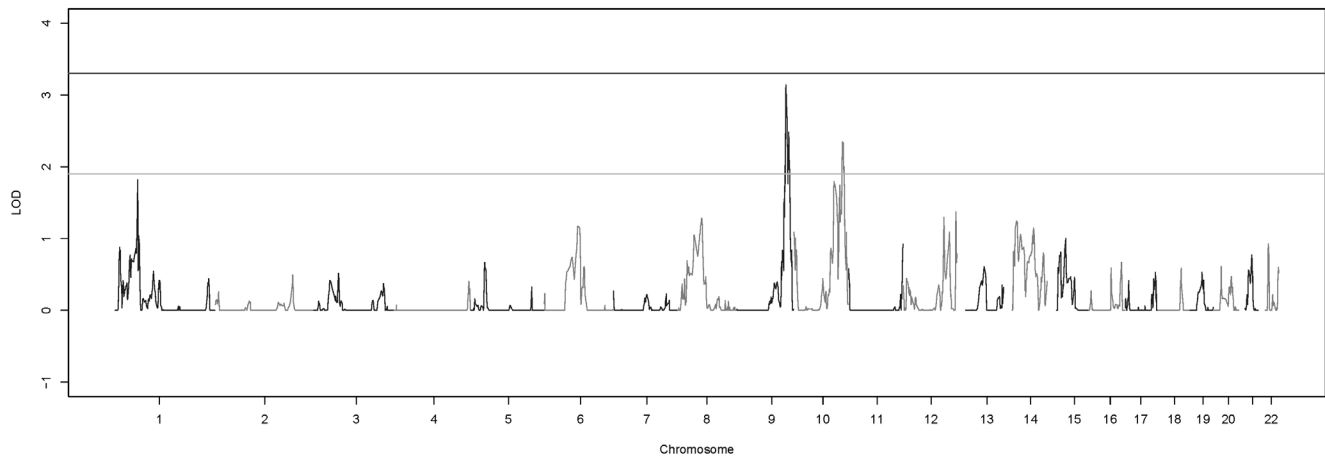


Figure 3. Genome-wide multipoint HLOD scores. Plot of the genome-wide multipoint heterogeneity logarithm of the odds (HLOD) scores produced by SimWalk2. The lines at 1.9 and 3.3 represent the suggestive and significant thresholds, respectively, recommended by Lander and Kruglyak.

## DISCUSSION

CHP two-point analysis identified genome-wide significant linkage at 10q26.13. Single-variant two-point analysis and multipoint analysis also identified highly suggestive linkage at that region. These additional suggestive signals in 10q26.13 are important, as they suggest that the significant signal is less likely to be a false positive. This signal appears to be novel with no prior reports of linkage to myopia in other populations. Terminal deletions of 10q26 have been shown to

cause myopia but also result in much more severe phenotypes, such as developmental disabilities, webbed neck, and musculoskeletal anomalies [67]. In addition, significant linkage has been published and replicated between myopia and a region much farther upstream at 10q21.11 (*MYP15* Gene ID 100294716, OMIM 612717) [68,69], although the causal gene at the *MYP15* locus remains unknown. All of this evidence taken together gives strong support for the location of a gene that is important in myopization in the 10q26 region.

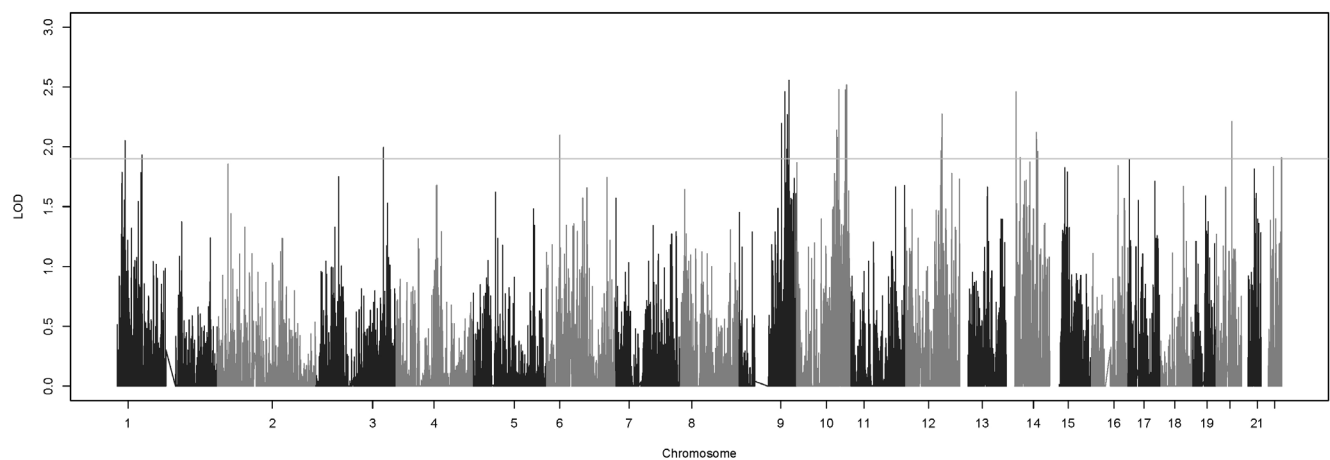


Figure 4. Genome-wide single-variant two-point HLOD scores. Plot showing the single-variant two-point heterogeneity logarithm of the odds (HLOD) scores produced by TwoPointLods. The lines at 1.9 and 3.3 represent the suggestive and significant thresholds, respectively, recommended by Lander and Kruglyak.

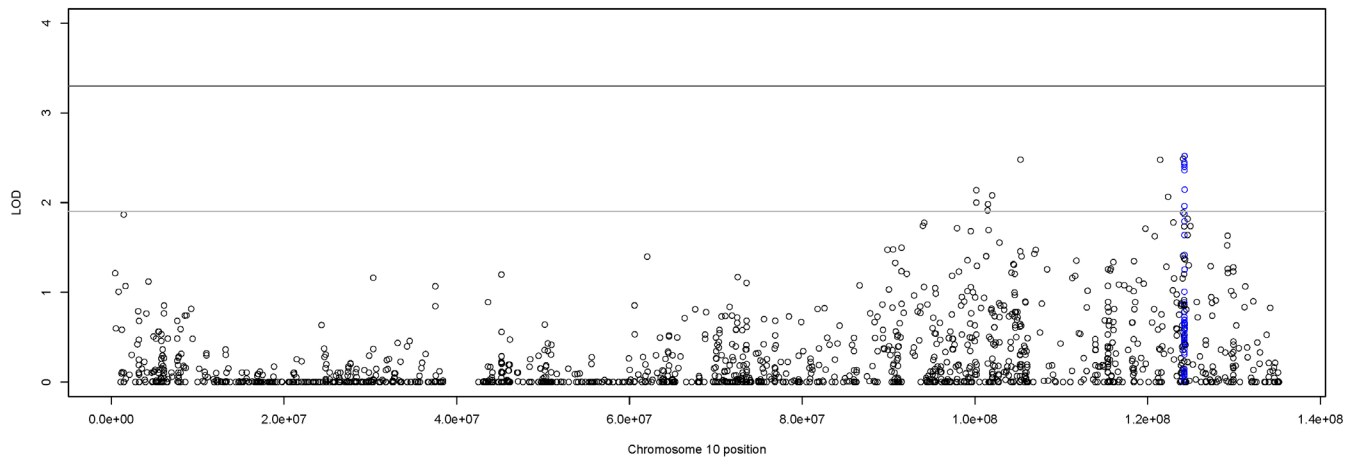


Figure 5. Chromosome 10 two-point single-variant LOD scores. Plot of the single-variant two-point heterogeneity logarithm of the odds (HLOD) scores produced by TwoPointLods for chromosome 10, which contains single nucleotide polymorphisms (SNPs) that exhibit suggestive evidence of linkage to myopia. The lines at 1.9 and 3.3 represent the respective suggestive and significant thresholds, respectively, recommended by Lander and Kruglyak. The SNPs in the age-related macular degeneration (AMD) risk locus *HTRA1* are marked in blue.

The significant signal in the CHP variant analysis localized at 10q26.13 at the transforming acidic coiled-coil containing protein 2 (*TACC2*) gene. *TACC2* interacts with the centrosome and microtubules and has been implicated in the proliferation of breast cancer [70]. This gene has not been previously implicated in myopia or any other eye diseases, although one alternatively spliced variant of *TACC2* was found to be expressed only in the retina and the brain in rats

[71]. However, just because *TACC2* has the highest HLOD score does not mean this is the causal gene. It is well-known that the maximum HLOD score is often not observed at the exact location of the causal gene. For this reason, it is standard practice to use the one LOD drop support interval when interpreting significant linkage results. The 95% support interval of the *TACC2* CHP HLOD score extends to approximately 20 cM on each side of the significant gene [72-75],

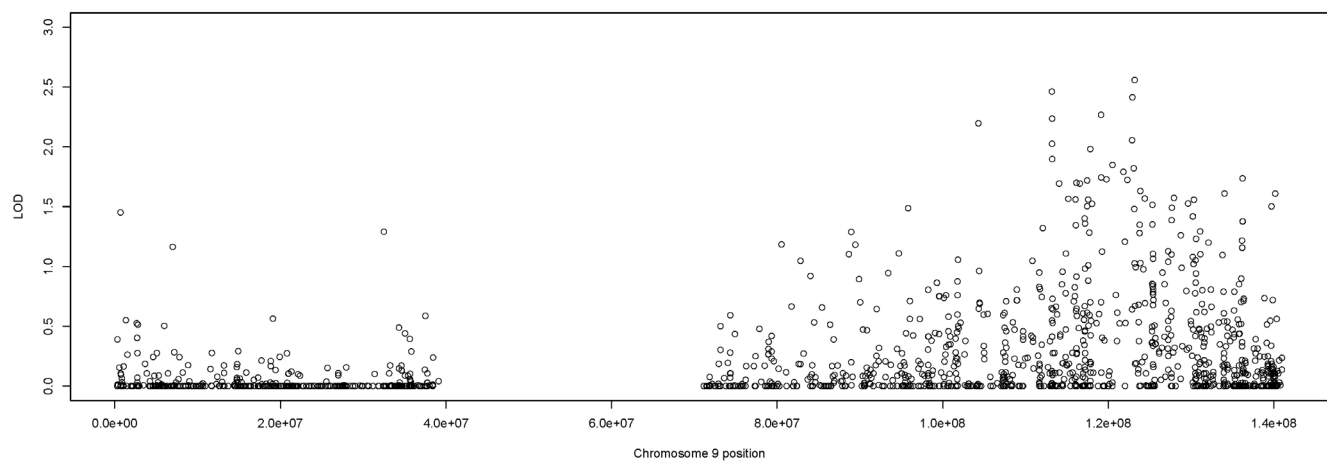


Figure 6. Chromosome 9 two-point single-variant LOD scores. Plot of the single-variant two-point heterogeneity logarithm of the odds (HLOD) scores produced by TwoPointLods for chromosome 9, which contains single nucleotide polymorphisms (SNPs) that exhibit suggestive evidence of linkage to myopia. The lines at 1.9 and 3.3 represent the respective suggestive and significant thresholds, respectively, recommended by Lander and Kruglyak.



encompassing the entire 10q26.13 region. Finally, we identified a large number of suggestive signals within 10q26.13; therefore, other genes in the region should be considered candidate genes.

The most biologically interesting candidate gene at 10q26.13 was *Htra serine peptidase 1 (HTRAI)*, a well-documented susceptibility locus for AMD in the general population [59-65] and in Chinese populations [65,76,77]. Despite the gene's known implications with AMD, *HTRAI* has not been previously reported to have any association with myopia [78,79]; this appears to be the first time the gene has been implicated in myopia. The present study identified *HTRAI* as suggestive for linkage in the single-variant two-point analysis and the multipoint analysis. CHP two-point analysis identified *HTRAI* as slightly below the genome-wide suggestive levels (HLOD = 1.6). Twenty-five SNPs in intronic regions of *HTRAI* were identified as suggestive in the single-variant two-point analysis. The variants could be part of a linked haplotype for myopia risk. The majority of the suggestive *HTRAI* SNPs are common, with an MAF of approximately 0.37, although in three of the individual families we observed linkage to rarer intronic SNPs (MAF ranging from 0.03 to 0.12). The fact that the overall linkage signal in *HTRAI* consisted of common SNPs explains why the magnitude of the *HTRAI* signal in the CHP analysis was lower; The CHP analysis was restricted to SNPs with an MAF <0.15. However, we caution that although *HTRAI* is an excellent candidate gene at this locus, this gene is only one of many candidate genes identified in this study, and we do not imply causality simply because of the gene's known function in AMD.

We also observed a suggestive signal at 9q33, which was identified as suggestive in all three analyses. This region has not been previously linked to myopia, although the nearby 9q34.11 region was identified as suggestive by nonparametric linkage methods that used high-grade myopia [80]. The present study seemed to replicate this finding. There are interesting candidate genes located at 9q33, including *TLR4*, which has previously been implicated in AMD and is expressed in the retina [81-83].

Many of the suggestive SNPs identified by the single-variant two-point analyses are common. Although this may seem to be somewhat counterintuitive for linkage studies that are designed to find rare, high penetrance variants, it actually makes sense within the context of this study. The exome-based array provided a good backbone of common variants located fairly evenly across the genome. This backbone was useful for tagging linked, haplotypes across the genome. It is possible that one of the common SNPs is causal

or that several of the common SNPs are causal with a low to moderate effect on disease risk. However, it is more likely that there is a rare variant on the segregating haplotype that is causal but was not on this limited exome-based chip (possibly different rare variants on different segregating haplotypes in different families; see Figure 1). Family-based studies can take advantage of long, linked haplotypes that have not been broken apart by many generations of meioses (as opposed to a population-based association study). This explains why the common variants identified here have not been identified in any association studies (including the family-based association studies we performed here); in the population, these common variants are not in linkage disequilibrium with a common risk variant. Follow-up studies involving targeted sequencing, which would sequence all variants (common/rare and coding/non-coding) in the 10q26 region, will give the opportunity to elucidate the true causal variant(s).

The two family-based association studies using these families found only one nominal significant result on chromosome 6 (not significant after adjusting for multiple testing). Again, this is not particularly surprising given the differences in what is detected with association and linkage-based tests (association tests have more power to detect common variants that tend to have small effects on risk, and linkage tests have more power to detect variants with larger effects on risk, which tend to be rare variants often detected on a common haplotype segregating with disease in a family). The rv-TDT was also likely underpowered, as only 34 trios were used (one selected from each extended family).

Myopia is a complex disease, with multiple genetic and environmental factors contributing to causation. Known locus heterogeneity within the disease makes identification of causal variants more difficult. There are various ways to approach complex traits. The approach in this study was to use exome-based genotypes from carefully ascertained multiplex families in linkage studies to find any high effect, autosomal dominant variants that have been predicted by segregation analysis and previous linkage studies in other populations. Association studies are another valid approach, and population-based case-control studies have been used successfully by us and others to detect common risk variants of small effect on myopia risk. It is important to remember that association studies and linkage studies should be used to complement, not supersede, each other, as they test different things. This is important with a trait like myopia, which is known to be affected by multiple genetic variants. Association studies are better at identifying common variants of moderate to low effect. They are also able to incorporate environmental covariates into the analyses, something that is

difficult to do in a linkage analysis. One hopes that as linkage studies regain popularity new methods to incorporate covariates will be developed. Linkage studies are better able to find rarer variants of larger effect and can use longer haplotypes to tag potential causal variants that were not genotyped. These haplotypes can allow us to zero in on a region of interest, such as we have done in this data set with 10q26.13. The fact that we were able to detect multiple suggestive linkage signals to this region in addition to the significant signal provides more confidence in the signal and less likelihood of a false positive (compared to the situation in which a single significant LOD score is observed surrounded by negative LOD scores at adjacent variant locations). Association and linkage analyses are important and necessary, and both approaches have distinct advantages and disadvantages when analyzing myopia.

In conclusion, we used a family-based approach to identify a novel significant linkage peak at 10q26.13 for myopia in Han Chinese families. Multiple suggestive signals in this region reinforce the likelihood of the region harboring a causal variant. There are excellent candidate genes in the region, such as *HTRAI*, a known AMD gene. Because these analyses were performed on an exome array and not DNA sequencing, it is possible that the true causal variants in each family that is linked to 10q26 were not genotyped on this chip. The linked common variants are likely tagging an ungenotyped causal (possibly rare) variant along the linked haplotype at 10q26.13 (see Figure 1 for an example of this phenomenon). However, it is important that these localizations, especially the significant linkage at 10q26.13, be published to inform all fellow researchers working on myopia of the candidate genes found in these linked regions. For the same reason, we also report some of the more suggestive linkage results, especially the suggestive signal at 9q33 identified by all three linkage analyses.

We plan to perform targeted sequencing of the families that show linkage at 10q26.13. Targeted sequencing will provide genotypes for all genomic variants in the region and give us the ability to find any rare causal variants that were tagged by the common variants along the linked haplotypes in this study, which, in turn, will increase our understanding of myopia genetics.

#### APPENDIX 1. GENOME-WIDE SUGGESTIVE VARIANTS FROM SINGLE VARIANT TWO-POINT ANALYSIS

Table displaying the genome-wide significant and suggestive linkage signals from the single variant two-point linkage analysis sorted by HLOD. The genome-wide significance threshold is 3.3 and the genome-wide suggestive

threshold is 1.9, as recommended by Lander and Kruglyak. CHR=chromosomal region, HLOD=heterogeneity LOD score, POS=position in basepairs of the SNP, MAF=minor allele frequency as calculated from the data set, FUNCTION=functional annotation of the SNP, GENE=genetic location of the SNP or closest genes in the case of intergenic SNPs. Annotations performed by ANNOVAR. To access the data, click or select the words “[Appendix 1](#)”

#### ACKNOWLEDGMENTS

The authors would like to thank all study participants and their families. This work was funded in part by the National Eye Institute Grant R01EY020483 and the Intramural Research Program of the National Human Genome Research Institute, National Institutes of Health. The authors have no competing, commercial, or conflicts of interest to disclose.

#### REFERENCES

1. Stambolian D. Genetic susceptibility and mechanisms for refractive error. *Clin Genet* 2013; 84:102-8. [PMID: 23647423].
2. Wojciechowski R, Hysi PG. Focusing in on the complex genetics of myopia. *PLoS Genet* 2013; 9:e1003442-[PMID: 23593034].
3. Simpson CL, Wojciechowski R, Oexle K, Murgia F, Portas L, Li X, Verhoeven VJ, Vitart V, Schache M, Hosseini SM, Hysi PG, Raffel LJ, Cotch MF, Chew E, Klein BE, Klein R, Wong TY, van Duijn CM, Mitchell P, Saw SM, Fossarello M, Wang JJ, Group DER, Polasek O, Campbell H, Rudan I, Oostra BA, Uitterlinden AG, Hofman A, Rivadeneira F, Amin N, Karssen LC, Vingerling JR, Doring A, Bettecken T, Bencic G, Gieger C, Wichmann HE, Wilson JF, Venturini C, Fleck B, Cumberland PM, Rahi JS, Hammond CJ, Hayward C, Wright AF, Paterson AD, Baird PN, Klaver CC, Rotter JI, Pirastu M, Meitinger T, Bailey-Wilson JE, Stambolian D. Genome-wide meta-analysis of myopia and hyperopia provides evidence for replication of 11 loci. *PLoS One* 2014; 9:e107110-[PMID: 25233373].
4. Klein AP, Duggal P, Lee KE, Cheng CY, Klein R, Bailey-Wilson JE, Klein BE. Linkage analysis of quantitative refraction and refractive errors in the Beaver Dam Eye Study. *Invest Ophthalmol Vis Sci* 2011; 52:5220-5. [PMID: 21571680].
5. Verhoeven VJ, Hysi PG, Saw SM, Vitart V, Mirshahi A, Guggenheim JA, Cotch MF, Yamashiro K, Baird PN, Mackey DA, Wojciechowski R, Ikram MK, Hewitt AW, Duggal P, Janmahasatian S, Khor CC, Fan Q, Zhou X, Young TL, Tai ES, Goh LK, Li YJ, Aung T, Vithana E, Teo YY, Tay W, Sim X, Rudan I, Hayward C, Wright AF. Large scale international replication and meta-analysis study confirms association of the 15q14 locus with myopia. The CREAM consortium. *Hum Genet* 2012; 131:1467-80. [PMID: 22665138].

6. Stambolian D, Wojciechowski R, Oexle K, Pirastu M, Li X, Raffel LJ, Cotch MF, Chew EY, Klein B, Klein R, Wong TY, Simpson CL, Klaver CC, van Duijn CM, Verhoeven VJ, Baird PN, Vitart V, Paterson AD, Mitchell P, Saw SM, Fossarello M, Kazmierkiewicz K, Murgia F, Portas L, Schache M, Richardson A, Xie J, Wang JJ, Rochtchina E, Group DER, Viswanathan AC, Hayward C, Wright AF, Polasek O, Campbell H, Rudan I, Oostra BA, Uitterlinden AG, Hofman A, Rivadeneira F, Amin N, Karssen LC, Vingerling JR, Hosseini SM, Doring A, Bettecken T, Vataavuk Z, Gieger C, Wichmann HE, Wilson JF, Fleck B, Foster PJ, Topouzis F, McGuffin P, Sim X, Inouye M, Holliday EG, Attia J, Scott RJ, Rotter JI, Meitinger T, Bailey-Wilson JE. Meta-analysis of genome-wide association studies in five cohorts reveals common variants in RBFOX1, a regulator of tissue-specific splicing, associated with refractive error. *Hum Mol Genet* 2013; 22:2754-64. [PMID: 23474815].
7. Kiefer AK, Tung JY, Do CB, Hinds DA, Mountain JL, Francke U, Eriksson N. Genome-wide analysis points to roles for extracellular matrix remodeling, the visual cycle, and neuronal development in myopia. *PLoS Genet* 2013; 9:e1003299-[PMID: 23468642].
8. Simpson CL, Wojciechowski R, Yee SS, Soni P, Bailey-Wilson JE, Stambolian D. Regional replication of association with refractive error on 15q14 and 15q25 in the Age-Related Eye Disease Study cohort. *Mol Vis* 2013; 19:2173-86. [PMID: 24227913].
9. Nussbaum RL, McInnes RR, Willard HF. *Thompson & Thompson Genetics in Medicine*. 8th ed. Philadelphia: Elsevier; 2016.
10. Elston RC. Linkage and association. *Genet Epidemiol* 1998; 15:565-76. [PMID: 9811419].
11. Ott J, Wang J, Leal SM. Genetic linkage analysis in the age of whole-genome sequencing. *Nat Rev Genet* 2015; 16:275-84. [PMID: 25824869].
12. Hall JM, Lee MK, Newman B, Morrow JE, Anderson LA, Huey B, King MC. Linkage of early-onset familial breast cancer to chromosome 17q21. *Science* 1990; 250:1684-9. [PMID: 2270482].
13. Hall JM, Friedman L, Guenther C, Lee MK, Weber JL, Black DM, King MC. Closing in on a breast cancer gene on chromosome 17q. *Am J Hum Genet* 1992; 50:1235-42. [PMID: 1598904].
14. Michailidou K, Lindstrom S, Dennis J, Beesley J, Hui S, Kar S, Lemacon A, Soucy P, Glubb D, Rostamianfar A, Bolla MK, Wang Q, Tyrer J, Dicks E, Lee A, Wang Z, Allen J, Keeman R, Eilber U, French JD, Qing Chen X, Fachal L, McCue K, McCart Reed AE. Association analysis identifies 65 new breast cancer risk loci. *Nature* 2017; 551:92-4. [PMID: 29059683].
15. Milne RL, Kuchenbaecker KB, Michailidou K, Beesley J, Kar S, Lindstrom S, Hui S, Lemacon A, Soucy P, Dennis J, Jiang X, Rostamianfar A, Finucane H, Bolla MK, McGuffog L, Wang Q, Aalfs CM, Investigators A, Adams M, Adlard J, Agata S. Identification of ten variants associated with risk of estrogen-receptor-negative breast cancer. *Nat Genet* 2017; 49:1767-78. [PMID: 29058716].
16. Caulfield M, Lavender P, Farrall M, Munroe P, Lawson M, Turner P, Clark AJ. Linkage of the angiotensinogen gene to essential hypertension. *N Engl J Med* 1994; 330:1629-33. [PMID: 8177268].
17. Liang J, Le TH, Edwards DRV, Tayo BO, Gaulton KJ, Smith JA, Lu Y, Jensen RA, Chen G, Yanek LR, Schwander K, Tajuddin SM, Sofer T, Kim W, Kayima J, McKenzie CA. Single-trait and multi-trait genome-wide association analyses identify novel loci for blood pressure in African-ancestry populations. *PLoS Genet* 2017; 13:e1006728-[PMID: 28498854].
18. Iyengar SK, Klein BE, Klein R, Jun G, Schick JH, Millard C, Liptak R, Russo K, Lee KE, Elston RC. Identification of a major locus for age-related cortical cataract on chromosome 6p12-q12 in the Beaver Dam Eye Study. *Proc Natl Acad Sci USA* 2004; 101:14485-90. [PMID: 15452352].
19. Ma X, Jiao X, Ma Z, Hejtmancik JF. Polymorphism rs7278468 is associated with Age-related cataract through decreasing transcriptional activity of the CRYAA promoter. *Sci Rep* 2016; 6:23206-[PMID: 26984531].
20. Guo H, Tong P, Liu Y, Xia L, Wang T, Tian Q, Li Y, Hu Y, Zheng Y, Jin X, Li Y, Xiong W, Tang B, Feng Y, Li J, Pan Q, Hu Z, Xia K. Mutations of P4HA2 encoding prolyl 4-hydroxylase 2 are associated with nonsyndromic high myopia. *Genet Med* 2015; 17:300-6. .
21. Zhou L, Li T, Song X, Li Y, Li H, Dan H. NYX mutations in four families with high myopia with or without CSNB1. *Mol Vis* 2015; 21:213-23. [PMID: 25802485].
22. Stambolian D, Ibay G, Reider L, Dana D, Moy C, Schlifka M, Holmes T, Ciner E, Bailey-Wilson JE. Genomewide linkage scan for myopia susceptibility loci among Ashkenazi Jewish families shows evidence of linkage on chromosome 22q12. *Am J Hum Genet* 2004; 75:448-59. [PMID: 15273935].
23. Wojciechowski R, Moy C, Ciner E, Ibay G, Reider L, Bailey-Wilson JE, Stambolian D. Genomewide scan in Ashkenazi Jewish families demonstrates evidence of linkage of ocular refraction to a QTL on chromosome 1p36. *Hum Genet* 2006; 119:389-99. [PMID: 16501916].
24. Andrew T, Maniatis N, Carbonaro F, Liew SH, Lau W, Spector TD, Hammond CJ. Identification and replication of three novel myopia common susceptibility gene loci on chromosome 3q26 using linkage and linkage disequilibrium mapping. *PLoS Genet* 2008; 4:e1000220-[PMID: 18846214].
25. Ibay G, Doan B, Reider L, Dana D, Schlifka M, Hu H, Holmes T, O'Neill J, Owens R, Ciner E, Bailey-Wilson JE, Stambolian D. Candidate high myopia loci on chromosomes 18p and 12q do not play a major role in susceptibility to common myopia. *BMC Med Genet* 2004; 5:20-[PMID: 15291966].
26. 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; 32:454-63. [PMID: 18293391].



27. Musolf AM, Simpson CL, Moiz BA, Long KA, Portas L, Murgia F, Ciner EB, Stambolian D, Bailey-Wilson JE. Caucasian Families Exhibit Significant Linkage of Myopia to Chromosome 11p. *Invest Ophthalmol Vis Sci* 2017; 58:3547-54. [PMID: 28715588].
28. Naiglin L, Gazagne C, Dallongeville F, Thalamas C, Idder A, Rascol O, Malecaze F, Calvas P. A genome wide scan for familial high myopia suggests a novel locus on chromosome 7q36. *J Med Genet* 2002; 39:118-24. [PMID: 11836361].
29. Stambolian D, Ciner EB, Reider LC, Moy C, Dana D, Owens R, Schlifka M, Holmes T, Ibay G, Bailey-Wilson JE. Genome-wide scan for myopia in the Old Order Amish. *Am J Ophthalmol* 2005; 140:469-76. [PMID: 16084785].
30. Naiglin L, Clayton J, Gazagne C, Dallongeville F, Malecaze F, Calvas P. Familial high myopia: evidence of an autosomal dominant mode of inheritance and genetic heterogeneity. *Ann Genet* 1999; 42:140-6. [PMID: 10526656].
31. Rong SS, Chen LJ, Pang CP. Myopia Genetics-The Asia-Pacific Perspective. *Asia-Pac J Ophthalmol* 2016; 5:236-44. [PMID: 27488065].
32. Jiao X, Wang P, Li S, Li A, Guo X, Zhang Q, Hejtmancik JF. Association of markers at chromosome 15q14 in Chinese patients with moderate to high myopia. *Mol Vis* 2012; 18:2633-46. [PMID: 23170057].
33. Yu Z, Zhou J, Chen X, Zhou X, Sun X, Chu R. Polymorphisms in the CTNND2 gene and 11q24.1 genomic region are associated with pathological myopia in a Chinese population. *Ophthalmologica Journal international d'ophtalmologie International journal of ophthalmology Z Augenheilkd* 2012; 228:123-9. .
34. Ho DW, Yap MK, Ng PW, Fung WY, Yip SP. Association of high myopia with crystallin beta A4 (CRYBA4) gene polymorphisms in the linkage-identified MYP6 locus. *PLoS One* 2012; 7:e40238-[PMID: 22792142].
35. Guo X, Xiao X, Li S, Wang P, Jia X, Zhang Q. Nonsyndromic high myopia in a Chinese family mapped to MYP1: linkage confirmation and phenotypic characterization. *Arch Ophthalmol* 2010; 128:1473-9. [PMID: 21060050].
36. Ma JH, Shen SH, Zhang GW, Zhao DS, Xu C, Pan CM, Jiang H, Wang ZQ, Song HD. Identification of a locus for autosomal dominant high myopia on chromosome 5p13.3-p15.1 in a Chinese family. *Mol Vis* 2010; 16:2043-54. [PMID: 21042559].
37. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ, Sham PC. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 2007; 81:559-75. [PMID: 17701901].
38. Sun L. Detecting pedigree relationship errors. *Methods Mol Biol* 2012; 850:25-46. [PMID: 22307692].
39. Matise TC, Chen F, Chen W, De La Vega FM, Hansen M, He C, Hyland FC, Kennedy GC, Kong X, Murray SS, Ziegler JS, Stewart WC, Buyske S. A second-generation combined linkage physical map of the human genome. *Genome Res* 2007; 17:1783-6. [PMID: 17989245].
40. Sobel E, Lange K. Descent graphs in pedigree analysis: applications to haplotyping, location scores, and marker-sharing statistics. *Am J Hum Genet* 1996; 58:1323-37. [PMID: 8651310].
41. Sobel E, Papp JC, Lange K. Detection and integration of genotyping errors in statistical genetics. *Am J Hum Genet* 2002; 70:496-508. [PMID: 11791215].
42. Sobel E, Sengul H, Weeks DE. Multipoint estimation of identity-by-descent probabilities at arbitrary positions among marker loci on general pedigrees. *Hum Hered* 2001; 52:121-31. [PMID: 11588394].
43. Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* 2005; 21:263-5. [PMID: 15297300].
44. Wang GT, Zhang D, Li B, Dai H, Leal SM. Collapsed haplotype pattern method for linkage analysis of next-generation sequence data. *European journal of human genetics Eur J Hum Genet* 2015; 23:1739-43. [PMID: 25873013].
45. Abecasis GR, Cherny SS, Cookson WO, Cardon LR. Merlin—rapid analysis of dense genetic maps using sparse gene flow trees. *Nat Genet* 2002; 30:97-101. [PMID: 11731797].
46. Chang X, Wang K. wANNOVAR: annotating genetic variants for personal genomes via the web. *J Med Genet* 2012; 49:433-6. [PMID: 22717648].
47. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res* 2010; 38:e164-[PMID: 20601685].
48. Zhao Y, Zhao F, Zong L, Zhang P, Guan L, Zhang J, Wang D, Wang J, Chai W, Lan L, Li Q, Han B, Yang L, Jin X, Yang W, Hu X, Wang X, Li N, Li Y, Petit C, Wang J, Wang HY, Wang Q. Exome sequencing and linkage analysis identified tenascin-C (TNC) as a novel causative gene in nonsyndromic hearing loss. *PLoS One* 2013; 8:e69549-[PMID: 23936043].
49. Kumar P, Henikoff S, Ng PC. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nat Protoc* 2009; 4:1073-81. [PMID: 19561590].
50. Ng PC, Henikoff S. SIFT: Predicting amino acid changes that affect protein function. *Nucleic Acids Res* 2003; 31:3812-4. [PMID: 12824425].
51. Sim NL, Kumar P, Hu J, Henikoff S, Schneider G, Ng PC. SIFT web server: predicting effects of amino acid substitutions on proteins. *Nucleic Acids Res* 2012; 40:W452-7-.
52. Vaser R, Adusumalli S, Leng SN, Sikic M, Ng PC. SIFT missense predictions for genomes. *Nat Protoc* 2016; 11:1-9. [PMID: 26633127].
53. Landrum MJ, Lee JM, Benson M, Brown G, Chao C, Chitipiralla S, Gu B, Hart J, Hoffman D, Hoover J, Jang W, Katz K, Ovetsky M, Riley G, Sethi A, Tully R, Villamarin-Salomon R, Rubinstein W, Maglott DR. ClinVar: public archive of interpretations of clinically relevant variants. *Nucleic Acids Res* 2016; 44:D1D862-8. [PMID: 26582918].

54. Adzhubei I, Jordan DM, Sunyaev SR. Predicting functional effect of human missense mutations using PolyPhen-2. *Current protocols in human genetics / editorial board, Jonathan L Haines [et al] 2013; Chapter 7:Unit7 20.*
55. De G, Yip WK, Ionita-Laza I, Laird N. Rare variant analysis for family-based design. *PLoS One* 2013; 8:e48495-[\[PMID: 23341868\]](#).
56. Xu X, Rakovski C, Xu X, Laird N. An efficient family-based association test using multiple markers. *Genet Epidemiol* 2006; 30:620-6. [\[PMID: 16868964\]](#).
57. He Z, O'Roak BJ, Smith JD, Wang G, Hooker S, Santos-Cortez RL, Li B, Kan M, Krumm N, Nickerson DA, Shendure J, Eichler EE, Leal SM. Rare-variant extensions of the transmission disequilibrium test: application to autism exome sequence data. *Am J Hum Genet* 2014; 94:33-46. [\[PMID: 24360806\]](#).
58. Lander E, Kruglyak L. Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. *Nat Genet* 1995; 11:241-7. [\[PMID: 7581446\]](#).
59. Yang Z, Camp NJ, Sun H, Tong Z, Gibbs D, Cameron DJ, Chen H, Zhao Y, Pearson E, Li X, Chien J, Dewan A, Harmon J, Bernstein PS, Shridhar V, Zabriskie NA, Hoh J, Howes K, Zhang K. A variant of the HTRA1 gene increases susceptibility to age-related macular degeneration. *Science* 2006; 314:992-3. [\[PMID: 17053109\]](#).
60. Cameron DJ, Yang Z, Tong Z, Zhao Y, Praggastis A, Brinton E, Harmon J, Chen Y, Pearson E, Bernstein PS, Brinton G, Li X, Jorgensen A, Schneider S, Gibbs D, Chen H, Wang C, Howes K, Camp NJ, Zhang K. 10q26 is associated with increased risk of age-related macular degeneration in the Utah population. *Adv Exp Med Biol* 2008; 613:253-8. [\[PMID: 18188952\]](#).
61. Gibbs D, Yang Z, Constantine R, Ma X, Camp NJ, Yang X, Chen H, Jorgensen A, Hau V, Dewan A, Zeng J, Harmon J, Buehler J, Brand JM, Hoh J, Cameron DJ, Dixit M, Tong Z, Zhang K. Further mapping of 10q26 supports strong association of HTRA1 polymorphisms with age-related macular degeneration. *Vision Res* 2008; 48:685-9. [\[PMID: 18207215\]](#).
62. Deangelis MM, Ji F, Adams S, Morrison MA, Harring AJ, Sweeney MO, Capone A Jr, Miller JW, Dryja TP, Ott J, Kim IK. Alleles in the HtrA serine peptidase 1 gene alter the risk of neovascular age-related macular degeneration. *Ophthalmology* 2008; 115:1209-15. .
63. Haines JL, Hauser MA, Schmidt S, Scott WK, Olson LM, Gallins P, Spencer KL, Kwan SY, Nouredine M, Gilbert JR, Schnetz-Boutaud N, Agarwal A, Postel EA, Pericak-Vance MA. Complement factor H variant increases the risk of age-related macular degeneration. *Science* 2005; 308:419-21. [\[PMID: 15761120\]](#).
64. Klein RJ, Zeiss C, Chew EY, Tsai JY, Sackler RS, Haynes C, Henning AK, SanGiovanni JP, Mane SM, Mayne ST, Bracken MB, Ferris FL, Ott J, Barnstable C, Hoh J. Complement factor H polymorphism in age-related macular degeneration. *Science* 2005; 308:385-9. [\[PMID: 15761122\]](#).
65. Xu Y, Guan N, Xu J, Yang X, Ma K, Zhou H, Zhang F, Sneltingen T, Jiao Y, Liu X, Wang N, Liu N. Association of CFH, LOC387715, and HTRA1 polymorphisms with exudative age-related macular degeneration in a northern Chinese population. *Mol Vis* 2008; 14:1373-81. [\[PMID: 18682812\]](#).
66. Suzuki H, Ota M, Meguro A, Katsuyama Y, Kawagoe T, Ishihara M, Asukata Y, Takeuchi M, Ito N, Shibuya E, Nomura E, Uemoto R, Nishide T, Namba K, Kitaichi N, Morimoto S, Kaburaki T, Ando Y, Takenaka S, Nakamura J, Saeki K, Ohno S, Inoko H, Mizuki N. Genetic characterization and susceptibility for sarcoidosis in Japanese patients: risk factors of BTNL2 gene polymorphisms and HLA class II alleles. *Invest Ophthalmol Vis Sci* 2012; 53:7109-15. [\[PMID: 22991420\]](#).
67. Lukusa T, Smeets E, Vermeesch JR, Fryns JP. Small terminal 10q26 deletion in a male patient with Noonan-like stigmata: diagnosis by cytogenetic and FISH analysis. *Genet Couns* 2002; 13:417-25. [\[PMID: 12558112\]](#).
68. Meng W, Butterworth J, Bradley DT, Hughes AE, Soler V, Calvas P, Malecaze F. A genome-wide association study provides evidence for association of chromosome 8p23 (MYP10) and 10q21.1 (MYP15) with high myopia in the French Population. *Invest Ophthalmol Vis Sci* 2012; 53:7983-8. [\[PMID: 23049088\]](#).
69. Nallasamy S, Paluru PC, Devoto M, Wasserman NF, Zhou J, Young TL. Genetic linkage study of high-grade myopia in a Hutterite population from South Dakota. *Mol Vis* 2007; 13:229-36. [\[PMID: 17327828\]](#).
70. Onodera Y, Takagi K, Miki Y, Takayama KI, Shibahara Y, Watanabe M, Ishida T, Inoue S, Sasano H, Suzuki T. TACC2 (transforming acidic coiled-coil protein 2) in breast carcinoma as a potent prognostic predictor associated with cell proliferation. *Cancer Med* 2016; 5:1973-82. [\[PMID: 27333920\]](#).
71. Padhi BK, Zigler JS Jr, Padhi P, Hose S, Sinha D. Expression pattern of an evolutionarily conserved splice variant in the rat Tacc2 gene. *Genesis* 2014; 52:378-86. [\[PMID: 24700488\]](#).
72. Ott J. *Analysis of Human Genetic Linkage*. 3rd ed. Baltimore: Johns Hopkins University Press; 1999.
73. Khoury MJ, Beaty TH, Cohen BH. *Fundamentals of Genetic Epidemiology*. 1st ed. New York: Oxford University Press; 1993.
74. Conneally PM, Edwards JH, Kidd KK, Lalouel J-M, Morton NE, Ott J, White R. Report on the committee on methods of linkage analysis and reporting. *Cytogenet Cell Genet* 1985; 40:356-9. [\[PMID: 3864600\]](#).
75. Edwards AWF. *Likelihood*. Baltimore: Johns Hopkins University Press; 1992.
76. Tian J, Yu W, Qin X, Fang K, Chen Q, Hou J, Li J, Chen D, Hu Y, Li X. Association of genetic polymorphisms and age-related macular degeneration in Chinese population. *Invest Ophthalmol Vis Sci* 2012; 53:4262-9. [\[PMID: 22618592\]](#).
77. Zhuang W, Li H, Liu Y, Zhao J, Ha S, Xiang W, Bai X, Li Z, Han Y, Sheng X. Association of specific genetic



- polymorphisms with age-related macular degeneration in a northern Chinese population. *Ophthalmic Genet* 2014; 35:156-61. [PMID: 24865190].
78. Woo SJ, Ahn J, Morrison MA, Ahn SY, Lee J, Kim KW, DeAngelis MM, Park KH. Analysis of Genetic and Environmental Risk Factors and Their Interactions in Korean Patients with Age-Related Macular Degeneration. *PLoS One* 2015; 10:e0132771-[PMID: 26171855].
79. Nakanishi H, Gotoh N, Yamada R, Yamashiro K, Otani A, Hayashi H, Tsujikawa A, Shimada N, Ohno-Matsui K, Mochizuki M, Saito M, Saito K, Iida T, Matsuda F, Yoshimura N. ARMS2/HTRA1 and CFH polymorphisms are not associated with choroidal neovascularization in highly myopic eyes of the elderly Japanese population. *Eye (Lond)* 2010; 24:1078-84. [PMID: 19680273].
80. Li YJ, Guggenheim JA, Bulusu A, Metlapally R, Abbott D, Malecaze F, Calvas P, Rosenberg T, Paget S, Creer RC, Kirov G, Owen MJ, Zhao B, White T, Mackey DA, Young TL. An international collaborative family-based whole-genome linkage scan for high-grade myopia. *Invest Ophthalmol Vis Sci* 2009; 50:3116-27. [PMID: 19324860].
81. Tu Z, Portillo JA, Howell S, Bu H, Subauste CS, Al-Ubaidi MR, Pearlman E, Lin F. Photoreceptor cells constitutively express functional TLR4. *J Neuroimmunol* 2011; 230:183-7. [PMID: 20801528].
82. Saraswathy S, Nguyen AM, Rao NA. The role of TLR4 in photoreceptor {alpha}a crystallin upregulation during early experimental autoimmune uveitis. *Invest Ophthalmol Vis Sci* 2010; 51:3680-6. [PMID: 20207969].
83. Zarepari S, Buraczynska M, Branham KE, Shah S, Eng D, Li M, Pawar H, Yashar BM, Moroi SE, Lichter PR, Petty HR, Richards JE, Abecasis GR, Elner VM, Swaroop A. Toll-like receptor 4 variant D299G is associated with susceptibility to age-related macular degeneration. *Hum Mol Genet* 2005; 14:1449-55. [PMID: 15829498].

Articles are provided courtesy of Emory University and the Zhongshan Ophthalmic Center, Sun Yat-sen University, P.R. China. The print version of this article was created on 14 January 2018. This reflects all typographical corrections and errata to the article through that date. Details of any changes may be found in the online version of the article.